



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΜΗΧΑΝΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ

Τομέας Ρευστών

Εργαστήριο Θερμικών Στροβιλομηχανών

Μονάδα Παράλληλης Υπολογιστικής Ρευστοδυναμικής και Βελτιστοποίησης

**ΕΠΙΛΥΣΗ
ΠΡΟΒΛΗΜΑΤΩΝ
ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ ΔΥΟ
ΣΤΟΧΩΝ ΜΕ ΒΑΣΗ ΤΗ
ΘΕΩΡΙΑ ΤΗΣ
ΣΥΝΕΧΙΣΗΣ ΤΟΥ
ΜΕΤΩΠΟΥ PARETO**

Διπλωματική εργασία
ΜΑΡΙΑΝΝΑΣ ΚΑΡΟΥΖΟΥ

Επιβλέπων: Κ. Χ. Γιαννάκογλου, Καθηγητής ΕΜΠ



Αθήνα, Σεπτέμβριος 2015

Ευχαριστίες

Κατά πρώτο λόγο, θα ήθελα να ευχαριστήσω μέσα από την καρδιά μου τη μητέρα μου, Σοφία, η οποία στέκεται δίπλα μου σαν ακλόνητος βράχος σε κάθε βήμα που κάνω. Κατά δεύτερο λόγο, θέλω να εκφράσω την ευγνωμοσύνη μου για το γεγονός ότι ήρθε στη ζωή μου το σκυλάκι μου, ο Teddy, καθώς σε κάθε δυσκολία που πέρασα ήταν η χαρά και η δύναμή μου και μου υπενθυμίζει καθημερινά τις έννοιες της αγνότητας και της αθωότητας. Κατά τρίτο λόγο, απευθύνω ένα μεγάλο ευχαριστώ στον ποδοσφαιρικό σύλλογο της Barcelona, η οποία δια χειρός (ή μάλλον ποδός) Lionel Messi κατέκτησε το treble για τη σεζόν 2014-15, διότι είναι περισσότερο από μία ομάδα (Més que un club) και πάντα με βοηθούσε να σταθώ στα πόδια μου όταν το είχα περισσότερο ανάγκη.

Επιπλέον, θα ήθελα να ευχαριστήσω θερμά τον καθηγητή μου κ. Κυριάκο Γιαννάκογλου για τις πολύτιμες γνώσεις και συμβουλές που μου προσέφερε κατά τη διάρκεια της παρούσας εργασίας αλλά και καθ' όλη τη φοίτηση μου στη Σχολή Μηχανολόγων Μηχανικών. Όλα αυτά τα χρόνια η πόρτα του ήταν πάντα ανοικτή για να θέσει τις γνώσεις, τη βοήθεια και το χρόνο του στη διάθεση των φοιτητών. Θυμάμαι, χαρακτηριστικά, ένα περιστατικό που συνέβη όταν βρισκόμουν στο 7^ο εξάμηνο της φοίτησής μου. Ο κ. Γιαννάκογλου μας είχε δώσει μία προαιρετική εργασία στο πλαίσιο του μαθήματος 'Μέθοδοι Αεροδυναμικής Βελτιστοποίησης' και είχε προκύψει ένα πρόβλημα στο πρόγραμμά μου το οποίο δεν μπορούσα να εντοπίσω. Ήταν Παρασκευή και μετά από πολλαπλές φορές που επισκέφτηκα το γραφείο του μέσα σε εκείνη την εβδομάδα φαινόταν να είχαμε βρει τη λύση αλλά λόγω του αγχωτικού μου χαρακτήρα φοβόμουν ότι κάτι θα προέκυπτε και επειδή ακολουθούσε Σαββατοκύριακο δεν θα μπορούσα να έρθω σε επαφή μαζί του για να λύσω την απορία και άρα δεν θα προλάβαινα να τελειώσω την εργασία. Η απάντηση που μου έδωσε ήταν αφοπλιστική: 'εμείς (εννοώντας ολόκληρη την ερευνητική ομάδα) είμαστε εδώ 365 ημέρες το χρόνο'. Δεν υπερέβαλε, όποτε και αν χρειάστηκα τη βοήθειά του, ακόμα και τα Χριστούγεννα ή το Πάσχα, μου την παρείχε απλόχερα.

Τέλος, θέλω να εκφράσω τις θερμές μου ευχαριστίες σε όλους τους ερευνητές της Μονάδας Παράλληλης Υπολογιστικής Ρευστοδυναμικής & Βελτιστοποίησης (ΜΠΥΡΒ/ΕΜΠ) για την προθυμία που έδειχναν να με βοηθήσουν σε κάθε πρόβλημα που μου παρουσιαζόταν κατά τη διάρκεια της φοίτησής μου. Ξεχωριστή αναφορά αξίζει η Διδάκτορας Βαρβάρα Ασούτη η οποία σε όλη την πορεία μου στη Σχολή Μηχανολόγων Μηχανικών διέθεσε πολλαπλές ώρες από το χρόνο της για να με βοηθήσει σε κάθε δυσκολία και πάντα στεκόταν δίπλα μου όχι μόνο σαν επαγγελματίας αλλά και σαν φίλη.

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΜΗΧΑΝΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ

Τομέας Ρευστών

Εργαστήριο Θερμικών Στροβιλομηχανών

Μονάδα Παράλληλης Υπολογιστικής Ρευστοδυναμικής και Βελτιστοποίησης

ΕΠΙΛΥΣΗ ΠΡΟΒΛΗΜΑΤΩΝ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ ΔΥΟ ΣΤΟΧΩΝ ΜΕ
ΒΑΣΗ ΤΗ ΘΕΩΡΙΑ ΤΗΣ ΣΥΝΕΧΙΣΗΣ ΤΟΥ ΜΕΤΩΠΟΥ PARETO

Διπλωματική εργασία **MARIANNAS KAPOYZOY**

Επιβλέπων: Κ. Χ. Γιαννάκογλου, Καθηγητής ΕΜΠ
Αθήνα, Σεπτέμβριος 2015

Περίληψη

Στην παρούσα διπλωματική εργασία επιχειρείται η γενίκευση υπάρχοντων επιλυτών για προβλήματα βελτιστοποίησης ενός στόχου στο πλαίσιο της πολυκριτηριακής βελτιστοποίησης. Πιο συγκεκριμένα, γίνεται προσπάθεια επίλυσης προβλημάτων βελτιστοποίησης δύο στόχων με τη χρήση αιτιοκρατικών μεθόδων, δηλαδή μεθόδων που βασίζονται στην κλίση της συνάρτησης στόχου. Κάτι τέτοιο καθίσταται δυνατό μέσω της θεωρίας της συνέχισης του μετώπου Pareto που αναπτύχθηκε από τους Schmidt και Schulz, [1]. Η εν λόγω θεωρία δηλώνει ότι αν είναι γνωστό ένα σημείο του μετώπου Pareto και οι πρώτες και δεύτερες παράγωγοι των συναρτήσεων στόχου ως προς τις μεταβλητές σχεδιασμού στο σημείο αυτό, μπορεί να προβλεφθεί ένα επόμενο σημείο του μετώπου που βρίσκεται σε σχετικά μικρή απόσταση από αυτό. Το προβλεφθέν σημείο, κατόπιν, 'διορθώνεται', διδόμενο ως σημείο εκκίνησης στην αιτιοκρατική μέθοδο με την οποία προέκυψε το πρώτο σημείο. Μία τέτοια προσέγγιση στοχεύει στη μείωση του υπολογιστικού κόστους σε σχέση λ.χ. με τις στοχαστικές μεθόδους. Στην παρούσα εργασία σε μία προσπάθεια περαιτέρω ελάττωσης του αριθμού των αξιολογήσεων που απαιτούνται για την εύρεση του μετώπου Pareto, δίνεται έμφαση στη διερεύνηση της αποτελεσματικότητας της μεθόδου χωρίς το στάδιο της διόρθωσης. Σε πρώτο στάδιο, η εργασία περιλαμβάνει θεωρητικά στοιχεία που αφορούν σε βασικές έννοιες της βελτιστοποίησης, στις μεθόδους βελτιστοποίησης και στη συνέχιση του μετώπου Pareto. Ακολούθως, εφαρμόζεται η μέθοδος που βασίζεται στη συνέχιση του μετώπου Pareto για την επίλυση τριών μαθηματικών προβλημάτων βελτιστοποίησης δύο στόχων με μοναδικούς περιορισμούς αυτούς που επιβάλλουν τα όρια των μεταβλητών σχεδιασμού. Για τα δύο πρώτα εξ αυτών το μέτωπο Pareto είναι συνεχές και αναζητείται με και χωρίς το στάδιο της διόρθωσης, ενώ για το τρίτο το μέτωπο Pareto είναι ασυνεχές/διακοπτόμενο και αποκτάται χωρίς τη συμπερίληψη του σταδίου της διόρθωσης. Έπειτα, ακολουθεί η βελτιστοποίηση των επιδόσεων υπερηχητικού επιβατικού αεροσκάφους μέσω της μεθόδου που βασίζεται στη συνέχιση του μετώπου Pareto χωρίς το στάδιο της διόρθωσης. Σημειώνεται ότι σε, αυτήν την περίπτωση, οι απαιτούμενες παράγωγοι υπολογίζονται μέσω της μεθόδου των μιγαδικών μεταβλητών. Για όλα τα προβλήματα τα αποτελέσματα συγκρίνονται με τα ανάλογα ενός αναπτυγμένου από τη ΜΠΥΡΒ/ΕΜΠ εξελικτικού αλγορίθμου. Τέλος, μέσω της εν λόγω σύγκρισης διαπιστώνεται ότι για τα προβλήματα που εξετάστηκαν η μέθοδος που βασίζεται στη συνέχιση του μετώπου Pareto υπερέχει σε 'ταχύτητα' σε σχέση με τη στοχαστική μέθοδο.

**NATIONAL TECHNICAL UNIVERSITY OF ATHENS
SCHOOL OF MECHANICAL ENGINEERING**

Fluids Section
Laboratory of Thermal Turbomachines
Parallel CFD & Optimization Unit

***SOLUTION OF TWO-OBJECTIVE OPTIMIZATION PROBLEMS
BASED ON PARETO CONTINUATION TECHNIQUES***

Diploma thesis

MARIANNA KAROYZOY

Advisor: K. C. Giannakoglou, Professor NTUA
Athens, September 2015

Summary

The goal of this diploma thesis is to investigate the generalization of existing solvers for the single objective optimization into the multiobjective framework. More specifically, there takes place an effort of solving two-objective optimization problems using deterministic/gradient-based methods. This is made possible through the Pareto-curve continuation theory which was developed by Schmidt and Schulz, [1]. This theory states that if a point on the Pareto front and the first- and second-order sensitivity derivatives at said point are known, then a next point that is at a relatively short distance from the first point can be predicted. The predicted point is then 'corrected' by being used as a starting value of the deterministic method that provided the first point. Such an approach aims at reducing the computational cost compared to e.g. stochastic methods. In the present study emphasis is placed on investigating the method's effectiveness without the correction step, in an attempt to further decrease the number of evaluations required to acquire the Pareto front. Initially, the thesis contains theoretical elements related to the basic concepts of optimization, the optimization methods and the continuation of the Pareto front. Subsequently, the method based on the continuation of the Pareto front is applied to the solution of three mathematical two-objective optimization problems whose only constraints are being imposed by the limits of the design variables. As far as the first two are concerned, the Pareto front is continuous and is sought with and without the correction step, while, as for the third problem, the Pareto front is discontinuous and is obtained without the inclusion of the correction step. Thereafter, follows the performance optimization of a supersonic civil aircraft by the method based on the Pareto-curve continuation without the correction step. Note that, in this case, the required sensitivity derivatives are calculated using the complex variable method. The results of all tested problems are compared to those of an evolutionary algorithm developed by the PCOpt/NTUA. To end with, through this comparison, it is demonstrated that, concerning the tested problems, the computational effort when using the method based on the continuation of the Pareto front as opposed to the stochastic method is significantly reduced.

Περιεχόμενα

1 Εισαγωγή	7
1.1 Γενικά περί βελτιστοποίησης – Σκοπός της εργασίας.....	7
1.2 Δομή της εργασίας.....	8
2 Θεωρητικό υπόβαθρο βελτιστοποίησης	10
2.1 Η έννοια της βελτιστοποίησης.....	10
2.2 Εισαγωγικές μαθηματικές έννοιες.....	10
2.2.1. Η έννοια της κυρτότητας.....	10
2.2.2. Ακρότατα συναρτήσεων.....	11
2.2.3. Αναλυτικός υπολογισμός ακροτάτων.....	11
2.3 Βελτιστοποίηση πραγματικών συναρτήσεων.....	15
2.3.1 Ορισμοί.....	15
2.3.2 Η βελτιστοποίηση ως υπολογιστική διαδικασία.....	16
2.3.3 Βελτιστοποίηση υπό περιορισμούς.....	18
2.4 Χειρισμός περιορισμών.....	19
2.4.1 Γενική στρατηγική.....	19
2.4.2 Διατύπωση με χρήση συνάρτησης Lagrange.....	20
2.4.2.1 Ελαχιστοποίηση με έναν περιορισμό ισότητας..	20
2.4.2.2 Ελαχιστοποίηση με έναν περιορισμό ανισότητας.....	22
2.4.2.3 Αναλυτική προσέγγιση – Συνθήκες Karush-Kuhn-Tucker.....	23
2.4.2.4 Ικανές συνθήκες.....	26
2.4.3 Σειριακή ελαχιστοποίηση για προβλήματα με περιορισμούς – Συναρτήσεις ποινής.....	26
2.4.4 Η επαυξημένη μέθοδος με πολλαπλασιαστές Lagrange.....	28
2.4.4.1 Η μέθοδος ALM για περιορισμούς ισότητας....	28
2.4.4.2 Η μέθοδος ALM για περιορισμούς ανισότητας..	29
2.4.4.3 Γενίκευση της μεθόδου ALM.....	30
2.5 Πολυκριτηριακή βελτιστοποίηση.....	33
2.5.1 Εισαγωγή.....	33
2.5.2 Διατύπωση του προβλήματος.....	34
2.5.3 Θεμελιώδεις έννοιες πολυκριτηριακής βελτιστοποίησης.....	35
2.5.3.1 Η έννοια της κυριαρχίας.....	35
2.5.3.2 Η έννοια των βέλτιστων λύσεων κατά Pareto....	37
3 Μέθοδοι βελτιστοποίησης	40
3.1 Εισαγωγή.....	40
3.2 Γενικές αρχές αλγορίθμων βελτιστοποίησης.....	43
3.3 Μέθοδοι ανίχνευσης κατά γραμμή.....	44
3.3.1 Βασική αρχή κάθε μεθόδου ανίχνευσης κατά γραμμή..	44

3.3.2	Η μέθοδος της απότομης καθόδου (steepest descent method).....	46
3.3.3	Η μέθοδος Newton.....	48
3.3.4	Η προσεγγιστική μέθοδος Newton.....	52
3.4	Εξελικτικοί αλγόριθμοι βελτιστοποίησης.....	53
3.4.1	Εισαγωγή.....	53
3.4.2	Περιγραφή ενός εξελικτικού αλγορίθμου.....	55
4	Συνέχιση του μετώπου Pareto στην πολυκριτηριακή βελτιστοποίηση (Pareto-curve continuation in multi-objective optimization)	61
4.1	Εισαγωγή.....	61
4.2	Μαθηματικό υπόβαθρο.....	61
4.2.1	Διατύπωση προβλήματος.....	61
4.2.2	Μέθοδος βαρών.....	62
4.2.3	Υπαρξη συνόλου μη-κυριαρχούμενων λύσεων (indifference set).....	65
4.2.4	Το θεώρημα πεπλεγμένης συνάρτησης (implicit function theorem).....	66
4.3	Συνέχιση μετώπου Pareto.....	67
4.3.1	Μέθοδοι πρόβλεψης-διόρθωσης (Predictor-Corrector Methods).....	67
4.3.2	Επαναπαραμετροποίηση – διαφορετική παραμετροποίηση (reparametrization).....	70
4.4	Μαθηματικό παράδειγμα.....	75
5	Επίλυση μαθηματικών προβλημάτων βελτιστοποίησης δύο στόχων με βάση τη θεωρία της συνέχισης του μετώπου Pareto	79
5.1	Εισαγωγή.....	79
5.2	Το λογισμικό βελτιστοποίησης EASY.....	80
5.3	Πρώτο μαθηματικό πρόβλημα.....	81
5.3.1	Διατύπωση προβλήματος.....	81
5.3.2	Εφαρμογή της θεωρίας της συνέχισης του μετώπου Pareto στο πρόβλημα.....	81
5.3.3	Περιγραφή του αλγορίθμου που αναπτύχθηκε για την επίλυση του προβλήματος.....	84
5.3.3.1	Περίπτωση αφαίρεσης σταδίου διόρθωσης.....	84
5.3.3.2	Μέθοδος πρόβλεψης-διόρθωσης.....	99
5.3.4	Παρουσίαση αποτελεσμάτων.....	104
5.3.4.1	Περίπτωση αφαίρεσης σταδίου διόρθωσης.....	104
5.3.4.2	Μέθοδος πρόβλεψης-διόρθωσης.....	127
5.4	Δεύτερο μαθηματικό πρόβλημα.....	134
5.4.1	Διατύπωση προβλήματος.....	134
5.4.2	Εφαρμογή της θεωρίας της συνέχισης του μετώπου Pareto στο πρόβλημα.....	135
5.4.3	Περιγραφή του αλγορίθμου που αναπτύχθηκε για την	

επίλυση του προβλήματος.....	135
5.4.4 Παρουσίαση αποτελεσμάτων.....	137
5.4.4.1 Περίπτωση αφαίρεσης σταδίου διόρθωσης.....	137
5.4.4.2 Μέθοδος πρόβλεψης-διόρθωσης.....	147
5.5 Τρίτο μαθηματικό πρόβλημα.....	151
5.5.1 Διατύπωση προβλήματος.....	151
5.5.2 Εφαρμογή της θεωρίας της συνέχισης του μετώπου Pareto στο πρόβλημα.....	152
5.5.3 Περιγραφή του αλγορίθμου που αναπτύχθηκε για την επίλυση του προβλήματος.....	152
5.5.4 Παρουσίαση αποτελεσμάτων.....	153
6 Βελτιστοποίηση επιδόσεων υπερηχητικού επιβατικού αεροσκάφους με βάση τη θεωρία της συνέχισης του μετώπου Pareto	171
6.1 Εισαγωγή.....	171
6.2 Αλγόριθμος ανάλυσης επιδόσεων υπερηχητικού επιβατικού αεροσκάφους – Διατύπωση προβλήματος.....	172
6.3 Εφαρμογή της θεωρίας της συνέχισης του μετώπου Pareto στο πρόβλημα.....	176
6.4 Περιγραφή του αλγορίθμου που αναπτύχθηκε για την επίλυση του προβλήματος.....	180
6.4.1 Γενικά.....	180
6.4.2 Η μέθοδος των μιγαδικών μεταβλητών (complex variable method).....	181
6.4.2.1 Θεωρητικά στοιχεία.....	181
6.4.2.2 Υλοποίηση σε προγραμματιστικό επίπεδο....	183
6.4.3 Η μέθοδος των πεπερασμένων διαφορών (finite difference method).....	187
6.4.3.1 Θεωρητικά στοιχεία.....	187
6.4.3.2 Υλοποίηση σε προγραμματιστικό επίπεδο....	188
6.4.4 Σύγκριση αποτελεσμάτων μεθόδου μιγαδικών μεταβλητών και κεντρικών πεπερασμένων διαφορών – Πιστοποίηση ορθότητας αποτελεσμάτων παραγών.....	193
6.5 Παρουσίαση αποτελεσμάτων.....	197
7 Επίλογος – Συζήτηση – Συμπεράσματα	204
Βιβλιογραφία	208

1 Εισαγωγή

1.1 Γενικά περί βελτιστοποίησης – Σκοπός της εργασίας

Μία τεράστια ποικιλία προβλημάτων βελτιστοποίησης που ανακύπτουν στη μηχανολογία αλλά και σε πολλούς άλλους επιστημονικούς τομείς περιλαμβάνουν εγγενώς περισσότερους από έναν στόχους σχεδιασμού. Η δυσκολία έγκειται στο γεγονός ότι εξαιρετικά συχνά επιβάλλεται να αντιμετωπιστεί η αντικρουόμενη φύση των διαφορετικών συναρτήσεων στόχου. Για παράδειγμα, στην περίπτωση κατασκευής μίας γέφυρας θα μπορούσαν να τεθούν ταυτοχρόνως ως στόχοι η μεγιστοποίηση της δυσκαμψίας και η ελαχιστοποίηση της συνολικής μάζας της εν λόγω γέφυρας. Ομοίως, το πρόβλημα ελαχιστοποίησης της οπισθέλκουσας μίας αεροτομής είναι συζευγμένο με την παραγόμενη άωση. Η μείωση της οπισθέλκουσας επιτυγχάνεται εύκολα με ‘θυσίες’ στην άωση. Ως εκ τούτου, η βελτιστοποίηση συνήθως περιλαμβάνει έναν περιορισμό που υπαγορεύει ότι η άωση θα πρέπει να έχει μία συγκεκριμένη σταθερή τιμή. Όμως, ο πραγματικός στόχος του σχεδιασμού θα ήταν η ελαχιστοποίηση της οπισθέλκουσας με ταυτόχρονη μεγιστοποίηση της άωσης. Καθώς είναι εξαιρετικά απίθανο οι συναρτήσεις των δύο αυτών αντικρουόμενων στόχων να λάβουν συγχρόνως τη βέλτιστη τιμή τους μέσω του ίδιου σχεδιασμού, η έννοια της μοναδικής βέλτιστης λύσης επιβάλλεται να αντικατασταθεί από αυτήν ενός συνόλου λύσεων που είναι μη-κυριαρχούμενες μεταξύ τους. Αυτό σημαίνει ότι στόχο αποτελεί η εύρεση όλων των πιθανών σχεδιασμών που οδηγούν σε μία τιμή οπισθέλκουσας, η οποία δεν δύναται να μειωθεί περαιτέρω δίχως, επίσης, να μειωθεί και η άωση. Ένας σχεδιασμός για τον οποίο ισχύει κάτι τέτοιο καλείται συνήθως βέλτιστος κατά Pareto (Pareto-optimal) και το σύνολο των βέλτιστων κατά Pareto λύσεων ονομάζεται σύνολο μη-κυριαρχούμενων λύσεων (indifference set) ή σύνολο Pareto.

Το ζήτημα που απασχολεί κάθε επιστήμονα που έρχεται αντιμέτωπος με ένα τέτοιο πρόβλημα βελτιστοποίησης είναι ποιά μέθοδο θα επιλέξει για την επίλυσή του. Οι δύο κυριότερες κατηγορίες στις οποίες διακρίνονται οι μέθοδοι βελτιστοποίησης είναι οι στοχαστικές (stochastic) και οι αιτιοκρατικές μέθοδοι (deterministic methods). Τα πλεονεκτήματα και τα μειονεκτήματα των διαφορετικών μεθόδων βελτιστοποίησης που έχουν αναπτυχθεί είναι αυτά που καθορίζουν τις επιλογές του κληθέντος να επιλύσει το πρόβλημα.

Εξαιρετικά συχνά, μείζονος σημασίας θέμα είναι το πόσο ‘γρήγορα’ μπορεί η εκάστοτε μέθοδος να αποδώσει το μέτωπο Pareto. Για το λόγο αυτό, η επιστημονική κοινότητα της βελτιστοποίησης προσανατολίζει μεγάλο μέρος των προσπαθειών της στην εύρεση

τρόπων ελάττωσης των αξιολογήσεων που απαιτούνται για την απόκτηση του μετώπου Pareto, δηλαδή του υπολογιστικού κόστους. Στο πλαίσιο αυτό, κινείται η μελέτη που πραγματοποιήθηκε από τους Schmidt και Schulz, [1]. Στην εν λόγω μελέτη, προτείνονται οι μέθοδοι πρόβλεψης-διόρθωσης οι οποίες καθιστούν δυνατή τη γενίκευση υπάρχοντων επιλυτών για προβλήματα βελτιστοποίησης ενός στόχου στο πλαίσιο της πολυκριτηριακής βελτιστοποίησης. Αυτές οι μέθοδοι βασίζονται στη θεωρία της συνέχισης του μετώπου Pareto, σύμφωνα με την οποία αν είναι γνωστό ένα σημείο του μετώπου Pareto και οι πρώτες και δεύτερες παράγωγοι των συναρτήσεων στόχου ως προς τις μεταβλητές σχεδιασμού στο σημείο αυτό, μπορεί να προβλεφθεί ένα επόμενο σημείο του μετώπου που βρίσκεται σε σχετικά μικρή απόσταση από αυτό. Στη συνέχεια, το προβλεφθέν σημείο 'διορθώνεται', χρησιμοποιούμενο ως σημείο εκκίνησης μίας μεθόδου προβλημάτων βελτιστοποίησης ενός στόχου.

Στόχο της παρούσας εργασίας αποτελεί η ανάπτυξη λογισμικού ικανού να προβεί στην επίλυση προβλημάτων βελτιστοποίησης δύο στόχων κάνοντας χρήση της θεωρίας της συνέχισης του μετώπου Pareto που παρουσιάζεται από τους Schmidt και Schulz. Τονίζεται ότι ιδιαίτερη έμφαση δίνεται στη διερεύνηση της αποτελεσματικότητας των μεθόδων που βασίζονται στη συνέχιση του μετώπου Pareto χωρίς το στάδιο της διόρθωσης, σε μία προσπάθεια περαιτέρω μείωσης του υπολογιστικού κόστους.

Αρχικά, μελετώνται τρία μαθηματικά προβλήματα βελτιστοποίησης δύο στόχων με μοναδικούς περιορισμούς αυτούς που επιβάλλουν τα όρια των μεταβλητών σχεδιασμού. Για τα δύο πρώτα εξ αυτών το μέτωπο Pareto είναι συνεχές και αναζητείται με και χωρίς το στάδιο της διόρθωσης, ενώ χρησιμοποιούνται ως μέθοδοι προβλημάτων βελτιστοποίησης ενός στόχου η μέθοδος της απότομης καθόδου και η μέθοδος Newton. Όσον αφορά το τρίτο μαθηματικό πρόβλημα, το μέτωπο Pareto είναι ασυνεχές ή διακοπτόμενο και αποκτάται χωρίς τη συμπερίληψη του σταδίου της διόρθωσης, ενώ ως μέθοδος βελτιστοποίησης προβλημάτων ενός στόχου χρησιμοποιείται η μέθοδος της απότομης καθόδου. Έπειτα, ακολουθεί η βελτιστοποίηση των επιδόσεων υπερηχητικού επιβατικού αεροσκάφους μέσω της μεθόδου που βασίζεται στη συνέχιση του μετώπου Pareto χωρίς το στάδιο της διόρθωσης, με τη μέθοδο Newton να επιλέγεται ως μέθοδος βελτιστοποίησης μονοκριτηριακών προβλημάτων. Σημειώνεται ότι σε, αυτήν την περίπτωση, οι απαιτούμενες παράγωγοι υπολογίζονται μέσω της μεθόδου των μιγαδικών μεταβλητών (ενώ για τα τρία μαθηματικά προβλήματα ο υπολογισμός έγινε αναλυτικά). Για όλα τα προβλήματα που εξετάζονται τα αποτελέσματα συγκρίνονται με τα ανάλογα ενός εξελικτικού αλγορίθμου αναπτυγμένου από τη ΜΠΥΡΒ/ΕΜΠ.

1.2 Δομή της εργασίας

Στο **Κεφάλαιο 2** περιλαμβάνονται θεωρητικά στοιχεία που αφορούν σε βασικές έννοιες της βελτιστοποίησης. Πιο συγκεκριμένα, γίνεται αναφορά σε εισαγωγικές μαθηματικές έννοιες, στον τρόπο χειρισμού των περιορισμών και στην πολυκριτηριακή βελτιστοποίηση.

Το **Κεφάλαιο 3** είναι αφιερωμένο στις μεθόδους βελτιστοποίησης. Αρχικά, πραγματοποιείται μία γενική ανασκόπηση και παρουσιάζονται οι γενικές αρχές των αλγορίθμων βελτιστοποίησης. Ακολούθως, περιγράφονται οι μέθοδοι ανίχνευσης κατά γραμμή και οι βασικές αρχές των εξελικτικών αλγορίθμων.

Στο **Κεφάλαιο 4** συναντάται η θεωρία της συνέχισης του μετώπου Pareto πάνω στην οποία είναι βασισμένη η παρούσα εργασία. Κατά πρώτο λόγο, αναπτύσσεται το μαθηματικό υπόβαθρο πάνω στο οποίο χτίζεται. Κατά δεύτερο λόγο, περιγράφονται οι μέθοδοι πρόβλεψης-διόρθωσης με βάση τη θεωρία της συνέχισης του μετώπου Pareto και κατά τρίτο λόγο παρατίθεται ένα μαθηματικό παράδειγμα που αντλήθηκε από τη βιβλιογραφία.

Στο **Κεφάλαιο 5** πραγματοποιείται η επίλυση τριών μαθηματικών προβλημάτων βελτιστοποίησης δύο στόχων (τα δύο πρώτα με και χωρίς το στάδιο της διόρθωσης, ενώ το τρίτο μόνο χωρίς το στάδιο της διόρθωσης). Αρχικά, γίνεται μία σύντομη αναφορά στο αναπτυχθέν από τη ΜΠΥΡΒ/ΕΜΠ λογισμικό βελτιστοποίησης το οποίο επιτρέπει την επαλήθευση των αποτελεσμάτων που προκύπτουν από τη μέθοδο που βασίζεται στη συνέχιση του μετώπου Pareto. Στη συνέχεια, για καθένα από τα τρία μαθηματικά προβλήματα περιγράφεται εκτενώς ο αλγόριθμος που αναπτύχθηκε για την επίλυσή του και παρουσιάζονται τα αποτελέσματα που προέκυψαν. Τέλος, τα εν λόγω αποτελέσματα σχολιάζονται και συγκρίνονται με αυτά του εξελικτικού αλγορίθμου.

Το **Κεφάλαιο 6** πραγματεύεται τη βελτιστοποίηση των επιδόσεων υπερηχητικού επιβατικού αεροσκάφους μέσω της μεθόδου που βασίζεται στη συνέχιση του μετώπου Pareto χωρίς το στάδιο της διόρθωσης. Αρχικά, περιγράφεται εν συντομία ο αλγόριθμος ανάλυσης επιδόσεων υπερηχητικού αεροσκάφους. Στη συνέχεια, περιγράφεται η μέθοδος των μιγαδικών μεταβλητών με βάση την οποία πραγματοποιείται ο υπολογισμός των πρώτων και δεύτερων παραγώγων των συναρτήσεων στόχου ως προς τις μεταβλητές σχεδιασμού. Ακολούθως, πιστοποιείται η ορθότητα των εν λόγω παραγώγων μέσω της μεθόδου των πεπερασμένων διαφορών. Τέλος, παρουσιάζεται το προκύπτον μέτωπο Pareto και συγκρίνεται με αυτό του εξελικτικού αλγορίθμου.

Η εργασία ολοκληρώνεται με το **Κεφάλαιο 7**, στο οποίο, μετά από μια σύντομη ανακεφαλαίωση, παρουσιάζονται τα συμπεράσματα και κάποιες μελλοντικές προτάσεις έρευνας.

2 Θεωρητικό υπόβαθρο βελτιστοποίησης

2.1 Η έννοια της βελτιστοποίησης

Όταν τίθεται ένα πρόβλημα προς επίλυση, η αντιμετώπισή του βασίζεται σε μια διαδοχή από εναλλακτικές αποφάσεις (decisions) και αξιολογήσεις (evaluations) των επιπτώσεων κάθε απόφασης. Αν καθεμία από τις εναλλακτικές αποφάσεις που ικανοποιούν τους περιορισμούς του προβλήματος μπορεί να περιγραφεί από ένα σύνολο $\{x_1, x_2, \dots, x_n\}$ που ονομάζεται διάνυσμα μεταβλητών σχεδιασμού και αν σε κάθε τέτοια περιγραφή μπορεί να αντιστοιχιστεί ένα πραγματικό μέτρο επίδοσης (performance measure), τότε ως βέλτιστη (optimal) νοείται η απόφαση που μεγιστοποιεί το εν λόγω μέτρο.

Στην πράξη, η βελτιστοποίηση (optimization) εφαρμόζεται σε προβλήματα σχεδιασμού ή λειτουργίας συστημάτων. Σύμφωνα με τον ορισμό του Pierre (1986), [58]:

Ένα σύστημα είναι βέλτιστο ως προς ένα δεδομένο μέτρο επίδοσης και ένα δεδομένο σύνολο περιορισμών εφόσον λειτουργεί/αποδίδει τουλάχιστον ίσα αν όχι καλύτερα από κάθε άλλο σύστημα που ικανοποιεί τους ίδιους περιορισμούς.

2.2 Εισαγωγικές μαθηματικές έννοιες

2.2.1 Η έννοια της κυρτότητας

Η κυρτότητα (convexity) αποτελεί μία από τις πλέον σημαντικές μαθηματικές έννοιες, με ιδιαίτερη σημασία στη βελτιστοποίηση. Συνεπώς, κρίνεται σκόπιμο να δοθεί, αρχικά, ο ορισμός της κυρτότητας για πεδία (sets) και στη συνέχεια, να γενικευθεί για συναρτήσεις πολλών μεταβλητών που συναντώνται διαρκώς σε προβλήματα βελτιστοποίησης.

Ένα n -διάστατο πεδίο S είναι κυρτό εφόσον, [2]:

$$\lambda \vec{x}_1 + (1 - \lambda) \vec{x}_2 \in S \quad \forall \vec{x}_1, \vec{x}_2 \in S \text{ και } \forall \lambda \in [0, 1].$$

Η παραπάνω σχέση ονομάζεται κυρτός συνδυασμός και υποδηλώνει ότι το ευθύγραμμο τμήμα που ενώνει το ζεύγος σημείων $\{\vec{x}_1, \vec{x}_2\} \in S$ κείται αποκλειστικά στο πεδίο S .

Έστω η πραγματική συνάρτηση $f(\vec{x})$ ορισμένη στο κυρτό πεδίο $X \subseteq R^n$. Η συνάρτηση f είναι, [2]:

- ✘ κυρτή (convex) στο πεδίο X αν $\forall \vec{x}_1, \vec{x}_2 \in X$ και $\forall \lambda \in [0, 1]$ ισχύει:

$$\lambda f(\vec{x}_1) + (1 - \lambda) f(\vec{x}_2) \geq f[\lambda \vec{x}_1 + (1 - \lambda) \vec{x}_2] \quad (2.1)$$

✘ κοίλη (concave) στο πεδίο X αν $\forall \vec{x}_1, \vec{x}_2 \in X$ και $\forall \lambda \in [0, 1]$ ισχύει:

$$\lambda f(\vec{x}_1) + (1 - \lambda) f(\vec{x}_2) \leq f[\lambda \vec{x}_1 + (1 - \lambda) \vec{x}_2]$$

Σε κάθε άλλη περίπτωση, η συνάρτηση είναι μη-κυρτή (non-convex). Εφόσον η συνάρτηση είναι κυρτή, τότε το ευθύγραμμο τμήμα που ενώνει δύο τυχαία σημεία του πεδίου ορισμού της δεν βρίσκεται ποτέ κάτω από το γράφημά της. Ομοίως, αν η συνάρτηση είναι κοίλη, τότε το ευθύγραμμο τμήμα δεν βρίσκεται ποτέ πάνω από το γράφημά της. Τέλος, αν η συνάρτηση είναι μη-κυρτή, τότε το ευθύγραμμο τμήμα βρίσκεται άλλοτε κάτω και άλλοτε πάνω από το γράφημά της. Κάθε κυρτή συνάρτηση είναι εξ ορισμού συνεχής.

2.2.2 Ακρότατα συναρτήσεων

Μια πραγματική συνάρτηση $f(\vec{x})$ ορισμένη στο $X \subseteq \mathbb{R}^n$ παρουσιάζει τοπικό ελάχιστο (local minimum) στο σημείο $\vec{x}^* \in X$ όταν υπάρχει περιοχή $X_0 \subset X$ του \vec{x}^* τέτοια ώστε για κάθε $x \in X_0$ να ισχύει, [4]:

$$f(\vec{x}^*) \leq f(\vec{x})$$

Αντίστοιχος είναι ο ορισμός και για το τοπικό μέγιστο (local maximum).

Κάθε σημείο τοπικού ελαχίστου ή τοπικού μεγίστου καλείται τοπικό ακρότατο (local extreme), ενώ όταν $X_0 \equiv X$, το ακρότατο χαρακτηρίζεται ως απόλυτο ή, συνηθέστερα, ολικό (global). Κατά συνέπεια, το ολικό ελάχιστο (global minimum) μιας συνάρτησης είναι το μικρότερο από τα τοπικά της ελάχιστα στο πεδίο ορισμού της. Αντίστοιχα, το ολικό μέγιστο (global maximum) είναι το μεγαλύτερο από τα τοπικά της μέγιστα.

2.2.3 Αναλυτικός υπολογισμός ακροτάτων

Κάθε σημείο μηδενισμού του διανύσματος κλίσης μιας συνάρτησης πολλών μεταβλητών $f(\vec{x})$, δηλαδή κάθε $\vec{x}^* \in X$ για το οποίο:

$$\nabla f(\vec{x}^*) = \text{grad } f(\vec{x}^*) = \left[\left. \frac{\partial f}{\partial x_1} \right|_{\vec{x}^*} \quad \left. \frac{\partial f}{\partial x_2} \right|_{\vec{x}^*} \quad \dots \quad \left. \frac{\partial f}{\partial x_n} \right|_{\vec{x}^*} \right]^T = 0$$

καλείται στάσιμο (stationary).

Το μητρώο των μερικών παραγώγων δεύτερης τάξης της συνάρτησης $f(\vec{x})$ ονομάζεται εσσιανό (Hessian) και έχει τη μορφή:

$$H(\vec{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

Έστω $H_i(\vec{x}^*)$ η i υπο-ορίζουσα του εσσιανού μητρώου H , η οποία προκύπτει με αφαίρεση των $n - i$ τελευταίων γραμμών και στηλών του. Διακρίνονται οι ακόλουθες περιπτώσεις, [2]:

- i. αν $H_i(\vec{x}^*) > 0$ για κάθε i , τότε το \vec{x}^* είναι τοπικό ελάχιστο,
- ii. αν $H_i(\vec{x}^*) \neq 0$ για κάθε i και $\text{sign}(H_i) = \text{sign}(-1)^i$, τότε το \vec{x}^* είναι τοπικό μέγιστο,
- iii. αν $H_n(\vec{x}^*) \neq 0$ και δεν ισχύει καμία από τις παραπάνω περιπτώσεις, τότε το \vec{x}^* είναι σημείο σέλας,
- iv. αν $H_n(\vec{x}^*) = 0$, δεν μπορεί να υπάρξει συμπέρασμα.

Η πρώτη συνθήκη υποδηλώνει ότι το εσσιανό μητρώο H είναι θετικά ορισμένο (positive definite) στο σημείο \vec{x}^* (δηλαδή $(\vec{x}^*)^T H \vec{x}^* > 0$), ενώ η δεύτερη συνθήκη υποδηλώνει ότι το H είναι αρνητικά ορισμένο στο \vec{x}^* (δηλαδή $(\vec{x}^*)^T H \vec{x}^* < 0$). Κατά συνέπεια, αν το μητρώο H είναι θετικά ορισμένο σε ένα στάσιμο σημείο \vec{x}^* , τότε αυτό είναι τοπικό ελάχιστο, ενώ αν είναι αρνητικά ορισμένο τότε το σημείο είναι τοπικό μέγιστο. Τέλος, αν το μητρώο H είναι μη-ορισμένο, τότε το \vec{x}^* δεν αντιστοιχεί σε ακρότατο και στην περίπτωση αυτή ονομάζεται σημείο σέλας (saddle point).

Στο σημείο αυτό κρίνεται σκόπιμο να παρατεθεί ένα παράδειγμα. Ας θεωρηθεί η συνάρτηση $f(x, y) = 3x^3 - 5y^2 - 225x + 70y + 23$. Τα στάσιμα σημεία της εν λόγω συνάρτησης υπολογίζονται μέσω των σχέσεων:

$$\left. \begin{array}{l} \frac{\partial f}{\partial x} = 9x^2 - 225 = 0 \\ \frac{\partial f}{\partial y} = -10y + 70 = 0 \end{array} \right\} \Rightarrow x = \pm 5, y = 7$$

Το εσσιανό μητρώο της ανωτέρω συνάρτησης είναι:

$$H(\vec{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 18x & 0 \\ 0 & -10 \end{bmatrix}$$

και οι τιμές του εσσιανού μητρώου στα στάσιμα σημεία είναι:

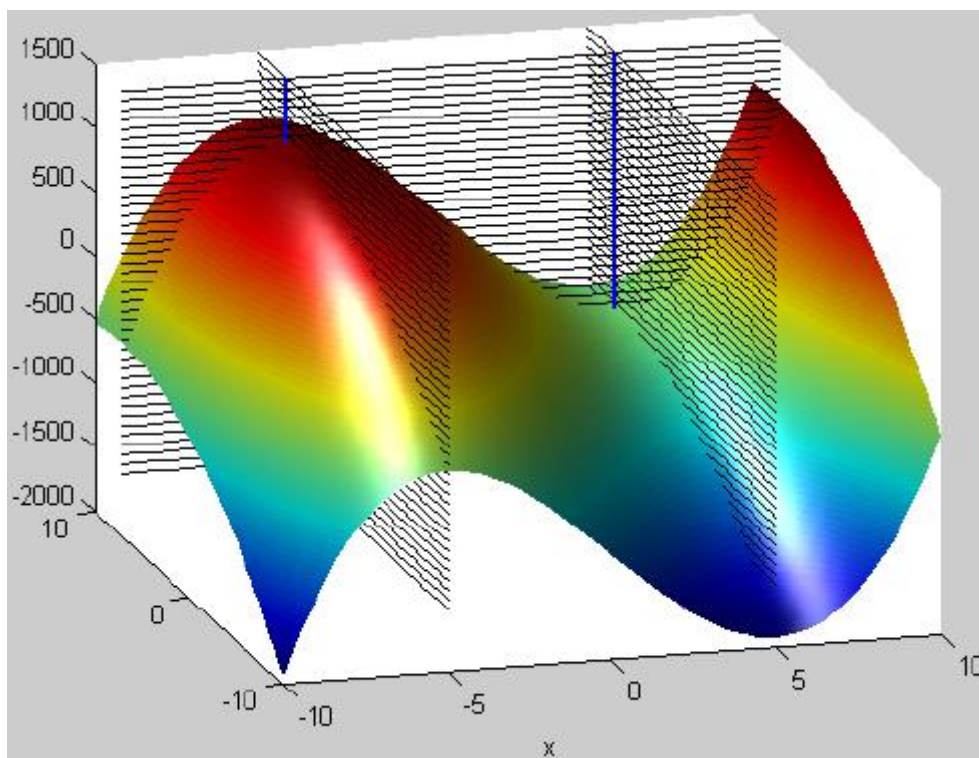
$$H|_{x=5,y=7} = \begin{bmatrix} 90 & 0 \\ 0 & -10 \end{bmatrix}, \quad H|_{x=-5,y=7} = \begin{bmatrix} -90 & 0 \\ 0 & -10 \end{bmatrix}$$

Τέλος, για τις $H_i(\vec{x}^*)$ έχουμε:

$$H_1|_{x=5,y=7} = 90, H_2|_{x=5,y=7} = -900$$

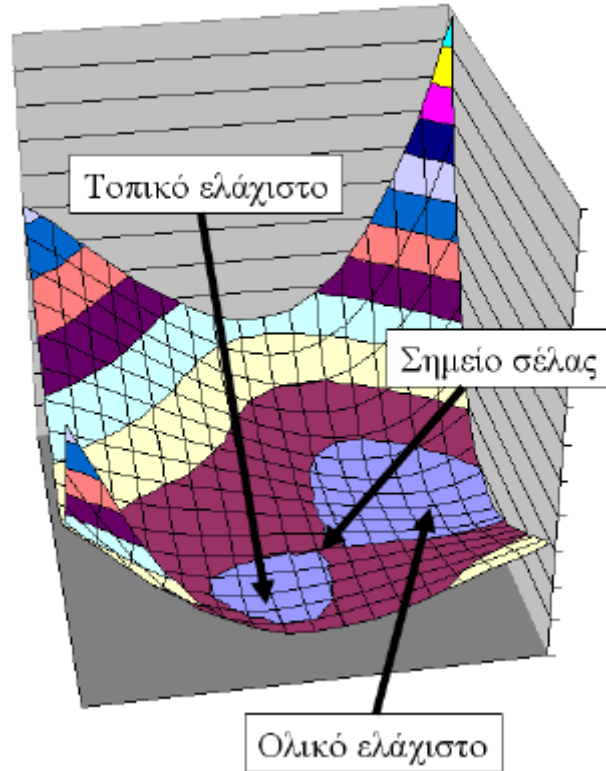
$$H_1|_{x=-5,y=7} = -90, H_2|_{x=-5,y=7} = 900$$

Για το στάσιμο σημείο $(x = 5, y = 7)$, λοιπόν, διαπιστώνουμε ότι ισχύει η περίπτωση (iii) και άρα πρόκειται για σημείο σέλας ενώ για το στάσιμο σημείο $(x = -5, y = 7)$ ισχύει η περίπτωση (ii) και άρα πρόκειται για μέγιστο. Στο σχήμα 2.1 φαίνεται η γραφική παράσταση της συνάρτησης με τα αντίστοιχα ακρότατα (στα σημεία τομής των κυανών ευθειών με την επιφάνεια).



Σχήμα 2.1 Γράφημα της $f(x, y) = 3x^3 - 5y^2 - 225x + 70y + 23$ με εμφανή τα ακρότατα της, [3].

Ακόμα ένα παράδειγμα φαίνεται στο σχήμα 2.2.



Σχήμα 2.2: Γραφική απεικόνιση της συνάρτησης $f(x_1, x_2) = 0,5(1,1x_1 - x_2)^4 + 0,5(x_1 - 0,5)(x_2 - 0,5)$. Η συνάρτηση είναι μη-κυρτή και έχει τρία στάσιμα σημεία (δύο τοπικά ελάχιστα και ένα σημείο σέλας), [2].

Αν η συνάρτηση $f(\vec{x})$ είναι κυρτή και παραγωγίσιμη στο κυρτό πεδίο ορισμού της X , τότε από τη σχέση (2.1) αποδεικνύεται ότι για κάθε $\{\vec{x}_1, \vec{x}_2\} \in X$ ισχύει (Marlow, 1993), [59]:

$$f(\vec{x}_2) \geq f(\vec{x}_1) + \nabla f(\vec{x}_1) (\vec{x}_2 - \vec{x}_1)$$

Στην περίπτωση αυτή, η $f(\vec{x})$ έχει μοναδικό ελάχιστο \vec{x}^* που αντιστοιχεί στο σημείο μηδενισμού της κλίσης $grad f(\vec{x})$. Κατά συνέπεια, αν μια συνάρτηση ικανοποιεί την αναγκαία συνθήκη $\nabla f(\vec{x}^*) = 0$ και την ικανή συνθήκη κυρτότητας, τότε παρουσιάζει ολικό ελάχιστο στο \vec{x}^* , αφού $\forall \vec{x}_2 \in X$ θα ισχύει $f(\vec{x}_2) \geq f(\vec{x}^*)$. Αντίστοιχα, αν η συνάρτηση είναι κοίλη, τότε έχει ένα μοναδικό στάσιμο σημείο που είναι ολικό μέγιστο. Διαφορετικά, αν η συνάρτηση είναι μη-κυρτή, τότε έχει περισσότερα του ενός στάσιμα σημεία, καθένα από τα οποία μπορεί να είναι τοπικό ελάχιστο ή τοπικό μέγιστο ή σημείο σέλας. Αυτό συνεπάγεται ότι η εύρεση του ολικού ακροτάτου μιας μη-κυρτής συνάρτησης προϋποθέτει τον εντοπισμό όλων των στάσιμων σημείων της και την επιλογή του ολικού ακροτάτου με σύγκριση των επιμέρους τιμών τους. Παραδείγματα μη-κυρτών συναρτήσεων φαίνονται στα σχήματα 2.1 και 2.2.

2.3 Βελτιστοποίηση πραγματικών συναρτήσεων

2.3.1 Ορισμοί

Έστω το βαθμωτό μέτρο επίδοσης ενός φυσικού ή μαθηματικού συστήματος:

$$P = f(x_1, x_2, \dots, x_n)$$

όπου $f(x_1, x_2, \dots, x_n)$ μια πραγματική συνάρτηση ορισμένη στο πεδίο $X \subseteq \mathbb{R}^n$ και $\vec{x} = [x_1, x_2, \dots, x_n]^T$ το διάνυσμα ανεξάρτητων μεταβλητών. Στη θεωρία βελτιστοποίησης, το μέτρο f καλείται συνάρτηση στόχου (objective function) ή και συνάρτηση κόστους (cost function), ενώ οι συνιστώσες της x_j καλούνται μεταβλητές ελέγχου (control variables) ή μεταβλητές απόφασης (decision variables) ή συνηθέστερα μεταβλητές σχεδιασμού (design variables).

Το πεδίο ορισμού της συνάρτησης καλείται εφικτή περιοχή (feasible region) ή εφικτός χώρος (feasible space) ή χώρος αναζήτησης (search space) ή χώρος πολιτικής (policy domain), και συμβολίζεται με X . Το πεδίο τιμών καλείται χώρος των αντικειμενικών συναρτήσεων (objective space) ή χώρος αξιολόγησης (evaluation space) και συμβολίζεται με F .

Σε αρχικό στάδιο, θεωρείται ότι το μέτρο επίδοσης είναι βαθμωτό μέγεθος, που σημαίνει ότι σε κάθε διάνυσμα \vec{x} αντιστοιχεί μια πραγματική τιμή $P = f(\vec{x})$. Η τιμή αυτή αντιπροσωπεύει ένα καθολικό κριτήριο του υπό μελέτη συστήματος ως προς τις μεταβλητές σχεδιασμού του. Το κριτήριο μπορεί να είναι ένα πραγματικό, μετρήσιμο μέγεθος (π.χ. οικονομικό όφελος, αξιοπιστία), ένας συνδυασμός πραγματικών μεγεθών ή ακόμη και ένας εμπειρικός αριθμητικός δείκτης χωρίς φυσική ερμηνεία, αρκεί να είναι μοναδικός για κάθε \vec{x} .

Με βάση τις παραπάνω παραδοχές, το πρόβλημα βελτιστοποίησης διατυπώνεται ως:

$$\text{opt } P := f(\vec{x}), \quad \vec{x} \in X$$

όπου ο τελεστής «opt» υποδηλώνει, ανάλογα βεβαίως με την φυσική ερμηνεία του μέτρου P , είτε την μεγιστοποίηση (maximization) είτε την ελαχιστοποίηση (minimization) της συνάρτησης f . Συνεπώς, η διατύπωση του προβλήματος προϋποθέτει τον προσδιορισμό του τύπου του τελεστή βελτιστοποίησης, της συνάρτησης f και του πεδίου X .

Επισημαίνεται ότι ως συνάρτηση στόχου δεν νοείται κατ' ανάγκη μια αναλυτική έκφραση. Στη γενική περίπτωση, πρόκειται για μια «διαδικασία» που επιστρέφει μια μοναδική τιμή συναρτήσεως των τιμών κάποιων ορισμάτων, δηλαδή μια συνάρτηση (function) με την υπολογιστική χροιά του όρου. Στο πλαίσιο της εν λόγω συνάρτησης μπορεί να υλοποιείται ένα σύνολο πολύπλοκων υπολογιστικών

διαδικασιών, όπως για παράδειγμα ένα μοντέλο προσομοίωσης, το αποτέλεσμα της οποίας είναι ένας πραγματικός αριθμητικός δείκτης που αποτιμά την επίδοση του συστήματος ως προς τα αντίστοιχα ορίσματα (μεταβλητές σχεδιασμού).

Αξιίζει να αναφέρουμε ότι οποιοδήποτε πρόβλημα μεγιστοποίησης μπορεί να μετατραπεί σε πρόβλημα ελαχιστοποίησης είτε με τον στοιχειώδη μετασχηματισμό:

$$\max f(\vec{x}) \equiv \min [-f(\vec{x})]$$

είτε με τον ισοδύναμό του:

$$\max f(\vec{x}) \equiv \min [1 / f(\vec{x})] \text{ εφόσον } f(\vec{x}) \neq 0 \text{ για κάθε } \vec{x} \in X.$$

2.3.2 Η βελτιστοποίηση ως υπολογιστική διαδικασία

Σύμφωνα με την έως τώρα διατύπωση του προβλήματος βελτιστοποίησης, ενώ το πεδίο αναζήτησης μπορεί να είναι διανυσματικό (εφόσον $n > 1$), το πεδίο αξιολόγησης είναι απαραίτητα βαθμωτό. Το χαρακτηριστικό αυτό επιτρέπει την αντικειμενική σύγκριση οιοδήποτε ζεύγους δυνατών λύσεων \vec{x}_1 και \vec{x}_2 , ελέγχοντας την σχέση μεταξύ των αντίστοιχων τιμών που λαμβάνει η συνάρτηση στόχου $f(\vec{x}_1)$ και $f(\vec{x}_2)$. Αν $f(\vec{x}_1) > f(\vec{x}_2)$, η πρώτη δυνατή λύση υπερτερεί σε σχέση με την δεύτερη, ενώ αν $f(\vec{x}_1) = f(\vec{x}_2)$, οι δύο δυνατές λύσεις είναι ισοδύναμες.

Αν, λοιπόν, υπήρχε η δυνατότητα συλλογής και αξιολόγησης όλων των συνδυασμών των μεταβλητών σχεδιασμού που ανήκουν στην εφικτή περιοχή, θα ήταν δυνατή η ταξινόμηση (ordering) των αντίστοιχων τιμών των συναρτήσεων στόχου και συνακόλουθα ο εντοπισμός της βέλτιστης τιμής της συνάρτησης στόχου αλλά και του βέλτιστου συνδυασμού μεταβλητών σχεδιασμού που επιτυγχάνουν την εν λόγω τιμή. Επειδή στην πράξη κάτι τέτοιο είναι πλήρως ανέφικτο (εκτός αν το πεδίο αναζήτησης είναι διακριτό, αποτελούμενο από πεπερασμένο και σχετικά μικρό πλήθος σημείων), κανείς καταφεύγει σε συστηματικές διαδικασίες (κωδικοποιημένες με τη μορφή αλγορίθμων) αναζήτησης της βέλτιστης λύσης ή μιας ικανοποιητικής προσέγγισης αυτής, με διερεύνηση ενός αντιπροσωπευτικού αριθμού σημείων του εφικτού χώρου. Η θεμελιώδης έννοια της συστηματικότητας έγκειται στην χρήση προδιαγεγραμμένων κανόνων μετάβασης (transition rules) από σημείο (ή σημεία) σε άλλο σημείο (ή σημεία) του εφικτού χώρου. Κάθε φορά που εντοπίζεται ένα τέτοιο σημείο, υπολογίζεται η αντίστοιχη τιμή της συνάρτησης στόχου και λαμβάνεται απόφαση σχετικά με την επόμενη κατεύθυνση. Μετά από τη διερεύνηση ικανοποιητικού αριθμού διανυσμάτων μεταβλητών σχεδιασμού, εντοπίζεται εύκολα το πλέον κατάλληλο, με βάση την τιμή της συνάρτησης, που θεωρείται εκτιμήτρια του βέλτιστου. Συνεπώς, αυτό που αναζητείται, στην πραγματικότητα, δεν είναι παρά

μια προσέγγιση της βέλτιστης λύσης και όχι η βέλτιστη λύση αυτή καθαυτή. Πράγματι, μόνο για εξειδικευμένες διατυπώσεις του προβλήματος βελτιστοποίησης (π.χ. σε γραμμικά προβλήματα) η συστηματική αυτή διαδικασία συγκλίνει με πλήρη βεβαιότητα στο πραγματικό βέλτιστο της συνάρτησης.

Με βάση τα παραπάνω, δίδεται ένας πρακτικός ορισμός της έννοιας της βελτιστοποίησης, [2]:

Βελτιστοποίηση ονομάζεται μια συστηματική διαδικασία αναζήτησης ενός πεπερασμένου πλήθους σημείων του εφικτού χώρου και αξιολόγησης αυτών με βάση ένα καθολικό βαθμωτό μέτρο επίδοσης του συστήματος, το οποίο αποτιμάται μέσω της συνάρτησης στόχου, με σκοπό τον εντοπισμό του σημείου εκείνου για το οποίο μεγιστοποιείται η εν λόγω επίδοση.

Τα ερωτήματα που απορρέουν από τον παραπάνω ορισμό είναι τα εξής:

- πώς υλοποιείται η συστηματική διαδικασία αναζήτησης, δηλαδή ο αλγόριθμος βελτιστοποίησης
- και πώς προσδιορίζεται ο αριθμός των σημείων που πρέπει να εξεταστούν ώστε να επιτευχθεί μια ικανοποιητική προσέγγιση της βέλτιστης λύσης, ποια είναι δηλαδή τα κριτήρια σύγκλισης ή/και τερματισμού του αλγορίθμου;

Η απάντηση στο πρώτο ερώτημα δεν είναι μονοσήμαντη, αφού για κάθε πρόβλημα βελτιστοποίησης ενδείκνυται μια διαφορετική στρατηγική αναζήτησης. Μάλιστα, όσο λιγότερο τυποποιημένο είναι το πρόβλημα (κάτι που είναι αναπόφευκτο όταν αναφερόμαστε στον πραγματικό, μη-γραμμικό, κόσμο), τόσο μεγαλύτερο εύρος μεθοδολογιών διατίθεται για τον χειρισμό του.

Όσον αφορά το δεύτερο ερώτημα, και με δεδομένο ότι για το συγκεκριμένο πρόβλημα δεν υπάρχει αλγόριθμος εγγυημένου εντοπισμού του βέλτιστου με πεπερασμένο πλήθος δοκιμών αλλά ούτε και αναλυτική λύση, η απάντηση εξαρτάται από τον χρόνο που διατίθεται για την πραγματοποίηση των αναγκαίων δοκιμών. Στα πραγματικά προβλήματα, το υπολογιστικό κόστος που απαιτείται για την αποτίμηση μιας λύσης αποτελεί τον πλέον δεσμευτικό παράγοντα για την διαδικασία της βελτιστοποίησης. Πολύ συχνά, ο παράγοντας αυτός οδηγεί σε απόφαση ελέγχου πολύ λιγότερων λύσεων σε σχέση ακόμα και με την τάξη μεγέθους που θα απαιτούσε το συγκεκριμένο πρόβλημα, κάτι που δημιουργεί προφανή κίνδυνο εντοπισμού μιας τελικής λύσης πολύ λιγότερο «αποδοτικής» σε σχέση με την θεωρητική βέλτιστη λύση. Κατά συνέπεια, το πρακτικό ζητούμενο της βελτιστοποίησης δεν είναι η ακριβέστερη δυνατή προσέγγιση του ολικού ακροτάτου, αλλά η εύρεση μιας πρόσφορης λύσης, μετά από εύλογο αριθμό δοκιμών, ανάλογα με τις ιδιαιτερότητες του εκάστοτε προβλήματος.

2.3.3 Βελτιστοποίηση υπό περιορισμούς

Το πρόβλημα βελτιστοποίησης που παρουσιάστηκε μέχρι τώρα είχε τη μορφή:

$$\min f(\vec{x}), \quad \vec{x} \in X.$$

Αν $X \equiv \mathbb{R}^n$, κάτι που σημαίνει ότι το πεδίο ορισμού ή ο χώρος αναζήτησης ή ο χώρος αποδεκτών λύσεων ταυτίζεται με τον n -διάστατο ευκλείδειο χώρο, το πρόβλημα βελτιστοποίησης διατυπώνεται χωρίς περιορισμούς (unconstrained optimization), ενώ αν $X \subset \mathbb{R}^n$, η διατύπωση γίνεται υπό περιορισμούς (constrained optimization).

Στη γενική περίπτωση βελτιστοποίησης με περιορισμούς, θεωρείται ότι το πεδίο X περιγράφεται από ένα πλήθος μαθηματικών περιορισμών (constraints) της μορφής:

$$c(x_1, x_2, \dots, x_n) \leq, =, \geq 0$$

Συνεπώς, το πρόβλημα βελτιστοποίησης μπορεί να επαναδιατυπωθεί με τους στόχους να είναι οι εξής, [4]:

- Ελαχιστοποίηση της τιμής της συνάρτησης στόχου $f(\vec{x})$, $f: \mathbb{R}^n \rightarrow \mathbb{R}$,
- Ικανοποίηση των περιορισμών ανισότητας (inequality) ή στην πραγματικότητα ανισοϊσότητα της μορφής:

$$c_i(\vec{x}) \leq 0, i \in I,$$

όπου I (εκ του Inequality) είναι το σύνολο τιμών του δείκτη i για τους περιορισμούς ανισότητας,

- Ικανοποίηση των περιορισμών ισότητας (equality) της μορφής:

$$c_i(\vec{x}) = 0, i \in E,$$

όπου E (εκ του Equality) είναι το σύνολο τιμών του δείκτη i για τους περιορισμούς ισότητας.

Πρέπει να σημειωθεί ότι στην κατηγορία των περιορισμών ανισότητας είναι λογικό να ενσωματωθούν και οι περιορισμοί που αφορούν τις ίδιες τις μεταβλητές σχεδιασμού που διατυπώνονται ως $x_i^L \leq x_i \leq x_i^U$, όπου ο άνω δείκτης L ή U συμβολίζει αντίστοιχα την κατώτερη (Lower) και ανώτερη (Upper) επιτρεπόμενη τιμή για την κάθε μεταβλητή i . Τέτοιοι περιορισμοί είναι δυνατό να επιβληθούν σε όλες τις μεταβλητές σχεδιασμού, σε κάποιες από αυτές ή και σε καμία, δηλαδή ο αριθμός τους θα είναι $\leq n$. Επιπλέον, δεν είναι απαραίτητο οι μεταβλητές να έχουν και άνω και κάτω όριο.

Από τα παραπάνω, διαπιστώνεται ότι ο χώρος αποδεκτών λύσεων για ένα δεδομένο πρόβλημα βελτιστοποίησης ορίζεται ως, [4]:

$$X = \{\vec{x} \mid c_i(\vec{x}) \leq 0, i \in I; c_i(\vec{x}) = 0, i \in E\}$$

Κάθε διάνυσμα $\vec{x} \in \mathbb{R}^n$ που δεν επαληθεύει τους περιορισμούς του προβλήματος βελτιστοποίησης θεωρείται μη-αποδεκτό (non-feasible). Σε προβλήματα με περιορισμούς, το θεωρητικό ολικό ακρότατο της συνάρτησης στόχου ενδέχεται να είναι μη-αποδεκτό, εφόσον βρίσκεται εκτός των ορίων του χώρου αναζήτησης. Στην περίπτωση αυτή, ζητείται το σημείο όπου μεγιστοποιείται/ελαχιστοποιείται η τιμή της συνάρτησης εντός του αποδεκτού (feasible) χώρου.

Η ύπαρξη περιορισμών εισάγει επιπλέον απαίτηση στη διαδικασία βελτιστοποίησης, αφού επιβάλλει την αναζήτηση αποδεκτών, αποκλειστικά, λύσεων. Αυτό σημαίνει ότι οι κανόνες μετάβασης πρέπει να διατυπωθούν με τρόπο τέτοιο ώστε να εξασφαλίζουν τη διερεύνηση αποδεκτών και μόνο περιοχών, μη επιτρέποντας την «παραβίαση» των ορίων του πεδίου X . Επιπλέον, η ύπαρξη περιορισμών αυξάνει σημαντικά την δυσχέρεια του προβλήματος λόγω του ότι συνεπάγεται πολλαπλάσιο αριθμό δοκιμών σε σχέση με την βελτιστοποίηση της ίδιας συνάρτησης, αλλά χωρίς περιορισμούς.

2.4 Χειρισμός περιορισμών

2.4.1 Γενική στρατηγική

Τα πραγματικά προβλήματα βελτιστοποίησης διέπονται από ένα πλήθος περιορισμών, τόσο φυσικών όσο και λειτουργικών. Αυτό, όπως αναφέρθηκε παραπάνω, εισάγει μια επιπλέον δυσκολία, δεδομένου ότι η βέλτιστη λύση πρέπει να αναζητηθεί στο χωρίο που συναληθεύει τους περιορισμούς. Για το λόγο αυτό, στρέφεται κανείς σε χαρακτηριστικές μεθόδους που έχουν αναπτυχθεί για το χειρισμό των περιορισμών. Σε όλες τις περιπτώσεις τέτοιων μεθόδων, η γενική στρατηγική έγκειται στον μετασχηματισμό του αρχικού προβλήματος βελτιστοποίησης υπό περιορισμούς, δηλαδή:

$$\min f(\vec{x}), \quad \vec{x} \in X \subset \mathbb{R}^n$$

σε ένα ισοδύναμο πρόβλημα χωρίς περιορισμούς, της μορφής:

$$\min \varphi(\vec{x}), \quad \vec{x} \in \mathbb{R}^n$$

το οποίο επιλύεται, στη συνέχεια, με οποιονδήποτε αλγόριθμο, κατάλληλο για την επίλυση προβλημάτων βελτιστοποίησης χωρίς περιορισμούς.

2.4.2 Διατύπωση με χρήση συνάρτησης Lagrange

2.4.2.1 Ελαχιστοποίηση με έναν περιορισμό ισότητας

Στο σημείο αυτό, κρίνεται σκόπιμο να διατυπωθεί η αναγκαία συνθήκη για τη βέλτιστη λύση σε ένα πρόβλημα ελαχιστοποίησης με έναν περιορισμό ισότητας, [4]:

Αναγκαία συνθήκη ώστε το διάνυσμα \vec{x}^* να αποτελεί βέλτιστη λύση στο πρόβλημα:

$$\min_{\vec{x} \in \mathbb{R}^n} f(\vec{x}), \quad \text{E} \quad c_1(\vec{x}) = 0 \quad (2.2)$$

είναι να μην υπάρχει κατεύθυνση \vec{p} για την οποία να ισχύουν συγχρόνως οι σχέσεις:

$$\begin{cases} \vec{p}^T \nabla c_1(\vec{x}^*) = 0 \\ \vec{p}^T \nabla f(\vec{x}^*) < 0 \end{cases}$$

ή εναλλακτικά είναι τα διανύσματα $\nabla f(\vec{x}^*)$ και $\nabla c_1(\vec{x}^*)$ να είναι παράλληλα, δηλαδή να ισχύει:

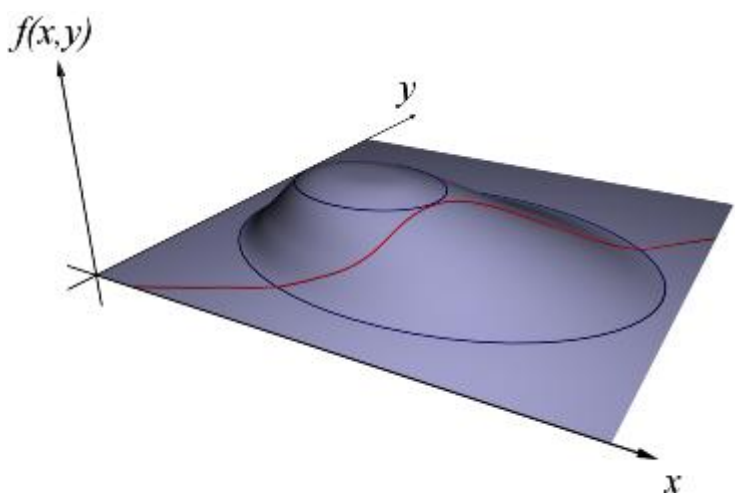
$$\nabla f(\vec{x}^*) = \lambda_1^* \nabla c_1(\vec{x}^*) \quad (2.3)$$

για κάποια τιμή του λ_1^* .

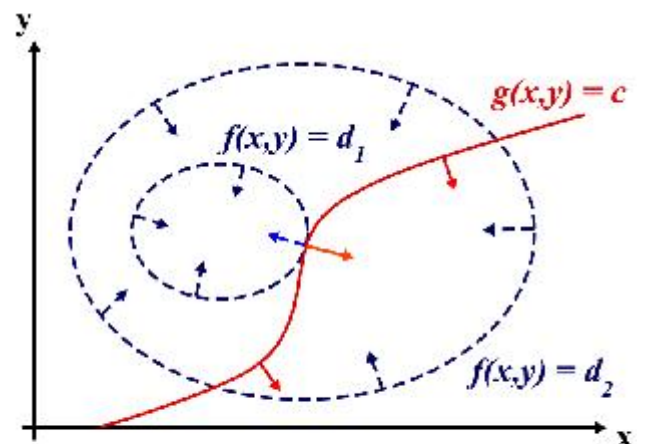
Για παράδειγμα ας θεωρηθεί το παρακάτω πρόβλημα:

$$\max f(x, y), \quad \text{E} \quad g(x, y) = c$$

Το πρόβλημα απεικονίζεται στο σχήμα 2.3.



(α)



(β)

Σχήμα 2.3: (α) 3D απεικόνιση προβλήματος (β) ο χάρτης των ισοϋψών καμπυλών του προβλήματος. Με κόκκινη γραμμή φαίνεται ο περιορισμός $g(x, y) = c$, ενώ με μπλε χρώμα οι ισο-γραμμές της $f(x, y)$. Το σημείο όπου η κόκκινη γραμμή εφάπτεται της μπλε αποτελεί τη λύση του προβλήματος και επειδή $d_1 > d_2$ η $f(x, y)$ μεγιστοποιείται, [5].

Αν υποθέσουμε ότι κινούμαστε πάνω στην καμπύλη $g(x, y) = c$, θα διαπιστώσουμε ότι όσο προχωράμε από το ένα σημείο της εν λόγω καμπύλης στο επόμενο, η τιμή της f δύναται να αλλάζει. Μόνο στην περίπτωση που η καμπύλη $g(x, y) = c$ συναντήσει εφαπτομενικά μία ισο-γραμμή της f , η τιμή της f παραμένει σταθερή. Οι δύο εν λόγω καμπύλες εφάπτονται όταν τα εφαπτομενικά σε αυτές διανύσματα είναι παράλληλα και επειδή το $grad$ μίας συνάρτησης είναι κάθετο στις ισο-γραμμές της αυτό ισοδυναμεί με το να ειπωθεί ότι τα διανύσματα ∇f και ∇g είναι παράλληλα.

Επόμενο βήμα της παρούσας ανάλυσης είναι η επαναδιατύπωση του προβλήματος (2.2) ως ένα πρόβλημα στο οποίο αναζητούνται τα στάσιμα σημεία μίας συνάρτησης Lagrange, η οποία ορίζεται ως, [4]:

$$L(\vec{x}, \lambda_1) = f(\vec{x}) - \lambda_1 c_1(\vec{x}) \quad (2.4)$$

Η εύρεση των στάσιμων σημείων της εν λόγω συνάρτησης Lagrange προκύπτει από το μηδενισμό της κλίσης της:

$$\nabla L(\vec{x}, \lambda_1) = \nabla f(\vec{x}) - \lambda_1 \nabla c_1(\vec{x}) = 0 \quad (2.5)$$

Η σχέση (2.5) παραπέμπει ευθέως στην ικανοποίηση της σχέσης (2.3) για κάποια τιμή του λ_1^* . Συνεπώς, οδηγείται κανείς στο συμπέρασμα ότι ένας τρόπος να επιλυθεί ένα πρόβλημα βελτιστοποίησης με ένα περιορισμό ισότητας είναι μέσω της αναζήτησης στάσιμων σημείων της συνάρτησης Lagrange (2.4). Πρέπει, όμως, να τονιστεί το γεγονός ότι η εν λόγω συνθήκη είναι απλά αναγκαία και όχι ικανή.

2.4.2.2 Ελαχιστοποίηση με έναν περιορισμό ανισότητας

Αρχικά, κρίνεται απαραίτητο να οριστούν οι έννοιες 'ενεργός περιορισμός' και 'ανενεργός περιορισμός', οι οποίες συναντώνται συχνά όταν γίνεται αναφορά σε προβλήματα βελτιστοποίησης με περιορισμούς ανισότητας, [4]:

Στην υποψήφια λύση $\vec{x} \in X$, ο i -ιστός περιορισμός ανισότητας είναι ενεργός αν και μόνο αν $c_i(\vec{x}) = 0$.

Στην υποψήφια λύση $\vec{x} \in X$, ο i -ιστός περιορισμός ανισότητας είναι ανενεργός αν και μόνο αν $c_i(\vec{x}) < 0$.

Στη συνέχεια, αντικαθίσταται ο περιορισμός ισότητας του προβλήματος (2.2) με έναν περιορισμό ανισότητας και διατυπώνεται η αναγκαία συνθήκη για τη βέλτιστη λύση σε ένα πρόβλημα ελαχιστοποίησης με έναν περιορισμό ανισότητας, [4]:

Αναγκαία συνθήκη ώστε το διάνυσμα \vec{x}^ να αποτελεί βέλτιστη λύση στο πρόβλημα:*

$$\begin{aligned} \min_{\vec{x} \in \mathbb{R}^n} f(\vec{x}), \quad & \text{εφ} \quad c_1(\vec{x}) \leq 0 \end{aligned}$$

είναι η:

$$\nabla f(\vec{x}^*) = \lambda_1^* \nabla c_1(\vec{x}^*), \quad \lambda_1^* \leq 0 \quad (2.6)$$

και παράλληλα πρέπει να ισχύει:

$$\lambda_1^* c_1(\vec{x}^*) = 0 \quad (2.7)$$

Η συνθήκη (2.6) σχετίζεται σαφέστατα με τις σχέσεις (2.4) και (2.5), δηλαδή με τη χρήση της συνάρτησης Lagrange. Διαφέρει δε από την αντίστοιχη συνθήκη για τη βελτιστοποίηση με περιορισμό ισότητας (2.3) ως προς την επιπλέον απαίτηση να είναι η ποσότητα λ_1^* μη-θετική. Αξίζει να επισημανθεί ότι, στην περίπτωση βελτιστοποίησης με περιορισμό ισότητας, είναι λογικό το πρόσημο του λ_1^* να μην έχει σημασία, καθώς ο περιορισμός ισότητας $c_1(\vec{x}) = 0$ μπορεί, επίσης, να γραφεί στη μορφή $-c_1(\vec{x}) = 0$. Συνεπώς, κάθε προσπάθεια άνιλησης οιασδήποτε πληροφορίας από το πρόσημο της ποσότητας λ_1^* είναι μάταιη, αφού και το $\nabla c_1(\vec{x})$ αλλάζει πρόσημο. Συγχρόνως με τη σχέση (2.6), όμως, στην περίπτωση που έχουμε περιορισμό ανισότητας, επιβάλλεται και η συμπληρωματική συνθήκη (2.7). Η εν λόγω συνθήκη μας πληροφορεί ότι θα ισχύει $\lambda_1^* < 0$ μόνο αν ο περιορισμός $c_1(\vec{x})$ είναι ενεργός (δηλαδή ισούται με 0), ενώ αν ο περιορισμός $c_1(\vec{x})$ είναι ανενεργός επιβάλλεται $\lambda_1^* = 0$.

2.4.2.3 Αναλυτική προσέγγιση - Συνθήκες Karush-Kuhn-Tucker

Στην ενότητα 2.4.2.1, έγινε διατύπωση της συνάρτησης Lagrange (σχέση 2.4) στην περίπτωση ενός προβλήματος βελτιστοποίησης με έναν περιορισμό ισότητας. Γενικεύοντας, ανεξάρτητα του πλήθους των περιορισμών αλλά και από το αν πρόκειται για περιορισμούς τύπου ισότητας ή ανισότητας, η συνάρτηση Lagrange γράφεται:

$$L(\vec{x}, \vec{\lambda}) = f(\vec{x}) - \sum_{i \in E \cup I} \lambda_i c_i(\vec{x}) \quad (2.8)$$

όπου με $\vec{\lambda}$ εκφράζεται το διάνυσμα των πολλαπλασιαστών Lagrange λ_i . Οι πρώτης τάξης αναγκαίες συνθήκες για την ύπαρξη βέλτιστης λύσης της συνάρτησης (2.8) διατυπώνονται στη μορφή του ακόλουθου θεωρήματος, [1, 4, 5, 6]:

Αν \vec{x}^* είναι μία τοπική λύση του προβλήματος:

$$\min f(\vec{x}), \quad \vec{x} \in X \subset \mathbb{R}^n \quad (2.9)$$

τότε υπάρχει ένα διάνυσμα πολλαπλασιαστών Lagrange $\vec{\lambda}^*$ με συνιστώσες $\lambda_i^*, i \in E \cup I$, τέτοιο ώστε στο $(\vec{x}^*, \vec{\lambda}^*)$ να ισχύουν οι παρακάτω συνθήκες:

$$\begin{aligned} \nabla_x L(\vec{x}^*, \vec{\lambda}^*) = 0 &\Leftrightarrow \nabla f(\vec{x}^*) - \sum_{i \in E \cup I} \lambda_i^* \nabla c_i(\vec{x}^*) = 0 \\ c_i(\vec{x}^*) &= 0, \quad \forall i \in E \\ c_i(\vec{x}^*) &\leq 0, \quad \forall i \in I \\ \lambda_i^* &\leq 0, \quad \forall i \in I \\ \lambda_i^* c_i(\vec{x}^*) &= 0, \quad \forall i \in E \cup I \end{aligned} \quad (2.10)$$

Οι συνθήκες (2.10), οι οποίες είναι αναγκαίες προϋποθέσεις ύπαρξης ακροτάτου ενός προβλήματος βελτιστοποίησης με περιορισμούς, είναι γνωστές ως συνθήκες Karush-Kuhn-Tucker (ή KKT). Πρέπει να αναφερθεί ότι οι εν λόγω συνθήκες είναι ικανές και αναγκαίες για την ύπαρξη ολικού ελαχίστου της f , εφόσον τόσο η συνάρτηση όσο και οι περιορισμοί ανισότητας είναι κυρτές συναρτήσεις, ενώ οι περιορισμοί ισότητας είναι γραμμικοί.

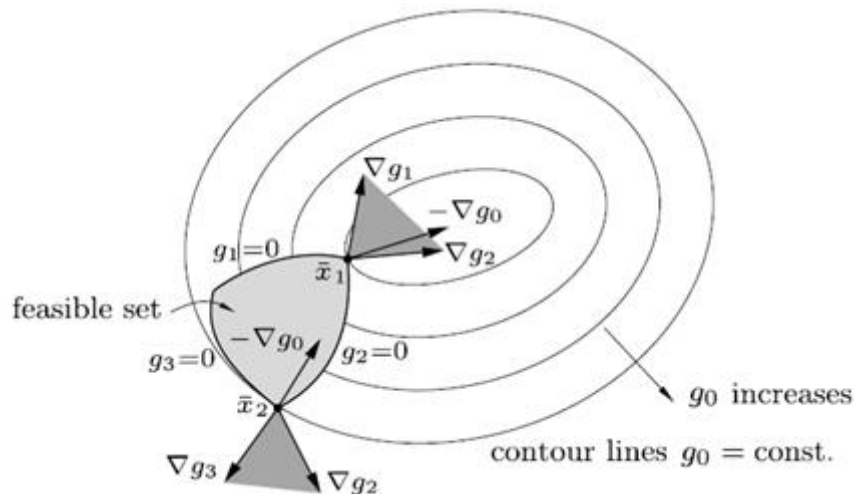
Η πρώτη συνθήκη KKT ονομάζεται συνθήκη στασιμότητας (stationarity) ή συνθήκη βελτιστότητας (optimality) και σε ένα πρόβλημα χωρίς περιορισμούς θα αντιστοιχούσε στη συνήθη απαίτηση $\nabla f(\vec{x}^*) = 0$. Η εν λόγω συνθήκη υποδηλώνει ότι η μετατόπιση από το σημείο \vec{x}^* προς οποιαδήποτε κατεύθυνση βελτίωσης της τιμής της συνάρτησης είναι μη-εφικτή και, ταυτόχρονα, οποιαδήποτε μετατόπιση εντός του εφικτού χώρου χειροτερεύει την τιμή της συνάρτησης. Σημειώνεται ότι η συνθήκη βελτιστότητας δίνει n εξισώσεις.

Η δεύτερη και η τρίτη εκ των συνθηκών (2.10) καλούνται συνθήκες πρωταρχικής εφικτότητας (primal feasibility) και απαιτούν, απλά, την τήρηση των περιορισμών.

Η τέταρτη συνθήκη ονομάζεται συνθήκη διπλής εφικτότητας (dual feasibility).

Η τελευταία από τις συνθήκες (2.10) καλείται συνθήκη συμπληρωματικής χαλαρότητας (complementary slackness condition). Η συνθήκη αυτή δηλώνει ότι στο βέλτιστο σημείο για κάθε περιορισμό ανισότητας, δηλαδή $\forall i \in I$, ο πολλαπλασιαστής Lagrange λ_i^* μπορεί να πάρει την τιμή 0 μόνο όταν ο εν λόγω περιορισμός ανισότητας είναι ανενεργός, ενώ αν ο περιορισμός ανισότητας είναι ενεργός, λαμβάνοντας υπόψη και την προτελευταία συνθήκη KKT (συνθήκη μη-θετικότητας), ο πολλαπλασιαστής Lagrange λ_i^* μπορεί να λάβει αρνητική τιμή. Η συνθήκη συμπληρωματικής χαλαρότητας ορίζει ότι η βέλτιστη λύση \vec{x}^* κείται πάνω στον γεωμετρικό τόπο που ορίζουν οι εξισώσεις $c_i(\vec{x}^*) = 0$. Οι περιορισμοί αυτοί είναι δεσμευτικοί, σε αντίθεση με τους υπόλοιπους περιορισμούς που είναι χαλαροί. Συνεπώς, αν στο πρόβλημα βελτιστοποίησης περιέχονται περιορισμοί ισότητας, αυτοί θα είναι εξ ορισμού δεσμευτικοί. Πρέπει, ακόμα, να αναφερθεί ότι, λόγω της συνθήκης συμπληρωματικής χαλαρότητας, η βέλτιστη λύση της συνάρτησης Lagrange (2.8) ταυτίζεται με αυτήν της $f(\vec{x})$, δηλαδή $L(\vec{x}^*, \vec{\lambda}^*) = f(\vec{x}^*)$. Συνεπώς, για να δύναται κανείς να διαχειριστεί τους περιορισμούς που ενδέχεται να υπάρχουν σε ένα πρόβλημα βελτιστοποίησης, εισάγονται οι επιπλέον μεταβλητές λ_i .

Ακολούθως, παρουσιάζεται μία γεωμετρική ερμηνεία των συνθηκών Karush-Kuhn-Tucker. Οι συνθήκες KKT δηλώνουν ότι σε ένα βέλτιστο σημείο \vec{x}^* το $-\nabla f(\vec{x}^*)$ περιέχεται στον κώνο που σχηματίζεται από τις κλίσεις (gradients) των ενεργών περιορισμών.



Σχήμα 2.4: Απεικόνιση συνθηκών KKT, [5].

Στο σχήμα 2.4 φαίνεται ένα παράδειγμα, όπου g_0 είναι η συνάρτηση στόχου και g_1 , g_2 και g_3 είναι οι ενεργοί περιορισμοί. Στο σημείο \bar{x}_1 του εν λόγω σχήματος ισχύει:

$$-\nabla g_0(\bar{x}_1) = -\lambda_1 \nabla g_1(\bar{x}_1) - \lambda_2 \nabla g_2(\bar{x}_1), \quad \lambda_1 \leq 0, \lambda_2 \leq 0$$

Συνεπώς, το σημείο \bar{x}_1 ικανοποιεί τις συνθήκες KKT, καθώς παρατηρείται ότι εκεί το διάνυσμα $-\nabla g_0(\bar{x}_1)$ ανήκει στον κώνο που σχηματίζουν οι δύο ενεργοί περιορισμοί g_1 και g_2 . Όσον αφορά το σημείο \bar{x}_2 , διαπιστώνεται ότι το διάνυσμα $-\nabla g_0(\bar{x}_2)$ δεν περιέχεται στον κώνο που σχηματίζεται από τους ενεργούς περιορισμούς στο \bar{x}_2 . Αυτό σημαίνει ότι δεν υπάρχουν $\lambda_2 \leq 0$ και $\lambda_3 \leq 0$ τέτοια ώστε να ισχύει:

$$-\nabla g_0(\bar{x}_2) = -\lambda_2 \nabla g_2(\bar{x}_2) - \lambda_3 \nabla g_3(\bar{x}_2)$$

δηλαδή το \bar{x}_2 δεν ικανοποιεί τις συνθήκες KKT και άρα δεν αποτελεί βέλτιστη λύση.

Τέλος, κρίνεται σκόπιμο να αναφερθεί ότι στην περίπτωση που οι περιορισμοί ανισότητας διατυπωθούν στη μορφή $c_i(\vec{x}^*) \geq 0$, αντί της μορφής $c_i(\vec{x}^*) \leq 0$ οι συνθήκες KKT γράφονται:

$$\nabla_x L(\vec{x}^*, \vec{\lambda}^*) = 0 \Leftrightarrow \nabla f(\vec{x}^*) - \sum_{i \in E \cup I} \lambda_i^* \nabla c_i(\vec{x}^*) = 0$$

$$c_i(\vec{x}^*) = 0, \quad \forall i \in E$$

$$c_i(\vec{x}^*) \geq 0, \quad \forall i \in I$$

$$\lambda_i^* \geq 0, \quad \forall i \in I$$

$$\lambda_i^* c_i(\vec{x}^*) = 0, \quad \forall i \in E \cup I$$

2.4.2.4 Ικανές συνθήκες

Οι ικανές συνθήκες (sufficient conditions) ώστε να είναι ένα σημείο \vec{x}^* αυστηρό τοπικό ελάχιστο του προβλήματος (2.9), όπου οι συναρτήσεις f και c_i , $i \in E \cup I$ είναι δύο φορές παραγωγίσιμες (twice differentiable) διατυπώνονται ως εξής, [1, 6]:

- να ικανοποιούνται οι αναγκαίες συνθήκες KKT και
- το εσσιανό μητρώο

$$\nabla_x^2 L(\vec{x}^*, \vec{\lambda}^*) = \nabla^2 f(\vec{x}^*) - \sum_{i \in E \cup I} \lambda_i^* \nabla^2 c_i(\vec{x}^*)$$

να είναι θετικά ορισμένο σε έναν υπόχωρο (subspace) του \mathbb{R}^n , όπως ορίζεται από τη συνθήκη:

$y^T \nabla_x^2 L(\vec{x}^*, \vec{\lambda}^*) y \geq 0$ να ισχύει για κάθε διάνυσμα y με διαστάσεις $n \times 1$ που ικανοποιεί τις συνθήκες:

- i. $\nabla c_j(\vec{x}^*) y = 0, j \in I_1 = \{j | c_j(\vec{x}^*) = 0, \lambda_j^* < 0\}$, δηλαδή το j ανήκει στο σύνολο των ενεργών περιορισμών,
- ii. για κάθε περιορισμό ισότητας $\nabla c_i(\vec{x}^*) y = 0, i \in E$
- iii. και $y \neq 0$.

2.4.3 Σειριακή ελαχιστοποίηση για προβλήματα με περιορισμούς – Συναρτήσεις ποινής

Για την αντιμετώπιση ενός προβλήματος βελτιστοποίησης το οποίο διέπεται από περιορισμούς ισότητας ή/και ανισότητας έχει αναπτυχθεί ένας αριθμός μεθόδων. Όπως αναφέρθηκε στην ενότητα 2.4.1, γενική στρατηγική τέτοιων μεθόδων αποτελεί ο μετασχηματισμός του αρχικού προβλήματος βελτιστοποίησης υπό περιορισμούς σε ένα αντίστοιχο πρόβλημα χωρίς περιορισμούς. Κατά συνέπεια, αν είναι διαθέσιμη μία οιαδήποτε μέθοδος βελτιστοποίησης ικανή να επιλύσει προβλήματα χωρίς περιορισμούς, αυτή δύναται να χρησιμοποιηθεί και για την επίλυση προβλημάτων που περιλαμβάνουν περιορισμούς. Το κόστος που απορρέει από την ανωτέρω στρατηγική είναι το γεγονός ότι καλείται κανείς να επιλύσει σειριακά περισσότερα από ένα προβλήματα βελτιστοποίησης χωρίς περιορισμούς.

Στην παρούσα ενότητα γίνεται αναφορά στις μεθόδους που για την ενσωμάτωση των περιορισμών στη συνάρτηση στόχου χρησιμοποιούν συναρτήσεις ποινής (penalty function methods). Σύμφωνα με τον Pierre (1986), [58]:

Συνάρτηση ποινής είναι μια αριθμητική έκφραση που εισάγεται στο μέτρο επίδοσης ενός συστήματος, με σκοπό την τεχνητή «χειροτέρευση» του εν λόγω μέτρου, στην περίπτωση παραβίασης των περιορισμών.

Σύμφωνα με τις εν λόγω μεθόδους για την επίλυση του ακόλουθου προβλήματος που εμπεριέχει περιορισμούς:

$$\min_{\vec{x} \in \mathbb{R}^n} f(\vec{x}), \quad \text{t.} \quad c_i(\vec{x}) = 0, i \in E, \quad c_i(\vec{x}) \leq 0, i \in I$$

καταφεύγει κανείς στον ορισμό της ψευδο-συνάρτησης στόχου (pseudo-objective function) $\Phi(\vec{x}, \omega_p)$ σύμφωνα με τη σχέση, [4]:

$$\Phi(\vec{x}, \omega_p) = f(\vec{x}) + \omega_p P(\vec{x})$$

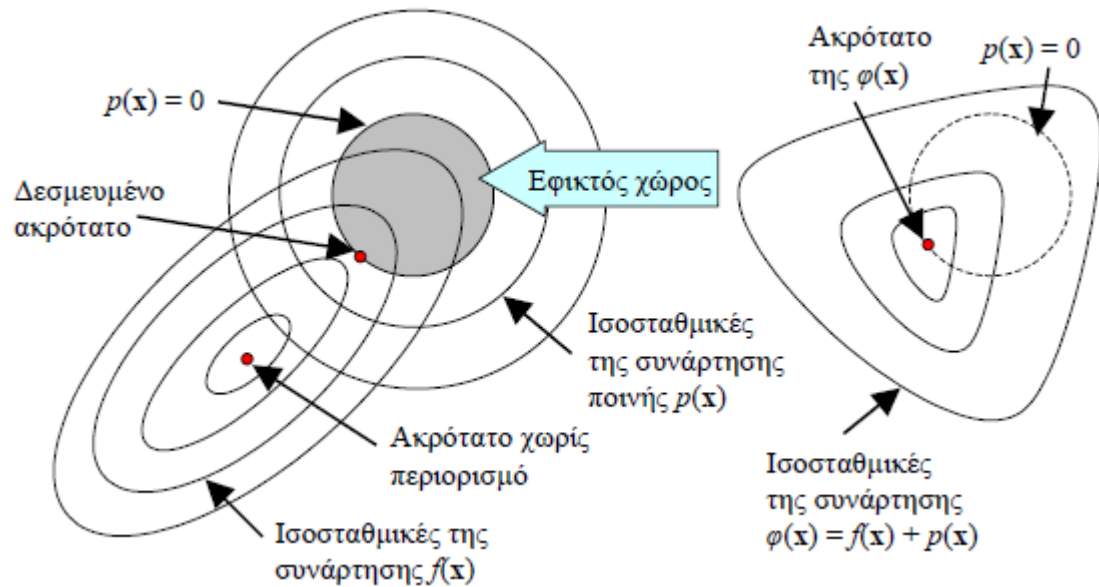
όπου ω_p ο λεγόμενος συντελεστής ή βάρος ποινής (penalty weight) και $P(\vec{x})$ η συνάρτηση ποινής, η έκφραση της οποίας εξαρτάται από τη παραλλαγή της μεθόδου που χρησιμοποιείται. Ο συντελεστής ω_p ,

συνήθως, διατηρεί σταθερή τιμή κατά τη διάρκεια μίας εκ των σειριακών κλήσεων της μεθόδου βελτιστοποίησης χωρίς περιορισμούς, η οποία καλείται κύκλος βελτιστοποίησης, και αλλάζει τιμή στην επόμενη, ενώ υπάρχει και η δυνατότητα να παραμείνει σταθερός σε όλους τους κύκλους βελτιστοποίησης.

Συνήθεις παραλλαγές μεθόδων που επιλύουν προβλήματα βελτιστοποίησης με περιορισμούς πραγματοποιώντας σειριακές κλήσεις μίας μεθόδου βελτιστοποίησης για προβλήματα χωρίς περιορισμούς και κάνοντας χρήση της $\Phi(\vec{x}, \omega_p)$ είναι οι εξής:

- Η μέθοδος της εξωτερικής ποινής, κατά την οποία η συνάρτηση ποινής $P(\vec{x})$ αποκτά μη-μηδενική τιμή και άρα 'τιμωρεί' τη συνάρτηση στόχου επαυξάνοντάς την, μόνο όταν δεν ικανοποιείται ένας περιορισμός.
- Η μέθοδος της εσωτερικής ποινής, κατά την οποία η συνάρτηση ποινής $P(\vec{x})$ αποκτά μη-μηδενική τιμή και άρα 'τιμωρεί' τη συνάρτηση στόχου επαυξάνοντάς την, όταν η υποψήφια λύση πλησιάζει τη μη-αποδεκτή περιοχή, έστω και αν αυτή είναι ακόμα αποδεκτή. Στην εν λόγω μέθοδο δεν επιτρέπεται να διαταραχθεί οποιοσδήποτε περιορισμός και η βέλτιστη λύση προκύπτει από τη διαδοχή αποδεκτών λύσεων (δηλαδή η βέλτιστη λύση προσεγγίζεται από την αποδεκτή περιοχή).
- Η μέθοδος της διευρυμένης εσωτερικής ποινής, η οποία συνδυάζει τα πλεονεκτήματα και αποφεύγει τα μειονεκτήματα των δύο ανωτέρω μεθόδων.

Στο σχήμα 2.5 παρουσιάζεται η γεωμετρική απεικόνιση της κυρτής συνάρτησης στόχου $f(\vec{x})$ στο πεδίο αναζήτησης X ταυτόχρονα με αυτήν της κυρτής συνάρτησης ποινής $P(\vec{x})$. Λόγω της κυρτότητας, η τομή των δύο επιφανειών δημιουργεί μια επίσης κυρτή επιφάνεια. Παρατηρείται ότι η ελάχιστη τιμή της συνάρτησης $\Phi(\vec{x}) = f(\vec{x}) + P(\vec{x})$ βρίσκεται στην τομή της αρχικής συνάρτησης $f(\vec{x})$ με την ισοσταθμική $P(\vec{x}) = 0$.



Σχήμα 2.5: Γραφική απεικόνιση της συνάρτησης στόχου f , της συνάρτησης ποινής P και της βοηθητικής συνάρτησης Φ , [2].

2.4.4 Η επαυξημένη μέθοδος με πολλαπλασιαστές Lagrange

Στην ενότητα 2.4.3 αναφέρθηκαν συνοπτικά οι μέθοδοι που κάνουν χρήση συναρτήσεων ποινής, οι οποίες είναι οι πιο κλασικές για την επίλυση προβλημάτων βελτιστοποίησης με περιορισμούς, χαρακτηρίζονται από εύκολο προγραμματισμό και είναι, γενικά, αρκετά αποδοτικές. Όμως, θα ήταν δυνατό να εξασφαλιστεί μεγαλύτερη αξιοπιστία και απόδοση, αν συμμετείχαν στον αλγόριθμο βελτιστοποίησης και οι παράμετροι-συνθήκες της βέλτιστης λύσης. Με άλλα λόγια, θα ήταν επιθυμητό η διαδικασία της βελτιστοποίησης να παύσει να εξαρτάται σε μεγάλο βαθμό από την επιλογή βοηθητικών παραμέτρων, όπως λ.χ. είναι οι συντελεστές ποινής. Κάτι τέτοιο επιτυγχάνεται με τις λεγόμενες επαυξημένες μεθόδους με χρήση πολλαπλασιαστών Lagrange ή συντομογραφικά μέθοδοι ALM (Augmented Lagrange Multiplier Methods).

2.4.4.1 Η μέθοδος ALM για περιορισμούς ισότητας

Αρχικά, διατυπώνεται η μέθοδος ALM στην περίπτωση που το πρόβλημα βελτιστοποίησης περιλαμβάνει μόνο περιορισμούς ισότητας. Έστω, λοιπόν, ότι τίθεται προς αντιμετώπιση το ακόλουθο πρόβλημα:

$$\min_{\vec{x} \in \mathbb{R}^n} f(\vec{x}), \quad \text{s.t.} \quad c_i(\vec{x}) = 0, i \in E$$

Επισημαίνεται ότι ο αριθμός των περιορισμών ισότητας που επιβάλλονται πρέπει να είναι μικρότερος του αριθμού n των μεταβλητών σχεδιασμού που είναι οι ελεύθερες παράμετροι του προβλήματος. Αν υπήρχαν ακριβώς n περιορισμοί ισότητας, για τον υπολογισμό των τιμών των ελεύθερων παραμέτρων του προβλήματος θα αρκούσε η επίλυση του συστήματος των n μη-γραμμικών, γενικά, εξισώσεων των περιορισμών. Στην περίπτωση που το πρόβλημα βελτιστοποίησης περιελάμβανε περισσότερους από n περιορισμούς ισότητας, η λύση του συστήματος θα ήταν απλά αδύνατη.

Για τη θεμελίωση της μεθόδου ALM, αφετηρία αποτελούν οι αναγκαίες συνθήκες Karush-Kuhn-Tucker (σχέσεις 2.10). Έτσι, η ψευδο-συνάρτηση στόχου στην οποία βασίζεται η μέθοδος ALM διατυπώνεται ως εξής, [4]:

$$\Phi^*(\vec{x}, \vec{\lambda}, \omega_p) = f(\vec{x}) + \sum_{i \in E} \{-\lambda_i c_i(\vec{x}) + \omega_p [c_i(\vec{x})]^2\}$$

Η $\Phi^*(\vec{x}, \vec{\lambda}, \omega_p)$ καλείται επαυξημένη συνάρτηση Lagrange.

Όσον αφορά τις τιμές των πολλαπλασιαστών Lagrange και του συντελεστή ποινής, υιοθετείται ένα επαναληπτικό σχήμα το οποίο ξεκινάει από τυχαίες τιμές λ_i (0 είναι μία συνηθισμένη επιλογή) και ω_p (συνήθως πρόκειται για μικρή τιμή) και προσεγγίζει τη λύση ανανεώνοντας τις τιμές των λ_i και ω_p σε κάθε κύκλο βελτιστοποίησης.

2.4.4.2 Η μέθοδος ALM για περιορισμούς ανισότητας

Στη συνέχεια, περιγράφεται η μέθοδος ALM στην περίπτωση που το πρόβλημα βελτιστοποίησης περιλαμβάνει μόνο περιορισμούς ανισότητας. Έστω, λοιπόν, ότι επιβάλλεται να επιλυθεί το ακόλουθο πρόβλημα:

$$\begin{aligned} \min f(\vec{x}), \quad & \text{ε} \quad c_i(\vec{x}) \leq 0, i \in I \\ \vec{x} \in \mathbb{R}^n \end{aligned}$$

Στην περίπτωση αυτή, το πρώτο βήμα είναι να μετατραπεί το ανωτέρω πρόβλημα σε ένα ισοδύναμο πρόβλημα βελτιστοποίησης με περιορισμούς ισότητας. Κάτι τέτοιο πραγματοποιείται αν εισάγουμε επιπλέον μεταβλητές σχεδιασμού και συγκεκριμένα μία για κάθε περιορισμό ανισότητας. Έτσι, κάθε περιορισμός της μορφής $c_i(\vec{x}) \leq 0$ μπορεί να γραφτεί:

$$c_i(\vec{x}) + z_i^2 = 0, \quad i \in I$$

Η νέα μεταβλητή z_i είναι υψωμένη στο τετράγωνο, καθώς για την μετατροπή του ανισοτικού περιορισμού $c_i(\vec{x}) \leq 0$ σε περιορισμό

ισότητας πρέπει η ποσότητα που προστίθεται στο αριστερό μέλος να είναι θετική.

Λαμβάνοντας υπόψη τα παραπάνω, η επαυξημένη συνάρτηση Lagrange, σε αυτή την περίπτωση, διατυπώνεται ως εξής, [4]:

$$\Phi^*(\vec{x}, \vec{\lambda}, \vec{z}, \omega_p) = f(\vec{x}) + \sum_{i \in I} \left\{ -\lambda_i [c_i(\vec{x}) + z_i^2] + \omega_p [c_i(\vec{x}) + z_i^2]^2 \right\}$$

Για τη διατύπωση της $\Phi^*(\vec{x}, \vec{\lambda}, \vec{z}, \omega_p)$ συναντάται η επιβάρυνση με ένα επιπλέον κόστος που είναι η εισαγωγή των ψευδο-μεταβλητών σχεδιασμού z_i .

Εναλλακτική διατύπωση της παραπάνω έκφρασης είναι η εξής, [4]:

$$\Phi^*(\vec{x}, \vec{\lambda}, \omega_p) = f(\vec{x}) + \sum_{i \in I} (-\lambda_i \psi_i + \omega_p \psi_i^2)$$

όπου

$$\psi_i = \max \left[c_i(\vec{x}), \frac{\lambda_i}{2\omega_p} \right]$$

Η σχέση αυτή μας δίνει την ψευδο-συνάρτηση στόχου, πάνω στην οποία βασίζεται η μέθοδος ALM με περιορισμούς ανισότητας. Σημειώνεται ότι η εν λόγω συνάρτηση έχει συνεχείς πρώτες παραγώγους ως προς \vec{x} αλλά ασυνεχείς δεύτερες παραγώγους στο σημείο $c_i = \lambda_i / (2\omega_p)$. Συνεπώς, απαιτείται προσοχή σε περίπτωση χρήσης κάποιας μεθόδου βελτιστοποίησης δεύτερης τάξης.

2.4.4.3 Γενίκευση της μεθόδου ALM

Στην παρούσα ενότητα, παρουσιάζεται η γενίκευση της επαυξημένης μεθόδου με χρήση πολλαπλασιαστών Lagrange (ALM) για την περίπτωση που το πρόβλημα βελτιστοποίησης περιλαμβάνει ταυτόχρονα περιορισμούς ισότητας και ανισότητας. Το πρόβλημα, λοιπόν, προς επίλυση είναι:

$$\min_{\vec{x} \in \mathbb{R}^n} f(\vec{x}), \quad \text{t. s.} \quad c_i(\vec{x}) = 0, i \in E, \quad c_i(\vec{x}) \leq 0, i \in I$$

Στην περίπτωση αυτή η επαυξημένη συνάρτηση Lagrange διατυπώνεται στη μορφή, [4]:

$$\begin{aligned} \Phi^*(\vec{x}, \vec{\lambda}, \omega_p) = & f(\vec{x}) + \sum_{i \in E} \left\{ -\lambda_i c_i(\vec{x}) + \omega_p [c_i(\vec{x})]^2 \right\} \\ & + \sum_{i \in I} (-\lambda_i \psi_i + \omega_p \psi_i^2) \end{aligned} \quad (2.11)$$

όπου

$$\psi_i = \max \left[c_i(\vec{x}), \frac{\lambda_i}{2\omega_p} \right] \quad (2.12)$$

Στη συνέχεια, παρατίθεται ο αλγόριθμος της μεθόδου ALM για προβλήματα βελτιστοποίησης με περιορισμούς ισότητας και ανισότητας, [4]:

- Βήμα 1: Επιλέγεται αρχική τιμή του διανύσματος των μεταβλητών σχεδιασμού \vec{x}^0 , αρχική τιμή του διανύσματος των πολλαπλασιαστών Lagrange $\vec{\lambda}^0$ (συνήθως τιμή είναι η μηδενική) και αρχική (συνήθως μικρή) τιμή του συντελεστή ποινής ω_p . Συγχρόνως, επιλέγεται τιμή του συντελεστή, έστω γ , με τον οποίο θα πολλαπλασιάζεται η τιμή του ω_p σε κάθε νέο κύκλο βελτιστοποίησης αλλά και η ανώτατη τιμή που θα επιτρέπεται να λάβει ο συντελεστή ποινής ω_p .
- Βήμα 2: Ελαχιστοποιείται η ψευδο-συνάρτηση στόχου $\Phi^*(\vec{x}, \vec{\lambda}, \omega_p)$ που δίνεται από τη σχέση (2.11), σε συνδυασμό με τη σχέση (2.12). Επιλύεται, δηλαδή, ένα πρόβλημα χωρίς περιορισμούς, επιλέγοντας οιαδήποτε κατάλληλη μέθοδο.
- Βήμα 3: Ελέγχεται η σύγκλιση της μεθόδου ALM. Τυπικό κριτήριο είναι η τιμή της συνάρτησης στόχου $f(\vec{x})$ να μεταβάλλεται πολύ λίγο (κρίνεται από το χρήστη) σε σχέση με τη τιμή της στον προηγούμενο κύκλο βελτιστοποίησης (αν δεν πρόκειται για τον πρώτο κύκλο) και συγχρόνως να ικανοποιούνται οι περιορισμοί του προβλήματος με την ανοχή που ορίζει ο χρήστης. Είναι φανερό ότι αν η μέθοδος έχει συγκλίνει ο αλγόριθμος σταματά. Σε αντίθετη περίπτωση, ο αλγόριθμος προχωράει στην εκτέλεση του επόμενου βήματος.
- Βήμα 4: Ανανεώνονται οι τιμές των πολλαπλασιαστών Lagrange σύμφωνα με τις σχέσεις:

$$\lambda_i \leftarrow \lambda_i - 2\omega_p c_i(\vec{x}), \quad i \in E$$

$$\lambda_i \leftarrow \lambda_i - 2\omega_p \max \left[c_i(\vec{x}), \frac{\lambda_i}{2\omega_p} \right], \quad i \in I$$

Αναφέρεται ότι οι σχέσεις αυτές βασίζονται σε θεωρητικές παραδοχές. Στη συνέχεια, ανανεώνεται και η τιμή του συντελεστή ποινής μέσω της σχέσης:

$$\omega_p \leftarrow \min(\gamma\omega_p, \omega_p^{max})$$

🌈 Βήμα 5: Επιστροφή στο Βήμα 2.

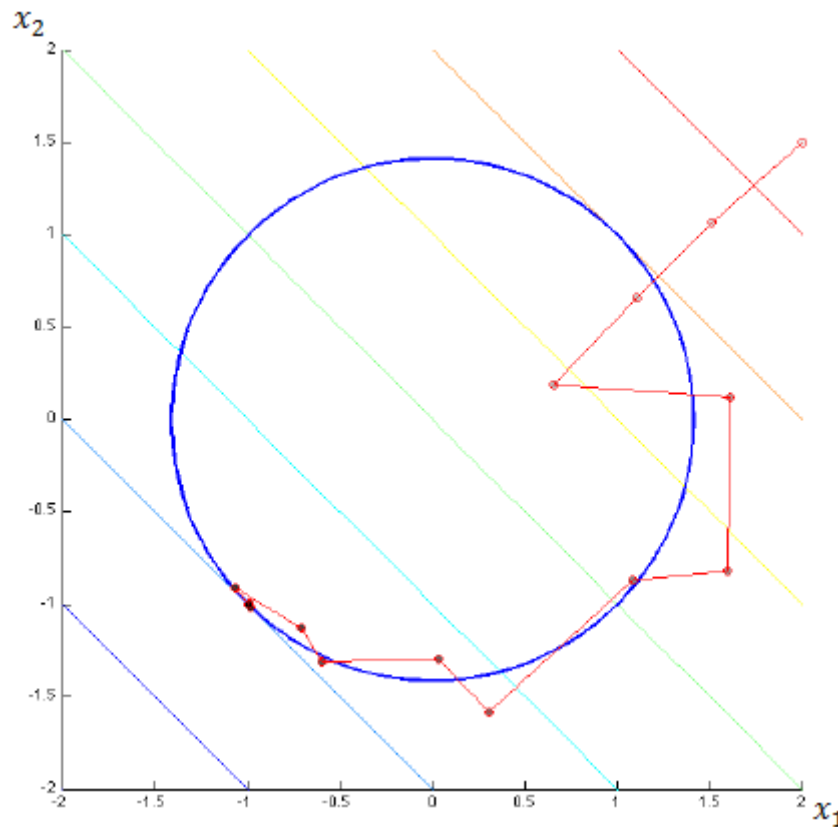
Ακολουθώντας, αναφέρονται κάποια πλεονεκτήματα της μεθόδου ALM:

- Είναι σχετικά ανεξάρτητη της τιμής του συντελεστή ποινής ω_p και δεν απαιτείται η προσεκτική αύξησή της τιμής αυτής από κύκλο σε κύκλο, με το ω_p να τείνει θεωρητικά στο ∞ .
- Παρέχει τη δυνατότητα εντοπισμού λύσεων που να ικανοποιούν ακριβώς τους περιορισμούς. Η βέλτιστη λύση που εντοπίζεται δίνει $c_i(\vec{x}) = 0, i \in E$ και ενδεχομένως, αν πράγματι ισχύει αυτό, $c_i(\vec{x}) = 0, i \in I$.
- Μπορεί να εξασφαλιστεί επιτάχυνση της σύγκλισης ανανεώνοντας τις τιμές των πολλαπλασιαστών Lagrange.
- Το σημείο εκκίνησης είναι δυνατό να βρίσκεται στη περιοχή των αποδεκτών ή των μη-αποδεκτών λύσεων.
- Φτάνοντας στη βέλτιστη λύση, η τιμή του λ_i για κάθε περιορισμό ανισότητας υποδεικνύει το αν ο εν λόγω περιορισμός ανισότητας είναι ενεργός ή όχι (ενεργός όταν $\lambda_i \neq 0$).

Τέλος, παρατίθεται ένα παράδειγμα στο οποίο γίνεται χρήση της μεθόδου ALM. Το πρόβλημα βελτιστοποίησης είναι το εξής:

$$\min_{\vec{x} \in \mathbb{R}^2} f(\vec{x}) = x_1 + x_2, \quad \text{t. s. } c(\vec{x}) = x_1^2 + x_2^2 - 2 = 0$$

Η αντικειμενική συνάρτηση του προβλήματος είναι ένα κεκλιμένο επίπεδο, ενώ ο περιορισμός ισότητας αναγκάζει τη λύση να βρίσκεται πάνω στον κύκλο με κέντρο το $(0,0)^T$ και ακτίνας ίσης με $\sqrt{2}$. Η βέλτιστη λύση είναι το σημείο $(-1,-1)^T$. Το σχήμα 2.6 καταδεικνύει της επαναλήψεις του αλγορίθμου της ALM, αν το σημείο εκκίνησης είναι το $(2,1.5)^T$. Ο αλγόριθμος τερματίζει μετά από 15 επαναλήψεις (κύκλους) προσεγγιστικά στη βέλτιστη λύση.



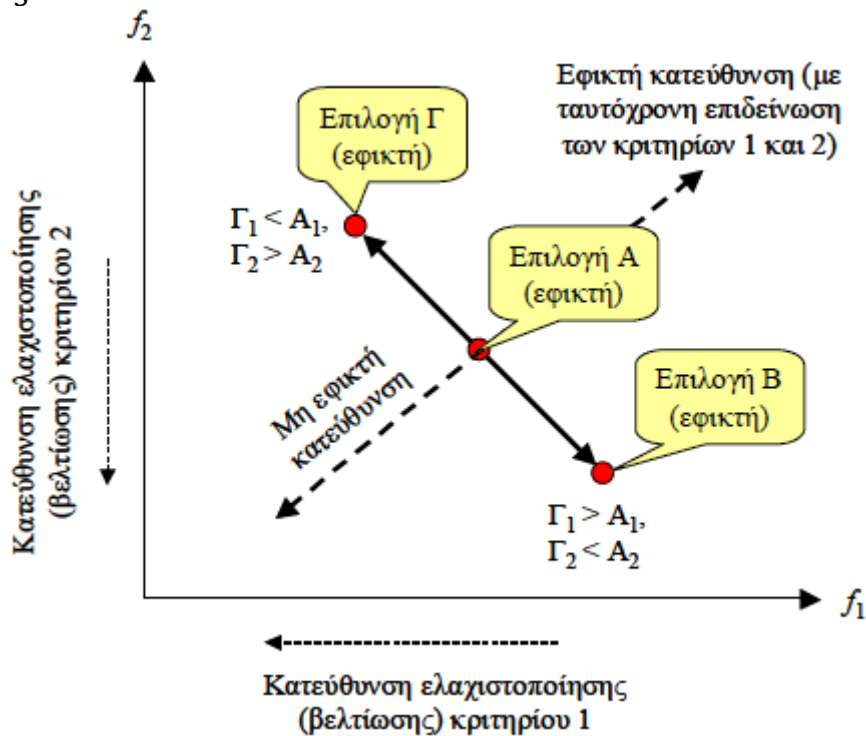
Σχήμα 2.6: Παράδειγμα χρήσης της μεθόδου ALM σε ένα πρόβλημα βελτιστοποίησης με ένα περιορισμό ισότητας, [7].

2.5 Πολυκριτηριακή βελτιστοποίηση

2.5.1 Εισαγωγή

Σε ένα τεράστιο φάσμα προβλημάτων βελτιστοποίησης που συναντώνται σε πολυάριθμες μηχανολογικές εφαρμογές αλλά και σε εφαρμογές πολλών άλλων επιστημονικών τομέων προκύπτει η ανάγκη ικανοποίησης περισσότερων του ενός και συχνά αντικρουόμενων στόχων σχεδιασμού. Έτσι, ‘αναδύεται’ η έννοια της πολυκριτηριακής βελτιστοποίησης (multi-objective optimization) ή διανυσματικής βελτιστοποίησης (vector optimization). Στην περίπτωση που συναντώνται προβλήματα πολλών ‘συγκρουόμενων’ στόχων, η λύση είναι ο συμβιβασμός, δηλαδή για να πραγματοποιηθεί βελτίωση ενός στόχου θα πρέπει να ‘θυσιάσει’ ένας άλλος (σχήμα 2.7). Για το λόγο αυτό, η ανάλυση των εν λόγω προβλημάτων αναφέρεται ως ανάλυση ανταλλαγής – συμβιβασμού (trade-off analysis). Στον τομέα της μηχανολογίας υπάρχουν αναρίθμητα παραδείγματα που υποδεικνύουν την ανάγκη εφαρμογής μίας τέτοιας ανάλυσης. Ένα από αυτά, είναι η τοποθέτηση περισσότερων λειτουργικών τμημάτων σε ένα τσιπ με ταυτόχρονη ελαχιστοποίηση της επιφάνειας του εν λόγω τσιπ και/ή της κατανάλωσης ισχύος. Ένα δεύτερο παράδειγμα

πολυκριτηριακής βελτιστοποίησης με αντικρουόμενους στόχους αποτελεί η κατασκευή ενός οχήματος που θα καλύπτει τη μέγιστη απόσταση ημερησίως, ενώ θα απαιτεί την ελάχιστη κατανάλωση ενέργειας.



Σχήμα 2.7: Γεωμετρική ερμηνεία της έννοιας των αντικρουόμενων στόχων. Αν π.χ. ληφθεί μια τρέχουσα επιλογή A, είναι δυνατή είτε η βελτίωση ενός κριτηρίου εις βάρος ενός άλλου (π.χ. επιλογές B και Γ), είτε η ταυτόχρονη επιδείνωση και των δύο κριτηρίων. Αντίθετα, η ταυτόχρονη βελτίωση του συνόλου των κριτηρίων δεν είναι πάντα εφικτή, [2].

Στη μονοκριτηριακή βελτιστοποίηση (single-objective optimization) ο χώρος των λύσεων είναι, συνήθως, εύκολα προσδιορίσιμος, ενώ υπάρχει μοναδική βέλτιστη λύση. Η εισαγωγή επιπλέον συναρτήσεων στόχου σε ένα πρόβλημα βελτιστοποίησης (πολυκριτηριακή βελτιστοποίηση) και η απαίτηση της ταυτόχρονης βελτιστοποίησης έκαστης, έχει σαν αποτέλεσμα τόσο την αύξηση του πλήθους των λύσεων (η βέλτιστη λύση δεν είναι μία αλλά πολλές) όσο και τη δυσχέρεια του ακριβούς προσδιορισμού του χώρου των λύσεων.

2.5.2 Διατύπωση του προβλήματος

Έστω το διανυσματικό μέτρο επίδοσης ενός συστήματος:

$$P = [P_1, P_2, \dots, P_m]^T = \vec{f}(\vec{x})$$

όπου $\vec{f}(\vec{x}) = [f_1(\vec{x}), f_2(\vec{x}), \dots, f_m(\vec{x})]^T$ διανυσματική συνάρτηση ορισμένη στο $X \subseteq \mathbb{R}^n$ και $\vec{x} = [x_1, x_2, \dots, x_n]^T$ διάνυσμα ανεξάρτητων

μεταβλητών. Οι συνιστώσες $P_i = f_i(\vec{x})$, $i = 1, \dots, m$ της διανυσματικής συνάρτησης είναι βαθμωτές συναρτήσεις, που αντιπροσωπεύουν m κριτήρια επίδοσης του συστήματος. Η συνάρτηση \vec{f} αντιστοιχεί το πεδίο ορισμού $X \subseteq \mathbb{R}^n$ στο διανυσματικό πεδίο τιμών $F \subseteq \mathbb{R}^m$, που υποδηλώνει το σύνολο όλων των εφικτών τιμών των επιμέρους συναρτήσεων. Το πεδίο F καλείται χώρος αξιολόγησης (evaluation space) ή χώρος των αντικειμενικών συναρτήσεων (objective space). Το γενικό πρόβλημα πολυκριτηριακής βελτιστοποίησης διατυπώνεται ως εξής:

$$\{opt f_1(\vec{x}), opt f_2(\vec{x}), \dots, opt f_m(\vec{x})\}, \quad x \in X$$

όπου ο τελεστής «opt» υποδηλώνει είτε τη μεγιστοποίηση είτε την ελαχιστοποίηση των συνιστωσών της συνάρτησης \vec{f} . Όπως και στα προβλήματα ενός στόχου, για τις συνιστώσες $f_i(\vec{x})$ θεωρείται, κατά σύμβαση, ως κοινός τελεστής η ελαχιστοποίηση.

2.5.3 Θεμελιώδεις έννοιες πολυκριτηριακής βελτιστοποίησης

2.5.3.1 Η έννοια της κυριαρχίας

Σε πραγματικές συναρτήσεις, το βαθμωτό πεδίο τιμών $F \subseteq R$ είναι πλήρως διατεταγμένο (ordered). Κατά συνέπεια, αν σε μία διαδικασία ελαχιστοποίησης ληφθούν δύο τυχόντα σημεία x και y από το πεδίο ορισμού X , τότε το x «υπερτερεί» σε σχέση με το y εφόσον $f(x) < f(y)$. Όταν, όμως, γίνεται αναφορά σε διανυσματικές συναρτήσεις και συνακολούθως σε διανυσματικά πεδία, η έννοια της διάταξης είναι ασαφής, δεδομένου ότι δεν μπορούν να εφαρμοστούν οι τελεστές σύγκρισης $\{<, \leq, \geq, >\}$ μεταξύ διανυσμάτων, όπως γίνεται με τους πραγματικούς αριθμούς. Σε μια διαδικασία πολυκριτηριακής βελτιστοποίησης, αυτό έχει ως συνέπεια την αδυναμία αντικειμενικής αξιολόγησης λύσεων και άρα την αδυναμία ορισμού της έννοιας του βέλτιστου.

Η εφαρμογή των τελεστών σύγκρισης γενικεύεται για διανυσματικά πεδία, εισάγοντας την έννοια της μερικής διάταξης (partial ordering). Για κάθε ζεύγος διανυσμάτων \vec{a} και \vec{b} , ορίζεται μια σχέση μερικής σύγκρισης ως εξής:

$$\vec{a} = \vec{b}, \text{ αν } a_i = b_i \forall i, \quad (i = 1, \dots, m)$$

$$\vec{a} \leq \vec{b}, \text{ αν } a_i \leq b_i \forall i, \quad (i = 1, \dots, m)$$

$$\vec{a} < \vec{b}, \text{ αν } a_i \leq b_i \forall i \text{ και } a_i < b_i \text{ για ένα τουλάχιστον } i, \quad (i = 1, \dots, m)$$

Επιπλέον, μπορεί να εμφανιστεί και η περίπτωση που δεν ικανοποιείται καμία από τις παραπάνω συνθήκες, δηλαδή ισχύει $a_i \leq b_i$ για κάποια κριτήρια και $a_i > b_i$ για τα υπόλοιπα. Τότε τα

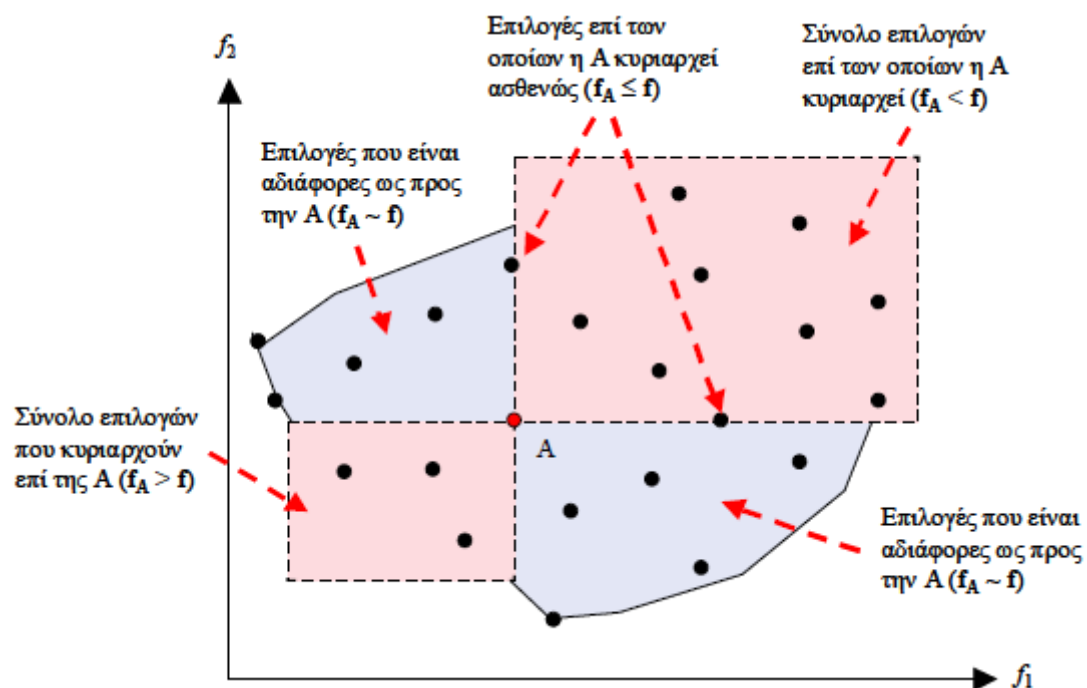
διανύσματα \vec{a} και \vec{b} αναφέρονται ως αδιάφορα (indifferent) μεταξύ τους, συσχέτιση που συμβολίζεται ως $\vec{a} \sim b$. Στα διανυσματικά πεδία, η εν λόγω συσχέτιση αποτελεί θεμελιώδη έννοια, καθώς υποδηλώνει ότι μεταξύ των διανυσμάτων υπάρχει ανταγωνιστική σχέση.

Με βάση τον προηγούμενο ορισμό της μερικής διάταξης διανυσμάτων, ορίζεται η λεγόμενη σχέση κυριαρχίας (dominance) μεταξύ δύο εφικτών λύσεων $\vec{x}, \vec{y} \in X$ ενός προβλήματος πολυκριτηριακής βελτιστοποίησης (συμβατικά ελαχιστοποίησης) ως εξής, [2, 4]:

Η λύση \vec{x} κυριαρχεί (ισχυρά) της λύσης \vec{y} , δηλαδή $\vec{x} < \vec{y}$, όταν και μόνο όταν ισχύει $f_i(\vec{x}) \leq f_i(\vec{y})$ για όλους τους στόχους ($i = 1, \dots, m$) και συγκρότως υπάρχει τουλάχιστον ένα i , δηλαδή ένας εξ αυτών των στόχων, για το οποίο ισχύει $f_i(\vec{x}) < f_i(\vec{y})$ (ή αλλιώς όταν και μόνο όταν $\vec{f}(\vec{x}) < \vec{f}(\vec{y})$).

Η λύση \vec{x} κυριαρχεί ασθενώς της λύσης \vec{y} , δηλαδή $\vec{x} \leq \vec{y}$, όταν και μόνο όταν ισχύει $f_i(\vec{x}) \leq f_i(\vec{y})$ για όλους τους στόχους ($i = 1, \dots, m$) (ή αλλιώς όταν και μόνο όταν $\vec{f}(\vec{x}) \leq \vec{f}(\vec{y})$).

Η λύση \vec{x} είναι αδιάφορη ως προς την λύση \vec{y} (ή αλλιώς αναφέρεται ότι οι δύο λύσεις είναι μη-κυριαρχούμενες μεταξύ τους) όταν $\vec{f}(\vec{x}) \sim \vec{f}(\vec{y})$.



Σχήμα 2.8: Γεωμετρική απεικόνιση των δυνατών σχέσεων κυριαρχίας ως προς ένα τυχαίο σημείο A του χώρου αξιολόγησης σε ένα πρόβλημα ελαχιστοποίησης δύο στόχων, [2].

Στο σχήμα 2.8 απεικονίζονται οι δυνατές σχέσεις κυριαρχίας σε ένα πεδίο αξιολόγησης $F \subseteq \mathbb{R}^m$ ενός υποθετικού προβλήματος ελαχιστοποίησης δύο στόχων. Αν θεωρηθεί ένα σημείο A που

αντιστοιχεί στην επιλογή αναφοράς, τότε, με βάση τον ορισμό της κυριαρχίας, το πεδίο F χωρίζεται σε τρεις υποπεριοχές:

- την υποπεριοχή κάτω αριστερά από το σημείο A , που περιλαμβάνει τις πιθανές λύσεις που κυριαρχούν επί της εν λόγω επιλογής αναφοράς·
- την υποπεριοχή πάνω δεξιά από το σημείο A , που περιλαμβάνει τις πιθανές λύσεις που κυριαρχούνται από την εν λόγω επιλογή αναφοράς·
- την υποπεριοχή πάνω αριστερά και κάτω δεξιά από το σημείο A , που περιλαμβάνει τις πιθανές λύσεις που είναι αδιάφορες ως προς την εν λόγω επιλογή αναφοράς.

Είναι προφανές ότι ο τρόπος διαχωρισμού του πεδίου αξιολόγησης εξαρτάται από το σημείο αναφοράς. Επιπλέον, ο προσανατολισμός των τριών υποπεριοχών εξαρτάται από τον τελεστή βελτιστοποίησης που ορίζεται για κάθε κριτήριο (\min ή \max). Με τη συμβατική θεώρηση της ταυτόχρονης ελαχιστοποίησης όλων των κριτηρίων, η βελτίωση σε σχέση με ένα σημείο αναφοράς επιτυγχάνεται αποκλειστικά προς την κάτω αριστερά περιοχή του πεδίου αξιολόγησης.

2.5.3.2 Η έννοια των βέλτιστων λύσεων κατά Pareto

Όπως αναφέρθηκε παραπάνω, στην περίπτωση ενός προβλήματος βελτιστοποίησης πολλών στόχων, η έννοια της μοναδικής βέλτιστης λύσης δεν υφίσταται αλλά επιβάλλεται να αντικατασταθεί από αυτήν ενός συνόλου λύσεων. Το σύνολο βέλτιστων λύσεων που προκύπτει από τη διαδικασία της πολυκριτηριακής βελτιστοποίησης έχει το εξής χαρακτηριστικό: οι λύσεις που εμπεριέχονται σε αυτό είναι μη-κυριαρχούμενες (non-dominated) μεταξύ τους. Αυτό σημαίνει ότι στοχεύουμε στην εύρεση όλων των πιθανών σχεδιασμών που οδηγούν σε μία τιμή οιοδήποτε στόχου, η οποία δεν δύναται να βελτιωθεί περαιτέρω δίχως, επίσης, να επιδεινωθεί η τιμή τουλάχιστον ενός εκ των άλλων στόχων. Ένας σχεδιασμός (δηλαδή ένα διάνυσμα μεταβλητών σχεδιασμού) για τον οποίο ισχύει κάτι τέτοιο καλείται συνήθως βέλτιστος κατά Pareto (Pareto-optimal) και το σύνολο των βέλτιστων κατά Pareto λύσεων ονομάζεται σύνολο μη-κυριαρχούμενων λύσεων (indifference set) ή σύνολο Pareto (Pareto set). Ο ορισμός της κατά Pareto βέλτιστης λύσης προκύπτει με βάση της έννοια της κυριαρχίας ως εξής, [4]:

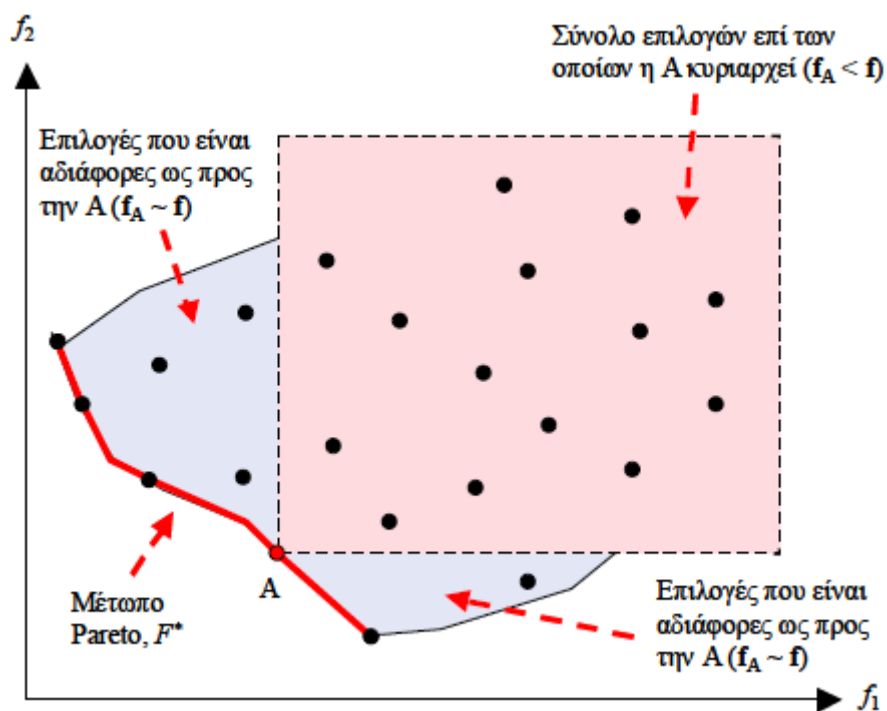
Το διάνυσμα $\vec{x}^ \in X$ αποτελεί κατά Pareto βέλτιστη λύση του προβλήματος ελαχιστοποίησης της \vec{f} αν και μόνο αν δεν υπάρχει κανένα άλλο διάνυσμα $\vec{x} \in X$, τέτοιο ώστε $f_i(\vec{x}) \leq f_i(\vec{x}^*)$ για όλους τους στόχους ($i = 1, \dots, m$) και συγχρόνως για τουλάχιστον ένα i , δηλαδή για τουλάχιστον έναν εξ αυτών των στόχων, να ισχύει $f_i(\vec{x}) < f_i(\vec{x}^*)$.*

Ισοδύναμη είναι και η διατύπωση, [2, 4]:

Το διάνυσμα $\vec{x}^* \in X$ αποτελεί κατά Pareto βέλτιστη λύση του προβλήματος ελαχιστοποίησης της \vec{f} αν και μόνο αν δεν υπάρχει κανένα άλλο διάνυσμα $\vec{x} \in X$ το οποίο να κυριαρχεί της \vec{x}^* , δηλαδή να μπορεί να βελτιώσει κάποιο στόχο χωρίς ταυτόχρονα να χειροτερέψει τουλάχιστον έναν άλλο στόχο.

Το σύνολο των μη-κυριαρχούμενων λύσεων συμβολίζεται με X^* , για το οποίο ισχύει $X^* \subset X$, δηλαδή πρόκειται για ένα υποσύνολο του πεδίου ορισμού X της συνάρτησης \vec{f} .

Η απεικόνιση $\vec{f}(\vec{x}^*)$ των μη-κυριαρχούμενων σημείων $\vec{x}^* \in X^*$ στο m -διάστατο πεδίο αξιολόγησης F ορίζει ένα υποσύνολό του, το F^* . Το $F^* \subset F$ καλείται μέτωπο Pareto (Pareto front). Αν θεωρηθεί ένα πρόβλημα ελαχιστοποίησης δύο στόχων, όπως φαίνεται στο σχήμα 2.9, το μέτωπο Pareto ταυτίζεται με το κάτω αριστερά όριο του πεδίου F . Κάθε σημείο του μετώπου Pareto είτε κυριαρχεί είτε είναι αδιάφορο ως προς κάθε τυχαίο σημείο του πεδίου F , ενώ είναι αυστηρά αδιάφορο σε σχέση με τα υπόλοιπα σημεία του μετώπου.



Σχήμα 2.9: Απεικόνιση του μετώπου Pareto σε ένα πρόβλημα ελαχιστοποίησης δύο κριτηρίων, [2].

Σε προβλήματα δύο στόχων ($m = 2$) το μέτωπο Pareto είναι μια καμπύλη, η οποία είναι γνωστή από τις οικονομικές επιστήμες και ως καμπύλη αντιστάθμισης (tradeoff curve), καθώς, όπως αναφέρθηκε, υποδηλώνει ένα συμβιβασμό μεταξύ των στόχων. Σε προβλήματα περισσότερων στόχων, το μέτωπο απεικονίζεται ως μια υπερεπιφάνεια

διάστασης $m - 1$, η οποία βρίσκεται πάντοτε πάνω στο όριο του διανυσματικού πεδίου τιμών F .

Η χάραξη του μετώπου Pareto F^* προϋποθέτει τον εντοπισμό του συνόλου X^* . Όταν το πεδίο αναζήτησης X είναι συνεχές, τότε το σύνολο X^* είναι επίσης συνεχές, δηλαδή έχει άπειρα σημεία. Σε αντίθεση, λοιπόν, με την μονοκριτηριακή βελτιστοποίηση, στην οποία αναζητείται ένα μοναδικό σημείο του εφικτού χώρου ως βέλτιστη λύση, στην πολυκριτηριακή βελτιστοποίηση αναζητείται, συνήθως (αλλά όχι πάντα), ένα συνεχές πεδίο, άγνωστης έκτασης και άγνωστης γεωμετρίας. Συνεπώς, το πρακτικό ζητούμενο είναι ο εντοπισμός ενός αντιπροσωπευτικού αριθμού μη-κυριαρχούμενων λύσεων, οι οποίες περιγράφουν με ικανοποιητική ακρίβεια το πεδίο X^* και με βάση τις οποίες είναι δυνατή μια προσεγγιστική χάραξη του μετώπου Pareto.

Στην πολυκριτηριακή βελτιστοποίηση, ο κατά Pareto ορισμός του βελτίστου ως ένα σύνολο μαθηματικά αποδεκτών λύσεων, δημιουργεί 'ασάφεια' στις πρακτικές εφαρμογές που αποσκοπούν στον εντοπισμό μίας και μόνο λύσης, η οποία αντιπροσωπεύει μια συγκεκριμένη απόφαση που πρέπει να ληφθεί. Πράγματι, από μαθηματική σκοπιά, όλα τα σημεία του μετώπου Pareto αντιπροσωπεύουν θεωρητικά ισοδύναμες εναλλακτικές επιλογές, δεδομένου ότι μεταξύ τους είναι μη-κυριαρχούμενα, ενώ κυριαρχούν έναντι κάθε άλλου εφικτού σημείου. Συνεπώς, η ανάγκη περαιτέρω διάκρισης αυτών επιβάλλει την εισαγωγή επιπλέον 'εξωγενών' κριτηρίων αξιολόγησης, πέρα από τα 'ενδογενή' μαθηματικά κριτήρια, δηλαδή τις συνιστώσες της διανυσματικής συνάρτησης στόχου. Τελικά, το ποια από τις λύσεις που αποτελούν το μέτωπο Pareto θα επιλεγεί εξαρτάται από τον κατασκευαστή ή από τον υπεύθυνο λήψης αποφάσεων (decision maker) μίας εταιρείας.

3 Μέθοδοι βελτιστοποίησης

3.1 Εισαγωγή

Η μέθοδος βελτιστοποίησης είναι το εργαλείο που ανιχνεύει τον χώρο των υποψηφίων λύσεων και προβαίνει στον εντοπισμό της βέλτιστης λύσης. Οι δύο κυριότερες κατηγορίες στις οποίες διακρίνονται οι μέθοδοι βελτιστοποίησης είναι οι **στοχαστικές** (stochastic) και οι **αιτιοκρατικές** μέθοδοι (deterministic methods).

Ο πιο διαδεδομένος εκπρόσωπος της κατηγορίας των στοχαστικών μεθόδων βελτιστοποίησης είναι οι εξελικτικοί αλγόριθμοι (evolutionary algorithms, EA), [4, 49-54], οι οποίοι περιλαμβάνουν τους γενετικούς αλγορίθμους (genetic algorithms), [47], και τις εξελικτικές στρατηγικές (evolution strategies), [48]. Οι εξελικτικοί αλγόριθμοι βασίζονται στη μίμηση φυσικών διεργασιών που σχετίζονται με την εξέλιξη των ειδών. Οι διεργασίες αυτές εφαρμόζονται στα άτομα ενός πληθυσμού, τα οποία αποτελούν τις υποψήφιες λύσεις του προβλήματος βελτιστοποίησης, παράγοντας τον πληθυσμό της επόμενης γενιάς, μέχρι την ικανοποίηση κάποιου κριτηρίου τερματισμού της εξέλιξης. Το γεγονός ότι οι εν λόγω αλγόριθμοι περιλαμβάνουν ένα πληθυσμό ατόμων που συμμετέχουν στην εξέλιξη και την καθορίζουν, καταδεικνύει τη δυνατότητά τους να εξερευνούν (σχεδόν) ολόκληρο το χώρο λύσεων, κάτι που σημαίνει ότι δεν εγκλωβίζονται σε τοπικά ακρότατα. Αυτό είναι και το μεγαλύτερο πλεονέκτημα των εξελικτικών αλγορίθμων αλλά και των στοχαστικών μεθόδων βελτιστοποίησης γενικότερα. Όμως, το ότι οι αλγόριθμοι αυτοί βασίζονται στην τυχαιότητα για να εξερευνήσουν το χώρο λύσεων σημαίνει ότι χρειάζονται πολύ 'χρόνο' για να βρουν το καθολικό ακρότατο. Με άλλα λόγια, λόγω της ανάγκης εκτίμησης της αντικειμενικής συνάρτησης (άρα σε πολλές περιπτώσεις της επίλυσης των εξισώσεων ροής) για όλα τα άτομα του πληθυσμού διαδοχικών γενεών, οι εξελικτικοί αλγόριθμοι έχουν υψηλό υπολογιστικό κόστος, κάτι που αποτελεί το μεγαλύτερο μειονέκτημά τους αλλά και αυτό των λοιπών στοχαστικών μεθόδων.

Μέλη της κατηγορίας των στοχαστικών μεθόδων είναι, [8], επίσης, η αναρρίχηση λόφων (hill climbing), η επαναληπτική αναρρίχηση λόφων (iterated hill climbing), η στοχαστική επαναληπτική αναρρίχηση λόφων (stochastic iterated hill climbing), η προσομοιωμένη απόπτηση (simulated annealing, [60]), η μέθοδος Nelder-Mead (SIMPLEX, [61]), η μέθοδος αποικιών μυρμηγκιών (ant colony optimization), κλπ.

Από την άλλη πλευρά, οι αλγόριθμοι των αιτιοκρατικών μεθόδων ξεκινούν από μια αρχική λύση και βασιζόμενοι σε μαθηματικά (και όχι στοχαστικά-τυχηματικά) κριτήρια οδεύουν προς τη βέλτιστη λύση.

Ουσιαστικά, οι εν λόγω αλγόριθμοι χρησιμοποιούν τις παραγώγους της αντικειμενικής συνάρτησης ως προς τις μεταβλητές σχεδιασμού, δηλαδή τις λεγόμενες παραγώγους ευαισθησίας, οι οποίες ‘δείχνουν’ προς την κατεύθυνση του (τοπικού ίσως) βελτίστου. Ακολουθώντας, λοιπόν, την κατεύθυνση της παραγώγου, η οποία υπολογίζεται σε κάθε επανάληψη, ένας τέτοιος αλγόριθμος οδεύει σταδιακά προς τη βέλτιστη λύση. Κάποιες από τις μεθόδους αυτής της κατηγορίας, για τον προσδιορισμό της βέλτιστης λύσης, καταφεύγουν και στον υπολογισμό (ακριβή ή προσεγγιστικό) των δεύτερων παραγώγων. Η φύση των αιτιοκρατικών μεθόδων που υπαγορεύει την επιλογή μιας αρχικής λύσης για την εκκίνησή τους και στερείται στοχαστικότητας έχει ως συνέπεια ο αλγόριθμος να εγκλωβίζεται σε τοπικά ακρότατα.

Η κυριότερη και πιο διαδεδομένη ‘ομάδα’ μεθόδων ανίχνευσης της βέλτιστης λύσης που ανήκουν στην κατηγορία των αιτιοκρατικών μεθόδων είναι οι μέθοδοι ανίχνευσης κατά γραμμή (line search methods). Στην εν λόγω υποκατηγορία των αιτιοκρατικών μεθόδων βελτιστοποίησης εντάσσονται, [4, 62], η μέθοδος της απότομης καθόδου (steepest descent method), η μέθοδος της (‘απλής’) καθόδου, η μέθοδος Newton, η προσεγγιστική μέθοδος Newton (Quasi-Newton method), η μέθοδος των συζυγών κατευθύνσεων (conjugate directions method), η μέθοδος των συζυγών κλίσεων (conjugate gradient method), κλπ. Μία, ακόμα, υποκατηγορία των αιτιοκρατικών μεθόδων βελτιστοποίησης που θεωρείται η εναλλακτική της υποκατηγορίας των μεθόδων της ανίχνευσης κατά γραμμή είναι οι μέθοδοι ανίχνευσης κατά περιοχές εμπιστοσύνης (trust region methods), [4].

Σημειώνεται, επιπλέον, ότι για τον υπολογισμό των παραγώγων της αντικειμενικής συνάρτησης ως προς τις μεταβλητές σχεδιασμού, δηλαδή των παραγώγων ευαισθησίας, χρησιμοποιούνται, μεταξύ άλλων, η μέθοδος των πεπερασμένων διαφορών (finite difference method), [23-27], η μέθοδος των μιγαδικών μεταβλητών (complex variable method), [22, 28-31], η μέθοδος της ευθείας διαφορίσης (direct differentiation method), [32-36], η μέθοδος της αυτόματης διαφορίσης (automatic/algorithmic differentiation), [37-40], και η συζυγής μέθοδος (adjoint method), [4, 41-46].

Τα πλεονεκτήματα και τα μειονεκτήματα των δύο διαφορετικών κατηγοριών μεθόδων βελτιστοποίησης είναι αυτά που καθορίζουν τις επιλογές του κάθε επιστήμονα που καλείται να λύσει ένα πρόβλημα.

Από τη μια πλευρά, οι εξελικτικοί αλγόριθμοι έχουν τη δυνατότητα να αποφεύγουν τον εγκλωβισμό σε τοπικά ακρότατα. Θεωρητικά, με την πάροδο του χρόνου-ενός επαρκούς αριθμού γενεών, οι αλγόριθμοι αυτοί δύνανται να συγκλίνουν στο καθολικό ακρότατο του κάθε προβλήματος. Ωστόσο, καταυτόν τον τρόπο, αυξάνεται κατά πολύ το υπολογιστικό κόστος, καθώς σε προβλήματα βελτιστοποίησης στην εξωτερική αεροδυναμική και τις στροβιλομηχανές, η κάθε εκτίμηση της τιμής της αντικειμενικής συνάρτησης προϋποθέτει την αριθμητική

επίλυση των εξισώσεων ροής. Το γεγονός αυτό συχνά αποτρέπει τη χρήση της εν λόγω μεθόδου σε μεγάλης κλίμακας προβλήματα αεροδυναμικού σχεδιασμού. Ένα πλεονέκτημα που διαθέτουν οι εξελικτικοί αλγόριθμοι και τους καθιστά ιδιαίτερα εύχρηστους σε προβλήματα βελτιστοποίησης είναι η ανεξαρτησία τους από το λογισμικό αξιολόγησης, δηλαδή το λογισμικό που υπολογίζει τις τιμές των συναρτήσεων στόχου. Ως επί το πλείστον, ο κώδικας βελτιστοποίησης και το λογισμικό αξιολόγησης είναι δύο διαφορετικά εργαλεία που συνδέονται μεταξύ τους μόνο μέσω των τιμών των μεταβλητών σχεδιασμού και αυτών των συναρτήσεων στόχου. Αυτό σημαίνει ότι αν ο εξελικτικός αλγόριθμος είναι διαθέσιμος, μπορεί να χρησιμοποιηθεί σε οιοδήποτε πρόβλημα βελτιστοποίησης χωρίς καμία σχεδόν επιπλέον επέμβαση στον κώδικα. Αξιοσημείωτη είναι, επιπλέον, η ευκολία με την οποία οι εξελικτικοί αλγόριθμοι χειρίζονται προβλήματα βελτιστοποίησης πολλών στόχων και παρέχουν το μέτωπο Pareto. Τέλος, εξίσου εύκολα, δύνανται οι εξελικτικοί αλγόριθμοι να αντιμετωπίσουν προβλήματα στα οποία επιβάλλονται περιορισμοί.

Αντίθετα, το σημαντικό πλεονέκτημα των αιτιοκρατικών μεθόδων συναντάται στη γρήγορη σύγκλισή τους στη βέλτιστη λύση. Οι μέθοδοι αυτοί, λόγω του ότι χρησιμοποιούν τις παραγώγους ευαισθησίας, απαιτούν 'μικρό' αριθμό επαναλήψεων και άρα εκτιμήσεων της τιμής της αντικειμενικής συνάρτησης, γεγονός που τις καθιστά ιδιαίτερα ελκυστικές για προβλήματα μεγάλης κλίμακας. Όπως αναφέρθηκε παραπάνω, βέβαια, έχουν το 'ελάττωμα' να εγκλωβίζονται σε τοπικά ακρότατα οπότε να κινδυνεύουν να μην φτάσουν στο ολικό ακρότατο. Κάτι τέτοιο δύναται να αντιμετωπιστεί καλώντας πολλαπλές φορές τον αλγόριθμο βελτιστοποίησης και δίνοντάς του διαφορετικές αρχικοποιήσεις, με αύξηση ασφαλώς του υπολογιστικού κόστους. Πρέπει να σημειωθεί ότι η 'επίδοση' των μεθόδων που βασίζονται στις παραγώγους της αντικειμενικής συνάρτησης ως προς τις μεταβλητές σχεδιασμού εξαρτάται σε μεγάλο βαθμό από τη μέθοδο υπολογισμού των παραγώγων αυτών. Έτσι, αν, για παράδειγμα, χρησιμοποιηθεί η μέθοδος των πεπερασμένων διαφορών, το κόστος υπολογισμού των παραγώγων ευαισθησίας θα είναι πολύ μεγάλο (ιδιαίτερα στην περίπτωση πολλών μεταβλητών σχεδιασμού). Ως αποτέλεσμα, το συνολικό κόστος της βελτιστοποίησης γίνεται απαγορευτικό για προβλήματα μεγάλης κλίμακας. Αντίθετα, η χρήση της συζυγούς μεθόδου για τον υπολογισμό των παραγώγων ευαισθησίας είναι ιδιαίτερα αποδοτική και το κόστος υπολογισμού των παραγώγων και, κατά συνέπεια, ολόκληρης της βελτιστοποίησης είναι σαφώς μικρότερο και κυρίως ανεξάρτητο από τον αριθμό των μεταβλητών σχεδιασμού.

Συμπερασματικά, η κάθε μέθοδος παρουσιάζει τα θετικά της αλλά μαζί ακολουθούν και κάποια αρνητικά. Το ιδανικό για έναν επιστήμονα θα ήταν να μπορέσει να εκμεταλλευτεί τα πλεονεκτήματα και των δύο κατηγοριών μεθόδων βελτιστοποίησης. Ο συγκερασμός

των πλεονεκτημάτων των στοχαστικών και των αιτιοκρατικών μεθόδων βελτιστοποίησης είναι δυνατό να επιτευχθεί με τον υβριδισμό τους και έτσι έχουμε την ανάπτυξη μεθόδων που συνδυάζουν 'λογικό' υπολογιστικό κόστος και αποφεύγουν τον εγκλωβισμό σε τοπικά ακρότατα.

3.2 Γενικές αρχές αλγορίθμων βελτιστοποίησης

Με δεδομένη την τεράστια ποικιλία σχημάτων αναζήτησης σε μη-γραμμικούς χώρους, είναι ιδιαίτερα δύσκολη η διατύπωση γενικών αρχών και κανόνων. Παρά ταύτα, στη δεδομένη ενότητα γίνεται η προσπάθεια να δοθούν κάποιες γενικές αρχές στις οποίες βασίζονται όλες οι μέθοδοι βελτιστοποίησης, ώστε να γίνει κατανοητός ο τρόπος με τον οποίο λειτουργεί ένας οιοσδήποτε αλγόριθμος βελτιστοποίησης για να φτάσει στη βέλτιστη λύση ανεξάρτητα από τη μέθοδο που ακολουθεί. Με άλλα λόγια εδώ παρουσιάζεται η 'φιλοσοφία' όλων των μεθόδων βελτιστοποίησης.

Το κοινό χαρακτηριστικό όλων των αλγορίθμων είναι πως πρόκειται για επαναληπτικές τεχνικές που ξεκινούν από ένα σύνολο αρχικών λύσεων $P^{[0]}$, το οποίο περιλαμβάνει ένα ή περισσότερα σημεία (δηλαδή n -διάστατα διανύσματα) από τον εφικτό χώρο X . Μετά από διαδοχικούς μετασχηματισμούς του αρχικού συνόλου, προκύπτει ένα τελικό σύνολο $P^{[N]}$, το οποίο θεωρείται πως περιέχει την εκτίμηση του ολικού βελτίστου.

Το σύνολο P περιγράφεται μαθηματικά ως ένα $n \times p$ μητρώο της μορφής:

$$P = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,p} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,p} \end{bmatrix}$$

όπου p το μέγεθος του συνόλου P .

Η διαδικασία αναζήτησης μπορεί να διατυπωθεί ως μια σειρά N το πλήθος μετασχηματισμών του αρχικού συνόλου $P^{[0]}$, δηλαδή:

$$P^{[0]} \rightarrow P^{[1]} \rightarrow \dots \rightarrow P^{[k-1]} \rightarrow P^{[k]} \rightarrow P^{[k+1]} \rightarrow \dots \rightarrow P^{[N]}$$

Η παραπάνω διαδικασία περιγράφει $p \times N$ το πλήθος κινήσεις σημείων πάνω στον χώρο αναζήτησης. Οι εν λόγω μετακινήσεις των σημείων πρέπει να εγγυώνται την *συστηματική* βελτίωση της τιμής της αντικειμενικής συνάρτησης, έτσι ώστε το σύνολο P να συγκλίνει στην περιοχή του πραγματικού ολικού βελτίστου. Η συστηματικότητα της διαδικασίας είναι αυτή που εξασφαλίζει ότι για ένα τουλάχιστον μέλος του τελικού συνόλου, έστω το $\vec{x}_l^{[N]}$, όπου $l = 1, 2, \dots, p$, ισχύει:

$$f(\vec{x}_l^{[N]}) < \min \{f(\vec{x}_1^{[0]}), f(\vec{x}_2^{[0]}), \dots, f(\vec{x}_p^{[0]})\}$$

οπότε το $\vec{x}_l^{[N]}$ λαμβάνεται ως η εκτίμηση του ολικού ελαχίστου x^* , χωρίς απαραίτητα να ταυτίζεται με αυτό, δεδομένου ότι πρόκειται για αριθμητικές-υπολογιστικές τεχνικές που εφαρμόζονται σε προβλήματα συνεχών μεταβλητών. Στόχος, δηλαδή, είναι η προσέγγιση της ακριβούς λύσης με ανοχή τέτοια που δεν επηρεάζει το αποτέλεσμα της βελτιστοποίησης στην πράξη.

Το πλήθος των μετασχηματισμών N είτε ορίζεται εκ των προτέρων (οπότε τίθεται ως κριτήριο τερματισμού της διαδικασίας) είτε ελέγχεται με βάση την πρόοδο της διαδικασίας αναζήτησης (εφαρμόζεται ένα προδιαγεγραμμένο κριτήριο σύγκλισης).

Η 'μετάβαση' από ένα τρέχον σύνολο λύσεων στο επόμενο γίνεται μέσω ενός κανόνα της μορφής:

$$P^{[k]} = P^{[k-1]} + \Delta P^{[k]}$$

όπου $\Delta P^{[k]}$ μια εφικτή «διαταραχή» στο X .

Όταν το σύνολο P είναι μονομελές ($p = 1$), η διαδικασία βελτιστοποίησης υποδηλώνει μια τροχιά στο n -διάστατο πεδίο αναζήτησης X , η συστηματικότητα της οποίας πρέπει να εγγυάται σύγκλιση στο ολικό ακρότατο.

Οι αλγόριθμοι μη-γραμμικής βελτιστοποίησης διαφοροποιούνται ως προς τις ακόλουθες συνιστώσες:

- ✘ τα χαρακτηριστικά του συνόλου P (μονομελές ή πολυμελές, αυτόματα οριζόμενο ή οριζόμενο από τον χρήστη)
- ✘ τη διαδικασία παραγωγής του αρχικού συνόλου $P^{[0]}$
- ✘ τους κανόνες μετάβασης (αιτιοκρατικοί, στοχαστικοί ή υβριδικοί)
- ✘ τα κριτήρια σύγκλισης και τερματισμού.

3.3 Μέθοδοι ανίχνευσης κατά γραμμή

3.3.1 Βασική αρχή κάθε μεθόδου ανίχνευσης κατά γραμμή

Η παρουσίαση που ακολουθεί καλύπτει ένα εύρος μεθόδων που αναζητούν τη βέλτιστη λύση πραγματοποιώντας μια διαδρομή ζιγκ-ζαγκ στο χώρο των λύσεων. Ήδη στην ενότητα 3.1 έχει γίνει μια ονομαστική αναφορά στις μεθόδους που ανήκουν στην κατηγορία των μεθόδων ανίχνευσης κατά γραμμή και εδώ θα αναλυθούν κάποιες από αυτές. Επίσης, αναφέρθηκε ότι οι δεδομένες μέθοδοι αποτελούν

το βασικότερο εκπρόσωπο της κατηγορίας των αιτιοκρατικών μεθόδων βελτιστοποίησης ή αλλιώς των μεθόδων που είναι βασισμένες στην κλίση της αντικειμενικής συνάρτησης (gradient-based methods). Αυτό σημαίνει ότι όλες οι μέθοδοι που θα σχολιαστούν απαιτούν τον υπολογισμό τουλάχιστον των τιμών των πρώτων παραγώγων της αντικειμενικής συνάρτησης ως προς τις μεταβλητές σχεδιασμού $\nabla f(\vec{x})$. Όλες οι εν λόγω μέθοδοι είναι επαναληπτικές ή βηματικές. Αυτό σημαίνει ότι κάθε ευθύγραμμο τμήμα, δηλαδή κάθε κλάδος της διαδρομής ζιγκ-ζαγκ, αποτελεί ένα βήμα ή μια πρόοδο της λύσης, η οποία ξεκίνησε από μια αρχική τιμή και εξελίσσεται προς τη βέλτιστη.

Για την παρουσίαση των μεθόδων ανίχνευσης κατά γραμμή θα θεωρηθεί ότι τα προβλήματα βελτιστοποίησης που απασχολούν αφορούν στην ελαχιστοποίηση της τιμής της αντικειμενικής συνάρτησης $f(\vec{x})$, $\vec{x} \in X$, στην οποία εμπλέκονται n μεταβλητές σχεδιασμού, χωρίς η λύση να υπόκειται σε περιορισμούς ισότητας ή/και ανισότητας. Με άλλα λόγια, γίνεται αναφορά σε προβλήματα ενός στόχου ή σε προβλήματα πολλών στόχων που έχουν διατυπωθεί στη μορφή μιας αντικειμενικής συνάρτησης προς ελαχιστοποίηση και, βέβαια, αμφότερα στερούνται περιορισμών.

Η βασική αρχή κάθε επαναληπτικής μεθόδου ανίχνευσης κατά γραμμή είναι, [4], να δημιουργείται η επόμενη προσέγγιση της λύσης \vec{x}^{N+1} χρησιμοποιώντας την τρέχουσα \vec{x}^N και μια κατεύθυνση ανίχνευσης \vec{p}^N (search direction), σύμφωνα με την παρακάτω σχέση:

$$\vec{x}^{N+1} = \vec{x}^N + \eta^N \vec{p}^N \quad (3.1)$$

όπου η^N είναι μια βαθμωτή ποσότητα που ρυθμίζει το μήκος του βήματος (step length) στην κατεύθυνση \vec{p}^N . Ο άνω δείκτης N θα συμβολίζει τον αριθμό της επανάληψης ή βήματος.

Οι μέθοδοι ανίχνευσης κατά γραμμή διαφέρουν, κυρίως, ως προς τον υπολογισμό της κατεύθυνσης ανίχνευσης \vec{p}^N και, δευτερευόντως, ως προς τον τρόπο υπολογισμού του η . Ανάλογα με την παραλλαγή της μεθόδου που χρησιμοποιείται, η τιμή του η μπορεί να είναι σταθερή με το πέρασμα των επαναλήψεων ή, αφού έχει ήδη επιλεγεί η κατεύθυνση ανίχνευσης \vec{p}^N , το η να υπολογίζεται εκ νέου σε κάθε βήμα (και γι' αυτό το λόγο έχει μπει ο δείκτης N που δηλώνει τον αριθμό της επανάληψης και στο η).

Έτσι, πλέον, το πρόβλημα μπορεί να διατυπωθεί μαθηματικά ως εξής:

$$\min_{\eta^N > 0} f(\vec{x}^N + \eta^N \vec{p}^N) \quad (3.2)$$

Έχοντας προεπιλέξει την κατεύθυνση ανίχνευσης \vec{p}^N , η σχέση (3.2) κρύβει ένα νέο μονοπαραμετρικό πρόβλημα βελτιστοποίησης, το οποίο σχετίζεται με την εύρεση της ελάχιστης τιμής της $f(\vec{x})$ σε μια συγκεκριμένη κατεύθυνση με άγνωστο το η^N .

Στο σημείο αυτό κρίνεται σκόπιμο να παρατεθούν δύο σχόλια σχετικά με τον τρόπο υπολογισμού της τιμής του η ή η^N στο επαναληπτικό σχήμα:

- Η περίπτωση σταθερής τιμής του η είναι σαφέστατα η απλούστερη. Προφανώς, όμως, δεν εξασφαλίζει την ταχύτερη σύγκλιση στη βέλτιστη λύση. Μάλιστα, αν υπάρχει τρόπος επιλογής της κατεύθυνσης ανίχνευσης \vec{p}^N , μια απλή (με ελάχιστο κόπο από τον προγραμματιστή) λύση είναι να επιλεγεί μια μικρή τιμή του η . Έτσι, ελπίζει κανείς ότι ο αλγόριθμος θα συγκλίνει αργά αλλά σταθερά. Είναι προφανές ότι κάτι τέτοιο δεν συνιστάται, αλλά παρά ταύτα κάποιες φορές χρησιμοποιείται αν λ.κ. πρόκειται για την επίλυση ενός απλού προβλήματος μία μόνο φορά.
- Η αντιδιαμετρικά ακραία περίπτωση είναι αυτή όπου το πρόβλημα (3.2) λύνεται πλήρως ως προς η^N σε κάθε νέα επανάληψη και πραγματικά (σε αντίθεση με την προηγούμενη περίπτωση) επιτυγχάνεται το μέγιστο κέρδος, καθώς η μέθοδος ανιχνεύει σαρώνοντας την ευθεία με αφετηρία το \vec{x}^N κατά την κατεύθυνση \vec{p}^N . Κάτι τέτοιο στοιχίζει ακριβά σε υπολογιστικό χρόνο. Πρέπει, επίσης, να τονιστεί ότι η λεπτομερής ανίχνευση κάθε φορά κατά την κατεύθυνση \vec{p}^N μπορεί να είναι αναίτια και επιβλαβής για την τελική σύγκλιση.

Συμπερασματικά, είναι δυνατό να υιοθετηθεί μια ενδιάμεση διαδικασία υπολογισμού του η^N , κατά την οποία ο αλγόριθμος δημιουργεί ένα μικρό αριθμό τιμών του η^N κατά \vec{p}^N , από αυτές επιλέγει εκείνη που δίνει μικρότερη τιμή στην $f(\vec{x}^{N+1})$ και συνεχίζει με το επόμενο βήμα, κοκ.

3.3.2 Η μέθοδος της απότομης καθόδου (steepest descent method)

Στη μέθοδο της απότομης καθόδου η εφαρμογή της σχέσης ανανέωσης της τιμής του διανύσματος \vec{x} , δηλαδή της σχέσης (3.1) γίνεται επιλέγοντας ως κατεύθυνση ανίχνευσης \vec{p}^N το αντίθετο του τοπικού διανύσματος κλίσης της συνάρτησης $f(\vec{x}^N)$, δηλαδή:

$$\vec{p}^N = -\nabla f(\vec{x}^N)$$

Έτσι, η σχέση (3.1) γράφεται ως εξής:

$$\vec{x}^{N+1} = \vec{x}^N - \eta^N \nabla f(\vec{x}^N)$$

Η κατεύθυνση $-\nabla f(\vec{x}^N)$ είναι ορθογώνια στην ισοσταθμική καμπύλη της $f(\vec{x})$ που διέρχεται από το σημείο \vec{x}^N .

Αν αναλογιστεί κανείς ότι η κλίση της $f(\vec{x})$, δηλαδή το διάνυσμα $\nabla f(\vec{x})$, δείχνει σε κάθε σημείο την κατεύθυνση της μέγιστης μεταβολής της τιμής της $f(\vec{x})$, μπορεί να κατανοηθεί η επιλογή του διανύσματος $-\nabla f(\vec{x}^N)$ ως κατεύθυνση ανίχνευσης της βέλτιστης λύσης.

Στο σημείο αυτό, για λόγους πληρότητας, κρίνεται σκόπιμο να παρατεθεί το εξής ανάπτυγμα Taylor:

$$f(\vec{x}^N + \eta^N \vec{p}^N) = f(\vec{x}^N) + \eta^N \vec{p}^{N^T} \nabla f(\vec{x}^N) + O(\eta^2)$$

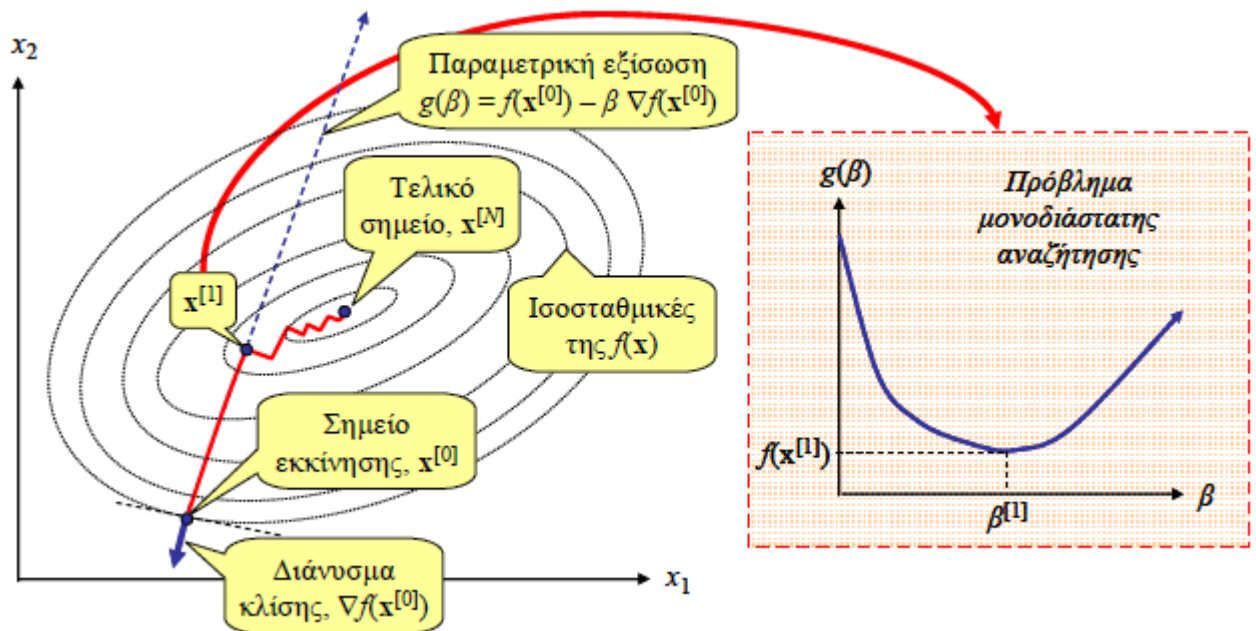
Το πρόσημο του δεύτερου όρου στο δεξιό μέλος (αρνητική ποσότητα για $\eta^N > 0$) δικαιολογεί και την επιλογή του διανύσματος $-\nabla f(\vec{x}^N)$ ως κατεύθυνση ανίχνευσης της βέλτιστης λύσης.

Χαρακτηριστικό της μεθόδου της απότομης καθόδου είναι ότι απαιτεί, σε κάθε νέα θέση, τον υπολογισμό της κλίσης της αντικειμενικής συνάρτησης, και πλεονέκτημά της ότι δεν απαιτεί πληροφορία για το εσσιανό μητρώο (που όπως παρουσιάστηκε στην ενότητα 2.2.3 περιέχει τις δεύτερες παραγώγους της αντικειμενικής συνάρτησης ως προς τις μεταβλητές σχεδιασμού). Η απόδοσή της εν λόγω μεθόδου εξαρτάται από την αρχική λύση \vec{x}^0 , το η^N και, προφανώς, τη μορφή της συνάρτησης $f(\vec{x})$.

Ο κανόνας μετάβασης διατυπώνεται παραμετρικά, συναρτήσει του μέτρου μετατόπισης η^N . Το εν λόγω μέτρο υπολογίζεται σε κάθε επαναληπτικό βήμα, με τρόπο ώστε να ελαχιστοποιείται η τιμή της συνάρτησης στη νέα θέση \vec{x}^{N+1} . Κατά συνέπεια, σε κάθε επανάληψη διαμορφώνεται η συνάρτηση στόχου μιας μεταβλητής:

$$g(\eta^N) = f[\vec{x}^N - \eta^N \nabla f(\vec{x}^N)]$$

Με αυτόν τον τρόπο, το αρχικό πρόβλημα ελαχιστοποίησης της συνάρτησης πολλών μεταβλητών $f(\vec{x})$ ανάγεται στην επίλυση διαδοχικών προβλημάτων ελαχιστοποίησης της συνάρτησης μίας μεταβλητής $g(\eta^N)$. Δηλαδή, αφού εντοπιστεί μια κατάλληλη διεύθυνση στο \mathbb{R}^n (που ταυτίζεται με το αντίθετο του διανύσματος κλίσης) απαιτείται μια διαδικασία *μονοδιάστατης αναζήτησης* (one-dimensional search) πάνω στην συνάρτηση $g(\eta^N)$, που απεικονίζει μια «τομή». Ο χειρισμός του προβλήματος γίνεται είτε αναλυτικά (εφόσον η έκφραση της συνάρτησης και των παραγώγων της είναι γνωστές) ή με χρήση αριθμητικών τεχνικών.



Σχήμα 3.1: Εφαρμογή της μεθόδου της απότομης καθόδου σε ένα υποθετικό πρόβλημα ελαχιστοποίησης μιας συνάρτησης δύο μεταβλητών (σημειώνεται ότι στο σχήμα το η συμβολίζεται ως β). Δεξιά, φαίνεται η αναγωγή του αρχικού προβλήματος σε διαδοχικά προβλήματα μονοδιάστατης αναζήτησης, [2].

Η πορεία αναζήτησης της μεθόδου της απότομης καθόδου περιγράφεται στο παράδειγμα του σχήματος 3.1. Σε κάθε βήμα, η διεύθυνση μετακίνησης είναι σχεδόν κάθετη στην προηγούμενη, ενώ είναι ακριβώς κάθετη στην περίπτωση που το ελάχιστο της συνάρτησης $g(\eta^N)$ μπορεί να εντοπιστεί με ακρίβεια. Λόγω του ανωτέρω χαρακτηριστικού, το βήμα διαταραχής επιβραδύνεται γύρω από το σημείο του ακροτάτου, καθώς ο αλγόριθμος επιβάλλει μια πορεία ζιγκ-ζαγκ.

Η μέθοδος δίνει καλά αποτελέσματα όταν χρησιμοποιείται σε απλά προβλήματα, αλλά γίνεται ιδιαίτερα αργή και αναποτελεσματική σε πολύπλοκα προβλήματα (με χαρακτηριστικό παράδειγμα τα προβλήματα βελτιστοποίησης στην αεροδυναμική).

3.3.3 Η μέθοδος Newton

Η μέθοδος Newton αποτελεί έναν από τους σημαντικότερους εκπροσώπους της κατηγορίας των μεθόδων ανίχνευσης κατά γραμμή. Για τη θεμελίωση της εν λόγω μεθόδου, [4], αρκεί να γραφτεί εκ νέου το ανάπτυγμα Taylor της $f(\vec{x}^N + \eta^N \vec{p}^N)$, όπως έγινε και στην προηγούμενη ενότητα κατά την παρουσίαση της μεθόδου της απότομης καθόδου. Σε αντίθεση, όμως, με την προηγούμενη περίπτωση, εδώ διατηρείται επιπλέον ο όρος δεύτερης τάξης και γίνεται η παραδοχή ότι $\eta^N = \eta = 1$. Έτσι, είναι:

$$f(\vec{x}^N + \vec{p}^N) \approx f(\vec{x}^N) + \vec{p}^{N^T} \nabla f(\vec{x}^N) + \frac{1}{2} \vec{p}^{N^T} \nabla^2 f(\vec{x}^N) \vec{p}^N$$

Η απαίτηση μηδενικής πρώτης παραγώγου της προσέγγισης της $f(\vec{x}^N + \vec{p}^N)$ ως προς \vec{p}^N δίνει:

$$\nabla f(\vec{x}^N) + \nabla^2 f(\vec{x}^N) \vec{p}^N = 0$$

Έτσι, καθορίζεται η κατεύθυνση ανίχνευσης για την τρέχουσα επανάληψη ως:

$$\vec{p}^N = -[\nabla^2 f(\vec{x}^N)]^{-1} \nabla f(\vec{x}^N)$$

Η τελευταία σχέση γράφεται με την προϋπόθεση ότι το εσσιανό μητρώο $\nabla^2 f(\vec{x}^N)$ είναι θετικά ορισμένο. Αν αυτή η προϋπόθεση δεν ικανοποιείται, η κατεύθυνση που προκύπτει από την παραπάνω σχέση δεν ορίζεται αφού δεν είναι δυνατό να υπολογιστεί ο αντίστροφος του εσσιανού μητρώου $\nabla^2 f(\vec{x}^N)$.

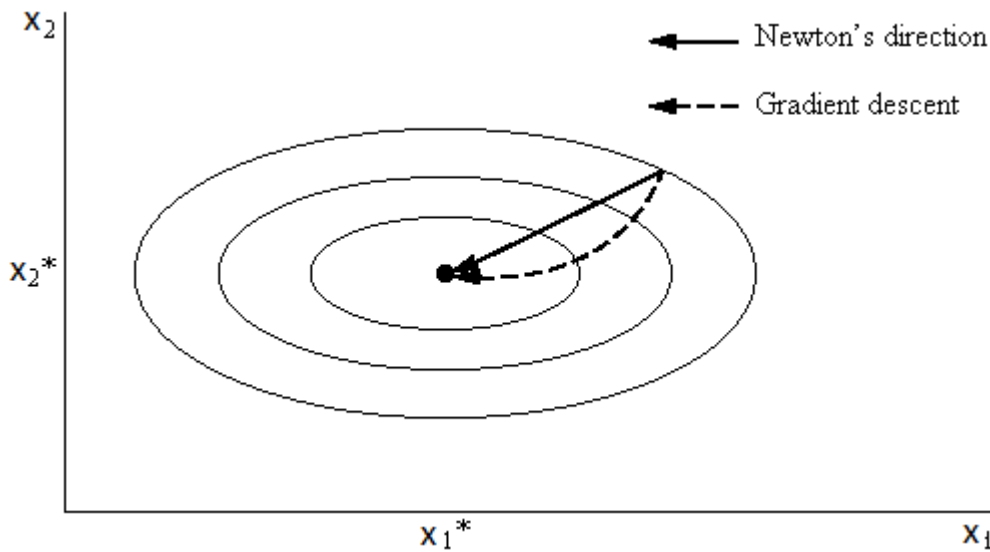
Έτσι, η σχέση (3.1) γράφεται ως εξής:

$$\vec{x}^{N+1} = \vec{x}^N - [\nabla^2 f(\vec{x}^N)]^{-1} \nabla f(\vec{x}^N)$$

Αν παρατηρηθεί η παραπάνω σχέση, θα διαπιστωθεί ότι υπάρχει μία προφανής διαφορά ανάμεσα στη μέθοδο της απότομης καθόδου και στη μέθοδο Newton. Η μέθοδος της απότομης καθόδου απαιτεί μόνο τον υπολογισμό των πρώτων παραγώγων της αντικειμενικής συνάρτησης ως προς τις μεταβλητές σχεδιασμού, ενώ η μέθοδος Newton απαιτεί και τον υπολογισμό (ή, σε ορισμένες παραλλαγές της που θα συζητηθούν παρακάτω, την προσέγγιση) του εσσιανού μητρώου της αντικειμενικής συνάρτησης. Ο υπολογισμός του ακριβούς εσσιανού μητρώου και η χρήση του σε μία 'ακριβή' μέθοδο Newton επιταχύνει, υπό προϋποθέσεις, σημαντικά τον αλγόριθμο βελτιστοποίησης. Είναι φανερό, όμως, ότι η ανάγκη αυτή εισάγει μεγαλύτερες υπολογιστικές δυσκολίες αλλά και αυξημένο υπολογιστικό κόστος κατά την εφαρμογή της δεδομένης μεθόδου. Για το λόγο αυτό, η μέθοδος Newton αξίζει να χρησιμοποιείται μόνο αν δύναται να προκαλέσει αισθητά ταχύτερη σύγκλιση σε σχέση με την μέθοδο της απότομης καθόδου.

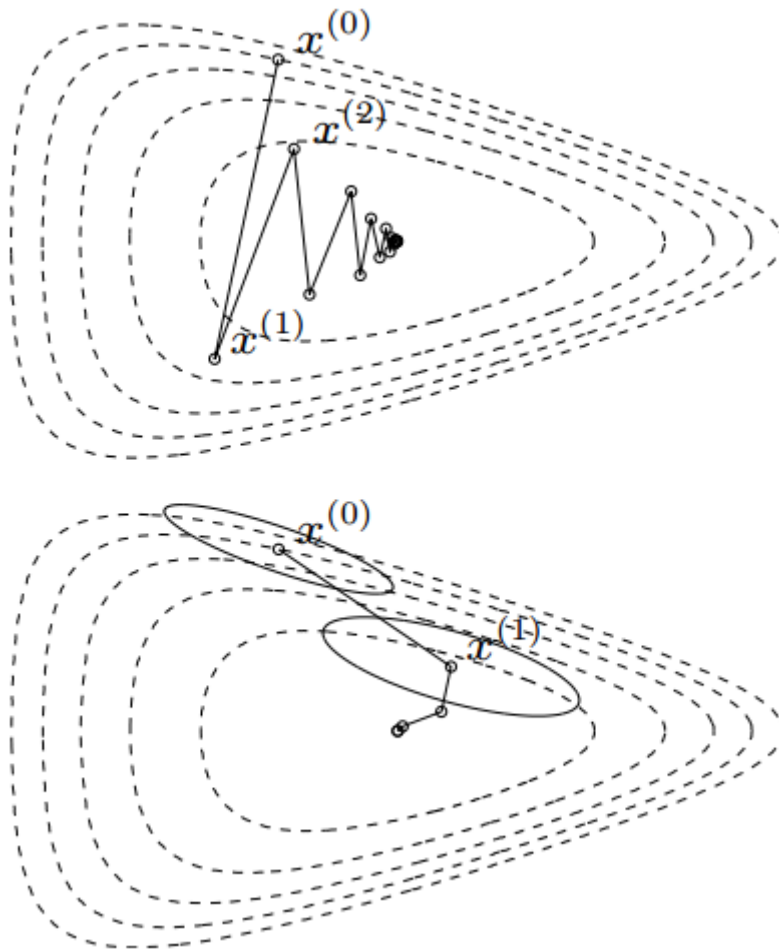
Αξίζει, επιπροσθέτως, να δοθεί έμφαση στο γεγονός ότι η μέθοδος Newton διορθώνει την κατεύθυνση της ανίχνευσης έτσι ώστε πάντα να δείχνει προς το ελάχιστο, ενώ, αντίθετα, η μέθοδος της απότομης καθόδου δείχνει προς την κατεύθυνση της μέγιστης αλλαγής. Οι κατευθύνσεις των δύο μεθόδων φαίνονται στο σχήμα 3.2. Οι δύο αυτές κατευθύνσεις μπορεί να συμπίπτουν αλλά μπορεί και όχι. Συμπίπτουν όταν οι ισοσταθμικές καμπύλες της αντικειμενικής συνάρτησης είναι κύκλοι, δηλαδή όταν η μεγαλύτερη και η μικρότερη ιδιοτιμή του εσσιανού μητρώου είναι ίδιες. Όταν, όμως, ο λόγος της

μεγαλύτερης προς τη μικρότερη ιδιοτιμή, δηλαδή η εξάπλωση των ιδιοτιμών (eigenvalue spread) αυξάνεται, η κλίση της επιφάνειας απόδοσης στις δύο κατευθύνσεις (για πρόβλημα δύο μεταβλητών σχεδιασμού) διαφέρει όλο και περισσότερο. Γι' αυτό το λόγο, για περιπτώσεις μεγάλης εξάπλωσης ιδιοτιμών, το 'μονοπάτι' βελτιστοποίησης που ακολουθεί η μέθοδος απότομης καθόδου είναι πολύ μακρύτερο από αυτό που ακολουθεί η μέθοδος Newton και άρα η μέθοδος Newton θα είναι πολύ γρηγορότερη. Φυσικά, το γεγονός αυτό δεν προκαλεί καμία έκπληξη αφού η μέθοδος Newton διαθέτει-χρησιμοποιεί πολύ περισσότερες πληροφορίες για την μορφή της αντικειμενικής συνάρτησης.



Σχήμα 3.2: Κατευθύνσεις των μεθόδων Newton και απότομης καθόδου, [9].

Στη συνέχεια, για να γίνει η σύγκριση της μεθόδου Newton με τη μέθοδο της απότομης καθόδου σε ένα πραγματικό πρόβλημα βελτιστοποίησης, παρατίθεται ένα παράδειγμα, το οποίο απεικονίζεται στο σχήμα 3.3.



Σχήμα 3.3: Πορεία σύγκλισης της μεθόδου της απότομης καθόδου (άνω σχήμα) και της μεθόδου Newton (κάτω σχήμα) για το πρόβλημα ελαχιστοποίησης της συνάρτησης $f(x_1, x_2) = e^{x_1+3x_2-0,1} + e^{x_1-3x_2-0,1} + e^{-x_1-0,1}$ (τα νούμερα δεν αναφέρονται διότι αυτό που ενδιαφέρει εδώ είναι η σύγκριση των πορειών σύγκλισης), [10].

Στο σημείο αυτό κρίνεται σκόπιμο να παρατεθεί το ακριβές ανάπτυγμα Taylor της $f(\vec{x}^N + \vec{p}^N)$:

$$f(\vec{x}^N + \vec{p}^N) = f(\vec{x}^N) + \vec{p}^{N^T} \nabla f(\vec{x}^N) + \frac{1}{2} \vec{p}^{N^T} \nabla^2 f(\vec{x}^N + t \vec{p}^N) \vec{p}^N, t \in (0,1)$$

Η κατεύθυνση ανίχνευσης που υπολογίζει η μέθοδος Newton είναι αξιόπιστη όταν η απόκλιση ανάμεσα στο προσεγγιστικό ανάπτυγμα Taylor και το ακριβές είναι μικρή. Πρέπει να τονιστεί ότι η προσεγγιστική σχέση είναι τετραγωνικής μορφής.

Για να κατασταθεί φανερό εάν η επιλογή κατεύθυνσης ανίχνευσης \vec{p}^N που πραγματοποιήθηκε παραπάνω οδηγεί σε μείωση της τιμής της αντικειμενικής συνάρτησης, δηλαδή εάν πραγματικά η μέθοδος Newton λειτουργεί ως μέθοδος καθόδου, πρέπει να υπολογιστεί το πρόσημο του γινομένου $\vec{p}^{N^T} \nabla f(\vec{x}^N)$. Έτσι, γράφεται η σχέση:

$$\vec{p}^{N^T} \nabla f(\vec{x}^N) = \vec{p}^{N^T} [-\nabla^2 f(\vec{x}^N) \vec{p}^N] = -\vec{p}^{N^T} \nabla^2 f(\vec{x}^N) \vec{p}^N$$

Το ότι το εσσιανό μητρώο $\nabla^2 f(\vec{x}^N)$ είναι θετικά ορισμένο οδηγεί στη βεβαιότητα ότι $\vec{p}^{N^T} \nabla^2 f(\vec{x}^N) \vec{p}^N > 0$. Άρα η μέθοδος Newton λειτουργεί ως μέθοδος καθόδου, αφού ισχύει $\vec{p}^{N^T} \nabla f(\vec{x}^N) < 0$.

Θεωρητικά, αλλά και πρακτικά, στις περισσότερες εφαρμογές της μεθόδου Newton, δεν υπεισέρχεται η βαθμωτή ποσότητα η^N (θεωρείται ότι $\eta^N = 1$) που καθορίζει το μέγεθος του βήματος κατά την κατεύθυνση \vec{p}^N . Στη βιβλιογραφία, ωστόσο, συναντώνται παραλλαγές της μεθόδου Newton οι οποίες λειτουργούν με $\eta^N = 1$ αλλά στην περίπτωση που δεν επιτυγχάνεται 'επαρκής' μείωση στην τιμή της συνάρτησης f , τροποποιείται η τιμή του η^N με εμπειρικές διορθωτικές σχέσεις.

3.3.4 Η προσεγγιστική μέθοδος Newton

Η προσεγγιστική μέθοδος Newton (Quasi-Newton), [4], αποτελεί παραλλαγή της κλασικής ('ακριβούς') μεθόδου Newton. Με την παραλλαγή αυτή αποφεύγεται το σημαντικό μειονέκτημα της μεθόδου Newton που είναι ο υπολογισμός του εσσιανού μητρώου της αντικειμενικής συνάρτησης, ενώ διατηρείται το πλεονέκτημα του πολύ καλού ρυθμού σύγκλισης (ιδίως καθώς πλησιάζει τη βέλτιστη λύση, αν βέβαια έχουν γίνει οι σωστές επιλογές).

Στο σημείο αυτό, κρίνεται αναγκαίο να παρουσιαστεί μία γενική σχέση για την επιλογή της κατεύθυνσης ανίχνευσης, η οποία είναι:

$$\vec{p}^N = -(B^N)^{-1} \nabla f(\vec{x}^N)$$

όπου το B^N είναι συμμετρικό και αντιστρέψιμο μητρώο. Στη μέθοδο της απότομης καθόδου έχει επιλεγθεί ως B^N το μοναδιαίο μητρώο, ενώ στη μέθοδο Newton το B^N ταυτίζεται με το εσσιανό μητρώο $\nabla^2 f(\vec{x}^N)$.

Κάθε προσεγγιστική μέθοδος Newton υπολογίζει και χρησιμοποιεί ως B^N μια προσέγγιση του εσσιανού μητρώου της αντικειμενικής συνάρτησης. Στην παρουσίαση της εν λόγω μεθόδου, λοιπόν, το σύμβολο B θα παριστάνει το προσεγγιστικό εσσιανό μητρώο.

Μία πρώτη ιδέα για τον υπολογισμό του μητρώου B βασίζεται σε 'ισοδύναμα' σχήματα πεπερασμένων διαφορών (άμεση συνέπεια αναπτυγμάτων κατά Taylor), σύμφωνα με τα οποία:

$$\nabla^2 f(\vec{x}^{N+1})(\vec{x}^{N+1} - \vec{x}^N) \approx \nabla f(\vec{x}^{N+1}) - \nabla f(\vec{x}^N)$$

(μέθοδος της τέμνουσας, secant method)

Οπότε το μητρώο B^{N+1} πρέπει να ικανοποιεί τη σχέση:

$$B^{N+1} \vec{s}^N = \vec{y}^N, \quad \text{όπου } \vec{s}^N = \vec{x}^{N+1} - \vec{x}^N \text{ και } \vec{y}^N = \nabla f(\vec{x}^{N+1}) - \nabla f(\vec{x}^N)$$

Πρακτικά, στο μητρώο B^{N+1} επιβάλλεται ο περιορισμός να είναι συμμετρικό, αφού συμμετρικό είναι και κάθε εσσιανό μητρώο.

Πρακτική αξία και ευρεία χρήση έχουν συναντήσει σχήματα τα οποία υλοποιούν αναδρομικούς τύπους, ώστε να υπολογίζεται το εσσιανό μητρώο στη νέα επανάληψη (έστω $N + 1$) από το ήδη γνωστό εσσιανό μητρώο της προηγούμενης (N) επανάληψης. Σε κάθε τέτοιο σχήμα, η αρχικοποίηση B^0 είναι, συνήθως, επιλογή του χρήστη που πρέπει, όμως, να ικανοποιεί τουλάχιστον την απαίτηση συμμετρίας.

Στη συνέχεια, παρατίθενται δύο από τις πιο γνωστές μεθόδους που υλοποιούν απευθείας τον υπολογισμό του B^{N+1} από το B^N . Αυτές είναι:

- Μέθοδος SR1 (Symmetric Rank One), κατά την οποία

$$B^{N+1} = B^N + \frac{(\vec{y}^N - B^N \vec{s}^N)(\vec{y}^N - B^N \vec{s}^N)^T}{(\vec{y}^N - B^N \vec{s}^N)^T \vec{s}^N}$$

- Μέθοδος BFGS (Broyden-Fletcher-Goldfarb-Shanno), κατά την οποία

$$B^{N+1} = B^N - \frac{B^N \vec{s}^N \vec{s}^{N^T} B^N}{\vec{s}^{N^T} B^N \vec{s}^N} + \frac{\vec{y}^N \vec{y}^{N^T}}{\vec{y}^{N^T} \vec{s}^N}$$

Και οι δύο αυτές μέθοδοι ικανοποιούν την εξίσωση της τέμνουσας και εξασφαλίζουν ότι κάθε παραγόμενο μητρώο B^{N+1} είναι, επιπλέον, συμμετρικό (αρκεί, βέβαια, να είναι συμμετρική η αρχικοποίηση B^0). Εναλλακτικά ως προς τις δύο ανωτέρω σχέσεις, οι οποίες εκ των πραγμάτων απαιτούν την αντιστροφή του ανανεωμένου μητρώου B^{N+1} , προτείνεται η παρακάτω σχέση που υπολογίζει και ανανεώνει απευθείας τον αντίστροφο του B που θα συμβολιστεί με H , δηλαδή $H^N = (B^N)^{-1}$ και $H^{N+1} = (B^{N+1})^{-1}$:

$$H^{N+1} = \left(I - \rho^N \vec{s}^N \vec{y}^{N^T} \right) H^N \left(I - \rho^N \vec{y}^N \vec{s}^{N^T} \right) + \rho^N \vec{s}^N \vec{s}^{N^T}, \quad \text{όπου } \rho^N = \frac{1}{\vec{y}^{N^T} \vec{s}^N}$$

Σε αυτή την περίπτωση, η κατεύθυνση ανίχνευσης προκύπτει άμεσα από τη σχέση:

$$\vec{p}^N = -H^N \nabla f(\vec{x}^N)$$

3.4 Εξελικτικοί αλγόριθμοι βελτιστοποίησης

3.4.1 Εισαγωγή

Από τη δεκαετία του 1990 και έκτοτε, η γεωμετρική αύξηση της υπολογιστικής ισχύος και το προσιτό κόστος των μεγάλων και γρήγορων υπολογιστών συνετέλεσαν στην ευρεία αποδοχή και χρήση των στοχαστικών μεθόδων βελτιστοποίησης, των οποίων ο βασικότερος εκπρόσωπος είναι οι εξελικτικοί αλγόριθμοι. Έχει ήδη γίνει αναφορά στους εν λόγω αλγόριθμους στην ενότητα 3.1, με ιδιαίτερη έμφαση στα πλεονεκτήματα και τα μειονεκτήματά τους. Συνεπώς, εδώ θα γίνει εστίαση στην περιγραφή του τρόπου λειτουργίας τους.

Βασικό γνώρισμα των εξελικτικών αλγορίθμων είναι ότι χειρίζονται *πληθυσμούς υποψηφίων λύσεων* και όχι μία μεμονωμένη λύση ανά επανάληψη, όπως άλλες στοχαστικές μέθοδοι για τις οποίες έγινε λόγος στην ενότητα 3.1 (λ.χ. η μέθοδος της προσομοιούμενης απόπτωσης). Οι εν λόγω αλγόριθμοι, αρχικά, αναπτύχθηκαν για την επίλυση προβλημάτων ενός στόχου, αλλά με κατάλληλες μετατροπές μπορούν να αντιμετωπίσουν προβλήματα πολλαπλών στόχων και μάλιστα υπερτερούν σε σχέση με άλλες μεθόδους λόγω του γεγονότος ότι χειρίζονται έναν πληθυσμό ατόμων και, με την ολοκλήρωσή τους, αποδίδουν ένα σύνολο λύσεων.

Με τον όρο «εξέλιξη» χαρακτηρίζουμε την αυτόματη διαδικασία της προσαρμογής κάθε συστήματος στο περιβάλλον. Ο όρος «περιβάλλον» περιλαμβάνει το σύνολο των εξωτερικών συνθηκών που επηρεάζουν το σύστημα. Οι εξελικτικοί αλγόριθμοι είναι υπολογιστικά μοντέλα τα οποία με χρήση αντίστοιχης με τη φύση διαδικασίας προσαρμογής στο περιβάλλον, επιλύουν προβλήματα. Για το σκοπό αυτό, χρησιμοποιούν στοχαστικούς μηχανισμούς εξέλιξης προερχόμενους από τη φύση, και βασίζονται στις αρχές εξέλιξης των ειδών όπως διατυπώθηκαν αρχικά από τον Δαρβίνο. Οι εν λόγω αλγόριθμοι προσομοιώνουν μερικώς και με απλουστευμένο τρόπο την εξέλιξη φυσικών πληθυσμών. Σύμφωνα με τη θεωρία του Δαρβίνου, τα άτομα ενός πληθυσμού ανταγωνίζονται για την απόκτηση πόρων όπως στέγη, τροφή καθώς και για την προσέλκυση συντρόφου. Τα επιτυχημένα άτομα έχουν μεγαλύτερη πιθανότητα να οδηγηθούν σε αναπαραγωγή και να διαιωισθούν. Αυτή η φυσική επιλογή έχει ως επακόλουθο τα γονίδια των επιτυχημένων ατόμων ή των ατόμων που έχουν προσαρμοστεί καλύτερα στο περιβάλλον να επιβιώνουν και με την πάροδο των γενεών τα χαρακτηριστικά τους να μεταφέρονται σε μεγαλύτερο αριθμό απογόνων. Ο συνδυασμός καλών χαρακτηριστικών από διαφορετικούς επιτυχημένους γονείς μπορεί να παράγει, ενδεχομένως, απογόνους με χαρακτηριστικά καλύτερα από αυτά των γονέων. Με τον τρόπο αυτό, οι φυσικοί πληθυσμοί εξελίσσονται και προσαρμόζονται στο εκάστοτε περιβάλλον.

Οι *γενετικοί αλγόριθμοι* είναι, μάλλον, η πιο διαδεδομένη κατηγορία εξελικτικών αλγορίθμων (ιδίως στον κλάδο της αεροδυναμικής και των στροβιλομηχανών), ενώ οι εξελικτικές στρατηγικές είναι μία λιγότερο χρησιμοποιούμενη τεχνική. Μία τρίτη κατηγορία, αρκετά διαφορετική από τις δύο προαναφερθείσες, είναι ο γενετικός προγραμματισμός,

[16, 17, 55, 56], ο οποίος αποτελεί μια σημαντική παραλλαγή των γενετικών αλγορίθμων όπου ως άτομα του πληθυσμού χρησιμοποιούνται προγράμματα υπολογιστή. Αν και καθεμία από τις παραπάνω κατηγορίες των εξελικτικών αλγορίθμων έχει τα δικά της διακριτά γνωρίσματα, η αυστηρή ταξινόμησή τους στερείται πλέον νοήματος, καθώς οι σύγχρονοι εξελικτικοί αλγόριθμοι δανείζονται στοιχεία (π.χ. τρόπο κωδικοποίησης, τελεστές εξέλιξης) από όλες τις κατηγορίες ανάλογα με το πρόβλημα.

Τα χαρακτηριστικά που πρέπει να έχει ένας αλγόριθμος βελτιστοποίησης για να χαρακτηριστεί εξελικτικός είναι τα εξής, [4]:

- Να χρησιμοποιεί πληθυσμούς ατόμων τα οποία εξελίσσονται συγχρόνως και όχι μεμονωμένα άτομα.
- Η εξέλιξη του πληθυσμού από γενιά σε γενιά (ορολογία αντίστοιχη των κύκλων/επαναλήψεων στις αιτιοκρατικές μεθόδους) να καθορίζεται από την τιμή καταλληλότητας (προβλήματα μεγιστοποίησης) ή την τιμή κόστους των ατόμων (προβλήματα ελαχιστοποίησης) με βάση κατάλληλη αντικειμενική συνάρτηση.
- Να λαμβάνει χώρα δυναμική διαδοχή πληθυσμών, με τη δημιουργία νέων ατόμων και εξαφάνιση άλλων με βάση την τιμή καταλληλότητάς/κόστους τους.
- Κατά την εξέλιξη από γενιά σε γενιά, να υπάρχει κληρονομικότητα χαρακτηριστικών από τους γονείς στους απογόνους, αλλά και στοχαστική εμφάνιση νέων χαρακτηριστικών.

3.4.2 Περιγραφή ενός εξελικτικού αλγορίθμου

Στην παρούσα ενότητα περιγράφεται ένας γενικευμένου εξελικτικού αλγορίθμου' με σύμμεικτα χαρακτηριστικά τα οποία έχουν ληφθεί από τους γενετικούς αλγορίθμους και τις εξελικτικές στρατηγικές (με απλές παρεμβάσεις ο αλγόριθμος μεταπίπτει στη μία κατηγορία ή την άλλη), [4].

Όπως αναφέρθηκε παραπάνω, οι εξελικτικοί αλγόριθμοι χειρίζονται πληθυσμούς λύσεων. Κατά την εξέλιξη, δηλαδή τη διαδικασία αναζήτησης της βέλτιστης λύσης, ένας πληθυσμός μ υποψήφιων λύσεων (γονείς, κατά τη γενετική ορολογία) εξελίσσεται στον πληθυσμό των λ απογόνων. Αυτοί οι απόγονοι είναι νέες λύσεις που προκύπτουν από τους μ γονείς και ενδεχομένως διαθέτουν καλύτερα χαρακτηριστικά. Από τους λ απογόνους, με κριτήριο την καταλληλότητά τους, επιλέγονται οι μ γονείς της επόμενης γενιάς. Αυτό συνεχίζεται, γενιά με γενιά, μέχρι να ικανοποιηθεί κάποιο κριτήριο σύγκλισης. Υπάρχουν αρκετά τέτοια κριτήρια και ένας εξελικτικός αλγόριθμος μπορεί να έχει σε ισχύ περισσότερα του ενός από αυτά. Μεταξύ αυτών είναι:

- ✗ Η μη περαιτέρω βελτίωση της λύσης για έναν αριθμό αξιολογήσεων γενεών
- ✗ Η ομογενοποίηση του πληθυσμού
- ✗ Η ανάλωση του υπολογιστικού χρόνου που επέτρεπε ο χρήστης

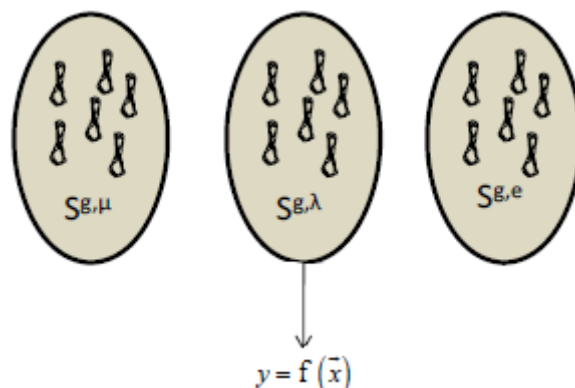
Σε κάθε γενιά g ενός εξελικτικού αλγορίθμου, μπορεί να θεωρηθεί ότι συνυπάρχουν τρία διακριτά σύνολα ατόμων. Συγκεκριμένα, έχουμε το σύνολο των γονέων $S^{g,\mu}$ με μ μέλη, των απογόνων $S^{g,\lambda}$ με λ μέλη και των επίλεκτων ή ελίτ ατόμων $S^{g,e}$ με e μέλη. Το σύνολο $S^{g,e}$ αποθηκεύει τα επίλεκτα-καλύτερα άτομα (δηλαδή υποψήφιες λύσεις) που έχουν προκύψει από τη διαδικασία εξέλιξης μέχρι την τρέχουσα γενιά. Ο σκοπός ύπαρξης του συνόλου των επίλεκτων είναι διπλός. Ο πρώτος λόγος είναι ότι, οποτεδήποτε σταματήσει η εξέλιξη, το σύνολο αυτό περιλαμβάνει την ή τις «βέλτιστες» λύσεις και ο δεύτερος λόγος είναι ότι, κατά την εξέλιξη, χρησιμοποιείται πληροφορία από το εν λόγω σύνολο για την ενίσχυση των καλών χαρακτηριστικών των μελών της νέας γενιάς που θα προκύψει. Η ενίσχυση αυτή αποτελεί τον λεγόμενο 'ελιτισμό' στους εξελικτικούς αλγορίθμους. Ακολούθως, θα περιγραφεί βήμα προς βήμα το τι λαμβάνει χώρα σε κάθε γενιά ενός εξελικτικού αλγορίθμου με την υπόθεση ότι αναφερόμαστε σε προβλήματα ενός στόχου.

Βήμα 1: Επιλέγονται από τον χρήστη βασικές παράμετροι της μεθόδου, όπως το μέγεθος των δύο βασικών πληθυσμών $S^{g,\mu}$ και $S^{g,\lambda}$, δηλαδή οι τιμές των μ και λ . Θέτουμε $g = 0$, συμβολίζοντας έτσι την αρχικοποίηση, τη λεγόμενη και μηδενική γενιά. Επιλέγονται τυχαία τα μέλη του αρχικού πληθυσμού $S^{0,\lambda}$.

Βήμα 2: Αξιολογούνται τα λ άτομα του συνόλου $S^{g,\lambda}$. Σχηματικά, το βήμα αυτό περιγράφεται από τη διαδικασία:

$$y(\vec{x}) = f(\vec{x}), \quad \forall \vec{x} \in S^{g,\lambda}$$

και φαίνεται στο σχήμα 3.4.



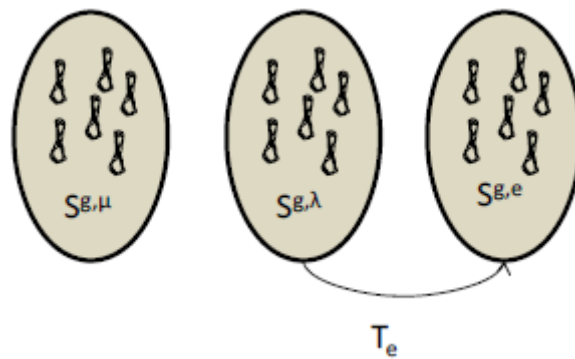
Σχήμα 3.4: Αξιολόγηση των μελών του $S^{g,\lambda}$, [11].

Αν πρόκειται για πρόβλημα βελτιστοποίησης στην αεροδυναμική, και για την αξιολόγηση απαιτούνται αριθμητικές αναλύσεις του πεδίου ροής, το βήμα αυτό πρακτικά απαιτεί λ κλήσεις του κώδικα ανάλυσης της ροής (λογισμικό αριθμητικής επίλυσης των εξισώσεων Navier-Stokes). Πρόκειται, δηλαδή, για το τμήμα του αλγορίθμου που φέρει το μεγαλύτερο (πρακτικά ολόκληρο το) υπολογιστικό κόστος.

Βήμα 3: Αναnevώνεται το σύνολο των επιλέκτων $S^{g,e}$ με όσα μέλη του $S^{g,\lambda}$ αξίζουν να ενταχθούν σε αυτό, με κριτήριο την τιμή της αντικειμενικής συνάρτησης. Σχηματικά, έχουμε:

$$S^{g+1,e} = T_e(S^{g,\lambda} \cup S^{g,e})$$

όπου με T_e συμβολίζεται ο τελεστής εντοπισμού των επίλεκτων μελών, όπως φαίνεται και στο σχήμα 3.5. Αν πρόκειται για πρόβλημα ενός στόχου, η απλούστερη εκδοχή αυτού του τελεστή είναι να εντοπίζει το (ένα) καλύτερο άτομο από τα δύο σύνολα που επεξεργάζεται ($S^{g,\lambda}$ και $S^{g,e}$). Προφανώς, υπάρχει πάντα η περίπτωση κανένας από τους λ να μην υπερτερεί της τρέχουσας βέλτιστης λύσης που αποτελεί και το μοναδικό στοιχείο του συνόλου των ελίτ. Τότε το σύνολο $S^{g,e}$ παραμένει ως έχει και αυτό είναι ένδειξη ότι ο εξελικτικός αλγόριθμος δεν κατάφερε να βελτιώσει την λύση κατά την τελευταία γενιά. Τονίζεται ότι ο εξελικτικός αλγόριθμος δεν είναι υποχρεωτικό να βελτιώνει την λύση σε κάθε γενιά. Σε προβλήματα πολλών στόχων (αλλά ακόμα και σε προβλήματα ενός στόχου αν έτσι το ορίσει ο χρήστης) το σύνολο $S^{g,e}$ μπορεί να περιέχει περισσότερα του ενός (πλήθος e) στοιχεία.

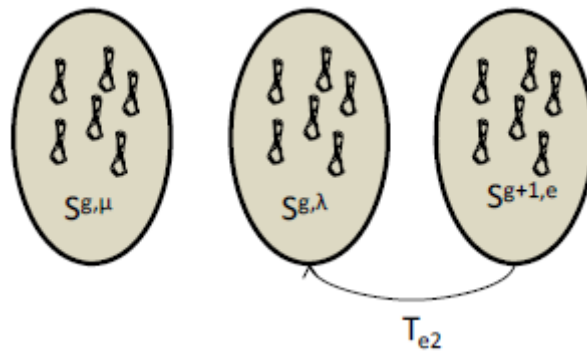


Σχήμα 3.5: Αναnevωση των μελών του $S^{g,e}$, [11].

Βήμα 4: Εφαρμόζεται ο τελεστής επιλεκτικότητας ή ελιτισμού, με τον οποίο (συνήθως τυχαία επιλεγόμενα) άτομα από το σύνολο των επιλέκτων $S^{g,e}$ αντικαθιστούν μέλη του συνόλου των απογόνων $S^{g,\lambda}$. Το ποια θα αντικατασταθούν εξαρτάται από την τιμή της αντικειμενικής συνάρτησης (συνήθως τα χειρότερα), αν και μέρος αυτών μπορεί να επιλέγεται τυχαία. Σχηματικά, έχουμε:

$$S^{g,\lambda} = T_{e2}(S^{g,\lambda} \cup S^{g+1,e})$$

όπου με T_{e2} συμβολίζεται ο τελεστής ελιτισμού. Το δεδομένο βήμα αναπαριστάται στο σχήμα 3.6. Με τον ελιτισμό αποφεύγεται μία νέα γενιά να δώσει βέλτιστη λύση χειρότερη από αυτή της προηγούμενης γενιάς.

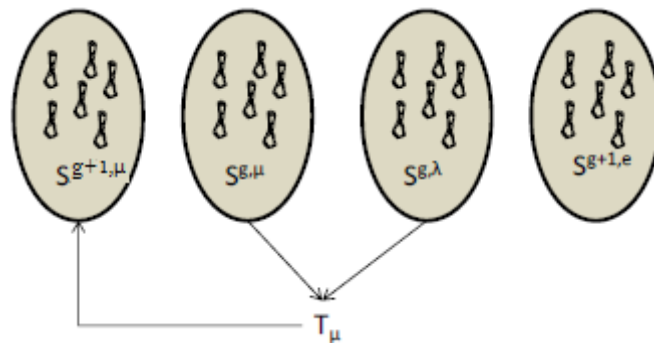


Σχήμα 3.6: Ανανέωση των μελών του $S^{g,\lambda}$, [11].

Βήμα 5: Εφαρμόζεται ο τελεστής επιλογής γονέων T_μ με τη βοήθεια του οποίου επιλέγεται το σύνολο των μελών του νέου πληθυσμού γονέων $S^{g+1,\mu}$. Για το σχηματισμό του χρησιμοποιούνται οι τρέχοντες πληθυσμοί γονέων $S^{g,\mu}$ και απογόνων $S^{g,\lambda}$, αν και είναι δυνατό οι μ γονείς να επιλέγονται μόνο από τους πρόσφατους λ απογόνους. Σχηματικά, έχουμε:

$$S^{g+1,\mu} = T_\mu(S^{g,\mu} \cup S^{g,\lambda})$$

όπου με T_μ συμβολίζεται ο τελεστής επιλογής γονέων. Το βήμα αυτό αναπαριστάται στο σχήμα 3.7.



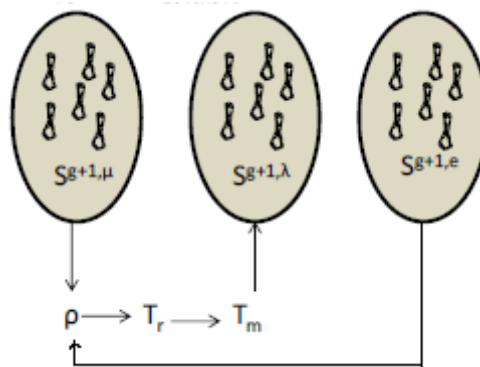
Σχήμα 3.7: Επιλογή μελών του νέου συνόλου γονέων $S^{g+1,\mu}$, [11].

Βήμα 6: Σε αυτό το βήμα εκτελείται η διαδικασία αναπαραγωγής για τη δημιουργία του συνόλου απογόνων της επόμενης γενιάς $S^{g+1,\lambda}$. Για τη δημιουργία καθενός από τα λ άτομα του επιλέγονται τυχαία ρ (συνήθως $\rho = 2$, αλλά δεν αποκλείεται η περίπτωση $\rho > 2$, π.χ όταν τρεις γονείς δημιουργούν τρεις απογόνους, κλπ) μέλη από το σύνολο γονέων $S^{g+1,\mu}$. Είναι, επίσης, δυνατό κάποιο ή κάποια από τα ρ να προέρχονται από το $S^{g+1,e}$, επιλεγόμενα συνήθως με μικρή

πιθανότητα για να μην εμφανίζει ο αλγόριθμος ισχυρά χαρακτηριστικά ελιτισμού. Για κάθε σύνολο ρ γονέων που επιλέχθηκαν, εφαρμόζονται διαδοχικά διάφοροι τελεστές, μεταξύ των οποίων οι πιο γνωστοί είναι ο τελεστής διασταύρωσης ή επανασυνδυασμού T_r (crossover/recombination operator) και ο τελεστής μετάλλαξης T_m (mutation operator). Σχηματικά, έχουμε:

$$S^{g+1,\lambda} = T_m[T_r(S^{g+1,\mu} \cup S^{g+1,e})]$$

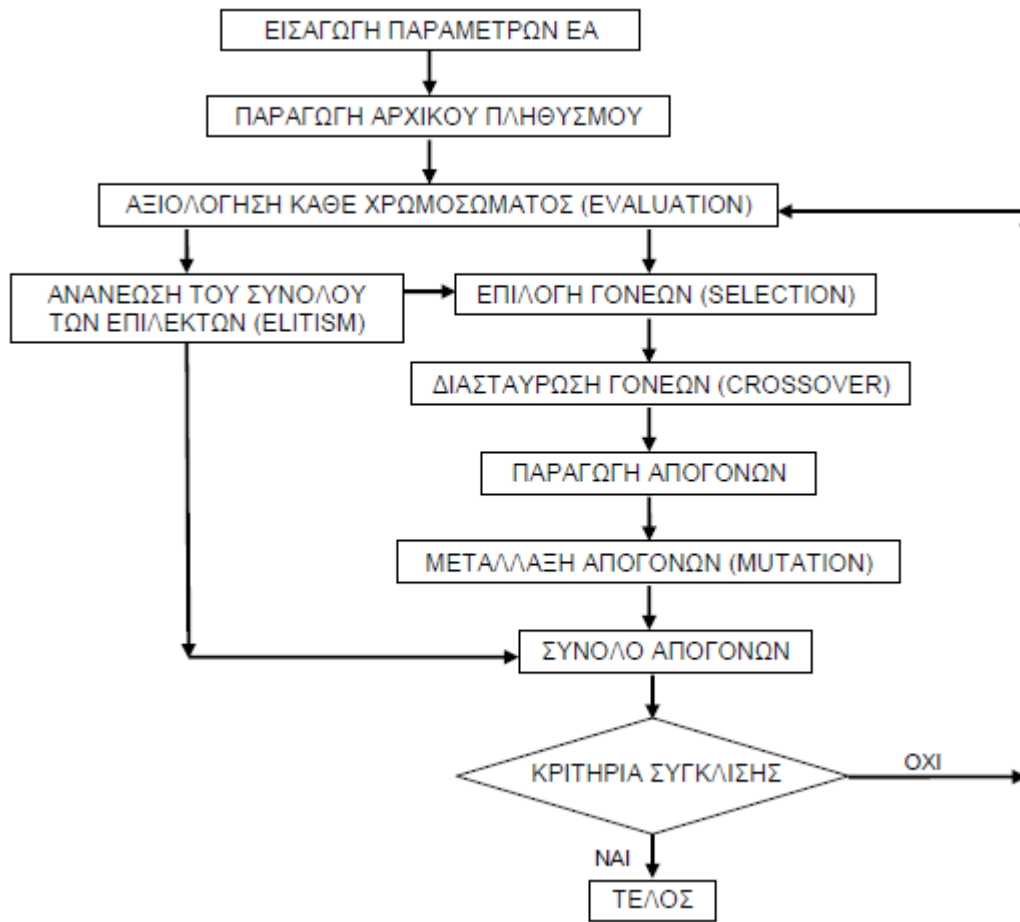
Το βήμα αυτό απεικονίζεται στο σχήμα 3.8.



Σχήμα 3.8: Επιλογή μελών του νέου συνόλου απογόνων $S^{g+1,\lambda}$, [11].

Βήμα 7: Εφαρμόζεται το κριτήριο σύγκλισης και αν η μέθοδος θεωρείται ότι δεν έχει συγκλίνει, αρχίζει μία νέα γενιά, δηλαδή επαναλαμβάνονται τα βήματα 2 έως 7, θέτοντας $g \leftarrow g + 1$.

Τα παραπάνω βήματα (1-7) απεικονίζονται στο διάγραμμα ροής του σχήματος 3.9.



Σχήμα 3.9: Διαγραμματική παρουσίαση της λειτουργίας ενός γενικευμένου εξελικτικού αλγορίθμου, [12].

4 Συνέχιση του μετώπου Pareto στην πολυκριτηριακή βελτιστοποίηση (Pareto-curve continuation in multi-objective optimization)

4.1 Εισαγωγή

Στόχος του παρόντος κεφαλαίου και πρωτεύων στόχος της παρούσας εργασίας είναι να ερευνηθεί η γενίκευση υπάρχόντων επιλυτών για προβλήματα βελτιστοποίησης ενός στόχου στο πλαίσιο της πολυκριτηριακής βελτιστοποίησης, [1], με κίνητρο την ελάττωση του αριθμού των αξιολογήσεων που απαιτείται να πραγματοποιηθούν για την εύρεση του μετώπου των μη-κυριαρχούμενων λύσεων, δηλαδή του υπολογιστικού κόστους. Σε πρώτο στάδιο, αναλύεται το μαθηματικό υπόβαθρο του προβλήματος. Στη συνέχεια, παρουσιάζεται η θεωρία της συνέχισης του μετώπου Pareto, πάνω στην οποία 'χτίζονται', έπειτα, οι μέθοδοι πρόβλεψης-διόρθωσης. Τέλος, παρατίθεται μία εφαρμογή που αντλήθηκε από τη βιβλιογραφία, [1].

4.2 Μαθηματικό υπόβαθρο

4.2.1 Διατύπωση προβλήματος

Ένα πρόβλημα βελτιστοποίησης πολλών στόχων με περιορισμούς διατυπώνεται ως εξής:

$$\min \vec{f}(\vec{x}) =: \min \begin{bmatrix} f_1(\vec{x}) \\ \vdots \\ f_m(\vec{x}) \end{bmatrix}$$

που υπόκειται σε (subject to – s.t.)

$$\begin{aligned} c_i(\vec{x}) &= 0, & i &= 1, \dots, m_1 \\ g_j(\vec{x}) &\geq 0, & j &= 1, \dots, m_2 \end{aligned}$$

όπου $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $c: \mathbb{R}^n \rightarrow \mathbb{R}^{m_1}$, $g: \mathbb{R}^n \rightarrow \mathbb{R}^{m_2}$ και είναι επαρκώς διαφορίσιμες (sufficiently differentiable).

Το σύνολο των εφικτών λύσεων (set of all feasible solutions) ορίζεται ως εξής:

$$X := \{ \vec{x} \in \mathbb{R}^n \mid c_i(\vec{x}) = 0, i = 1, \dots, m_1, g_j(\vec{x}) \geq 0, j = 1, \dots, m_2 \}$$

Όπως αναφέρθηκε παραπάνω (ενότητα 2.5), στην περίπτωση ενός προβλήματος βελτιστοποίησης πολλών στόχων, η έννοια της μοναδικής βέλτιστης λύσης $\vec{x}^* \in X$ δεν υφίσταται αλλά επιβάλλεται να αντικατασταθεί από αυτήν ενός συνόλου λύσεων, με τις λύσεις αυτές να αποτελούν το σύνολο – μέτωπο Pareto. Επίσης, αναφέρθηκε ότι το σύνολο βέλτιστων λύσεων που προκύπτει από τη διαδικασία της πολυκριτηριακής βελτιστοποίησης έχει το εξής χαρακτηριστικό: οι λύσεις που εμπεριέχονται σε αυτό είναι μη-κυριαρχούμενες μεταξύ τους (η έννοια αναλύθηκε στην ενότητα 2.5.3.1).

4.2.2 Μέθοδος βαρών

Ο ορισμός του ότι μία λύση \vec{x}^* λέγεται βέλτιστη κατά Pareto αν είναι μη-κυριαρχούμενη, είναι ιδιαίτερα χρήσιμος στις εξελικτικές προσεγγίσεις, όπου μόνο οι μη-κυριαρχούμενες λύσεις επιτρέπεται να μεταφερθούν στην επόμενη γενιά. Όμως, στις στρατηγικές συνέχισης καμπύλης (curve continuation strategies), όπως θα αναλυθεί παρακάτω, πολύ πιο χρήσιμος είναι ο χαρακτηρισμός που παρέχεται από τη μέθοδο των βαρών. Η προσέγγιση αυτή συνίσταται στη στάθμιση των συναρτήσεων στόχου με χρήση προεπιλεγμένων συντελεστών βάρους, οπότε η αντικειμενική συνάρτηση προκύπτει ως γραμμικός συνδυασμός των συνιστωσών της διανυσματικής συνάρτησης στόχου. Συνεπώς, το πρόβλημα βελτιστοποίησης πολλών στόχων μετασχηματίζεται σε βαθμωτό της μορφής:

$$\min_{\vec{x} \in X} \sum_{i=1}^m w_i f_i(\vec{x}) =: \min_{\vec{x} \in X} \vec{w}^T \vec{f}(\vec{x}) =: \min_{\vec{x} \in X} f_w(\vec{x}) \quad (4.1)$$

όπου $\vec{w} \in \mathbb{R}^m$ με τις συνιστώσες του w_i να είναι μη-αρνητικές και να αποτελούν τους συντελεστές βάρους που υποδηλώνουν την σχετική σημασία των επιμέρους κριτηρίων.

Αν \vec{x}^* είναι η λύση του παραπάνω προβλήματος μπορεί εύκολα να αποδειχθεί διά της ατόπου (by contradiction) ότι η εν λόγω λύση είναι μη-κυριαρχούμενη και άρα ανήκει στο μέτωπο Pareto του προβλήματος βελτιστοποίησης πολλών στόχων. Αυτό δηλώνει ότι η ελαχιστοποίηση του γραμμικού συνδυασμού (για κάποιες τιμές των w_i) εξασφαλίζει ότι θα πάρουμε ένα σημείο του μετώπου Pareto αλλά το αντίστροφο γενικά δεν ισχύει.

Συνήθως θεωρείται ότι:

$$\sum_{i=1}^m w_i = 1$$

Με την προσθήκη του ανωτέρω επιπλέον περιορισμού καθίσταται δυνατό να μειώσουμε κατά έναν πλεονάζοντα τους βαθμούς

ελευθερίας στην παραμετροποίηση που έχει ως αποτέλεσμα τον κυρτό συνδυασμό (convex combination) των συναρτήσεων στόχου.

Όπως αναφέρθηκε, αν επιλεγούν κάποιες τιμές για τα βάρη w_i και επιλυθεί το ανωτέρω μετασχηματισμένο πρόβλημα βελτιστοποίησης θα προκύψει μία λύση \vec{x}^* η οποία θα αποτελεί μέλος του Pareto του αρχικού προβλήματος βελτιστοποίησης πολλών στόχων. Είναι φανερό, λοιπόν, ότι αν πραγματοποιηθεί η επίλυση του μετασχηματισμένου προβλήματος βελτιστοποίησης πολλές φορές με διαφορετικές τιμές στα βάρη w_i δύναται να επιτευχθεί η εύρεση του μετώπου Pareto του αρχικού προβλήματος βελτιστοποίησης πολλών στόχων.

Πρέπει, όμως, να ειπωθεί ότι επειδή η βέλτιστη λύση του εν λόγω προβλήματος εξαρτάται δραστικά από τις τιμές των συντελεστών βάρους και καθώς δεν υπάρχει αντικειμενικός τρόπος επιλογής τους, ενδέχεται να προκύψουν μη-αποδεκτές λύσεις στην περίπτωση που δεν υπάρχει επαρκής γνώση του προβλήματος. Ακόμα, αν επιχειρηθεί η απόκτηση του μετώπου Pareto μέσω της εν λόγω μεθόδου, είναι εξαιρετικά πιθανό να παρουσιαστεί πολύ κακή διανομή των σχεδιασμών – σημείων πάνω στο μέτωπο που θα προκύψει. Επιπλέον, η ενέργεια της επίλυσης του μετασχηματισμένου προβλήματος πολλές φορές αυξάνει πολύ το υπολογιστικό κόστος της βελτιστοποίησης.

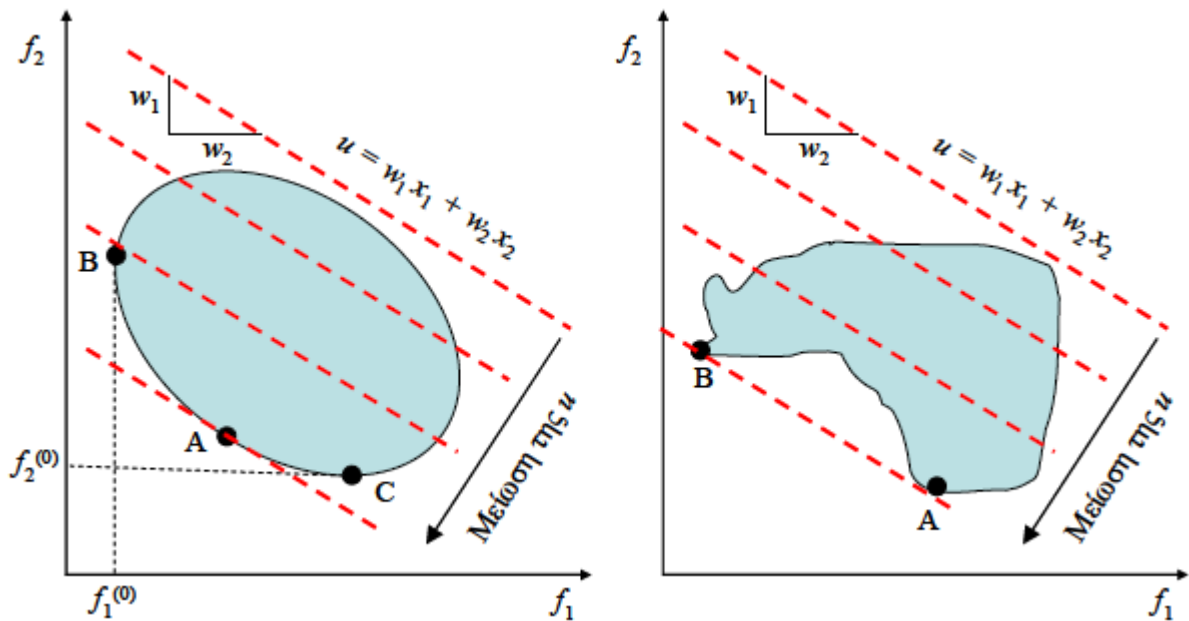
Η γεωμετρική ερμηνεία της μεθόδου απεικονίζεται στο σχήμα 4.1. Η βέλτιστη λύση κείται στο σημείο όπου η αντικειμενική συνάρτηση:

$$u(\vec{x}) = \sum_{i=1}^m w_i f_i(\vec{x})$$

εφάπτεται του πεδίου τιμών. Εφόσον το εν λόγω πεδίο είναι κυρτό, τότε μεταβάλλοντας τις τιμές των συντελεστών βάρους, δηλαδή την κλίση της ευθείας $u(\vec{x})$, εντοπίζονται διαφορετικές βέλτιστες λύσεις που είναι εξ ορισμού μη-κυριαρχούμενες.

Στο αριστερό τμήμα του σχήματος 4.1 απεικονίζεται ένα υποθετικό κυρτό πεδίο τιμών για την περίπτωση ενός προβλήματος ελαχιστοποίησης δύο στόχων, f_1 και f_2 . Το μέτωπο Pareto αντιστοιχεί στο κάτω όριο του πεδίου και περιλαμβάνει το σύνολο των μη-κυριαρχούμενων λύσεων που κείνται μεταξύ των σημείων B και C. Η ελαχιστοποίηση της συνάρτησης $u = w_1 x_1 + w_2 x_2$, η οποία στον χώρο των δύο διαστάσεων έχει την μορφή κεκλιμένου επιπέδου, οδηγεί στον εντοπισμό των μη-κυριαρχούμενων λύσεων του προβλήματος με την καθεμία να αντιστοιχεί μονοσήμαντα σε συγκεκριμένο συνδυασμό συντελεστών βάρους. Σε ένα πρόβλημα πολλών στόχων συνεχών μεταβλητών υπάρχουν άπειροι συνδυασμοί τιμών των συντελεστών βάρους. Συνεπώς, άπειρο είναι και το πλήθος των μη-κυριαρχούμενων λύσεων, με τις ακραίες να αντιστοιχούν σε μηδενικά βάρη για όλα πλην ενός κριτηρίου. Στο υποθετικό πρόβλημα του

σχήματος, οι ακραίες λύσεις B και C προκύπτουν για τα ζεύγη βαρών (1, 0) και (0, 1), αντίστοιχα.



Σχήμα 4.1: Γεωμετρική ερμηνεία της μεθόδου βαρών σε ένα πρόβλημα ελαχιστοποίησης δύο συναρτήσεων για την περίπτωση κυρτού (αριστερά) και μη-κυρτού (δεξιά) μετώπου Pareto. Το σκιασμένο πεδίο είναι ο εφικτός χώρος, ενώ με διακεκομμένη γραμμή παριστάνονται οι ισοσταθμικές της αντικειμενικής συνάρτησης, [2].

Πέραν των ήδη αναφερθέντων σε παραπάνω τμήμα της ενότητας προβλημάτων που συναντώνται στην περίπτωση επιλογής της μεθόδου των βαρών για την εύρεση του μετώπου Pareto, υπάρχει ένα ακόμα σοβαρό μειονέκτημα που σχετίζεται με την ανεπάρκεια της μεθόδου στην περίπτωση που το μέτωπο Pareto είναι μη-κυρτό. Όταν αντιμετωπίζεται ένα πρόβλημα στο οποίο ισχύει κάτι τέτοιο, είναι αδύνατος ο εντοπισμός των μη-κυριαρχούμενων λύσεων του εν λόγω προβλήματος, δηλαδή είναι αδύνατο να αποκτηθεί το μέτωπο Pareto. Μάλιστα, αποδεικνύεται ότι αυτό ισχύει για οποιοδήποτε συνδυασμό συντελεστών βάρους, [63]. Η γεωμετρική ερμηνεία της παραπάνω παρατήρησης παρέχεται στο δεξί τμήμα του σχήματος 4.2, όπου απεικονίζεται ένα υποθετικό μη-κυρτό πεδίο τιμών για την περίπτωση ενός προβλήματος ελαχιστοποίησης δύο στόχων, f_1 και f_2 . Στην περίπτωση αυτή, η ελαχιστοποίηση της γραμμικής αντικειμενικής συνάρτησης $u = w_1 x_1 + w_2 x_2$ οδηγεί στα όρια A και B του μη-κυρτού πεδίου, όπου η καμπύλη της u εφάπτεται του εφικτού χώρου. Ωστόσο, καμία άλλη ενδιαμέση λύση δεν είναι δυνατόν να εντοπιστεί, δηλαδή δεν υπάρχει κανένας συνδυασμός βαρών με τη χρήση των οποίων δύναται να αποκτηθεί κάποια ενδιαμέση λύση αλλά πάντα θα λαμβάνεται μία από τις δύο ακραίες.

4.2.3 Ύπαρξη συνόλου μη-κυριαρχούμενων λύσεων (indifference set)

Όπως καταδείχθηκε στην προηγούμενη ενότητα, με τη χρήση του κυρτού συνδυασμού των επιμέρους συναρτήσεων στόχου (γραμμικός συνδυασμός συναρτήσεων στόχου όπου όλοι οι συντελεστές είναι μη-αρνητικοί και το άθροισμά τους ισούται με 1), το πρόβλημα εύρεσης μίας μη-κυριαρχούμενης λύσης, δηλαδή ενός σημείου του Pareto, ανάγεται στην επίλυση ενός προβλήματος βελτιστοποίησης ενός στόχου στο οποίο εμπεριέχονται περιορισμοί για κάποιες συγκεκριμένες τιμές των συντελεστών w_i ή για ένα δεδομένο διάνυσμα \vec{w} . Σε συνδυασμό με το θεώρημα πεπλεγμένης συνάρτησης (implicit function theorem), το οποίο θα παρουσιασθεί αναλυτικά παρακάτω, η ύπαρξη της λύσης του εν λόγω μετασχηματισμένου προβλήματος βελτιστοποίησης οδηγεί στην ύπαρξη του συνόλου των μη-κυριαρχούμενων λύσεων ή μέτωπο Pareto. Για το λόγο αυτό, η συγκεκριμένη ενότητα επικεντρώνεται στη θεωρία επίλυσης ενός προβλήματος βελτιστοποίησης ενός στόχου με περιορισμούς, η ανάλυση της οποίας έχει πραγματοποιηθεί διεξοδικά στο κεφάλαιο 2. Το πρόβλημα, λοιπόν, που τίθεται προς αντιμετώπιση είναι το εξής:

$$\min f(\vec{x})$$

που υπόκειται σε (subject to – s.t.)

$$\begin{aligned} c_i(\vec{x}) &= 0, & i &= 1, \dots, m_1 \\ g_j(\vec{x}) &\geq 0, & j &= 1, \dots, m_2 \end{aligned}$$

όπου $f \in C_2(\mathbb{R}^n, \mathbb{R})$, $c \in C_2(\mathbb{R}^n, \mathbb{R}^{m_1})$, $g \in C_2(\mathbb{R}^n, \mathbb{R}^{m_2})$.

Το ανωτέρω πρόβλημα είναι στην ουσία η σχέση 4.1 αν παραλείψουμε τους συντελεστές w_i (και βέβαια εδώ προστέθηκαν και οι περιορισμοί).

Στην ενότητα 2.4.2.3 διατυπώθηκε η συνάρτηση Lagrange (σχέση 2.8), η οποία για το δεδομένο πρόβλημα γράφεται ως εξής:

$$L(\vec{x}, \vec{\lambda}, \vec{\mu}) := f(\vec{x}) - \vec{\lambda}^T \vec{c}(\vec{x}) - \vec{\mu}^T \vec{g}(\vec{x}) = f(\vec{x}) - \sum_{i=1}^{m_1} \lambda_i c_i(\vec{x}) - \sum_{j=1}^{m_2} \mu_j g_j(\vec{x})$$

όπου $\vec{x} \in \mathbb{R}^n$, $\vec{\lambda} \in \mathbb{R}^{m_1}$, $\vec{\mu} \in \mathbb{R}^{m_2}$.

Στην ενότητα 2.4.2.3 διατυπώθηκαν, επίσης, οι πρώτης τάξης αναγκαίες συνθήκες για την ύπαρξη βέλτιστης λύσης ενός προβλήματος βελτιστοποίησης ενός στόχου με περιορισμούς που είναι γνωστές ως συνθήκες Karush-Kuhn-Tucker (ή KKT). Επιπλέον,

στην ενότητα 2.4.2.4 παρουσιάσθηκαν οι δεύτερης τάξης ικανές συνθήκες για την ύπαρξη βέλτιστης λύσης του εν λόγω προβλήματος. Για το δεδομένο πρόβλημα, οι συνθήκες KKT γράφονται ως εξής:

$$\nabla_x L(\vec{x}, \vec{\lambda}, \vec{\mu}) = 0 \Leftrightarrow \nabla_x f(\vec{x}) - \sum_{i=1}^{m_1} \lambda_i \nabla_x c_i(\vec{x}) - \sum_{j=1}^{m_2} \mu_j \nabla_x g_j(\vec{x}) = 0$$

$$c_i(\vec{x}) = 0, \quad \forall i \in m_1$$

$$g_j(\vec{x}) \geq 0, \quad \forall j \in m_2$$

$$\mu_j \geq 0, \quad \forall j \in m_2$$

$$\mu_j g_j(\vec{x}) = 0, \quad \forall j \in m_2$$

Στόχος είναι να εφαρμοστεί το θεώρημα πεπλεγμένης συνάρτησης (το οποίο θα παρουσιαστεί παρακάτω) στις ανωτέρω συνθήκες KKT. Δυστυχώς, όμως, το εν λόγω θεώρημα δεν 'επιτρέπει' την ύπαρξη ανισοτικών περιορισμών. Ωστόσο, έχει αποδειχθεί ότι το παραπάνω σύστημα που αποτελεί τις συνθήκες KKT δύναται να επαναδιατυπωθεί χωρίς την παρουσία ανισοτήτων, [13]. Για λόγους απλότητας, στο παρόν κεφάλαιο γίνεται χρήση της προσέγγισης του συνόλου των ενεργών περιορισμών ανισότητας (active set) και υποτίθεται ότι οι περιορισμοί ανισότητας είναι πάντα ενεργοί (δηλαδή, όπως αναφέρθηκε στην ενότητα 2.4.2.2, ισχύει $g_j(\vec{x}) = 0, \forall j \in m_2$), ώστε να μπορούν να παραλειφθούν περαιτέρω αναφορές σε περιορισμούς ανισότητας. Με άλλα λόγια, δύναται να ειπωθεί ότι γίνεται αναφορά σε προβλήματα που περιέχουν μόνο ισοτικούς περιορισμούς. Κάτι τέτοιο δεν αποτελεί κενό στη θεωρία, αφού, όπως αναφέρθηκε, το ανωτέρω σύστημα μπορεί να ξαναγραφεί με την αποκλειστική χρήση ισοτικών σχέσεων.

Στην ειδική περίπτωση χρήσης μόνο ισοτικών περιορισμών και λαμβάνοντας υπόψη τη διατύπωση του προβλήματος με τη χρήση της μεθόδου των βαρών, οι συνθήκες KKT μπορούν να γραφούν ως εξής:

$$\vec{H}(\vec{x}, \vec{\lambda}, \vec{w}) := \begin{bmatrix} \nabla_x f_w(\vec{x}) - \sum_{i=1}^{m_1} \lambda_i \nabla_x c_i(\vec{x}) \\ \vec{c}(\vec{x}) \end{bmatrix} = \begin{bmatrix} \nabla_x f_w(\vec{x}) - \vec{\lambda}^T \nabla_x \vec{c}(\vec{x}) \\ \vec{c}(\vec{x}) \end{bmatrix} = 0$$

4.2.4 Το θεώρημα πεπλεγμένης συνάρτησης (implicit function theorem)

Ορίζεται το διάνυσμα $\vec{z} := (\vec{x}, \vec{\lambda})^T$.

Ας υποθέσουμε ότι το σύνολο $\Omega \subset \mathbb{R}^d \times \mathbb{R}^k$ είναι ανοιχτό (open) και ότι $\vec{H} \in C_1(\Omega, \mathbb{R}^d)$. Επιπλέον, ας υποθέσουμε ότι $(\vec{z}, \vec{w}) \in \Omega$, έτσι ώστε $\vec{H}(\vec{z}, \vec{w}) = 0$ και $\det \frac{\partial \vec{H}}{\partial \vec{z}}(\vec{z}, \vec{w}) \neq 0$.

Τότε ισχύουν τα εξής, [1]:

- Υπάρχουν ανοιχτές γειτονιές (open neighborhoods) $U(\vec{w}) \subset \mathbb{R}^k$ και $W(\vec{z}, \vec{w}) \subset \Omega$, καθώς και μία συνάρτηση $\vec{h} : U \rightarrow \mathbb{R}^d$, η οποία για $(\vec{z}, \vec{w}) \in W$ ικανοποιεί:

$$\vec{H}(\vec{z}, \vec{w}) = 0 \Leftrightarrow \vec{z} = \vec{h}(\vec{w})$$

- Επιπροσθέτως, $\vec{h} \in C_1(U, \mathbb{R}^d)$ και:

$$\frac{\partial \vec{h}}{\partial \vec{w}}(\vec{w}) = - \left\{ \frac{\partial \vec{H}}{\partial \vec{z}}[\vec{h}(\vec{w}), \vec{w}] \right\}^{-1} \frac{\partial \vec{H}}{\partial \vec{w}}[\vec{h}(\vec{w}), \vec{w}]$$

Η απόδειξη αυτού του θεωρήματος μπορεί να βρεθεί σε διάφορες πηγές, λ.χ. [14]. Παρουσιάζει ενδιαφέρον το γεγονός ότι, οι προϋποθέσεις – απαιτήσεις για την ισχύ του θεωρήματος πεπλεγμένης συνάρτησης ικανοποιούνται αυτόματα όταν ικανοποιούνται οι πρώτης τάξης συνθήκες KKT και οι δεύτερης τάξης ικανές συνθήκες, [15]. Στο σημείο αυτό, λοιπόν, μπορούμε να θεωρήσουμε την ύπαρξη του συνόλου μη-κυριαρχούμενων λύσεων εφόσον το πρόβλημα ελαχιστοποίησης του κυρτού συνδυασμού είναι καλά τοποθετημένο (well posed).

4.3 Συνέχιση μετώπου Pareto

4.3.1 Μέθοδοι πρόβλεψης-διόρθωσης (Predictor-Corrector Methods)

Όπως αναφέρθηκε στην ενότητα 4.2.2, αν επιλεγθούν κάποιες τιμές για τα βάρη w_i και επιλυθεί το μετασχηματισμένο πρόβλημα βελτιστοποίησης, θα προκύψει μία λύση η οποία θα αποτελεί μέλος του μετώπου Pareto του αρχικού προβλήματος βελτιστοποίησης πολλών στόχων. Δύναται, λοιπόν, απλά να πραγματοποιηθεί η επίλυση του κυρτού συνδυασμού για έναν επαρκή αριθμό διαφορετικών τιμών των συντελεστών w_i , ώστε να αποκτηθεί το μέτωπο Pareto του προβλήματος βελτιστοποίησης πολλών στόχων. Όμως, κάτι τέτοιο, όπως σχολιάστηκε παραπάνω, συνήθως έχει ως αποτέλεσμα μία πολύ κακή διακριτοποίηση του μετώπου μη-κυριαρχούμενων λύσεων. Επιπλέον, η εν λόγω προσέγγιση έχει πολύ μεγάλο υπολογιστικό κόστος.

Κρίνεται, λοιπόν, σκόπιμο στην παρούσα εργασία να ερευνηθεί η γενίκευση υπαρχόντων επιλυτών για προβλήματα βελτιστοποίησης ενός στόχου στο πλαίσιο της πολυκριτηριακής βελτιστοποίησης, γεγονός που καθίσταται δυνατό μέσω των μεθόδων πρόβλεψης-διόρθωσης, [1].

Στο σημείο αυτό, ας υποθέσουμε ότι το διάνυσμα $\vec{z}^k := \vec{h}(\vec{w}^k)$ είναι το k -οστό σημείο πάνω στο μέτωπο των μη-κυριαρχούμενων λύσεων ή μέτωπο Pareto. Χωρίς να λαμβάνονται υπόψη ιδιομορφίες λόγω μίας αλλαγής στα πρόσημα των ιδιοτιμών του εσσιανού μητρώου ή μίας αλλαγής του συνόλου των ενεργών περιορισμών, μία εφαπτομενική προσέγγιση (tangential approximate) \vec{z}^{k+1} του επόμενου σημείου \vec{z}^{k+1} πάνω στο μέτωπο Pareto παρέχεται από την παρακάτω σχέση:

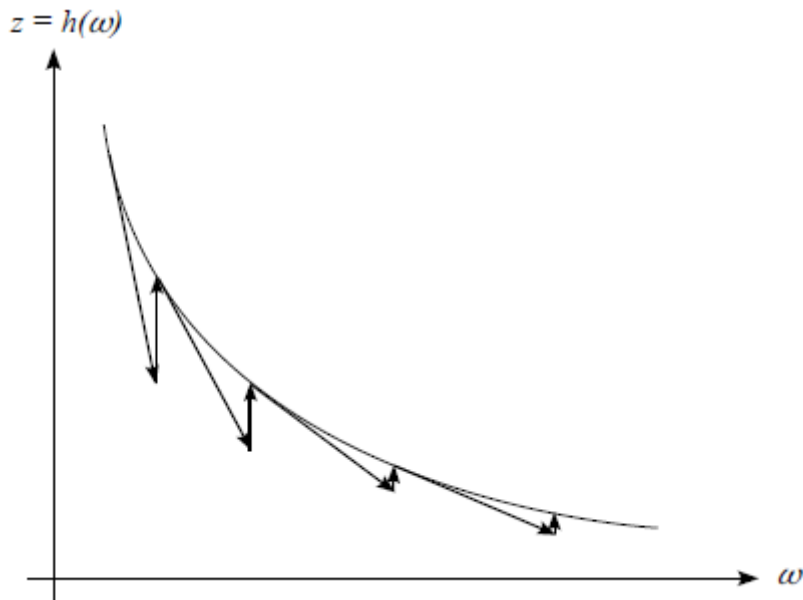
$$\vec{z}^{k+1} := \vec{z}^k + \frac{\partial}{\partial \vec{w}} \vec{h}(\vec{w}^k)(\vec{w}^{k+1} - \vec{w}^k) \quad (4.2)$$

Έντευθεν συνάγεται ότι το διάνυσμα \vec{z}^{k+1} αποτελεί την πρόβλεψη (prediction) του διανύσματος \vec{z}^{k+1} μετά από ένα βήμα μήκους (step length) $\delta^k := \|\vec{w}^{k+1} - \vec{w}^k\|$.

Όσον αφορά το βήμα διόρθωσης (correction step), το προβλεφθέν σημείο (predicted point) \vec{z}^{k+1} δύναται, μετέπειτα, να χρησιμοποιηθεί ως τιμή εκκίνησης – αρχική τιμή (starting value) για οποιοδήποτε αλγόριθμο του οποίου έγινε χρήση όταν επιδιώχθηκε η εύρεση του σημείου \vec{z}^k . Το γεγονός αυτό μετατρέπει τον αλγόριθμο πρόβλεψης-διόρθωσης (predictor-corrector algorithm) σε έναν πολύ καλό υποψήφιο όταν γίνεται η μετάβαση από ένα πρόβλημα βελτιστοποίησης ενός στόχου στο πλαίσιο της πολυκριτηριακής βελτιστοποίησης.

Στο σχήμα 4.2 απεικονίζεται διαγραμματικά το πώς λειτουργεί ο αλγόριθμος πρόβλεψης-διόρθωσης. Βλέπουμε, ουσιαστικά, ότι η μέθοδος πραγματοποιεί ένα ζγκ-ζαγκ. Δηλαδή, ο αλγόριθμος πρόβλεψης-διόρθωσης ξεκινάει από την εύρεση ενός σημείου του Pareto κάνοντας χρήση ενός οιοδήποτε αλγορίθμου βελτιστοποίησης για προβλήματα ενός στόχου. Στη συνέχεια, προχωράει στην εύρεση μίας προσέγγισης ενός επόμενου σημείου του Pareto με βάση την σχέση (4.2) (πρόβλεψη), το οποίο θα απέχει κάποια ‘μικρή’ απόσταση από το μέτωπο Pareto. Έπειτα, χρησιμοποιεί την εν λόγω προσέγγιση ως αρχική τιμή στον ίδιο αλγόριθμο βελτιστοποίησης που πραγματοποίησε τον υπολογισμό του πρώτου σημείου του μετώπου των μη-κυριαρχούμενων λύσεων (διόρθωση). Έτσι φθάνει στην εύρεση ενός νέου σημείου του μετώπου Pareto αλλά, σε αυτή την περίπτωση, με πολύ λιγότερες επαναλήψεις (άρα και πολύ μικρότερο υπολογιστικό κόστος) του επιλεκθέντος αλγορίθμου βελτιστοποίησης για προβλήματα ενός στόχου, καθώς η τιμή εκκίνησης βρίσκεται πολύ κοντά στο μέτωπο Pareto σε αντίθεση με την τυχαία αρχική τιμή που επιλέχθηκε για την εύρεση του πρώτου σημείου. Πλέον, η

μέθοδος έχει στην κατοχή της ένα δεύτερο σημείο του μετώπου Pareto το οποίο μπορεί να χρησιμοποιήσει για να υπολογίσει την προσέγγιση μίας τρίτης μη-κυριαρχούμενης λύσης. Η περιγραφείσα διαδικασία επαναλαμβάνεται για όσα σημεία του μετώπου Pareto επιθυμεί να αποκτήσει ο χρήστης.



Σχήμα 4.2: Διαγραμματική απεικόνιση του αλγορίθμου πρόβλεψης-διόρθωσης, [1].

Στη σχέση (4.2), η οποία μας παρέχει την προσέγγιση του $(k + 1)$ -οστού σημείου του μετώπου Pareto όταν γνωρίζουμε ήδη το k -οστό, εμφανίζεται ο όρος $\frac{\partial}{\partial \vec{w}} \vec{h}(\vec{w}^k)$. Ο υπολογισμός του εν λόγω όρου πραγματοποιείται μέσω του θεωρήματος πεπλεγμένης συνάρτησης που παρουσιάστηκε στην ενότητα 4.2.4. Έτσι, είναι:

$$\frac{\partial \vec{h}}{\partial \vec{w}}(\vec{w}^k) = - \left[\frac{\partial \vec{H}}{\partial \vec{z}}(\vec{z}^k, \vec{w}^k) \right]^{-1} \frac{\partial \vec{H}}{\partial \vec{w}}(\vec{z}^k, \vec{w}^k)$$

Το πρόβλημα βελτιστοποίησης πολλών στόχων που τελεί υπό διαχείριση, όπως διατυπώθηκε στην ενότητα 4.2.1, υπόκειται σε περιορισμούς ισότητας και ανισότητας. Όμως, όπως αναφέρθηκε στην ενότητα 4.2.3, το θεώρημα πεπλεγμένης συνάρτησης δεν 'επιτρέπει' την ύπαρξη ανισοτικών σχέσεων. Έτσι, για λόγους απλότητας, κατά την παρουσίαση της μεθόδου πρόβλεψης-διόρθωσης, γίνεται χρήση της προσέγγισης του συνόλου των ενεργών περιορισμών ανισότητας (active set) και υποτίθεται ότι οι περιορισμοί ανισότητας είναι πάντα ενεργοί. Αν, λοιπόν, ο δείκτης (a) δηλώνει μόνο αυτές τις συνιστώσες που βρίσκονται στο τρέχον σύνολο ενεργών ανισοτικών περιορισμών, ισχύουν:

$$\frac{\partial \vec{H}^\alpha}{\partial \vec{w}}(\vec{x}, \vec{\lambda}, \vec{\mu}, \vec{w}) = \begin{bmatrix} [\nabla_x f_j(\vec{x}) - \nabla_x f_m(\vec{x})]_{j=1, \dots, m-1} \\ 0 \\ 0 \end{bmatrix} \in \mathbb{R}^{(n+m_1+m_2^a) \times (m-1)}$$

$$\frac{\partial \vec{H}^\alpha}{\partial \vec{z}}(\vec{x}, \vec{\lambda}, \vec{\mu}, \vec{w}) = \begin{bmatrix} Hess_x L(\vec{x}, \vec{\lambda}, \vec{\mu}, \vec{w}) & - \left[\frac{\partial \vec{c}}{\partial \vec{x}}(\vec{x}) \right]^T & - \left[\frac{\partial \vec{g}^a}{\partial \vec{x}}(\vec{x}) \right]^T \\ \frac{\partial \vec{c}}{\partial \vec{x}}(\vec{x}) & 0 & 0 \\ \frac{\partial \vec{g}^a}{\partial \vec{x}}(\vec{x}) & 0 & 0 \end{bmatrix} \in \mathbb{R}^{(n+m_1+m_2^a) \times (n+m_1+m_2^a)}$$

Παραμένει, ωστόσο, ένα πρόβλημα. Ο επιλυτής που διατίθεται, δηλαδή ο αλγόριθμος επίλυσης προβλημάτων βελτιστοποίησης ενός στόχου, πρέπει να προσαρμοστεί ώστε τώρα να λύνει τον κυρτό συνδυασμό των συναρτήσεων στόχου ο οποίος βασίζεται στις παραμέτρους w_i . Ωστόσο, αν προβούμε σε μία συγκεκριμένη επαναπαραμετροποίηση (reparametrization), είναι δυνατό να μειωθούν οι αλλαγές που πρέπει να γίνουν στον υπάρχοντα αλγόριθμο επίλυσης προβλημάτων βελτιστοποίησης ενός στόχου. Ουσιαστικά, τα βάρη w_i μπορούν να θεωρηθούν ως οι συζυγείς μεταβλητές ενός συγκεκριμένου προβλήματος βελτιστοποίησης με περιορισμούς που θα παρουσιαστεί παρακάτω.

4.3.2 Επαναπαραμετροποίηση – διαφορετική παραμετροποίηση (reparametrization)

Στην παρούσα ενότητα πραγματοποιείται μία διαφορετική παραμετροποίηση από αυτή που παρουσιάστηκε παραπάνω με τα βάρη w_i .

Ας θεωρήσουμε το ακόλουθο πρόβλημα βελτιστοποίησης:

$$\min f_l(\vec{x})$$

που υπόκειται σε (subject to – s.t.)

$$\begin{aligned} f_k(\vec{x}) &= \hat{f}_k, & \forall k \in \{1, \dots, m\} \setminus \{l\} \\ c_i(\vec{x}) &= 0, & i = 1, \dots, m_1 \\ g_j(\vec{x}) &\geq 0, & j = 1, \dots, m_2 \end{aligned}$$

Η παραπάνω διατύπωση δηλώνει ότι αντί να γίνει η μετάβαση στην ελαχιστοποίηση του κυρτού συνδυασμού των επιμέρους συναρτήσεων στόχου, δύναται να διατηρηθεί η αρχική διατύπωση του προβλήματος (όπως παρουσιάστηκε στην ενότητα 4.2.3) και να παραμετροποιηθεί το σύνολο των μη-κυριαρχούμενων λύσεων κάνοντας χρήση όχι των

τιμών των βαρών w_i αλλά των τιμών των περιορισμών \hat{f}_k . Αν πραγματοποιηθεί σύγκριση των αναγκαίων συνθηκών για την ύπαρξη βέλτιστης λύσης των δύο διατυπώσεων, θα διαπιστωθεί ότι οι εν λόγω διατυπώσεις είναι ισοδύναμες. Ισχύει:

$$\begin{aligned} & \sum_{j=1}^{m-1} w_j \nabla f_j(\vec{x}) + \left(1 - \sum_{j=1}^{m-1} w_j\right) \nabla f_m(\vec{x}) - \left[\frac{\partial \vec{c}(\vec{x})}{\partial \vec{x}}\right]^T \lambda - \left[\frac{\partial \vec{g}(\vec{x})}{\partial \vec{x}}\right]^T \mu = 0 \Leftrightarrow \\ \Leftrightarrow & \nabla f_l(\vec{x}) - \sum_{\substack{j=1 \\ j \neq l}}^{m-1} \frac{-w_j}{w_l} \nabla f_j(\vec{x}) - \frac{-\left(1 - \sum_{j=1}^{m-1} w_j\right)}{w_l} \nabla f_m(\vec{x}) - \frac{1}{w_l} \left[\frac{\partial \vec{c}(\vec{x})}{\partial \vec{x}}\right]^T \lambda - \\ & - \frac{1}{w_l} \left[\frac{\partial \vec{g}(\vec{x})}{\partial \vec{x}}\right]^T \mu = 0 \end{aligned}$$

Όπως γίνεται φανερό μέσω των παραπάνω σχέσεων, η επαναπαραμετροποίηση αποτελεί περίπου μία διαίρεση με το w_l και τα αρχικά βάρη γίνονται, ελαφρώς μετασχηματισμένα, οι συζυγείς μεταβλητές.

Περαιτέρω συγκρίσεις αποκαλύπτουν:

- Μία παραμετροποίηση που χρησιμοποιεί τον κυρτό συνδυασμό εξαρτάται γραμμικά από το \vec{w} . Αυτή η γραμμικότητα, γενικά, εξαλείφεται μετά την επαναπαραμετροποίηση.
- Αν μελετήσουμε τον κυρτό συνδυασμό, καθίσταται φανερό το γεγονός ότι η παράμετρος \vec{w} θα εισέλθει στο εσσιανό μητρώο. Αυτό σημαίνει ότι στην περίπτωση χρήσης του κυρτού συνδυασμού, απαιτούνται περισσότερες αλλαγές στον αλγόριθμο επίλυσης προβλημάτων ενός στόχου που διατίθεται από ότι στην περίπτωση της επαναπαραμετροποίησης. Στην τελευταία περίπτωση, οι τιμές των \hat{f}_k δεν εισέρχονται στο εσσιανό μητρώο αλλά μόνο στο δεξί μέλος του συστήματος, γεγονός που οδηγεί σε λιγότερες αλλαγές. Επιπλέον, όπως θα φανεί παρακάτω, η νέα παραμετροποίηση απλοποιεί τους υπολογισμούς στο βήμα πρόβλεψης της μεθόδου.
- Όταν κάνουμε χρήση του κυρτού συνδυασμού, τα διαστήματα στα οποία κινείται το \vec{w} είναι σχετικά σαφή, ενώ αν προβούμε σε επαναπαραμετροποίηση, οι πιθανές αρχικές τιμές των \hat{f}_k είναι πολύ δυσκολότερο να προσδιοριστούν. Αυτό είναι λιγότερο περιοριστικό από πρακτικής άποψης, όπου ένας σχεδιαστής μπορεί να δώσει μία προσέγγιση, αλλά μία κακή επιλογή μπορεί να έχει ως αποτέλεσμα το πρόβλημα να μη δύναται να επιλυθεί.

Στη δεδομένη περίπτωση, αν εφαρμοστεί το θεώρημα πεπλεγμένης συνάρτησης που παρουσιάστηκε στην ενότητα 4.2.4, θα έχουμε:

$$\left\{ \frac{\partial \vec{H}}{\partial \vec{z}} [\vec{h}(\hat{f}_k), \hat{f}_k] \right\} \frac{\partial \vec{h}}{\partial \hat{f}_k}(\hat{f}_k) = - \frac{\partial \vec{H}}{\partial \hat{f}_k} [\vec{h}(\hat{f}_k), \hat{f}_k] \quad (4.3)$$

Ισχύει, ακόμα, το εξής:

$$\frac{\partial \vec{H}}{\partial \vec{f}}(\vec{x}, \vec{\lambda}, \vec{\mu}, \vec{f}) = \begin{bmatrix} 0 & 0 & \dots & 0 \\ -1 & 0 & \dots & 0 \\ 0 & -1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -1 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \end{bmatrix} \begin{matrix} \rightarrow n \\ \rightarrow k = 2 \\ \rightarrow k = 3 \\ \vdots \\ \rightarrow k = m \\ \rightarrow m_1 \\ \rightarrow m_2^a \end{matrix}$$

Όπως φαίνεται από την παραπάνω σχέση, το μητρώο $\partial \vec{H} / \partial \vec{f}$ περιέχει $n + (m - 1) + m_1 + m_2^a$ γραμμές. Επιπλέον, αποτελείται από $(m - 1)$ στήλες. Συνεπώς, $\partial \vec{H} / \partial \vec{f} \in \mathbb{R}^{(n+m-1+m_1+m_2^a) \times (m-1)}$.

Συμπερασματικά, η παράγωγος $\partial \vec{H} / \partial \hat{f}_k$ που απαιτείται από το θεώρημα πεπλεγμένης συνάρτησης για τον υπολογισμό της $\partial \vec{h} / \partial \hat{f}_k$, όπως υποδεικνύει η σχέση 4.3, ουσιαστικά είναι ένας πίνακας στήλη με όλα του τα στοιχεία μηδενικά εκτός από το $(n + k)$ -οστό στοιχείο που έχει την τιμή -1 . Έτσι, το σύστημα που επιβάλλεται να επιλυθεί για την απόκτηση της παραγώγου $\partial \vec{h} / \partial \hat{f}_k$ είναι το εξής:

$$\begin{bmatrix} Hess_x L(\vec{x}, \vec{\lambda}, \vec{\mu}, \vec{f}) & - \left[\frac{\partial f_{2, \dots, m}(\vec{x})}{\partial \vec{x}} \right]^T & - \left[\frac{\partial \vec{c}(\vec{x})}{\partial \vec{x}} \right]^T & - \left[\frac{\partial \vec{g}^a(\vec{x})}{\partial \vec{x}} \right]^T \\ \frac{\partial f_{2, \dots, m}(\vec{x})}{\partial \vec{x}} & 0 & 0 & 0 \\ \frac{\partial \vec{c}(\vec{x})}{\partial \vec{x}} & 0 & 0 & 0 \\ \frac{\partial \vec{g}^a(\vec{x})}{\partial \vec{x}} & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{\partial \vec{x}}{\partial \hat{f}_k} \\ \frac{\partial \lambda_{f_2, \dots, f_m}}{\partial \hat{f}_k} \\ \frac{\partial \vec{\lambda}_c}{\partial \hat{f}_k} \\ \frac{\partial \vec{\mu}}{\partial \hat{f}_k} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Η γνώση της παραγώγου $\partial \vec{h} / \partial \hat{f}_k$ που παρέχεται αν επιλυθεί το παραπάνω σύστημα δύναται να χρησιμοποιηθεί διττά:

- ✘ Οι προσαυξήσεις (increments) της παραμετροποίησης μπορούν να βασιστούν στην κλίση της καμπύλης του μετώπου Pareto, γεγονός που οδηγεί σε καλύτερη διακριτοποίηση των σημείων πάνω στο μέτωπο Pareto.
- ✘ Η τιμή εκκίνησης του βήματος διόρθωσης μπορεί να προεκβληθεί με βάση την εν λόγω παράγωγο, επιτυγχάνοντας, γενικά, καλύτερη σύγκλιση.

Στο σημείο αυτό, κρίνεται σκόπιμο να παρουσιαστεί η εν λόγω μέθοδος στην απλή περίπτωση που έχουμε, λόγου χάρη, δύο συναρτήσεις στόχου (δηλαδή $m = 2$) και τρεις μεταβλητές σχεδιασμού (δηλαδή $n = 3$), ώστε να επιδειχθεί σαφέστερα η λειτουργία της. Επίσης, υποθέτουμε ότι δεν υπάρχουν περιορισμοί. Έτσι, τίθεται προς επίλυση το ακόλουθο πρόγραμμα βελτιστοποίησης:

$$\min f_1(\vec{x})$$

που υπόκειται σε (subject to – s.t.)

$$f_2(\vec{x}) = \hat{f}_2$$

Από την παραπάνω διατύπωση είναι φανερό ότι επιλέχθηκε $l = 1$. Θα ήταν δυνατό, εναλλακτικά, να επιλεγεί η ελαχιστοποίηση της δεύτερης συνάρτησης στόχου με την πρώτη να αποτελεί τον περιορισμό ισότητας.

Για το δεδομένο πρόβλημα, η συνάρτηση Lagrange γράφεται ως εξής:

$$L(\vec{x}, \hat{\lambda}_2) = f_1(\vec{x}) - \hat{\lambda}_2 [f_2(\vec{x}) - \hat{f}_2]$$

Οι συνθήκες KKT για το εν λόγω πρόβλημα διατυπώνονται ως εξής:

$$\nabla_x L(\vec{x}, \hat{\lambda}_2) = 0 \Leftrightarrow \nabla_x f_1(\vec{x}) - \hat{\lambda}_2 \nabla_x f_2(\vec{x}) = 0$$

$$f_2(\vec{x}) - \hat{f}_2 = 0$$

Εναλλακτικά, το ανωτέρω σύστημα εξισώσεων γράφεται:

$$\vec{H}(\vec{x}, \hat{\lambda}_2, \hat{f}_2) = 0$$

όπου

$$H_1 = \frac{\partial f_1(\vec{x})}{\partial x_1} - \hat{\lambda}_2 \frac{\partial f_2(\vec{x})}{\partial x_1} = 0$$

$$H_2 = \frac{\partial f_1(\vec{x})}{\partial x_2} - \hat{\lambda}_2 \frac{\partial f_2(\vec{x})}{\partial x_2} = 0$$

$$H_3 = \frac{\partial f_1(\vec{x})}{\partial x_3} - \hat{\lambda}_2 \frac{\partial f_2(\vec{x})}{\partial x_3} = 0$$

$$H_4 = f_2(\vec{x}) - \hat{f}_2 = 0$$

Ορίζεται το διάνυσμα $\vec{z} = (\vec{x}, \hat{\lambda}_2)^T$, το οποίο είναι φανερό ότι έχει διαστάσεις $(n + m - 1) \times 1$, δηλαδή 4×1 .

Σύμφωνα με το θεώρημα πεπλεγμένης συνάρτησης είναι:

$$\vec{H}(\vec{z}, \hat{f}_2) = 0 \Leftrightarrow \vec{z} = \vec{h}(\hat{f}_2)$$

δηλαδή

$$z_1 = x_1 = h_1(\hat{f}_2)$$

$$z_2 = x_2 = h_2(\hat{f}_2)$$

$$z_3 = x_3 = h_3(\hat{f}_2)$$

$$z_4 = \hat{\lambda}_2 = h_4(\hat{f}_2)$$

Επιπροσθέτως, λόγω του θεωρήματος αυτού ισχύει:

$$\frac{\partial \vec{h}}{\partial \hat{f}_2}(\hat{f}_2) = - \left\{ \frac{\partial \vec{H}}{\partial \vec{z}} [\vec{h}(\hat{f}_2), \hat{f}_2] \right\}^{-1} \frac{\partial \vec{H}}{\partial \hat{f}_2} [\vec{h}(\hat{f}_2), \hat{f}_2]$$

Από την ανωτέρω σχέση, καθίσταται φανερό το γεγονός ότι για τον υπολογισμό της παραγώγου $\partial \vec{h} / \partial \hat{f}_2$ απαιτείται ο υπολογισμός των παραγώγων $\partial \vec{H} / \partial \vec{z}$ και $\partial \vec{H} / \partial \hat{f}_2$. Έτσι, είναι:

$$\frac{\partial \vec{H}}{\partial \hat{f}_2} = \begin{bmatrix} \partial H_1 / \partial \hat{f}_2 \\ \partial H_2 / \partial \hat{f}_2 \\ \partial H_3 / \partial \hat{f}_2 \\ \partial H_4 / \partial \hat{f}_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \end{bmatrix}$$

$$\frac{\partial \vec{H}}{\partial \vec{z}} = \begin{bmatrix} \partial H_1 / \partial x_1 & \partial H_1 / \partial x_2 & \partial H_1 / \partial x_3 & \partial H_1 / \partial \hat{\lambda}_2 \\ \partial H_2 / \partial x_1 & \partial H_2 / \partial x_2 & \partial H_2 / \partial x_3 & \partial H_2 / \partial \hat{\lambda}_2 \\ \partial H_3 / \partial x_1 & \partial H_3 / \partial x_2 & \partial H_3 / \partial x_3 & \partial H_3 / \partial \hat{\lambda}_2 \\ \partial H_4 / \partial x_1 & \partial H_4 / \partial x_2 & \partial H_4 / \partial x_3 & \partial H_4 / \partial \hat{\lambda}_2 \end{bmatrix} =$$

$$= \begin{bmatrix} \frac{\partial^2 f_1}{\partial x_1^2} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_1^2} & \frac{\partial^2 f_1}{\partial x_1 \partial x_2} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_1 \partial x_2} & \frac{\partial^2 f_1}{\partial x_1 \partial x_3} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_1 \partial x_3} & -\frac{\partial f_2}{\partial x_1} \\ \frac{\partial^2 f_1}{\partial x_1 \partial x_2} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_1 \partial x_2} & \frac{\partial^2 f_1}{\partial x_2^2} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_2^2} & \frac{\partial^2 f_1}{\partial x_2 \partial x_3} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_2 \partial x_3} & -\frac{\partial f_2}{\partial x_2} \\ \frac{\partial^2 f_1}{\partial x_1 \partial x_3} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_1 \partial x_3} & \frac{\partial^2 f_1}{\partial x_2 \partial x_3} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_2 \partial x_3} & \frac{\partial^2 f_1}{\partial x_3^2} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_3^2} & -\frac{\partial f_2}{\partial x_3} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_3} & 0 \end{bmatrix}$$

Στη συνέχεια, καθίσταται δυνατό, μέσω της ήδη υπολογισθείσας παραγώγου $\partial \vec{h} / \partial \hat{f}_2$, να πραγματοποιηθεί η εύρεση των νέων τιμών των μεταβλητών σχεδιασμού και του $\hat{\lambda}_2$, δηλαδή της προσέγγισης του επόμενου σημείου πάνω στο μέτωπο Pareto, ως εξής:

$$\vec{z}^{new} = \vec{z}^{old} + \frac{\partial \vec{h}}{\partial \hat{f}_2} (\hat{f}_2^{new} - \hat{f}_2^{old})$$

4.4 Μαθηματικό παράδειγμα

Το ακόλουθο μαθηματικό παράδειγμα επιλέχθηκε από την βιβλιογραφία, [1], ως παράδειγμα για την ‘επίδειξη’ της μεθόδου. Ας υποθεθεί ότι ο συντελεστής $a > 0$ αποτελεί μία κανονικοποίηση (regularization) και ότι το \vec{n} είναι το εξωτερικό κάθετο διάνυσμα. Οι ακόλουθες συναρτήσεις ορίζονται στο πεδίο $\Omega := [0,1]^2$:

$$f_1(\varphi, q) := \int_{y=1} \left[\frac{\partial \varphi}{\partial \vec{n}}(x, 1) - g(x) \right]^2 dx + a \int_{y=1} \left\{ q^2(x) + \left[\frac{\partial q}{\partial x}(x) \right]^2 \right\} dx$$

$$f_2(\varphi, q) := \int_{\Omega} [\varphi^2(x, y)] dx dy$$

Στη συνέχεια, ας θεωρηθεί το ακόλουθο πολυκριτηριακό πρόβλημα βελτιστοποίησης με περιορισμούς:

$$\min_{(\varphi, q)} \begin{bmatrix} f_1(\varphi, q) \\ f_2(\varphi, q) \end{bmatrix}$$

που υπόκειται σε (subject to – s.t.)

$$\begin{aligned} \Delta \varphi(x, y) &= 0, & \text{στο } \Omega \\ \varphi(x, 1) &= q(x), & \text{στο } y = 1 \\ \varphi(x, 0) &= \varphi_0(x), & \text{στο } y = 0 \\ \varphi(0, y) &= \varphi(1, y) & (\text{περιοδικότητα}) \end{aligned}$$

Αρχικά, το πρόβλημα επιλύεται με τη χρήση της μεθόδου των βαρών, όπως αυτή περιγράφηκε στην ενότητα 4.2.2. Συνεπώς, επιχειρείται η ελαχιστοποίηση του κυρτού συνδυασμού:

$$\min_{(\varphi, q)} w f_1 + (1 - w) f_2$$

Στην περίπτωση αυτή, οι αναγκαίες συνθήκες για την ύπαρξη βέλτιστης λύσης, όπως παρουσιάζονται στη βιβλιογραφία, [1], είναι οι εξής:

‘Συζυγείς εξισώσεις (costate/adjoint)’ $dL/d\varphi$:

$$\begin{aligned} 2w \frac{\partial \varphi}{\partial \vec{n}}(x, 1) - 2wg(x) + \lambda_1(x, 1) &= 0, \quad \text{στο } y = 1 \\ 2(1 - w)\varphi + \Delta \lambda_1 &= 0, \quad \text{στο } \Omega \\ \lambda_1 &= 0, \quad \text{στο } y = 0 \\ \lambda_2 &= \partial \lambda_1 / \partial \vec{n}, \quad \text{στο } y = 1 \\ \lambda_3 &= \partial \lambda_1 / \partial \vec{n}, \quad \text{στο } y = 0 \end{aligned}$$

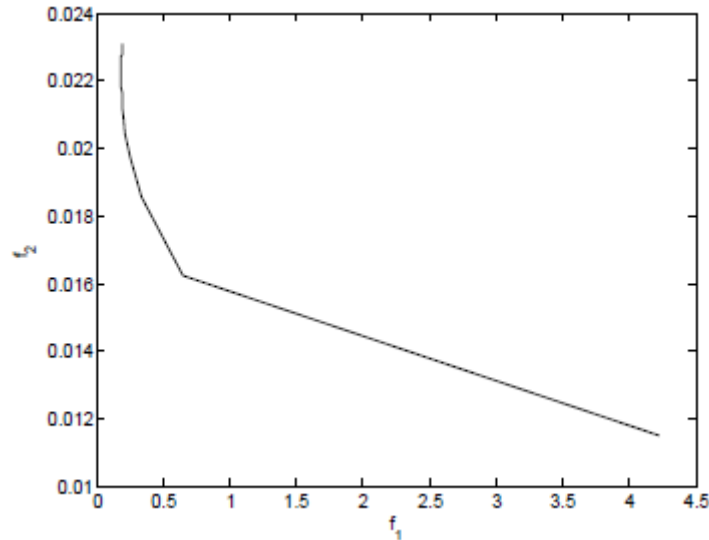
‘Εξισώσεις σχεδιασμού (design)’ dL/dq :

$$2wa(q - \ddot{q}) = \lambda_2 = \partial \lambda_1 / \partial \vec{n}, \quad \text{στο } y = 1$$

‘Εξισώσεις κατάστασης (state)’ $dL/d\lambda$:

$$\begin{aligned} \Delta \varphi &= 0, \quad \text{στο } \Omega \\ \varphi(x, 1) &= q, \quad \text{στο } y = 1 \\ \varphi(x, 0) &= \varphi_0, \quad \text{στο } y = 0 \end{aligned}$$

Οι ανωτέρω εξισώσεις διακριτοποιήθηκαν με τη χρήση πεπερασμένων διαφορών και επιλύθηκαν για 250 ισαπέχουσες τιμές του w σε ένα πλέγμα 20×21 , χωρίς να γίνει καμία χρήση της εφαπτομενικής προσέγγισης που αναλύθηκε στην ενότητα 4.3. Το αποτέλεσμα που προέκυψε για το μέτωπο Pareto φαίνεται στο σχήμα 4.3. Αν παρατηρηθεί το εν λόγω σχήμα, καθίσταται φανερό το ότι η διακριτοποίηση της καμπύλης είναι ανεπαρκής, δηλαδή, αν και οι προσαυξήσεις στο w ήταν σταθερές, τα σημεία που προέκυψαν συναντώνται συγκεντρωμένα στο πάνω μέρος της καμπύλης.



Σχήμα 4.3: Μέτωπο Pareto για πλέγμα 20×21 και 250 ισαπέχοντα βήματα στο w , [1].

Ακολουθως, πραγματοποιείται η επαναπαραμετροποίηση της καμπύλης. Στην περίπτωση αυτή, πρέπει να αποφασιστεί αν ο περιορισμός με βάση τον οποίο θα γίνει η παραμετροποίηση θα αποτελείται από τη συνάρτηση f_1 ή την f_2 . Ανάλογα με τη ληφθείσα απόφαση, είτε ο παράγοντας $v_2 := w$ είτε ο $v_1 := 1 - w$ θα γίνει η συζυγής μεταβλητή για το νέο περιορισμό, π.χ. μία παραμετροποίηση με βάση την f_1 θα προσέθετε το v_1 ως έναν επιπλέον βαθμό ελευθερίας στις αναγκαίες συνθήκες και ταυτοχρόνως θα έθετε $v_2 \equiv 1$. Έχοντας υπόψη τα ανωτέρω, οι νέες αναγκαίες συνθήκες για την ύπαρξη βελτίστου είναι οι εξής, [1]:

‘Συζυγείς εξισώσεις (costate/adjoint)’ $dL/d\varphi$:

$$\begin{aligned}
 2v_2 \frac{\partial \varphi}{\partial \vec{n}}(x, 1) - 2v_2 g(x) + \lambda_1(x, 1) &= 0, \quad \text{στο } y = 1 \\
 2v_1 \varphi + \Delta \lambda_1 &= 0, \quad \text{στο } \Omega \\
 \lambda_1 &= 0, \quad \text{στο } y = 0 \\
 \lambda_2 &= \partial \lambda_1 / \partial \vec{n}, \quad \text{στο } y = 1 \\
 \lambda_3 &= \partial \lambda_1 / \partial \vec{n}, \quad \text{στο } y = 0
 \end{aligned}$$

‘Εξισώσεις σχεδιασμού (design)’ dL/dq :

$$2v_2 a(q - \ddot{q}) = \lambda_2 = \partial \lambda_1 / \partial \vec{n}, \quad \text{στο } y = 1$$

‘Εξισώσεις κατάστασης (state)’ $dL/d\lambda$:

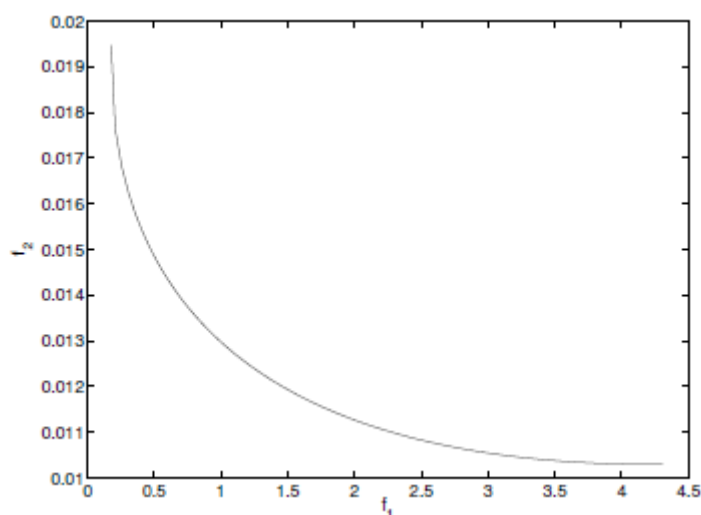
$$\begin{aligned}
 \Delta \varphi &= 0, \quad \text{στο } \Omega \\
 \varphi(x, 1) &= q, \quad \text{στο } y = 1 \\
 \varphi(x, 0) &= \varphi_0, \quad \text{στο } y = 0 \\
 \left. \begin{aligned} f_1(\varphi, q) &= \hat{f}_1 \\ f_2(\varphi, q) &= \hat{f}_2 \end{aligned} \right\} \text{ισχύει είτε το ένα είτε το άλλο}
 \end{aligned}$$

Το σύστημα αυτό διακριτοποιήθηκε εκ νέου με τη χρήση πεπερασμένων διαφορών και έπειτα επιλύθηκε μέσω διαδοχικού τετραγωνικού προγραμματισμού (sequential quadratic programming – SQP). Επίσης, στη δεύτερη αυτή περίπτωση δεν χρησιμοποιήθηκαν σταθερές προσαυξήσεις για την παραμετροποίηση αλλά έγινε χρήση της σχέσης:

$$\hat{f}_k^{i+1} = \hat{f}_k^i + \delta \frac{\partial f_l}{\partial \hat{f}_k}(\hat{f}_k^i)$$

όπου το δ δηλώνει το μήκος βήματος, το i είναι ο μετρητής των επαναλήψεων, το k δηλώνει τη συνάρτηση που χρησιμοποιήθηκε για την παραμετροποίηση/ως περιορισμός και το l δηλώνει τη συνάρτηση που χρησιμοποιήθηκε ως στόχος βελτιστοποίησης.

Η καμπύλη των μη-κυριαρχούμενων λύσεων που προκύπτει σαν αποτέλεσμα στη δεύτερη περίπτωση όπου έγινε χρήση της μεθόδου πρόβλεψης-διόρθωσης φαίνεται στο σχήμα 4.4. Η τιμή που επιλέχθηκε για το δ έδωσε ως αποτέλεσμα 86 σημεία στην καμπύλη.



Σχήμα 4.4: Μέτωπο Pareto για την περίπτωση χρήσης της μεθόδου πρόβλεψης-διόρθωσης, [1].

Αν γίνει σύγκριση των δύο καμπυλών που προέκυψαν, γίνονται προφανείς οι τεράστιες βελτιώσεις από πλευράς διακριτοποίησης της καμπύλης. Επιπλέον, στη δεύτερη περίπτωση, με τη χρήση της μεθόδου πρόβλεψης-διόρθωσης έχει επιτευχτεί μικρότερο υπολογιστικό κόστος.

5 Επίλυση μαθηματικών προβλημάτων βελτιστοποίησης δύο στόχων με βάση τη θεωρία της συνέχισης του μετώπου Pareto

5.1 Εισαγωγή

Στο κεφάλαιο 4 παρουσιάστηκαν οι μέθοδοι πρόβλεψης-διόρθωσης, οι οποίες επιτρέπουν την επίλυση προβλημάτων βελτιστοποίησης πολλών στόχων με περιορισμούς με βάση τη θεωρία της συνέχισης της καμπύλης των μη-κυριαρχούμενων λύσεων. Συγκεκριμένα, στην ενότητα 4.3.1 περιγράφηκε το πώς λειτουργεί ο αλγόριθμος πρόβλεψης-διόρθωσης. Όπως εκτέθηκε στην εν λόγω ενότητα, ο αλγόριθμος πρόβλεψης-διόρθωσης ξεκινάει από την εύρεση ενός σημείου του μετώπου Pareto κάνοντας χρήση ενός οιαδήποτε αλγορίθμου βελτιστοποίησης για προβλήματα ενός στόχου. Στη συνέχεια, προχωράει στην εύρεση μίας προσέγγισης ενός επόμενου σημείου του Pareto με βάση τη θεωρία της συνέχισης της καμπύλης των μη-κυριαρχούμενων λύσεων (πρόβλεψη), το οποίο θα απέχει κάποια 'μικρή' απόσταση από το μέτωπο Pareto. Έπειτα, χρησιμοποιεί την εν λόγω προσέγγιση ως αρχική τιμή στον ίδιο αλγόριθμο βελτιστοποίησης που πραγματοποίησε τον υπολογισμό του πρώτου σημείου του μετώπου μη-κυριαρχούμενων λύσεων (διόρθωση). Έτσι φθάνει στην εύρεση ενός νέου σημείου του μετώπου Pareto αλλά, σε αυτήν την περίπτωση, με πολύ λιγότερες επαναλήψεις (άρα και πολύ μικρότερο υπολογιστικό κόστος) του επιλεχθέντος αλγορίθμου βελτιστοποίησης για προβλήματα ενός στόχου, καθώς η τιμή εκκίνησης βρίσκεται πολύ κοντά στο μέτωπο Pareto σε αντίθεση με την τυχαία αρχική τιμή που επιλέχθηκε για την εύρεση του πρώτου σημείου. Πλέον, η μέθοδος έχει στην κατοχή της ένα δεύτερο σημείο του μετώπου μη-κυριαρχούμενων λύσεων το οποίο μπορεί να χρησιμοποιήσει για να υπολογίσει την προσέγγιση ενός τρίτου σημείου του Pareto. Η περιγραφείσα διαδικασία επαναλαμβάνεται για όσα σημεία του μετώπου Pareto επιθυμεί να αποκτήσει ο χρήστης.

Στο παρόν κεφάλαιο, θα πραγματοποιηθεί η παρουσίαση της εφαρμογής της προαναφερθείσας μεθόδου που αναπτύχθηκε στο κεφάλαιο 4 σε μαθηματικά προβλήματα βελτιστοποίησης δύο στόχων, ώστε αρχικά να επιδειχθεί με σαφήνεια ο τρόπος λειτουργίας της και στη συνέχεια να εξεταστούν τα πλεονεκτήματα και τα μειονεκτήματά της. Στο σημείο αυτό, πρέπει να τονιστεί ότι, στο πλαίσιο αυτού του κεφαλαίου αλλά και της παρούσας εργασίας στην ολότητά της, δίνεται έμφαση στη χρήση της μεθόδου που βασίζεται στη συνέχιση του μετώπου Pareto με μία βασική διαφοροποίηση. Συγκεκριμένα, λήφθηκε η απόφαση να διερευνηθεί η ποιότητα των αποτελεσμάτων που θα αποκτηθούν στην περίπτωση της μη-χρήσης του τμήματος της

διόρθωσης. Με άλλα λόγια, ο αλγόριθμος θα προβαίνει, κατά πρώτο λόγο, στην εύρεση ενός σημείου του μετώπου Pareto με τη χρήση ενός αλγορίθμου βελτιστοποίησης για προβλήματα ενός στόχου και, κατά δεύτερο λόγο, στην εύρεση των επόμενων σημείων του Pareto με βάση τη θεωρία για τη συνέχιση της καμπύλης των μη-κυριαρχούμενων λύσεων, χωρίς τα εν λόγω σημεία να 'διορθώνονται' δηλαδή να χρησιμοποιούνται ως αρχική τιμή στον αλγόριθμο βελτιστοποίησης που πραγματοποίησε τον υπολογισμό του πρώτου σημείου του μετώπου μη-κυριαρχούμενων λύσεων. Καθίσταται φανερό ότι κίνητρο για μία τέτοια απόφαση αποτελεί η περαιτέρω μείωση του υπολογιστικού κόστους, καθώς αποφεύγονται οι επαναλήψεις που πραγματοποιεί ο αλγόριθμος βελτιστοποίησης ενός στόχου, όταν δεχθεί ως σημείο εκκίνησής του την προσέγγιση της μη-κυριαρχούμενης λύσης που προκύπτει με βάση τη θεωρία της συνέχισης του μετώπου Pareto, για να συγκλίνει στην επόμενη μη-κυριαρχούμενη λύση.

Αναφέρεται ότι παρουσιάζονται τρία μαθηματικά προβλήματα βελτιστοποίησης δύο στόχων με μοναδικούς περιορισμούς αυτούς που επιβάλλουν τα όρια των μεταβλητών σχεδιασμού. Για τα δύο πρώτα εξ αυτών το μέτωπο Pareto είναι συνεχές και αναζητείται με και χωρίς το στάδιο της διόρθωσης, ενώ χρησιμοποιούνται ως μέθοδοι προβλημάτων βελτιστοποίησης ενός στόχου η μέθοδος της απότομης καθόδου και η μέθοδος Newton. Όσον αφορά το τρίτο μαθηματικό πρόβλημα, το μέτωπο Pareto είναι ασυνεχές και αποκτάται χωρίς τη συμπερίληψη του σταδίου της διόρθωσης, ενώ ως μέθοδος βελτιστοποίησης προβλημάτων ενός στόχου χρησιμοποιείται η μέθοδος της απότομης καθόδου. Ο υπολογισμός των παραγώγων και για τα τρία προβλήματα γίνεται αναλυτικά.

5.2 Το λογισμικό βελτιστοποίησης EASY

Μετά την επίλυση των προβλημάτων που πραγματεύεται η παρούσα εργασία με βάση τη θεωρία της συνέχισης του μετώπου Pareto, κρίνεται απαραίτητη η επίλυσή τους και με μία δεύτερη μέθοδο, ώστε να κατασταθεί δυνατή η επαλήθευση και η σύγκριση των αποτελεσμάτων. Αυτός είναι ο ρόλος που έρχεται να υιοθετήσει το λογισμικό βελτιστοποίησης EASY, [19]. Ανεπτυγμένο από την Μονάδα Παράλληλης Υπολογιστικής Ρευστοδυναμικής & Βελτιστοποίησης του ΕΜΠ (ΜΠΥΡΒ/ΕΜΠ), το EASY (Evolutionary Algorithm System) είναι ένα πλήρες λογισμικό βελτιστοποίησης γραμμένο σε C++ με περιβάλλον χρήσης προγραμματισμένο σε Java. Προσφέρει μια πληθώρα εργαλείων και επιλογών που δίνουν τη δυνατότητα επίλυσης προβλημάτων βελτιστοποίησης ενός ή περισσότερων στόχων τα οποία δύναται να περιλαμβάνουν ή όχι περιορισμούς. Αν και το ακρωνύμιο του παραπέμπει στο γεγονός ότι οι εξελικτικοί αλγόριθμοι χρησιμοποιούνται ως η μοναδική μέθοδος βελτιστοποίησης, το λογισμικό EASY υποστηρίζει και στοχαστικές αλλά και αιτιοκρατικές

μεθόδους βελτιστοποίησης, καθώς και τον υβριδισμό τους. Πρέπει, βέβαια, να σημειωθεί ότι για τις ανάγκες της δεδομένης εργασίας χρησιμοποιούνται οι εξελικτικοί αλγόριθμοι. Πρόκειται για ένα λογισμικό στα αδιαμφισβήτητα πλεονεκτήματα του οποίου συγκαταλέγεται η εύκολη χρήση του αλλά και ο απλός τρόπος με τον οποίο πραγματοποιείται η επικοινωνία του με τα υπόλοιπα προγράμματα που εκτελούνται κατά τη διαδικασία της βελτιστοποίησης και αφορούν στην αξιολόγηση των υποψήφιων λύσεων. Επίσης, περιλαμβάνει ποικίλα εργαλεία των οποίων η χρήση μπορεί να μειώσει σημαντικά το υπολογιστικό κόστος μιας διαδικασίας βελτιστοποίησης, ενώ προσφέρει και τη δυνατότητα παράλληλης εκτέλεσης σε δίκτυα υπολογιστών.

5.3 Πρώτο μαθηματικό πρόβλημα

5.3.1 Διατύπωση προβλήματος

Στην παρούσα ενότητα, γίνεται η παρουσίαση του πρώτου μαθηματικού προβλήματος που επιλέχθηκε να επιλυθεί με βάση τη θεωρία της συνέχισης του μειώπου Pareto. Πρόκειται για ένα πρόβλημα ελαχιστοποίησης δύο στόχων (f_1 και f_2) με δύο μεταβλητές σχεδιασμού (x_1 και x_2). Το πρόβλημα επιβάλλει τόσο άνω όσο και κάτω όρια στις δύο μεταβλητές σχεδιασμού. Αν εξαιρεθούν οι περιορισμοί που υποχρεώνουν τις δύο μεταβλητές σχεδιασμού να κινούνται ανάμεσα στο κάτω και το άνω όριό τους, το πρόβλημα δεν υπόκειται σε κανέναν άλλο περιορισμό, ισότητας ή ανισότητας. Οι δύο συναρτήσεις στόχου που συναντώνται στο εν λόγω πρόβλημα είναι οι εξής:

$$f_1(x_1, x_2) = x_1^2$$

$$f_2(x_1, x_2) = \frac{x_2^2 - 5x_2 + \sin x_2 + 15}{x_1}$$

Έτσι, το πρόβλημα βελτιστοποίησης που αποτελεί αντικείμενο της δεδομένης ενότητας διατυπώνεται ως εξής:

$$\min \vec{f}(\vec{x}) = \min \begin{bmatrix} f_1(\vec{x}) \\ f_2(\vec{x}) \end{bmatrix}$$

που υπόκειται σε (subject to – s.t.)

$$0,5 \leq x_1 \leq 15$$

$$0,5 \leq x_2 \leq 10$$

5.3.2 Εφαρμογή της θεωρίας της συνέχισης του μετώπου Pareto στο πρόβλημα

Σύμφωνα με τα όσα αναπτύχθηκαν στην ενότητα 4.3.2, το παραπάνω πρόβλημα βελτιστοποίησης επαναδιατυπώνεται ως εξής:

$$\min f_1(\vec{x})$$

που υπόκειται σε (subject to – s.t.)

$$f_2(\vec{x}) = \hat{f}_2 \Leftrightarrow f_2(\vec{x}) - \hat{f}_2 = 0$$

Από την παραπάνω διατύπωση είναι φανερό ότι επιλέχθηκε $l = 1$. Θα ήταν δυνατό, εναλλακτικά, να επιλεγεί η ελαχιστοποίηση της δεύτερης συνάρτησης στόχου με την πρώτη να αποτελεί τον περιορισμό ισότητας.

Στο σημείο αυτό είναι απαραίτητο να γίνει ένα σχόλιο σε σχέση με τους περιορισμούς που επιβάλλουν άνω και κάτω όρια στις δύο μεταβλητές σχεδιασμού. Στην ενότητα 4.2 αναφέρθηκε ότι το θεώρημα πεπλεγμένης συνάρτησης που εφαρμόζεται στις συνθήκες KKT δεν 'επιτρέπει' την ύπαρξη ανισοτικών περιορισμών. Επιπλέον, δηλώθηκε ότι το σύστημα που αποτελεί τις συνθήκες KKT δύναται να επαναδιατυπωθεί χωρίς την παρουσία ανισοτήτων, [13]. Για λόγους απλότητας, όμως, στην παρούσα εργασία λήφθηκε η απόφαση τέτοιου είδους περιορισμοί που αφορούν στα όρια των μεταβλητών σχεδιασμού να λαμβάνουν διαφορετικής 'μεταχείρισης'. Αποφασίστηκε, δηλαδή, να μην υπεισέρχονται στη συνάρτηση Lagrange άρα ούτε και στις συνθήκες KKT αλλά να εκτελείται από τον αλγόριθμο ένας απλός έλεγχος για το αν ικανοποιούνται κάθε φορά που μεταβάλλονται οι μεταβλητές σχεδιασμού. Στην περίπτωση που μία μεταβλητή σχεδιασμού έχει λάβει μία τιμή μικρότερη από το κάτω όριό της, ο αλγόριθμος της δίνει την τιμή του κάτω ορίου, ενώ εάν έχει λάβει μεγαλύτερη τιμή από το άνω όριό της, ο αλγόριθμος της παρέχει την τιμή του άνω ορίου. Αυτό θα διαφανεί καλύτερα παρακάτω όταν θα πραγματοποιηθεί η περιγραφή του αλγορίθμου.

Λαμβάνοντας υπόψη το τελευταίο σχόλιο, για το δεδομένο πρόβλημα, η συνάρτηση Lagrange γράφεται ως εξής:

$$L(\vec{x}, \hat{\lambda}_2) = f_1(\vec{x}) - \hat{\lambda}_2 [f_2(\vec{x}) - \hat{f}_2]$$

Οι συνθήκες KKT για το εν λόγω πρόβλημα διατυπώνονται ως εξής:

$$\nabla_x L(\vec{x}, \hat{\lambda}_2) = 0 \Leftrightarrow \nabla_x f_1(\vec{x}) - \hat{\lambda}_2 \nabla_x f_2(\vec{x}) = 0$$

$$f_2(\vec{x}) - \hat{f}_2 = 0$$

Εναλλακτικά, το ανωτέρω σύστημα εξισώσεων γράφεται:

$$\vec{H}(\vec{x}, \hat{\lambda}_2, \hat{f}_2) = 0$$

όπου

$$H_1 = \frac{\partial f_1(\vec{x})}{\partial x_1} - \hat{\lambda}_2 \frac{\partial f_2(\vec{x})}{\partial x_1} = 0$$

$$H_2 = \frac{\partial f_1(\vec{x})}{\partial x_2} - \hat{\lambda}_2 \frac{\partial f_2(\vec{x})}{\partial x_2} = 0$$

$$H_3 = f_2(\vec{x}) - \hat{f}_2 = 0$$

Ορίζεται το διάνυσμα $\vec{z} = (\vec{x}, \hat{\lambda}_2)^T = \begin{bmatrix} x_1 \\ x_2 \\ \hat{\lambda}_2 \end{bmatrix}$, το οποίο είναι φανερό ότι έχει

διαστάσεις $(n + m - 1) \times 1$, δηλαδή 3×1 .

Σύμφωνα με το θεώρημα πεπλεγμένης συνάρτησης είναι:

$$\vec{H}(\vec{z}, \hat{f}_2) = 0 \Leftrightarrow \vec{z} = \vec{h}(\hat{f}_2)$$

δηλαδή

$$z_1 = x_1 = h_1(\hat{f}_2)$$

$$z_2 = x_2 = h_2(\hat{f}_2)$$

$$z_3 = \hat{\lambda}_2 = h_3(\hat{f}_2)$$

Επιπροσθέτως, λόγω του θεωρήματος αυτού ισχύει:

$$\frac{\partial \vec{h}}{\partial \hat{f}_2}(\hat{f}_2) = - \left\{ \frac{\partial \vec{H}}{\partial \vec{z}} [\vec{h}(\hat{f}_2), \hat{f}_2] \right\}^{-1} \frac{\partial \vec{H}}{\partial \hat{f}_2} [\vec{h}(\hat{f}_2), \hat{f}_2]$$

Από την ανωτέρω σχέση, καθίσταται φανερό το γεγονός ότι για τον υπολογισμό της παραγώγου $\partial \vec{h} / \partial \hat{f}_2$ απαιτείται ο υπολογισμός των παραγώγων $\partial \vec{H} / \partial \vec{z}$ και $\partial \vec{H} / \partial \hat{f}_2$. Έτσι, ισχύουν:

$$\frac{\partial \vec{H}}{\partial \hat{f}_2} = \begin{bmatrix} \partial H_1 / \partial \hat{f}_2 \\ \partial H_2 / \partial \hat{f}_2 \\ \partial H_3 / \partial \hat{f}_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}$$

$$\frac{\partial \vec{H}}{\partial \vec{z}} = \begin{bmatrix} \partial H_1 / \partial x_1 & \partial H_1 / \partial x_2 & \partial H_1 / \partial \hat{\lambda}_2 \\ \partial H_2 / \partial x_1 & \partial H_2 / \partial x_2 & \partial H_2 / \partial \hat{\lambda}_2 \\ \partial H_3 / \partial x_1 & \partial H_3 / \partial x_2 & \partial H_3 / \partial \hat{\lambda}_2 \end{bmatrix} \Leftrightarrow$$

$$\Leftrightarrow \frac{\partial \vec{H}}{\partial \vec{z}} = \begin{bmatrix} \frac{\partial^2 f_1}{\partial x_1^2} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_1^2} & \frac{\partial^2 f_1}{\partial x_1 \partial x_2} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_1 \partial x_2} & -\frac{\partial f_2}{\partial x_1} \\ \frac{\partial^2 f_1}{\partial x_1 \partial x_2} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_1 \partial x_2} & \frac{\partial^2 f_1}{\partial x_2^2} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_2^2} & -\frac{\partial f_2}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & 0 \end{bmatrix}$$

Στη συνέχεια, καθίσταται δυνατό, αφού εκτελεστεί ο υπολογισμός της παραγώγου $\partial \vec{h} / \partial \hat{f}_2$, να πραγματοποιηθεί η εύρεση των νέων τιμών των μεταβλητών σχεδιασμού και του $\hat{\lambda}_2$, δηλαδή του επόμενου σημείου πάνω στο μέτωπο Pareto (στην πραγματικότητα αποτελεί προσέγγισή του, καθώς απέχει μία 'μικρή' απόσταση από αυτό, αλλά, όπως αναφέρθηκε παραπάνω, στην περίπτωση αφαίρεσης του σταδίου της διόρθωσης θα λαμβάνεται ως η επόμενη μη-κυριαρχούμενη λύση χωρίς να πραγματοποιείται διόρθωσή της), ως εξής:

$$\vec{z}^{new} = \vec{z}^{old} + \frac{\partial \vec{h}}{\partial \hat{f}_2} (\hat{f}_2^{new} - \hat{f}_2^{old})$$

δηλαδή

$$z_1^{new} = x_1^{new} = x_1^{old} + \frac{\partial h_1}{\partial \hat{f}_2} (\hat{f}_2^{new} - \hat{f}_2^{old})$$

$$z_2^{new} = x_2^{new} = x_2^{old} + \frac{\partial h_2}{\partial \hat{f}_2} (\hat{f}_2^{new} - \hat{f}_2^{old})$$

$$z_3^{new} = \hat{\lambda}_2^{new} = \hat{\lambda}_2^{old} + \frac{\partial h_3}{\partial \hat{f}_2} (\hat{f}_2^{new} - \hat{f}_2^{old})$$

5.3.3 Περιγραφή του αλγορίθμου που αναπτύχθηκε για την επίλυση του προβλήματος

5.3.3.1 Περίπτωση αφαίρεσης σταδίου διόρθωσης

Στην ενότητα αυτή, επιδεικνύεται ο αλγόριθμος που αναπτύχθηκε για την επίλυση του ανωτέρω παρουσιασθέντος μαθηματικού προβλήματος μέσω της μεθόδου που βασίζεται στη θεωρία του μετώπου Pareto χωρίς τη συμπερίληψη του σταδίου της διόρθωσης.

Το πρώτο βήμα που απαιτείται να εκτελεστεί από έναν τέτοιο αλγόριθμο είναι μία αρχική εκτίμηση που αφορά στην τιμή του \hat{f}_2 (\hat{f}_2^{up}). Είναι, δηλαδή, ζωτικής σημασίας να προσδιοριστεί μία τιμή της δεύτερης συνάρτησης στόχου η οποία θα αποτελεί το άνω όριο του μετώπου Pareto στον άξονα των f_2 . Αυτή θα είναι η τιμή του f_2 για το πρώτο σημείο του Pareto που θα αποκτηθεί ώστε να είναι δυνατή αργότερα η κίνηση προς τα κάτω. Όπως αναφέρθηκε στο κεφάλαιο 4, για να είναι εφικτή η παροχή της αρχικής αυτής εκτίμησης για το \hat{f}_2 , θα πρέπει να υπάρχει καλή γνώση του προβλήματος, δηλαδή ο κληθείς να επιλύσει το πρόβλημα οφείλει να γνωρίζει τα μεγέθη των τιμών των συναρτήσεων στόχου που αναμένονται στο μέτωπο Pareto. Σε αντίθετη περίπτωση, ο χρήστης οφείλει να κάνει δοκιμές οι οποίες, βέβαια, ανεβάζουν το υπολογιστικό κόστος. Αυτό αποτελεί ένα μειονέκτημα της μεθόδου με την οποία ασχολείται η παρούσα εργασία. Ωστόσο πρέπει να αντιπαρατεθεί το γεγονός ότι όταν πρόκειται για ένα πραγματικό πρόβλημα, όπως λ.χ. είναι η ελαχιστοποίηση του c_D συνοδευόμενη από τη μεγιστοποίηση του c_L σε μία αεροτομή, ο μηχανικός γνωρίζει κατά προσέγγιση τα όρια ανάμεσα στα οποία κινούνται οι τιμές των συναρτήσεων στόχου που αναμένονται στο μέτωπο Pareto.

Διαθέτοντας, πλέον, την αρχική εκτίμηση για το \hat{f}_2 , ο αλγόριθμος επιδίδεται στην εύρεση του πρώτου σημείου του Pareto (go to Pareto). Για να γίνει αυτό επιβάλλεται να επιλυθεί ένα μονοκριτηριακό πρόβλημα βελτιστοποίησης στο οποίο περιλαμβάνεται ένας περιορισμός ισότητας (οι περιορισμοί που αφορούν στα όρια ανάμεσα στα οποία κινούνται οι μεταβλητές σχεδιασμού αντιμετωπίζονται διαφορετικά, όπως αναφέρθηκε στην προηγούμενη ενότητα). Η μέθοδος που επιλέχθηκε για την επίλυση του αναφερθέντος προβλήματος είναι η επαυξημένη μέθοδος με πολλαπλασιαστές Lagrange ή συντομογραφικά ALM, η οποία αναλύθηκε στην ενότητα 2.4.4. Στην περίπτωση του προβλήματος που πραγματεύεται η παρούσα ενότητα, η επαυξημένη συνάρτηση Lagrange διατυπώνεται ως εξής:

$$\Phi^*(\vec{x}, \hat{\lambda}_2, \omega_p) = f_1(\vec{x}) - \hat{\lambda}_2 [f_2(\vec{x}) - \hat{f}_2] + \omega_p [f_2(\vec{x}) - \hat{f}_2]^2$$

Στην ενότητα 2.4.4.3 παρουσιάσθηκε αναλυτικά ο αλγόριθμος της μεθόδου ALM για προβλήματα βελτιστοποίησης με περιορισμούς ισότητας και ανισότητας. Για το λόγο αυτό, κρίνεται σκόπιμο να αναφερθούν εδώ τα βασικά στοιχεία προσαρμοσμένα στην περίπτωση που διαχειρίζεται η παρούσα ενότητα. Για να προβεί η ALM στην εύρεση του πρώτου σημείου του Pareto, χρειάζεται μία αρχική τιμή του διανύσματος των μεταβλητών σχεδιασμού \vec{x}^0 , μία αρχική τιμή του πολλαπλασιαστή Lagrange $\hat{\lambda}_2^0$ (συνήθως τιμή είναι η μηδενική), μία αρχική (συνήθως μικρή) τιμή του συντελεστή ποινής ω_p^0 , μία τιμή του συντελεστή γ , με τον οποίο θα πολλαπλασιάζεται η τιμή του ω_p σε

κάθε νέο κύκλο βελτιστοποίησης, και μία ανώτατη τιμή που θα επιτρέπεται να λάβει ο συντελεστής ποινής ω_p^{max} . Με δεδομένα όλα τα παραπάνω, το επόμενο βήμα είναι η ελαχιστοποίηση της ψευδο-συνάρτησης στόχου $\Phi^*(\vec{x}, \hat{\lambda}_2, \omega_p)$. Επιλύεται, δηλαδή, ένα πρόβλημα χωρίς περιορισμούς, επιλέγοντας οιαδήποτε κατάλληλη μέθοδο. Στη συγκεκριμένη περίπτωση, επιλέχθηκαν δύο τέτοιες μέθοδοι: η μέθοδος της απότομης καθόδου (αναλυτική περιγραφή της συναντάται στην ενότητα 3.3.2) και η μέθοδος Newton (αναλυτική περιγραφή της συναντάται στην ενότητα 3.3.3). Κατόπιν, ελέγχεται η σύγκλιση της ALM. Το κριτήριο που εφαρμόζεται είναι η ικανοποίηση του περιορισμού ισότητας του προβλήματος με την επιθυμητή ανοχή. Είναι φανερό ότι αν η μέθοδος έχει συγκλίνει ο αλγόριθμος σταματά. Σε αντίθετη περίπτωση, ο αλγόριθμος προχωράει στην εκτέλεση του επόμενου βήματος που είναι η ανανέωση της τιμής του $\hat{\lambda}_2$ και αυτής του ω_p σύμφωνα με τις σχέσεις:

$$\hat{\lambda}_2 \leftarrow \hat{\lambda}_2 - 2\omega_p [f_2(\vec{x}) - \hat{f}_2]$$

$$\omega_p \leftarrow \min(\gamma\omega_p, \omega_p^{max})$$

Μετά την ανανέωση των τιμών των $\hat{\lambda}_2$ και ω_p , ο αλγόριθμος επιστρέφει στο βήμα όπου επιλύεται το πρόβλημα ελαχιστοποίησης της $\Phi^*(\vec{x}, \hat{\lambda}_2, \omega_p)$ μέχρι να επιτευχθεί η σύγκλιση της ALM.

Έχοντας αποκτήσει, πλέον, το πρώτο σημείο του μετώπου Pareto, ο αλγόριθμος προχωράει στην εύρεση όσων επιπλέον μη-κυριαρχούμενων λύσεων επιθυμεί ο χρήστης (move on Pareto), στηριζόμενος στη θεωρία της συνέχισης του μετώπου Pareto όπως αναλύθηκε στην προηγούμενη ενότητα.

Για να υλοποιηθούν όλα τα παραπάνω και να επιτευχθεί η επίλυση του μαθηματικού προβλήματος βελτιστοποίησης που παρουσιάστηκε στην ενότητα 5.3.1, προγραμματίστηκε λογισμικό σε γλώσσα FORTRAN 77. Στη συνέχεια, πραγματοποιείται η περιγραφή των βημάτων που εκτελεί ο αναφερθείς κώδικας με τελικό στόχο την εύρεση του μετώπου Pareto, ώστε να διαφανεί καθαρότερα η ακριβής λειτουργία του αναπτυχθέντος αλγορίθμου του οποίου τα βήματα παρουσιάστηκαν παραπάνω στη γενικότητά τους.

Αρχικά, επιβάλλεται διπλή ακρίβεια στις πραγματικές μεταβλητές. Ακολούθως, ο κώδικας διαβάζει τις πληροφορίες που χρειάζεται από έναν αριθμό αρχείων εισόδου. Συγκεκριμένα, τροφοδοτούνται στο πρόγραμμα η αρχική εκτίμηση για το \hat{f}_2 (\hat{f}_2^{up}), η αρχική τιμή του πολλαπλασιαστή Lagrange $\hat{\lambda}_2^0$, η αρχική τιμή του συντελεστή ποινής ω_p^0 , η τιμή του συντελεστή γ , η ανώτατη τιμή που επιτρέπεται να λάβει ο συντελεστής ποινής ω_p^{max} , ένας αριθμός που θα πληροφορεί για το αν επιθυμείται να χρησιμοποιηθεί η μέθοδος της απότομης καθόδου

ή η μέθοδος Newton (δίνεται 1 για την πρώτη και 2 για τη δεύτερη) ως μέθοδος βελτιστοποίησης προβλημάτων ενός στόχου δίχως περιορισμούς, το μέτρο μετατόπισης η (σταθερό σε κάθε επανάληψη) για τη μέθοδο της απότομης καθόδου, ο αριθμός των σημείων που είναι επιθυμητό να αποκτηθούν πάνω στο μέτωπο Pareto n_{elit} , μία εκτίμηση \hat{f}_2^{down} η οποία αφορά στο κάτω όριο του μετώπου Pareto στον άξονα των f_2 (εναλλακτικά θα μπορούσε να δίνεται ένα βήμα μετατόπισης $\delta\hat{f}_2$), ένας αριθμός που δηλώνει το πόσες είναι οι μεταβλητές σχεδιασμού στο πρόβλημα, μία αρχική τιμή του διανύσματος των μεταβλητών σχεδιασμού \vec{x}^0 και, τέλος, τα κάτω και άνω όρια των μεταβλητών σχεδιασμού.

Στη συνέχεια, ο κώδικας εισέρχεται στο τμήμα που είναι υπεύθυνο για την εύρεση του πρώτου σημείου πάνω στο μέτωπο Pareto (go to Pareto). Το πρώτο πράγμα που λαμβάνει χώρα σε αυτό το τμήμα είναι ο υπολογισμός των τιμών των συναρτήσεων στόχου f_1 και f_2 για $\vec{x} = \vec{x}^0$. Στο σημείο αυτό, εισάγεται στον κώδικα ο μετρητής *ievals* ο οποίος θα είναι υπεύθυνος για την παροχή στο χρήστη μετά το πέρας της εκτέλεσης του προγράμματος του αριθμού των αξιολογήσεων που πραγματοποιήθηκαν για την απόκτηση του μετώπου Pareto. Μετά, λοιπόν, τον υπολογισμό των τιμών των συναρτήσεων στόχου για $\vec{x} = \vec{x}^0$ ο εν λόγω μετρητής λαμβάνει την τιμή 1. Ακολουθώντας, υπολογίζεται η τιμή της συνάρτησης του περιορισμού ισότητας $c(\vec{x}) = f_2(\vec{x}) - \hat{f}_2$ για $\vec{x} = \vec{x}^0$. Έπειτα, ο κώδικας δημιουργεί δύο αρχεία, ένα που αφορά στην 'ιστορία' της μεθόδου της απότομης καθόδου ή της Newton και ένα που αφορά σε αυτή της ALM. Στο παρόν σημείο που βρίσκεται το πρόγραμμα, 'γεμίζει' την πρώτη γραμμή των αναφερθέντων αρχείων γράφοντας τις ίδιες πληροφορίες και στα δύο οι οποίες είναι οι τιμές των μεταβλητών σχεδιασμού (στην παρούσα φάση $\vec{x} = \vec{x}^0$), οι τιμές των συναρτήσεων στόχου που υπολογίστηκαν βάσει των προαναφερθέντων τιμών των μεταβλητών σχεδιασμού και η τιμή που έχει λάβει η συνάρτηση περιορισμού $c(\vec{x})$.

Κατόπιν, εισάγεται ένας νέος μετρητής j ο οποίος θα είναι υπεύθυνος για τον αριθμό των κύκλων που πραγματοποιεί η ALM. Σε πρώτο στάδιο ο εν λόγω μετρητής λαμβάνει την τιμή 0 και έπειτα με την εκτέλεση κάθε κύκλου της ALM η τιμή του θα αυξάνεται κατά 1. Το επόμενο βήμα αφορά στην εισαγωγή ενός ακόμα μετρητή k ο οποίος είναι επιφορτισμένος με την παροχή του αριθμού των επαναλήψεων που εκτελούνται από τη μέθοδο της απότομης καθόδου ή τη μέθοδο Newton (ανάλογα με το ποια επιλέγεται). Ακολουθώντας, υπολογίζονται οι πρώτες παράγωγοι των συναρτήσεων στόχου f_1 και f_2 ως προς τις μεταβλητές σχεδιασμού για τις τρέχουσες τιμές των μεταβλητών σχεδιασμού. Στη συγκεκριμένη περίπτωση πρόκειται για ένα απλό μαθηματικό πρόβλημα οπότε οι εν λόγω παράγωγοι υπολογίζονται αναλυτικά ως εξής:

$$\frac{\partial f_1}{\partial x_1} = 2x_1, \quad \frac{\partial f_1}{\partial x_2} = 0$$

$$\frac{\partial f_2}{\partial x_1} = -\frac{x_2^2 - 5x_2 + \sin x_2 + 15}{x_1^2}, \quad \frac{\partial f_2}{\partial x_2} = \frac{2x_2 - 5 + \cos x_2}{x_1}$$

Στο σημείο αυτό, ο μετρητής που διαχειρίζεται τον αριθμό των αξιολογήσεων που εκτελούνται από το πρόγραμμα αυξάνει την τιμή του κατά 1. Αυτό συμβαίνει διότι, αν και, όπως προλέχθηκε, στη συγκεκριμένη περίπτωση οι παράγωγοι υπολογίζονται αναλυτικά, ο μετρητής θεωρεί ότι για την εύρεσή τους έγινε χρήση της συζυγούς μεθόδου (adjoint method) και είναι γνωστό, [4], ότι η εν λόγω μέθοδος χρειάζεται 1+1 αξιολογήσεις ανεξάρτητα του πλήθους των μεταβλητών σχεδιασμού (η μία έχει ήδη ληφθεί υπόψη κατά την εύρεση των τιμών των συναρτήσεων στόχου) για να υπολογίσει τις πρώτες παραγώγους. Ο λόγος για να θεωρηθεί κάτι τέτοιο είναι ότι αν επρόκειτο για ένα πραγματικό πρόβλημα, η συζυγής μέθοδος θα ήταν ενδεχομένως η πρώτη επιλογή για την απόκτηση των παραγώγων (λόγω της 'ταχύτητάς' της), οπότε ενδιαφέρει ο αριθμός των αξιολογήσεων να υπολογιστεί με βάση αυτή τη μέθοδο.

Έπειτα, ο κώδικας προχωράει στον υπολογισμό των πρώτων παραγώγων της επαυξημένης συνάρτησης Lagrange $\Phi^*(\vec{x}, \hat{\lambda}_2, \omega_p)$ ως προς τις μεταβλητές σχεδιασμού, ως εξής:

$$\frac{\partial \Phi^*}{\partial x_i} = \frac{\partial f_1}{\partial x_i} - \hat{\lambda}_2 \frac{\partial f_2}{\partial x_i} + 2\omega_p c \frac{\partial f_2}{\partial x_i}, \quad i = 1, 2$$

Στην περίπτωση που έχει επιλεγεί η μέθοδος Newton, όπως αναφέρθηκε στην ενότητα 3.3.3, επιτάσσεται η απόκτηση και των δευτέρων παραγώγων. Έτσι, ο κώδικας, αρχικά, προβαίνει στον υπολογισμό του εσσιανού μητρώου για κάθε μία από τις συναρτήσεις στόχου. Όπως συνέβη και με τις πρώτες παραγώγους, ο εν λόγω υπολογισμός στο απλό αυτό μαθηματικό πρόβλημα πραγματοποιείται αναλυτικά ως εξής:

$$H_1 = \begin{bmatrix} \frac{\partial^2 f_1}{\partial x_1^2} & \frac{\partial^2 f_1}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f_1}{\partial x_1 \partial x_2} & \frac{\partial^2 f_1}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$

$$H_2 = \begin{bmatrix} \frac{\partial^2 f_2}{\partial x_1^2} & \frac{\partial^2 f_2}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f_2}{\partial x_1 \partial x_2} & \frac{\partial^2 f_2}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} \frac{2(x_2^2 - 5x_2 + \sin x_2 + 15)}{x_1^3} & -\frac{2x_2 - 5 + \cos x_2}{x_1^2} \\ -\frac{2x_2 - 5 + \cos x_2}{x_1^2} & \frac{2 - \sin x_2}{x_1} \end{bmatrix}$$

Στο σημείο αυτό, ο μετρητής που διαχειρίζεται τον αριθμό των αξιολογήσεων που εκτελούνται από το πρόγραμμα αυξάνει την τιμή

του κατά n . Αυτό συμβαίνει διότι, αν και, όπως ειπώθηκε, στη συγκεκριμένη περίπτωση οι δεύτερες παράγωγοι υπολογίζονται αναλυτικά, ο μετρητής θεωρεί ότι για την εύρεσή τους έγινε χρήση της μεθόδου της ευθείας διαφορίσης (direct differentiation method) και είναι γνωστό, [4], ότι η εν λόγω μέθοδος χρειάζεται n επιπλέον αξιολογήσεις (οι $1+1$ έχουν ήδη συνυπολογιστεί στο στάδιο υπολογισμού των συναρτήσεων στόχου και των πρώτων παραγώγων), όσες είναι και οι μεταβλητές σχεδιασμού, για να υπολογίσει τις δεύτερες παραγώγους.

Το επόμενο βήμα που εκτελείται σε περίπτωση επιλογής της μεθόδου Newton, είναι η εύρεση του εσσιανού μητρώου της επαυξημένης συνάρτησης Lagrange $\Phi^*(\vec{x}, \hat{\lambda}_2, \omega_p)$ ως εξής:

$$\frac{\partial^2 \Phi^*}{\partial x_i \partial x_j} = \frac{\partial^2 f_1}{\partial x_i \partial x_j} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_i \partial x_j} + 2\omega_p c \frac{\partial^2 f_2}{\partial x_i \partial x_j} + 2\omega_p \frac{\partial f_2}{\partial x_j} \frac{\partial f_2}{\partial x_i}, \quad i, j = 1, 2$$

Το τελευταίο επιπλέον βήμα που εκτελείται σε περίπτωση χρήσης της μεθόδου Newton, είναι η εύρεση της κατεύθυνσης ανίχνευσης για την τρέχουσα επανάληψη. Όπως διατυπώθηκε στην ενότητα 3.3.3, αυτό γίνεται μέσω της σχέσης:

$$\vec{p} = -[\nabla^2 \Phi^*(\vec{x})]^{-1} \nabla \Phi^*(\vec{x})$$

Για την επίλυση του ανωτέρω συστήματος και την απόκτηση του \vec{p} , ο κώδικας καλεί την υπορουτίνα gauss. Η αναφερθείσα υπορουτίνα περιέχει τον αλγόριθμο που υλοποιεί τη γνωστή από τη γραμμική άλγεβρα μέθοδο απαλοιφής κατά Gauss για την επίλυση γραμμικών συστημάτων, [18]. Το γραμμικό σύστημα που επιλύεται από την υπορουτίνα gauss είναι της μορφής $Ay = B$. Στη συγκεκριμένη περίπτωση, είναι $A = \nabla^2 \Phi^*(\vec{x})$, $y = \vec{p}$ και $B = -\nabla \Phi^*(\vec{x})$. Η εν λόγω υπορουτίνα, λοιπόν, λαμβάνει ως πληροφορίες από το πρόγραμμα τον αριθμό των μεταβλητών σχεδιασμού, το εσσιανό μητρώο της Φ^* και το αντίθετο του μητρώου που περιέχει τις πρώτες παραγώγους της Φ^* ως προς τις μεταβλητές σχεδιασμού και επιστρέφει σε αυτό την κατεύθυνση ανίχνευσης \vec{p} για την τρέχουσα επανάληψη.

Το ακόλουθο βήμα που εκτελείται από το πρόγραμμα είναι η ανανέωση των τιμών των μεταβλητών σχεδιασμού. Αν η επιλεγθείσα μέθοδος βελτιστοποίησης για προβλήματα ενός στόχου χωρίς περιορισμούς είναι η μέθοδος της απότομης καθόδου η εν λόγω ανανέωση γίνεται ως εξής:

$$x_i^{k+1} = x_i^k - \eta \left. \frac{\partial \Phi^*}{\partial x_i} \right|_k, \quad i = 1, 2$$

Στην περίπτωση που επιλέγεται η μέθοδος Newton, η ανανέωση των μεταβλητών σχεδιασμού πραγματοποιείται μέσω της παρακάτω σχέσης:

$$\vec{x}^{k+1} = \vec{x}^k + \vec{p}^k$$

Στη συνέχεια, πραγματοποιείται έλεγχος για το αν οι μεταβλητές σχεδιασμού στη δεδομένη επανάληψη k είναι εκτός των ορίων τους. Σε περίπτωση που η μεταβλητή x_i έχει αποκτήσει στην τρέχουσα επανάληψη τιμή μικρότερη του κάτω ορίου της, το πρόγραμμα της επιβάλλει την τιμή του κάτω αυτού ορίου, ενώ αν έχει αποκτήσει τιμή μεγαλύτερη του άνω ορίου της, της επιβάλλεται η τιμή του άνω αυτού ορίου.

Το επόμενο βήμα που λαμβάνει χώρα είναι ο υπολογισμός των τιμών των συναρτήσεων στόχου f_1 και f_2 για τις νέες τιμές των μεταβλητών σχεδιασμού που προέκυψαν. Στο σημείο αυτό η τιμή του μετρητή *inevals* που παρέχει τον αριθμό των αξιολογήσεων που εκτελούνται από το πρόγραμμα αυξάνεται κατά 1. Ακολούθως, υπολογίζεται η τιμή της συνάρτησης του περιορισμού ισότητας $c(\vec{x}) = f_2(\vec{x}) - \hat{f}_2$ για τις νέες τιμές των μεταβλητών σχεδιασμού. Κατόπιν, το πρόγραμμα γράφει στη γραμμή $k + 1$ του αρχείου που περιλαμβάνει την 'ιστορία' της μεθόδου της απότομης καθόδου ή της Newton τις νέες τιμές των μεταβλητών σχεδιασμού, τις νέες τιμές των συναρτήσεων στόχου που υπολογίστηκαν βάσει των προαναφερθεισών τιμών των μεταβλητών σχεδιασμού και την τιμή που έχει λάβει η συνάρτηση περιορισμού $c(\vec{x})$.

Στη συνέχεια, ο κώδικας προβαίνει στον έλεγχο της σύγκλισης της μεθόδου της απότομης καθόδου ή της μεθόδου Newton (ανάλογα με το ποια έχει επιλεγεί). Αυτό γίνεται μέσω ενός επαναληπτικού βρόχου για τις μεταβλητές σχεδιασμού μέσα στον οποίο αυξάνεται η τιμή ενός δείκτη (*iconv*) κατά 1 για κάθε μεταβλητή που είτε ικανοποιεί το κριτήριο μηδενικής παραγώγου $\partial\Phi^*/\partial x_i$ με την επιθυμητή ανοχή (την οποία επιβάλλει ο χρήστης ανάλογα με τις ανάγκες του) είτε έχει λάβει την τιμή κάποιου ορίου της. Αν η μέθοδος βελτιστοποίησης για προβλήματα ενός στόχου χωρίς περιορισμούς δεν έχει συγκλίνει, δηλαδή αν $iconv < n$, ο μετρητής k αυξάνει την τιμή του κατά 1 και ο κώδικας επιστρέφει στο σημείο όπου υπολογίζονται οι πρώτες παράγωγοι των συναρτήσεων στόχου f_1 και f_2 ως προς τις μεταβλητές σχεδιασμού με τη διαδικασία που περιγράφηκε από το εν λόγω σημείο και πέρα να επαναλαμβάνεται για τις νέες τιμές των μεταβλητών σχεδιασμού. Σημειώνεται ότι σε περίπτωση που το k ξεπεράσει μία πολύ μεγάλη τιμή (που ορίζεται από το χρήστη) το πρόγραμμα θεωρεί αδύνατη τη σύγκλιση και σταματά την εκτέλεσή του. Αν, αντίθετα, στην τρέχουσα επανάληψη ισχύει $iconv = n$, έχει επιτευχθεί η σύγκλιση της μεθόδου βελτιστοποίησης για προβλήματα ενός στόχου χωρίς περιορισμούς. Τότε, ο κώδικας γράφει στη γραμμή $j + 1$ του αρχείου που περιέχει την 'ιστορία' της ALM τις τρέχουσες

τιμές των μεταβλητών σχεδιασμού (προέκυψαν μετά από τη σύγκλιση της μεθόδου της απότομης καθόδου/Newton), τις τρέχουσες τιμές των συναρτήσεων στόχου που υπολογίστηκαν βάσει των προαναφερθεισών τιμών των μεταβλητών σχεδιασμού, την τιμή που έχει λάβει η συνάρτηση περιορισμού $c(\vec{x})$ και τις τιμές των $\hat{\lambda}_2$ και ω_p . Κατόπιν, το πρόγραμμα προχωράει στον έλεγχο της σύγκλισης της ALM. Το κριτήριο που εφαρμόζεται είναι η ικανοποίηση του περιορισμού ισότητας του προβλήματος με την επιθυμητή ανοχή. Αν η ALM δεν έχει συγκλίνει, δηλαδή αν η τρέχουσα απόλυτη τιμή του c είναι μεγαλύτερη από έναν ‘πολύ μικρό’ αριθμό που ορίζει ο χρήστης βασιζόμενος στην ακρίβεια που απαιτείται από το πρόβλημα, ο κώδικας προχωράει στην ανανέωση της τιμής του $\hat{\lambda}_2$ και αυτής του ω_p σύμφωνα με τις σχέσεις:

$$\hat{\lambda}_2^{j+1} = \hat{\lambda}_2^j - 2\omega_p [f_2(\vec{x})|_j - \hat{f}_2]$$

$$\omega_p^{j+1} = \min(\gamma\omega_p^j, \omega_p^{max})$$

Μετά την ανανέωση των τιμών των $\hat{\lambda}_2$ και ω_p , το πρόγραμμα επιστρέφει στο βήμα όπου ο μετρητής j που είναι υπεύθυνος για τον αριθμό των κύκλων που πραγματοποιεί η ALM αυξάνει την τιμή του κατά 1 με την εκτέλεση κάθε κύκλου της ALM. Έτσι, ξεκινάει άλλος ένας κύκλος της ALM και εκτελούνται εκ νέου όλα όσα περιγράφηκαν παραπάνω από το εν λόγω σημείο και μετά. Σημειώνεται ότι σε περίπτωση που το j ξεπεράσει μία πολύ μεγάλη τιμή (που ορίζεται από το χρήστη) το πρόγραμμα θεωρεί αδύνατη τη σύγκλιση της ALM και η εκτέλεσή του τερματίζεται. Αν, αντίθετα, στον τρέχοντα κύκλο, έχει επιτευχθεί η σύγκλιση της ALM, δηλαδή αν ικανοποιείται ο περιορισμός ισότητας με την επιθυμητή ανοχή, το πρόγραμμα, εξερχόμενο από τον επαναληπτικό βρόχο της ALM, έχει, πλέον, φτάσει στην εύρεση του πρώτου σημείου πάνω στο μέτωπο Pareto.

Μία από τις πληροφορίες που διάβασε ο κώδικας στο ξεκίνημα της εκτέλεσής του ήταν ο αριθμός των σημείων που είναι επιθυμητό να αποκτηθούν πάνω στο μέτωπο Pareto n_{elit} . Έχοντας ήδη εκτελέσει τη διαδικασία που περιγράφηκε ενδελεχώς παραπάνω, ο κώδικας διαθέτει το πρώτο σημείο του μετώπου Pareto και εισέρχεται στο τμήμα που είναι υπεύθυνο για την εύρεση των υπόλοιπων $n_{elit} - 1$ μη-κυριαρχούμενων λύσεων που είναι επιθυμητό να αποκτηθούν. Ο κώδικας, λοιπόν, θα αρχίσει να ‘κινείται’ πάνω στο μέτωπο μη-κυριαρχούμενων λύσεων (move on Pareto) με βάση τη θεωρία της συνέχισης του μετώπου Pareto, η εφαρμογή της οποίας στο δεδομένο μαθηματικό πρόβλημα παρουσιάσθηκε στην ενότητα που προηγήθηκε.

Το επόμενο βήμα που εκτελείται από τον κώδικα είναι ο υπολογισμός του βήματος μετατόπισης $\delta\hat{f}_2$. Κάτι τέτοιο επιτυγχάνεται αν αφαιρεθεί

το κάτω όριο του μετώπου Pareto που παρασχεόθηκε στο πρόγραμμα από το άνω όριο και η διαφορά διαιρεθεί με $n_{elit} - 1$.

Μετά από τα παραπάνω, λαμβάνει χώρα η είσοδος στον επιφορτισμένο με την παροχή των υπόλοιπων $n_{elit} - 1$ μη-κυριαρχούμενων λύσεων βρόχο. Ακολουθεί η περιγραφή των εντολών που περιέχονται σε αυτόν και επαναλαμβάνονται κυκλικά για τα $n_{elit} - 1$ σημεία του Pareto που αναζητούνται, από το δεύτερο μέχρι το τελευταίο. Το πρώτο βήμα που περιλαμβάνεται στον εν λόγω βρόχο είναι η εύρεση της νέας τιμής του \hat{f}_2 ως εξής:

$$\hat{f}_2^{new} = \hat{f}_2^{old} - \delta \hat{f}_2$$

Στη συνέχεια, υπολογίζονται οι πρώτες παράγωγοι των συναρτήσεων στόχου f_1 και f_2 ως προς τις μεταβλητές σχεδιασμού για τις τρέχουσες τιμές των μεταβλητών σχεδιασμού (όταν πρόκειται για την πρώτη επανάληψη είναι οι τιμές που προέκυψαν μετά το τέλος της ALM). Ο τρόπος υπολογισμού τους σχολιάσθηκε παραπάνω. Στο σημείο αυτό, ο μετρητής που διαχειρίζεται τον αριθμό των αξιολογήσεων που εκτελούνται από το πρόγραμμα *ieval*s αυξάνει την τιμή του κατά 1. Αυτό συμβαίνει διότι, όπως προλέχθηκε, ο μετρητής θεωρεί ότι για την εύρεσή τους έγινε χρήση της συζυγούς μεθόδου. Ακολουθώντας, ο κώδικας προβαίνει στον υπολογισμό του εσσιανού μητρώου για κάθε μία από τις συναρτήσεις στόχου. Μετά τον υπολογισμό των δεύτερων παραγώγων, ο μετρητής *ieval*s αυξάνει την τιμή του κατά n . Αυτό συμβαίνει διότι, όπως ειπώθηκε, ο εν λόγω μετρητής θεωρεί ότι για την εύρεσή τους έγινε χρήση της μεθόδου της ευθείας διαφόρισης.

Το επόμενο βήμα που εκτελείται είναι η κατασκευή των μητρώων $\partial \vec{h} / \partial \vec{z}$ και $\partial \vec{h} / \partial \hat{f}_2$, όπως αυτά παρουσιάστηκαν στην προηγούμενη ενότητα. Έπειτα, πραγματοποιείται ο υπολογισμός της παραγώγου $\partial \vec{h} / \partial \hat{f}_2$ μέσω της σχέσης:

$$\frac{\partial \vec{h}}{\partial \hat{f}_2}(\hat{f}_2) = - \left\{ \frac{\partial \vec{H}}{\partial \vec{z}} [\vec{h}(\hat{f}_2), \hat{f}_2] \right\}^{-1} \frac{\partial \vec{H}}{\partial \hat{f}_2} [\vec{h}(\hat{f}_2), \hat{f}_2]$$

Για την επίλυση του ανωτέρω συστήματος εξισώσεων και την απόκτηση του $\partial \vec{h} / \partial \hat{f}_2$, ο κώδικας καλεί την υπορουτίνα *gauss*, η λειτουργία της οποίας περιγράφηκε ανωτέρω. Πρέπει να τονιστεί ότι, επειδή, όπως αναφέρθηκε, το γραμμικό σύστημα που επιλύεται από την υπορουτίνα *gauss* είναι της μορφής $Ay = B$, όπου, στη συγκεκριμένη περίπτωση, $A = \partial \vec{H} / \partial \vec{z}$, $y = \partial \vec{h} / \partial \hat{f}_2$ και $B = - \partial \vec{H} / \partial \hat{f}_2$, το μητρώο που στην πραγματικότητα κατασκευάστηκε ήταν το $B = - \partial \vec{H} / \partial \hat{f}_2$ και όχι το $\partial \vec{H} / \partial \hat{f}_2$. Η εν λόγω υπορουτίνα, λοιπόν, λαμβάνει ως πληροφορίες από το πρόγραμμα τον αριθμό των μεταβλητών σχεδιασμού προστεθειμένο με τον αριθμό των περιορισμών (στη συγκεκριμένη περίπτωση είναι ένας), καθώς πρόκειται για τον αριθμό

των εξισώσεων του συστήματος, το μητρώο $\partial \vec{H} / \partial \vec{z}$ και το αντίθετο του μητρώου $\partial \vec{H} / \partial \hat{f}_2$ και επιστρέφει σε αυτό την παράγωγο $\partial \vec{h} / \partial \hat{f}_2$. Στη συνέχεια, καθίσταται δυνατό, αφού, πλέον, διατίθεται η παράγωγος $\partial \vec{h} / \partial \hat{f}_2$, να πραγματοποιηθεί η εύρεση των νέων τιμών των μεταβλητών σχεδιασμού και του $\hat{\lambda}_2$, δηλαδή του επόμενου σημείου πάνω στο μέτωπο Pareto (στην πραγματικότητα αποτελεί προσέγγισή του, καθώς απέχει μία ‘μικρή’ απόσταση από αυτό, αλλά, όπως αναφέρθηκε, στην περίπτωση αφαίρεσης του σταδίου της διόρθωσης λαμβάνεται ως η επόμενη μη-κυριαρχούμενη λύση χωρίς να πραγματοποιείται διόρθωσή της), ως εξής:

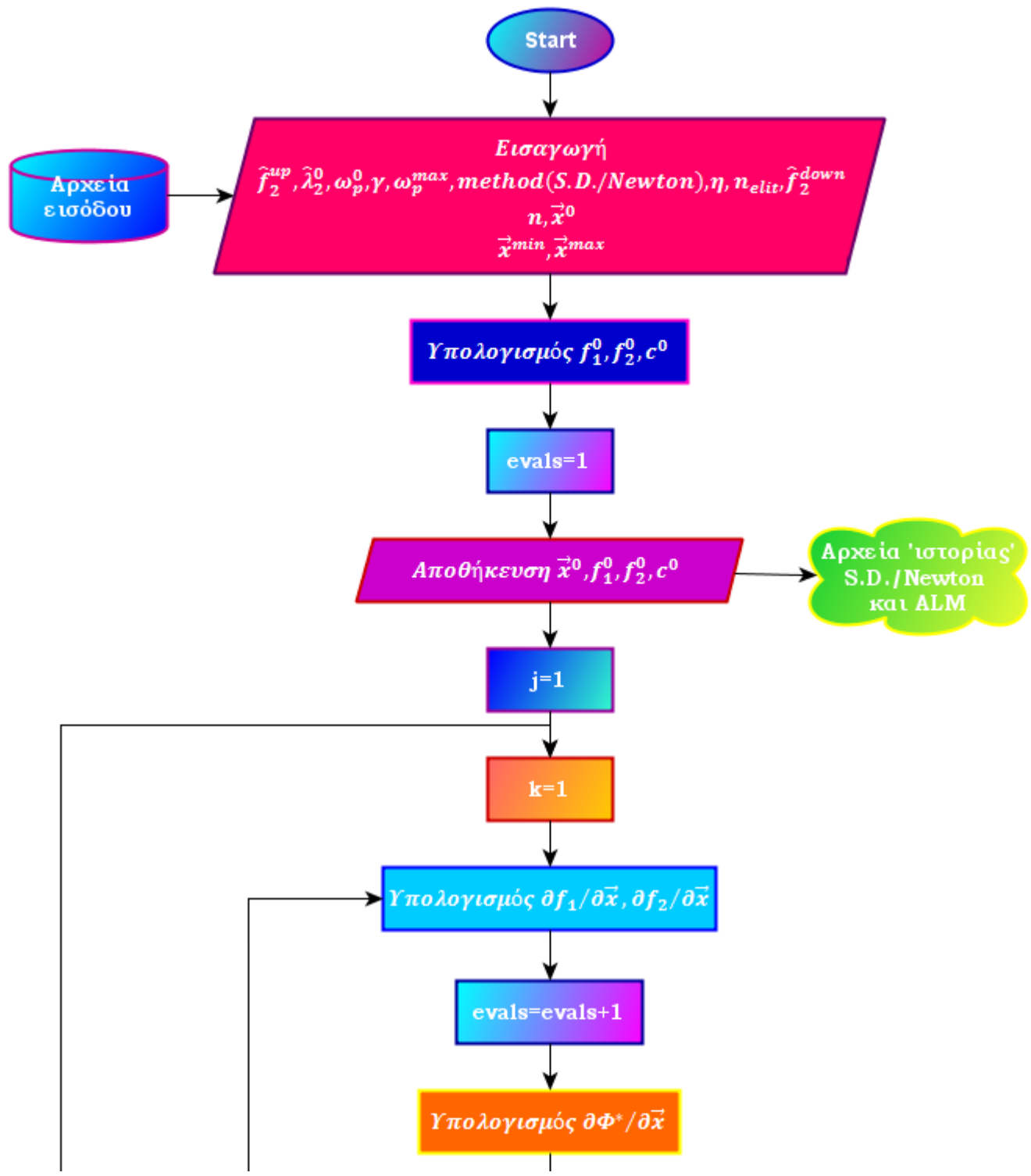
$$\vec{z}^{new} = \vec{z}^{old} + \frac{\partial \vec{h}}{\partial \hat{f}_2} (\hat{f}_2^{new} - \hat{f}_2^{old})$$

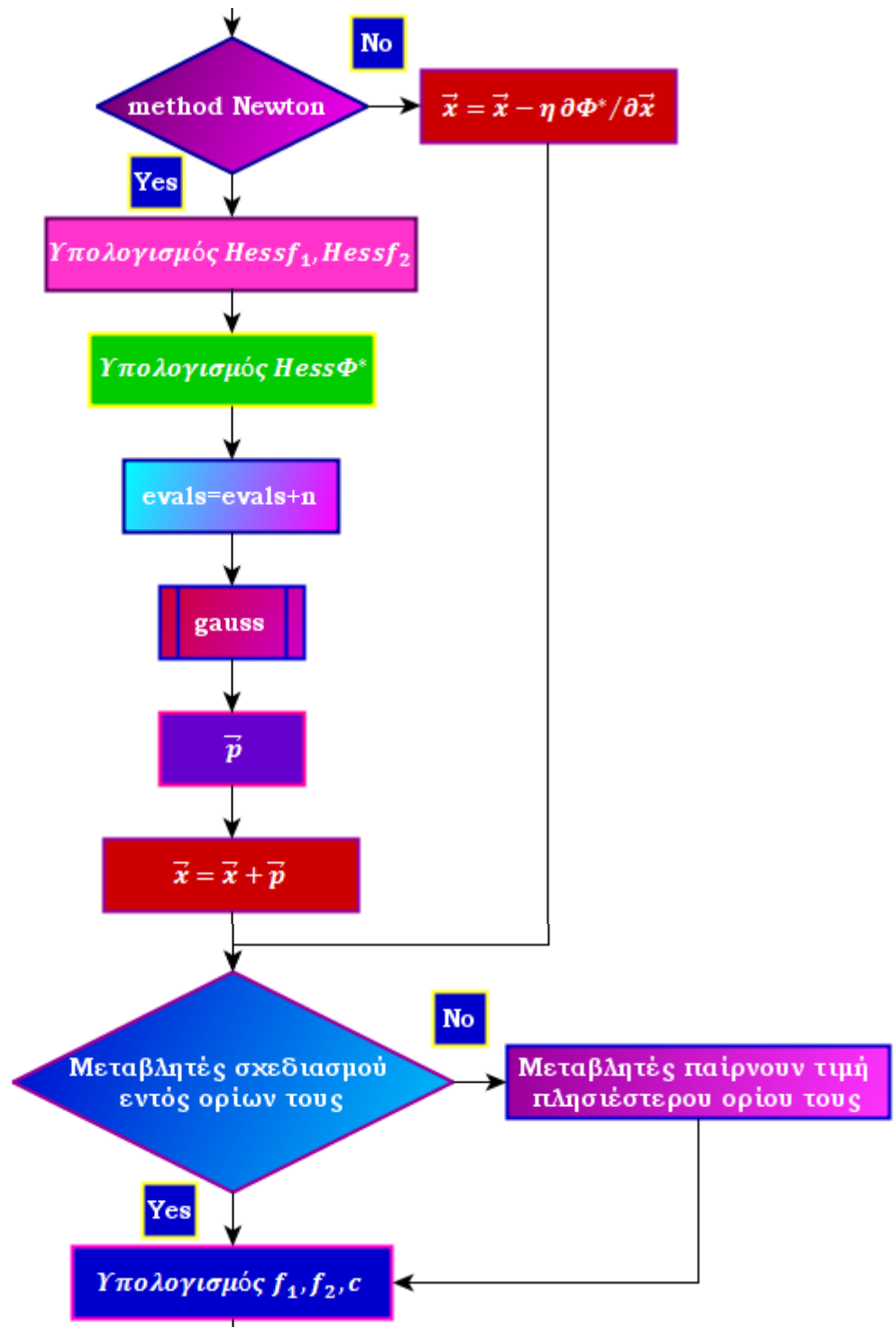
Ακολούθως, πραγματοποιείται έλεγχος για το αν οι μεταβλητές σχεδιασμού στη δεδομένη επανάληψη είναι εκτός των ορίων τους. Σε περίπτωση που η μεταβλητή x_i έχει αποκτήσει στην τρέχουσα επανάληψη τιμή μικρότερη του κάτω ορίου της, το πρόγραμμα της επιβάλλει την τιμή του κάτω αυτού ορίου, ενώ αν έχει αποκτήσει τιμή μεγαλύτερη του άνω ορίου της, της επιβάλλεται η τιμή του άνω αυτού ορίου.

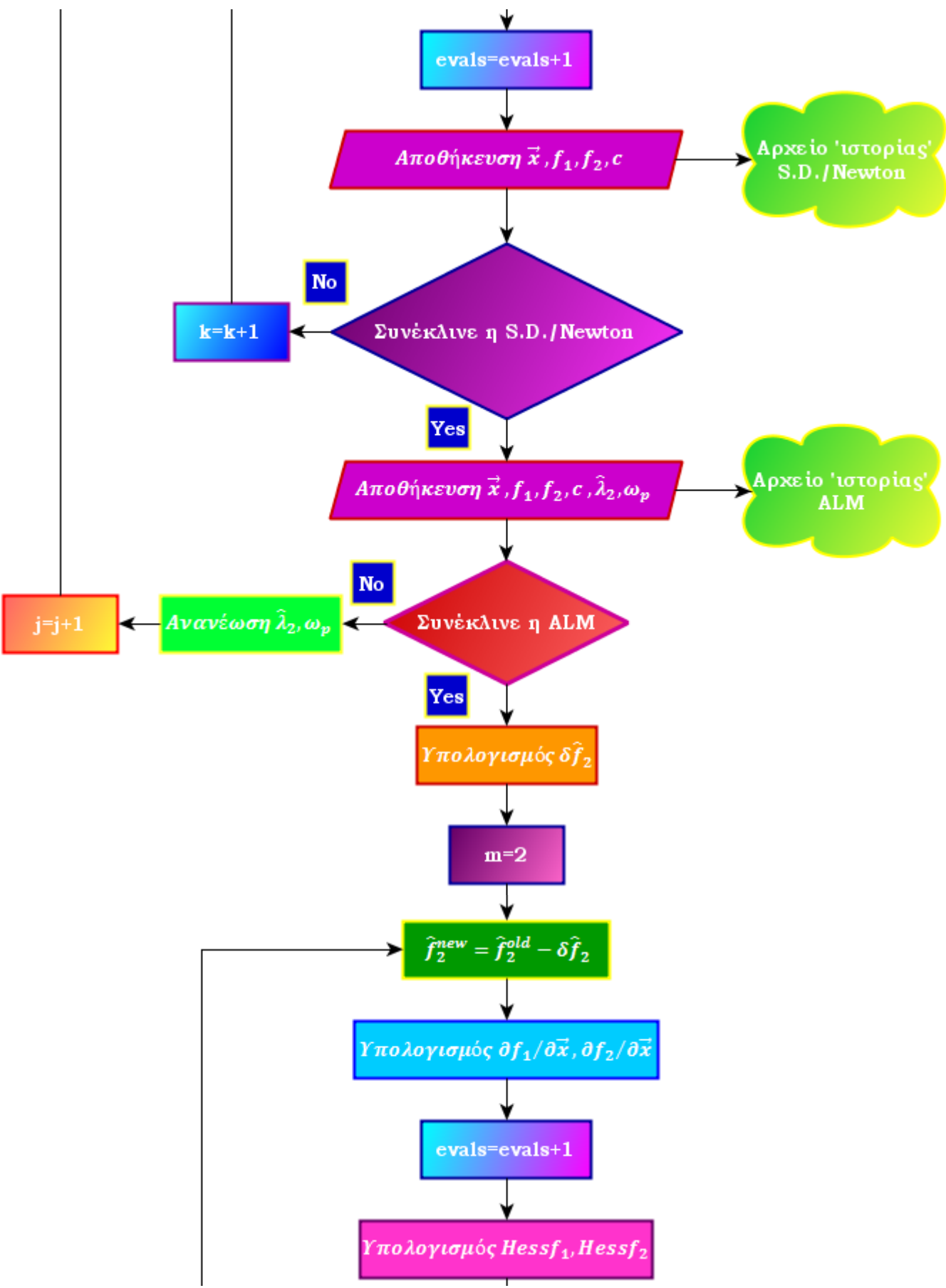
Κατόπιν, υπολογίζονται οι τιμές των συναρτήσεων στόχου f_1 και f_2 με βάση τις προαναφερθείσες τιμές των μεταβλητών σχεδιασμού. Στο παρόν σημείο, το πρόγραμμα υποχρεούται να αυξήσει εκ νέου την τιμή του μετρητή *itvals* κατά 1. Η τελευταία εντολή που περιλαμβάνεται στο βρόχο εύρεσης των $n_{elit} - 1$ μη-κυριαρχούμενων λύσεων είναι η εκχώρηση της τιμής \hat{f}_2^{new} στη μεταβλητή \hat{f}_2^{old} .

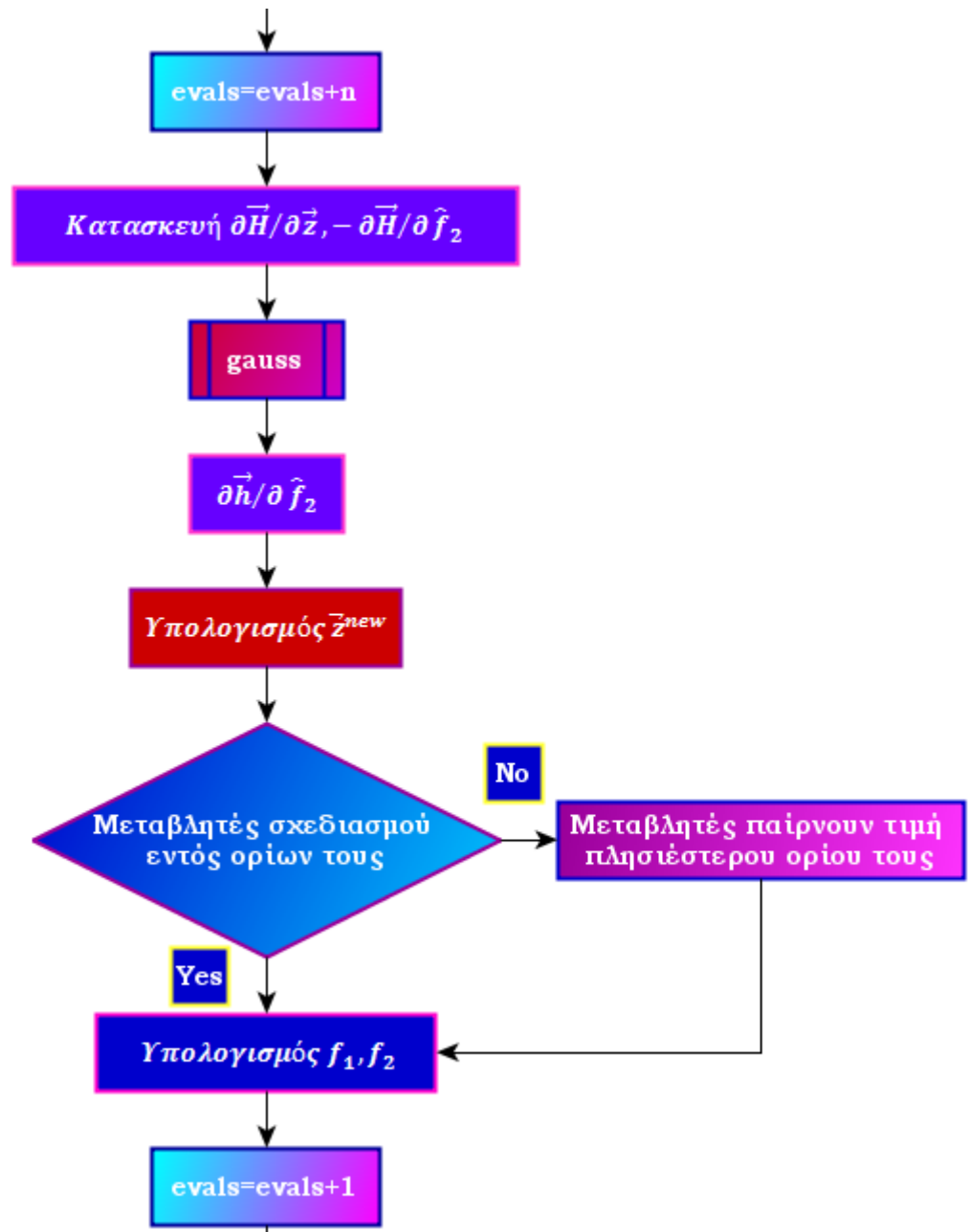
Μετά την έξοδο του προγράμματος από τον ανωτέρω περιγραφέντα βρόχο, αποθηκεύονται σε αρχεία εξόδου ο αριθμός των συνολικών αξιολογήσεων που πραγματοποιήθηκαν, οι τιμές που έχει λάβει το \hat{f}_2 , οι τιμές των συναρτήσεων στόχου και οι τιμές των μεταβλητών σχεδιασμού για όλα τα σημεία του Pareto που αποκτήθηκαν.

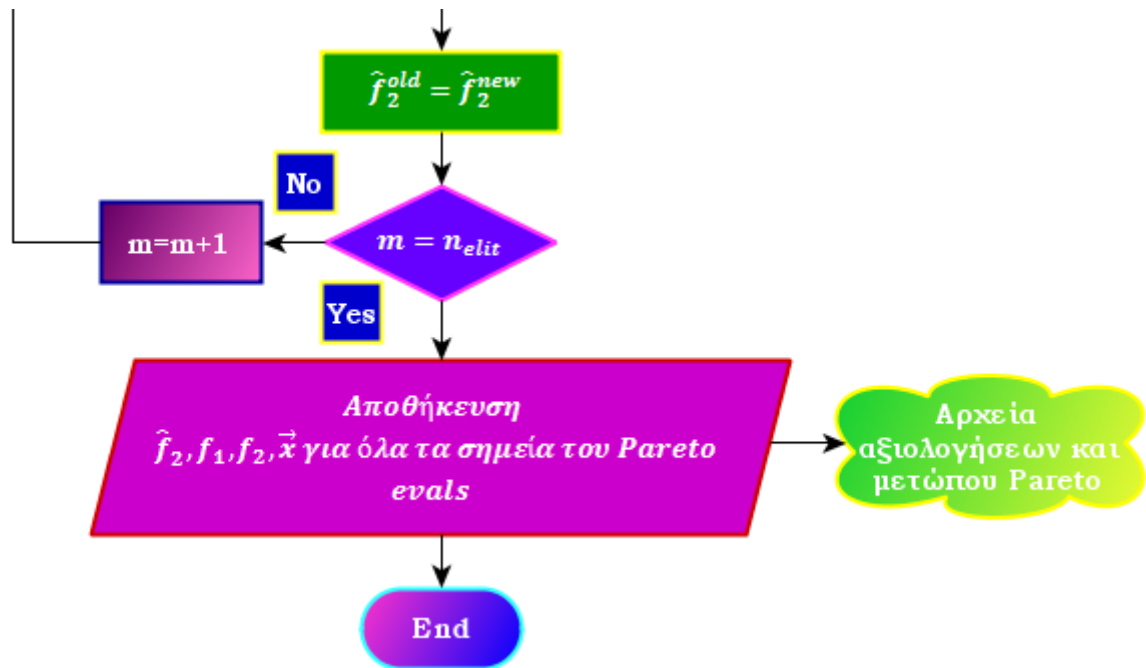
Στο δεδομένο σημείο, κρίνεται σκόπιμο να παρατεθεί το διάγραμμα ροής του ανωτέρω περιγραφέντος αλγορίθμου που αναπτύχθηκε για την επίλυση του προβλήματος που διαχειρίζεται η παρούσα ενότητα (σχήμα 5.1).







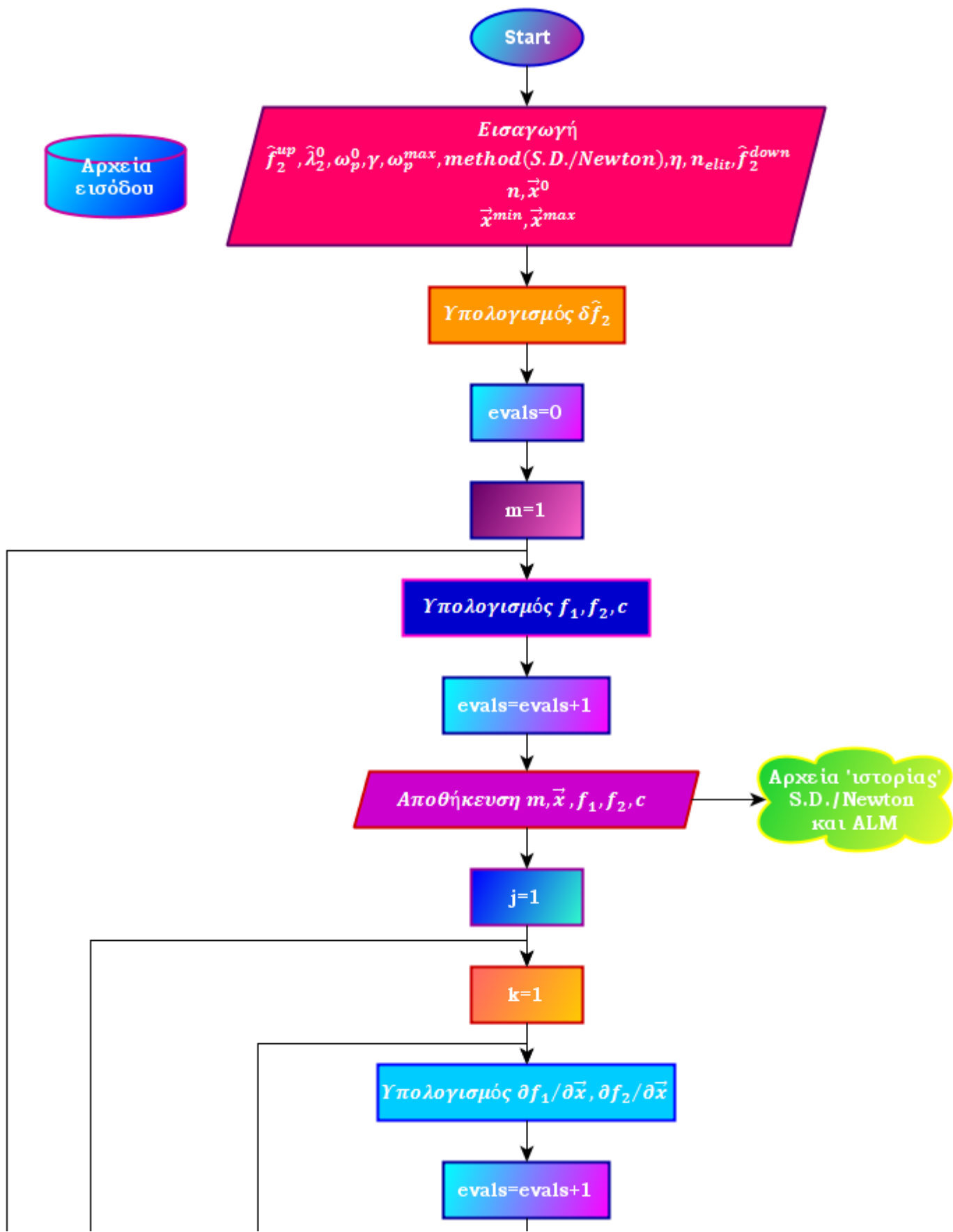


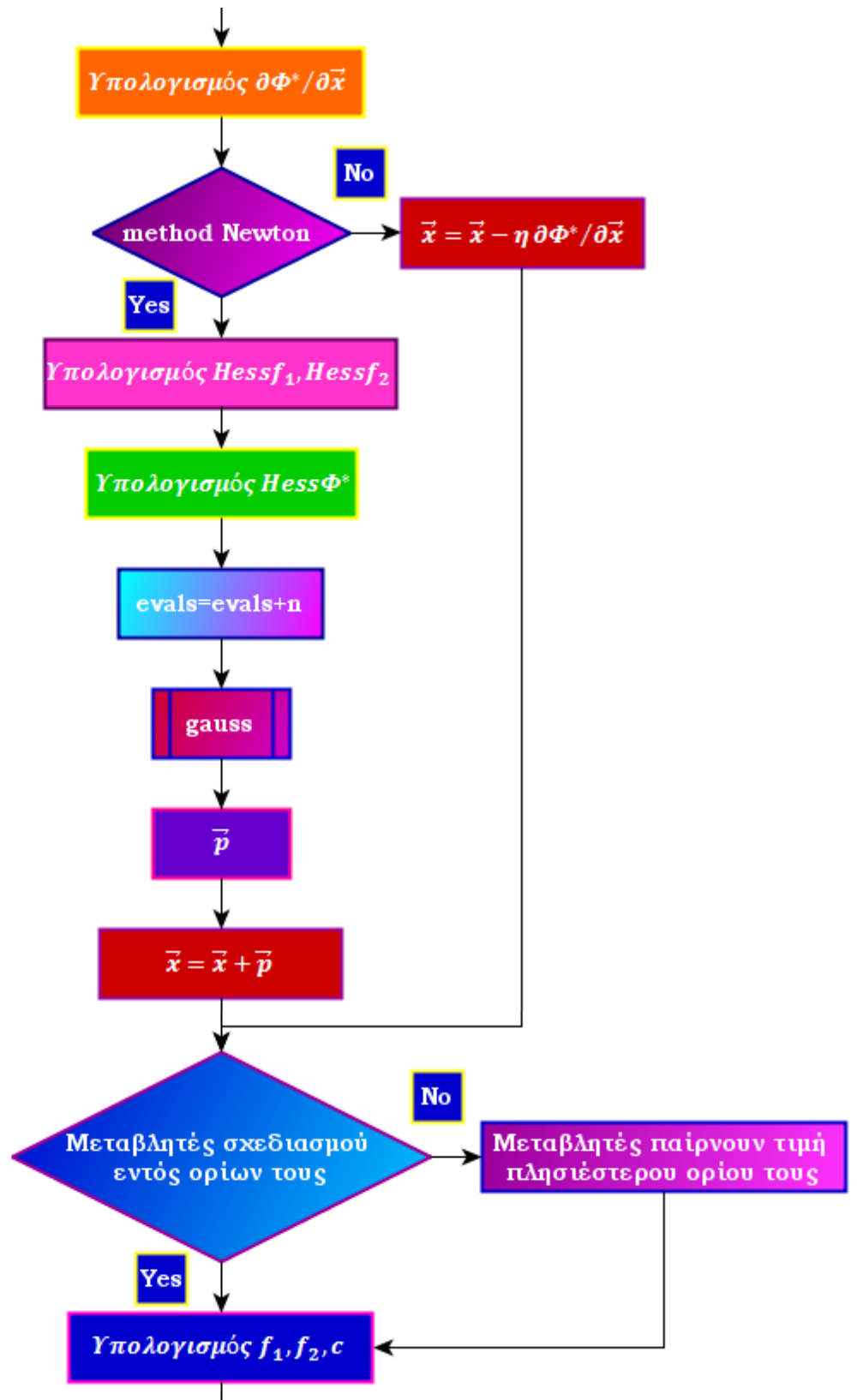


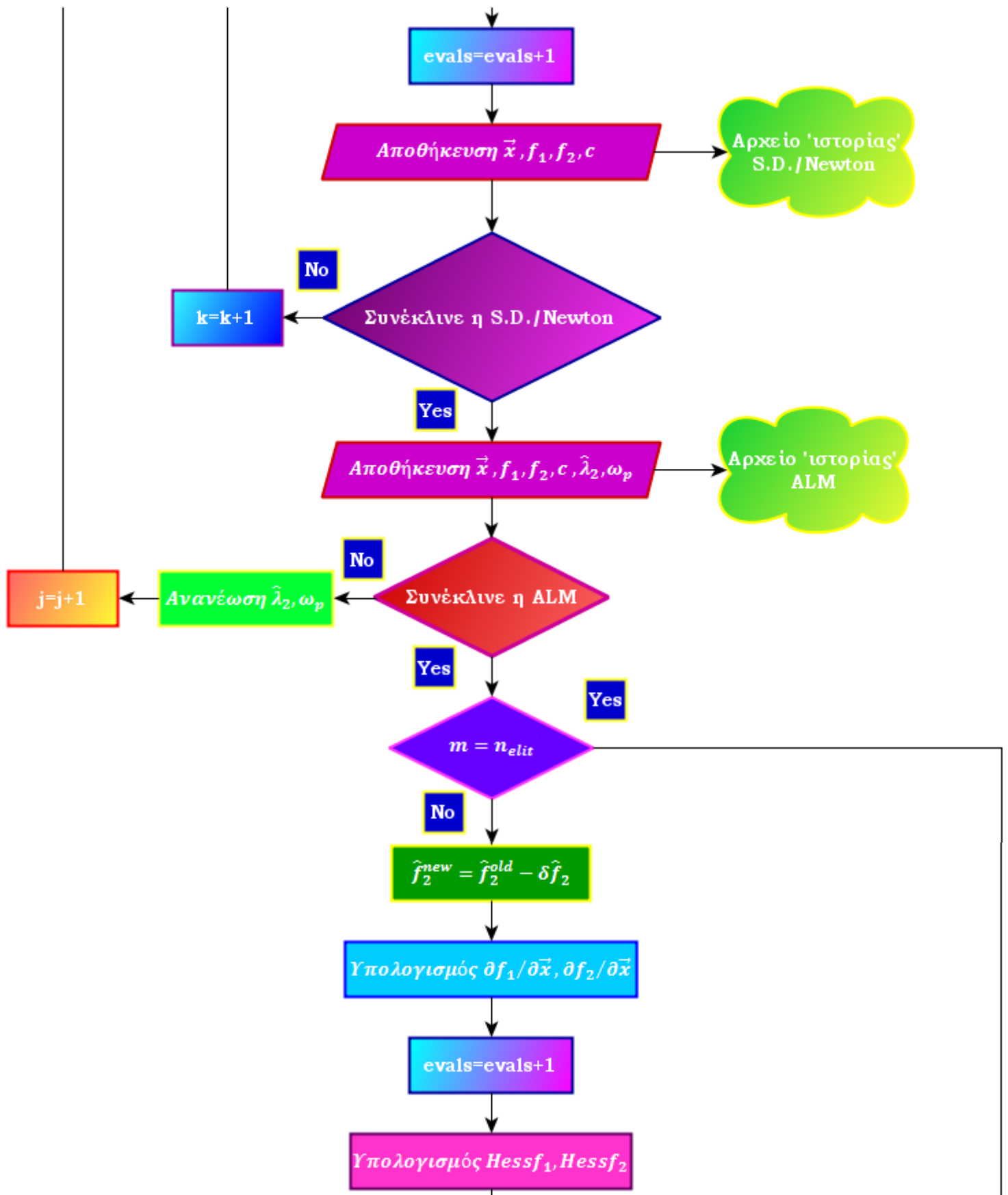
Σχήμα 5.1: Διάγραμμα ροής του αναπτυχθέντος αλγορίθμου για την επίλυση προβλημάτων βελτιστοποίησης δύο στόχων με τη μέθοδο που βασίζεται στη συνέχιση του μετώπου Pareto χωρίς το στάδιο της διόρθωσης.

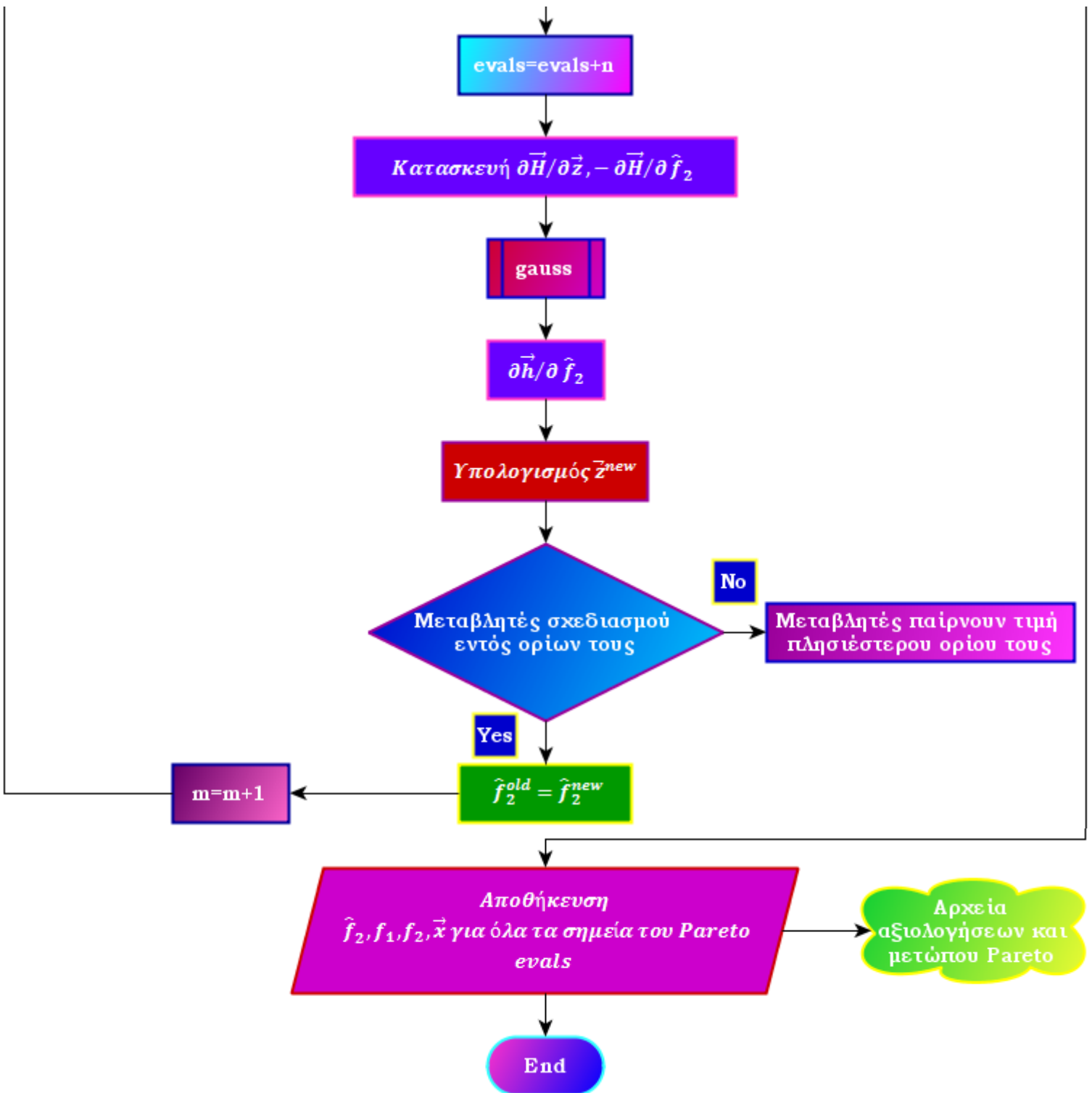
5.3.3.2 Μέθοδος πρόβλεψης-διόρθωσης

Για να υλοποιηθεί η μέθοδος πρόβλεψης-διόρθωσης συνεγράφη ένας νέος κώδικας σε γλώσσα FORTRAN 77, ανάλογος με αυτόν που αναλύθηκε στην ενότητα 5.3.3.1. Η διαφορά με προηγουμένως είναι ότι πλέον υπάρχει ένας εξωτερικός βρόχος που αφορά στα σημεία του μετώπου Pareto που απαιτείται να αποκτηθούν. Ο βρόχος αυτός περιλαμβάνει το τμήμα της ALM και αυτό της πρόβλεψης των μη-κυριαρχούμενων λύσεων (όπως αυτά περιγράφηκαν παραπάνω). Αυτό σημαίνει ότι το πρόγραμμα προβαίνει στην εύρεση του πρώτου σημείου του μετώπου Pareto μέσω της ALM, προχωράει στην πρόβλεψη του δεύτερου σημείου, το διορθώνει με χρήση εκ νέου της ALM (πλέον το $\hat{\lambda}_2^0$ θα έχει την τιμή του $\hat{\lambda}_2$ που προέκυψε από την πρόβλεψη του εν λόγω σημείου και το ω_p^0 θα έχει την τιμή του ω_p για το οποίο συνέκλινε η ALM στο προηγούμενο σημείο), προβλέπει το τρίτο σημείο κ.ο.κ. μέχρι να διορθώσει και το τελευταίο σημείο που ζητήθηκε οπότε εξέρχεται από το βρόχο και γράφει τα αποτελέσματα. Η περιγραφή των επιμέρους τμημάτων του εν λόγω βρόχου πραγματοποιήθηκε αναλυτικά στην προηγούμενη ενότητα, οπότε στο παρόν σημείο κρίνεται σκόπιμο να παρατεθεί μόνο το διάγραμμα ροής του αλγορίθμου ώστε να καταδειχθούν οι διαφορές μεταξύ των δύο περιπτώσεων (σχήμα 5.2).









Σχήμα 5.2: Διάγραμμα ροής του αναπτυχθέντος αλγορίθμου για την επίλυση προβλημάτων βελτιστοποίησης δύο στόχων με τη μέθοδο πρόβλεψης-διόρθωσης.

5.3.4 Παρουσίαση αποτελεσμάτων

5.3.4.1 Περίπτωση αφαίρεσης σταδίου διόρθωσης

Αρχικά, όπως αναφέρθηκε παραπάνω, ο σχεδιαστής, έχοντας καλή γνώση του προβλήματος προς επίλυση και άρα των ορίων των αποκρίσεων που αναμένει (ή επιθυμεί να αποκτήσει), καλείται να λάβει μια απόφαση σχετικά με την αρχική εκτίμηση για το \hat{f}_2 (η οποία, όπως σχολιάσθηκε παραπάνω, θα αποτελεί το άνω όριο του μετώπου Pareto στον άξονα των f_2). Έστω, λοιπόν, ότι σε πρώτη φάση τροφοδοτείται στο πρόγραμμα ως αρχική εκτίμηση για το \hat{f}_2 η τιμή 15. Με δεδομένη την τιμή του \hat{f}_2^{up} , το πρόγραμμα απαιτεί κατάλληλες τιμές των παραμέτρων που σχετίζονται με την ALM (παρουσιάστηκαν στην ενότητα που προηγήθηκε) ώστε να συγκλίνει στο πρώτο σημείο του Pareto. Είναι ευνόητο ότι ανάλογα με τις παραμέτρους που επιλέγονται, το πρόγραμμα εκτελεί περισσότερες ή λιγότερες επαναλήψεις για τη σύγκλιση της ALM στο πρώτο σημείο του Pareto, ενώ η επιλογή αυτή έχει άμεσο αντίκτυπο στον υπολογισμό των αξιολογήσεων που πραγματοποιούνται στο εν λόγω τμήμα του προγράμματος και κατ' επέκταση στο σύνολό του. Μία αρχική επιλογή των παραμέτρων που σχετίζονται με την ALM δύναται να είναι η εξής: $\lambda_2^0 = 0$ (συνηθίζεται η αρχική τιμή του πολλαπλασιαστή Lagrange να δίνεται 0), $\omega_p^0 = 0,001$ (συνηθίζεται η αρχική τιμή του συντελεστή ποινής να είναι μικρή), $\gamma = 1,5$, $\omega_p^{max} = 500$, $method = 1$ (δηλαδή επιλέγεται ως μέθοδος βελτιστοποίησης για προβλήματα ενός στόχου χωρίς περιορισμούς η μέθοδος της απότομης καθόδου) και $\eta = 0,01$ (σταθερό σε κάθε επανάληψη της μεθόδου της απότομης καθόδου). Επιπλέον, έστω ότι στην αρχική αυτή φάση επιλέγεται $n_{elit} = 30$ (ο αριθμός των σημείων που είναι επιθυμητό να αποκτηθούν πάνω στο μέτωπο Pareto) και $\hat{f}_2^{down} = 2$ (εκτίμηση για το \hat{f}_2 η οποία αφορά στο κάτω όριο του μετώπου Pareto στον άξονα των f_2). Τέλος, ως σημείο εκκίνησης \tilde{x}^0 επιλέγεται το διάνυσμα $[10 \ 5]^T$.

Ένα επιπλέον ζήτημα που τίθεται είναι η επιλογή της επιθυμητής ανοχής στην ικανοποίηση των κριτηρίων σύγκλισης της μεθόδου βελτιστοποίησης προβλημάτων ενός στόχου χωρίς περιορισμούς και της ALM. Στο παρόν σημείο, κρίνεται σκόπιμο να παρατεθούν τα αποτελέσματα του προγράμματος, για τις προαναφερθείσες τιμές των παραμέτρων, στο πρώτο τμήμα του όπου επιτυγχάνεται η εύρεση του πρώτου σημείου του μετώπου Pareto για διάφορες ανοχές. Οι τιμές των μεταβλητών σχεδιασμού και των συναρτήσεων στόχου για το πρώτο σημείο του μετώπου Pareto (στρογγυλοποιημένες στο έκτο δεκαδικό ψηφίο) για διάφορες επιθυμητές ανοχές συνοδευόμενες από τις συνολικές αξιολογήσεις που πραγματοποιούνται από το πρόγραμμα (δηλαδή συνυπολογίζονται και οι αξιολογήσεις που γίνονται κατά τη διάρκεια της εύρεσης των υπολοίπων σημείων του Pareto) σε κάθε περίπτωση φαίνονται στον πίνακα 5.1.

Ανοχή	x_1	x_2	f_1	f_2	Αξιολογήσεις
1,00E-06	0,609406	2,994614	0,371376	15,000000	86567
1,00E-05	0,609406	2,994664	0,371376	15,000001	67517
1,00E-04	0,609406	2,995201	0,371375	15,000016	48505
1,00E-03	0,609403	3,000908	0,371372	15,000138	29619
1,00E-02	0,609611	3,058788	0,371626	15,001286	12127
1,00E-01	0,628105	3,551837	0,394516	15,057240	2525

Πίνακας 5.1: Πρώτη μη-κυριαρχούμενη λύση και συνολικές αξιολογήσεις που πραγματοποιούνται για την εύρεση του μετώπου Pareto για διαφορετικές τιμές των ανοχών με τις οποίες ικανοποιούνται τα κριτήρια σύγκλισης της μεθόδου της απότομης καθόδου και της ALM όταν $f_2^{up} = 15$, $\vec{x}^0 = [10 \ 5]^T$, $\hat{\lambda}_2^0 = 0$, $\omega_p^0 = 0,001$, $\gamma = 1,5$, $\omega_p^{max} = 500$, $\eta = 0,01$ και $n_{elit} = 30$.

Η επιλογή της επιθυμητής ανοχής εξαρτάται σαφέστατα από τις απαιτήσεις ακρίβειας που επιβάλλει το πρόβλημα, ενώ είναι ευνόητο ότι όσο μεγαλύτερη ακρίβεια απαιτείται τόσο περισσότερες αξιολογήσεις πρέπει να πραγματοποιηθούν. Στη δεδομένη περίπτωση, έστω ότι ως ικανοποιητική ανοχή για τα δύο κριτήρια σύγκλισης (μετά από μελέτη των δεδομένων του πίνακα 5.1) θεωρείται η τιμή 10^{-3} .

Με τη δεδομένη επιλογή για την ανοχή αλλά και για τις ανωτέρω αναφερθείσες παραμέτρους, όπως φαίνεται στον ανωτέρω πίνακα, οι συνολικές αξιολογήσεις που χρειάζονται για την απόκτηση του μετώπου Pareto είναι 29619. Είναι φανερό ότι πρόκειται για ένα πολύ μεγάλο αριθμό που επιβάλλεται να μειωθεί σημαντικά. Για να εξετασθεί αν δύναται να επιλυθεί το πρόβλημα με λιγότερες αξιολογήσεις είναι αναγκαίο να μεταβληθούν οι σχετιζόμενες με την ALM παράμετροι. Στο σημείο αυτό, λοιπόν, κρίθηκε απαραίτητη η πραγματοποίηση μίας διερεύνησης με στόχο να επιδειχθεί το πώς συμπεριφέρεται το πρόγραμμα όσον αφορά το υπολογιστικό κόστος με τη μεταβολή των παραμέτρων της ALM αλλά και του σημείου εκκίνησης \vec{x}^0 . Πρέπει να σημειωθεί ότι, καθ' όλη τη διάρκεια της δεδομένης διερεύνησης, οι αξιολογήσεις που πραγματοποιούνται στο δεύτερο τμήμα του προγράμματος που αφορά στη συνέχιση του μετώπου των μη-κυριαρχούμενων λύσεων είναι σταθερές και ίσες με $(n_{elit} - 1)(1 + n + 1) = 116$. Στον πίνακα 5.2 επιδεικνύονται τα αποτελέσματα της αναφερθείσας διερεύνησης δηλαδή οι συνολικές αξιολογήσεις που εκτελεί το πρόγραμμα για την απόκτηση του μετώπου Pareto για ποικίλες τιμές των παραμέτρων $\hat{\lambda}_2^0$, ω_p^0 , γ , η και για διάφορα σημεία εκκίνησης \vec{x}^0 .

$\hat{\lambda}_2^0$	ω_p^0	γ	η	$(\bar{x}^0)^T$	Αξιολογήσεις
0	0,001	1,5	0,01	[10 5]	29619
0	0,001	1,5	0,02	[10 5]	14863
0	0,001	1,5	0,03	[10 5]	9951
0	0,001	1,5	0,04	[10 5]	δεν συγκλίνει
0	0,001	1,5	0,03	[0,5 0,5]	12215
0	0,001	1,5	0,03	[15 10]	10681
0	0,001	1,5	0,03	[2 2]	10541
0	0,001	1,5	0,03	[5 5]	9927
0	0,001	1,5	0,03	[4 4]	9329
0	0,001	1,5	0,03	[2 4]	9297
0	0,001	1,5	0,03	[1 3]	281
0	0,003	1,5	0,03	[10 5]	4485
0	0,003	1,5	0,04	[10 5]	3439
0	0,003	1,5	0,05	[10 5]	δεν συγκλίνει
0	0,004	1,5	0,03	[10 5]	3703
0	0,004	1,5	0,04	[10 5]	δεν συγκλίνει
0	0,01	1,5	0,02	[10 5]	4593
0	0,01	1,5	0,03	[10 5]	δεν συγκλίνει
0	0,02	1,5	0,02	[10 5]	4267
0	0,03	1,5	0,01	[10 5]	8113
0	0,03	1,5	0,02	[10 5]	δεν συγκλίνει
0	0,04	1,5	0,01	[10 5]	7991
0	0,05	1,5	0,01	[10 5]	δεν συγκλίνει
0	0,003	1,5	0,04	[2 4]	3151
0	0,003	1,1	0,04	[10 5]	3447
0	0,003	1,1	0,05	[10 5]	2779
0	0,003	1,1	0,1	[10 5]	1473
0	0,003	1,1	0,12	[10 5]	1303
0	0,003	1,1	0,13	[10 5]	δεν συγκλίνει
0	0,003	1,1	0,12	[2 4]	1209
0	0,003	1,01	0,12	[10 5]	1265
0	0,003	1,01	0,17	[10 5]	1071
0	0,003	1,01	0,19	[10 5]	δεν συγκλίνει
0	0,003	1,01	0,17	[2 4]	1003
2	0,003	1,01	0,16	[10 5]	707
2	0,003	1,01	0,17	[10 5]	717
4	0,003	1,01	0,14	[10 5]	583
4	0,003	1,01	0,15	[10 5]	585
4	0,003	1,01	0,14	[2 4]	579

Πίνακας 5.2: Συνολικές αξιολογήσεις που πραγματοποιούνται για την εύρεση του μετώπου Pareto για διαφορετικές τιμές των παραμέτρων της ALM όταν οι ανοχές για τα κριτήρια σύγκλισης της μεθόδου της απότομης καθόδου και της ALM λαμβάνουν την τιμή 10^{-3} , $\hat{f}_2^{up} = 15$ και $n_{elit} = 30$.

Η προσπάθεια για ελάττωση των αξιολογήσεων που απαιτούνται για τη σύγκλιση της ALM (άρα κατά συνέπεια και του προγράμματος στο σύνολό του), όπως αυτή απεικονίζεται στον πίνακα 5.2, ξεκίνησε με σταδιακή αύξηση του η , ενώ οι τιμές $\hat{\lambda}_2^0$, ω_p^0 και γ διατηρήθηκαν σταθερές. Έγινε, όμως, αντιληπτό ότι το η δεν ήταν δυνατό να αυξηθεί κατά πολύ, καθώς η μέθοδος δεν συνέκλινε. Αργότερα, έγινε απόπειρα αύξησης του ω_p^0 , με τα καλύτερα αποτελέσματα να παρουσιάζονται για $\omega_p^0 = 0,003$. Στη συνέχεια, πραγματοποιήθηκε μείωση του γ , με τη γρηγορότερη σύγκλιση να επιτυγχάνεται στην περίπτωση όπου $\gamma = 1,01$. Σημειώνεται ότι με τη μείωση του γ επιτράπη η περαιτέρω αύξηση του η χωρίς να αποκλίνει η μέθοδος. Τέλος, επιχειρήθηκε η αύξηση του $\hat{\lambda}_2^0$ με την περίπτωση $\hat{\lambda}_2^0 = 4$ να δίνει τα γρηγορότερα αποτελέσματα. Ο καλύτερος συνδυασμός παραμέτρων που επιτεύχθη, όπως φαίνεται από τον ανωτέρω πίνακα, είναι $\hat{\lambda}_2^0 = 4$, $\omega_p^0 = 0,003$, $\gamma = 1,01$ και $\eta = 0,14$. Αν, λ.χ. επιλεγεί $\vec{x}^0 = [2 \ 4]^T$ οι συνολικές αξιολογήσεις που απαιτούνται για την εύρεση του μετώπου Pareto είναι 579. Στην περίπτωση αυτή, μετά τη σύγκλιση της ALM λαμβάνεται το πρώτο σημείο του Pareto για το οποίο ισχύουν (με στρογγυλοποίηση στα 6 δεκαδικά ψηφία) $x_1 = 0,609397$, $x_2 = 2,995000$, $f_1 = 0,371365$ και $f_2 = 15,000216$.

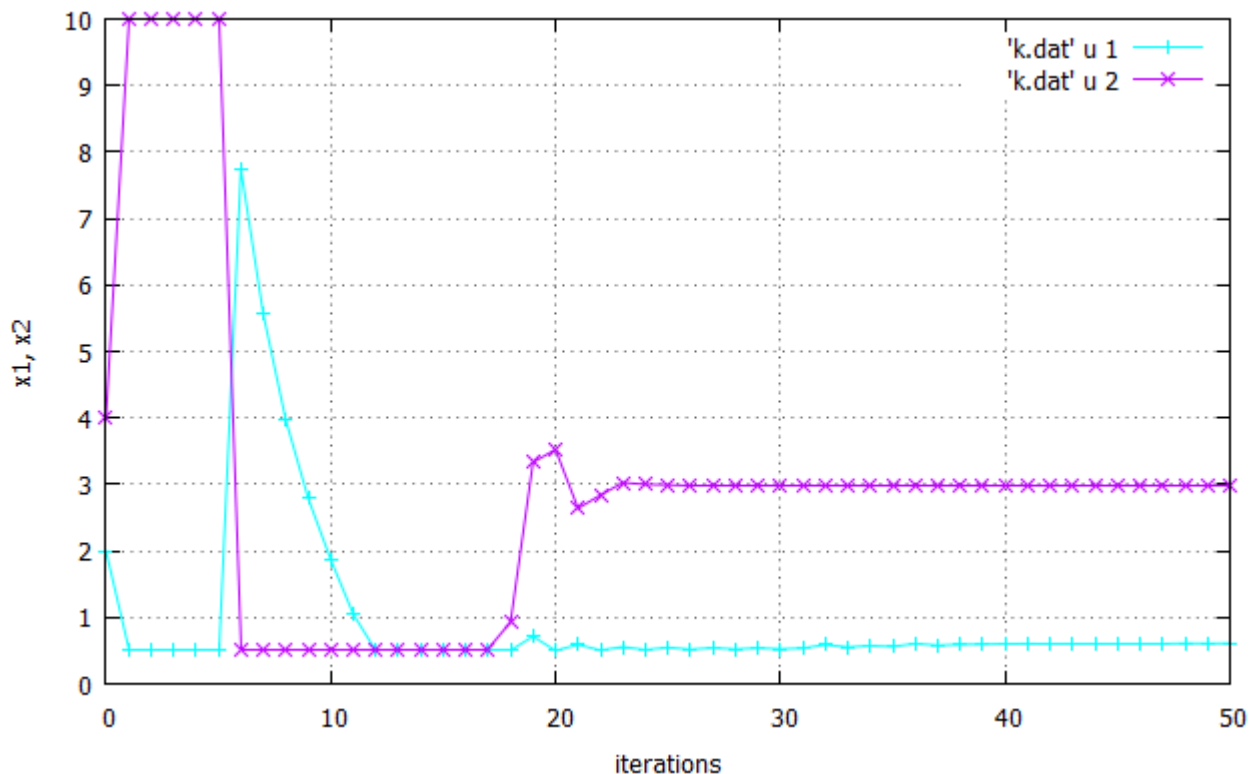
Στην προσπάθεια περαιτέρω μείωσης των αξιολογήσεων, εξετάζεται η περίπτωση το κριτήριο σύγκλισης της απότομης καθόδου να ικανοποιείται με επιθυμητή ανοχή 10^{-2} , ενώ αυτό της ALM συνεχίζει με 10^{-3} . Στην προαναφερθείσα περίπτωση, για το πρώτο σημείο του Pareto ισχύουν $x_1 = 0,609435$, $x_2 = 3,003142$, $f_1 = 0,371411$ και $f_2 = 14,999409$, ενώ οι συνολικές αξιολογήσεις που πραγματοποιούνται από το πρόγραμμα είναι 419. Είναι φανερό ότι έχει επιτευχθεί μείωση των αξιολογήσεων, ενώ τα αποτελέσματα για το πρώτο σημείο του Pareto είναι πρακτικά ίδια. Αν η επιθυμητή ανοχή στην περίπτωση του κριτηρίου σύγκλισης της μεθόδου της απότομης καθόδου μειωθεί περαιτέρω σε 10^{-1} προκύπτουν για το πρώτο σημείο του Pareto $x_1 = 0,610246$, $x_2 = 3,107591$, $f_1 = 0,372401$ και $f_2 = 14,999127$, ενώ οι συνολικές αξιολογήσεις φτάνουν τις 303. Τα αποτελέσματα και σε αυτήν την περίπτωση είναι πολύ καλά με τις συνολικές αξιολογήσεις να έχουν μειωθεί σημαντικά.

Στο σημείο αυτό κρίνεται απαραίτητο να αναφερθεί ότι κατά τη διάρκεια της εκτέλεσης του προγράμματος, παρατηρήθηκε ότι η μεταβλητή σχεδιασμού x_2 μεταβάλλεται με πολύ πιο αργούς ρυθμούς από ότι η x_1 . Για το λόγο αυτό, δοκιμάστηκε ένα 'τέχνασμα' το οποίο είναι η χρήση διαφορετικού η (προφανώς μεγαλύτερο) για τη μεταβλητή x_2 . Τα αποτελέσματα που προκύπτουν από το εν λόγω 'τέχνασμα' παρουσιάζονται στον πίνακα 5.3. Σημειώνεται ότι για όλες τις περιπτώσεις που απεικονίζονται στον εν λόγω πίνακα έχουν ληφθεί $\hat{\lambda}_2^0 = 4$, $\omega_p^0 = 0,003$, $\gamma = 1,01$, $\eta_1 = 0,14$ και $\vec{x}^0 = [2 \ 4]^T$, ενώ η επιθυμητή ανοχή με την οποία ικανοποιείται το κριτήριο σύγκλισης της ALM είναι 10^{-3} .

Ανοχή S.D.	η_2	x_1	x_2	f_1	f_2	Αξιολογήσεις
1,00E-03	10	0,609389	2,994609	0,371355	15,000429	317
1,00E-02	10	0,609413	2,994609	0,371385	14,999826	237
1,00E-01	9	0,609378	2,994609	0,371341	15,000698	217

Πίνακας 5.3: Πρώτη μη-κυριαρχούμενη λύση και συνολικές αξιολογήσεις που πραγματοποιούνται για την εύρεση του μετώπου Pareto για διαφορετικές τιμές της ανοχής με την οποία ικανοποιείται το κριτήριο σύγκλισης της μεθόδου της απότομης καθόδου όταν η ανοχή για την ALM παίρνει την τιμή 10^{-3} , $\hat{f}_2^{up} = 15$, $\bar{x}^0 = [2 \ 4]^T$, $\hat{\lambda}_2^0 = 4$, $\omega_p^0 = 0,003$, $\gamma = 1,01$, $\omega_p^{max} = 500$, $\eta_1 = 0,14$ και $n_{elit} = 30$.

Στο σημείο αυτό, αξίζει να αναφερθεί ότι στην τελευταία περίπτωση που φαίνεται στον πίνακα 5.3, για να συγκλίνει η ALM πραγματοποιήθηκαν 28 κύκλοι, ενώ η μέθοδος της απότομης καθόδου 'κλήθηκε' 50 φορές συνολικά σε αυτούς τους 28 κύκλους. Σημειώνεται ότι η ALM, στην εν λόγω περίπτωση, πραγματοποιεί $1 + 50(1 + 1) = 101$ αξιολογήσεις για την εύρεση του πρώτου σημείου του Pareto και 116 αξιολογήσεις απαιτούνται για την εύρεση των υπόλοιπων 29 σημείων που ζητήθηκαν. Επίσης, πρέπει να ειπωθεί ότι η ALM συνέκλινε για $\hat{\lambda}_2 = 0,049480$ και $\omega_p = 0,003925$. Στο σχήμα 5.3 παριστάνεται η πορεία που ακολούθησαν οι μεταβλητές σχεδιασμού κατά τη διάρκεια της εκτέλεσης των 50 επαναλήψεων της μεθόδου της απότομης καθόδου 'μέσα' στην ALM.



Σχήμα 5.3: Διάγραμμα των μεταβλητών σχεδιασμού x_1 (γαλάζια καμπύλη) και x_2 (μωβ καμπύλη) συναρτήσει των επαναλήψεων που εκτελέστηκαν από τη μέθοδο της απότομης καθόδου 'μέσα' στην ALM μέχρι την εύρεση του

πρώτου σημείου του μετώπου Pareto όταν οι ανοχές με τις οποίες ικανοποιούνται τα κριτήρια σύγκλισης της μεθόδου της απότομης καθόδου και της ALM είναι 10^{-1} και 10^{-3} αντίστοιχα, $\hat{f}_2^{up} = 15$, $\vec{x}^0 = [2 \ 4]^T$, $\hat{\lambda}_2^0 = 4$, $\omega_p^0 = 0,003$, $\gamma = 1,01$, $\omega_p^{max} = 500$, $\eta_1 = 0,14$ και $\eta_2 = 9$.

Ακολουθώς, κρίνεται σκόπιμο να επιχειρηθεί η εύρεση του πρώτου σημείου του μετώπου Pareto με τη χρήση της μεθόδου Newton (*method* = 2) αντί της μεθόδου της απότομης καθόδου στο τμήμα της ALM όπου πραγματοποιείται η βελτιστοποίηση ενός μονοκριτηριακού προβλήματος χωρίς περιορισμούς. Όπως και στην περίπτωση χρήσης της μεθόδου της απότομης καθόδου, σε πρώτη φάση, επιλέγονται $\hat{f}_2^{up} = 15$, $\omega_p^{max} = 500$, $n_{elit} = 30$ και $\hat{f}_2^{down} = 2$. Επίσης, επιλέγεται ως επιθυμητή ανοχή για τα δύο κριτήρια σύγκλισης η τιμή 10^{-3} . Όπως έγινε και στην περίπτωση επιλογής της μεθόδου της απότομης καθόδου, έλαβε χώρα μία διερεύνηση με στόχο να επιδειχθεί το πώς συμπεριφέρεται το πρόγραμμα όσον αφορά το υπολογιστικό κόστος με τη μεταβολή των παραμέτρων της ALM αλλά και του σημείου εκκίνησης \vec{x}^0 . Πρέπει να σημειωθεί ότι, όπως και προηγουμένως, καθ' όλη τη διάρκεια της δεδομένης διερεύνησης, οι αξιολογήσεις που πραγματοποιούνται στο δεύτερο τμήμα του προγράμματος που αφορά στη συνέχιση του μετώπου των μη-κυριαρχούμενων λύσεων είναι σταθερές και ίσες με $(n_{elit} - 1)(1 + n + 1) = 116$. Στον πίνακα 5.4 φαίνονται οι συνολικές αξιολογήσεις που εκτελεί το πρόγραμμα για την απόκτηση του μετώπου Pareto για διάφορες τιμές των παραμέτρων $\hat{\lambda}_2^0$, ω_p^0 , και γ , ενώ γίνεται επιλογή και διαφορετικών σημείων εκκίνησης \vec{x}^0 .

$\hat{\lambda}_2^0$	ω_p^0	γ	$(\bar{x}^0)^T$	Αξιολογήσεις
0	0,001	1,5	[10 5]	δεν συγκλίνει
0	0,002	1,5	[10 5]	221
0	0,004	1,5	[10 5]	205
0	0,006	1,5	[10 5]	197
0	0,008	1,5	[10 5]	201
0	0,01	1,5	[10 5]	189
0	0,04	1,5	[10 5]	185
0	0,06	1,5	[10 5]	173
0	0,1	1,5	[10 5]	173
0	0,2	1,5	[10 5]	177
0	0,3	1,5	[10 5]	173
0	0,8	1,5	[10 5]	181
0	1,2	1,5	[10 5]	225
0	0,3	1,01	[10 5]	181
0	0,3	1,1	[10 5]	181
0	0,3	2	[10 5]	173
0	0,3	10	[10 5]	173
0,1	0,3	1,5	[10 5]	181
1	0,3	1,5	[10 5]	193
0	0,01	1,5	[2 4]	185
0	0,4	1,5	[2 4]	173
0	0,4	1,5	[0,5 0,5]	173

Πίνακας 5.4: Συνολικές αξιολογήσεις που πραγματοποιούνται για την εύρεση του μετώπου Pareto για διαφορετικές τιμές των παραμέτρων της ALM όταν οι ανοχές για τα κριτήρια σύγκλισης της μεθόδου Newton και της ALM λαμβάνουν την τιμή 10^{-3} , $\hat{f}_2^{up} = 15$ και $n_{elit} = 30$.

Η διερεύνηση που απεικονίζεται στον πίνακα 5.4 ξεκίνησε με σταδιακή αύξηση του ω_p^0 , ενώ οι τιμές $\hat{\lambda}_2^0$ και γ διατηρήθηκαν σταθερές. Διαπιστώθηκε ότι, αρχικά, καθώς αυξανόταν η τιμή του ω_p^0 , ο αριθμός των απαιτούμενων αξιολογήσεων μειωνόταν, ενώ από ένα σημείο και μετά η αύξηση του ω_p^0 οδηγούσε σε αύξηση των αξιολογήσεων που πραγματοποιούσε η ALM (άρα και το πρόγραμμα συνολικά). Τα ταχύτερα αποτελέσματα, στην περίπτωση που το σημείο εκκίνησης είναι το $[10\ 5]^T$, φαίνεται να επιτυγχάνονται όταν το ω_p^0 κυμαίνεται μεταξύ των τιμών 0,06 και 0,1 αλλά και όταν $\omega_p^0 = 0,3$. Στη συνέχεια, πραγματοποιήθηκε μείωση και αύξηση του γ . Παρατηρήθηκε ότι η μείωση του γ οδηγεί σε μικρή αύξηση των αξιολογήσεων, ενώ η αύξηση του γ αφήνει τις αξιολογήσεις αμετάβλητες. Επιχειρήθηκε, έπειτα, η αύξηση του $\hat{\lambda}_2^0$. Οποιαδήποτε αύξηση του $\hat{\lambda}_2^0$ οδηγεί σε αύξηση των αξιολογήσεων που απαιτούνται. Τέλος, εξετάστηκε και η περίπτωση εκκίνησης από διαφορετικό σημείο. Κρίνεται αναγκαίο να τονιστεί ότι στην περίπτωση χρήσης της μεθόδου Newton, όπως δηλώνουν τα αποτελέσματα που φαίνονται

στον ανωτέρω πίνακα, μία αλλαγή σε οιαδήποτε από τις παραμέτρους που αναφέρθηκαν δεν επιφέρει μεγάλη αυξομείωση στον αριθμό των αξιολογήσεων που πραγματοποιούνται, σε αντίθεση με την περίπτωση επιλογής της μεθόδου της απότομης καθόδου όπου παρατηρήθηκαν εξαιρετικά σημαντικές αυξομειώσεις.

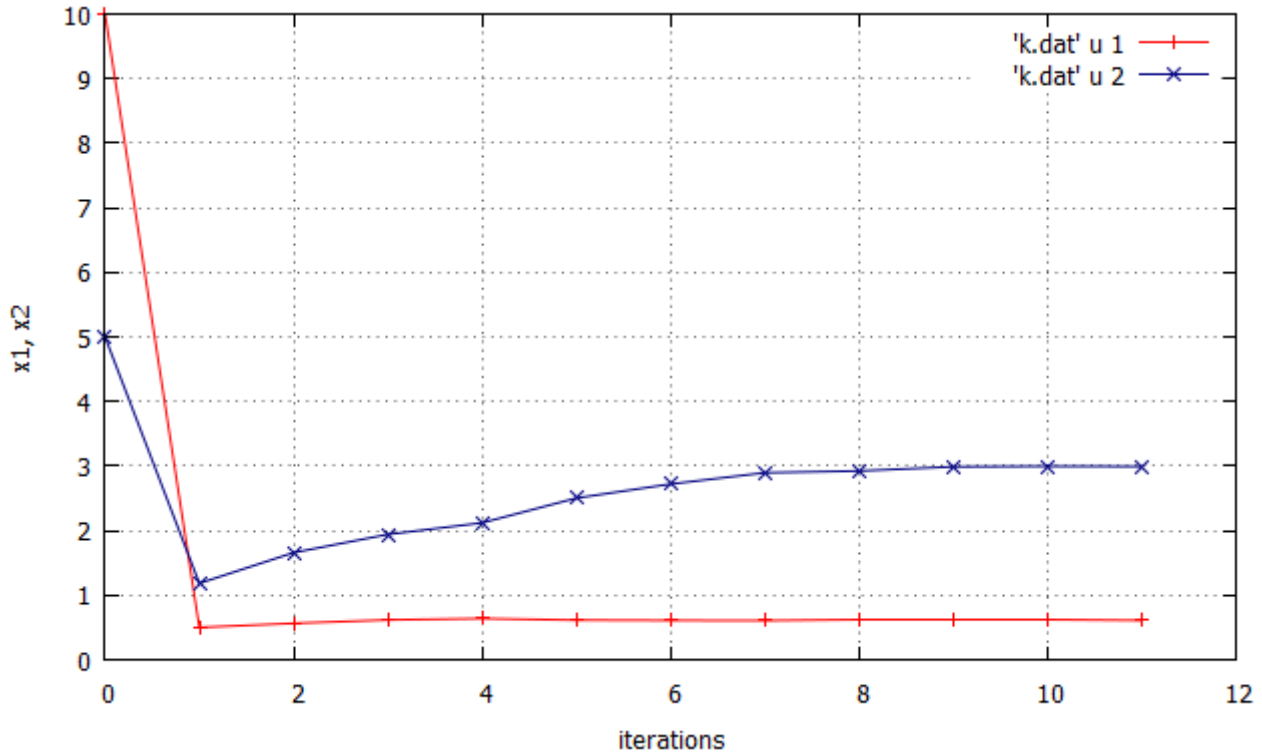
Έστω ότι από τους συνδυασμούς παραμέτρων που παρουσιάζονται στον πίνακα 5.4 επιλέγεται αυτός για τον οποίο ισχύουν $\hat{\lambda}_2^0 = 0$, $\omega_p^0 = 0,3$, $\gamma = 1,5$ και $\vec{x}^0 = [10 \ 5]^T$. Τότε, όπως καθίσταται φανερό από τον εν λόγω πίνακα, οι συνολικές αξιολογήσεις που πραγματοποιούνται για την εύρεση του μετώπου Pareto είναι 173 και άρα οι αξιολογήσεις που απαιτεί η ALM για να συγκλίνει είναι 57. Στην εν λόγω περίπτωση, για το πρώτο σημείο του Pareto που λαμβάνεται μετά τη σύγκλιση της ALM ισχύουν $x_1 = 0,609371$, $x_2 = 2,994609$, $f_1 = 0,371333$ και $f_2 = 15,000874$. Όπως και προηγουμένως στην περίπτωση επιλογής της μεθόδου της απότομης καθόδου, εξετάζεται η περίπτωση το κριτήριο σύγκλισης της μεθόδου Newton να ικανοποιείται με επιθυμητές ανοχές που είναι μικρότερες από 10^{-3} , ενώ αυτό της ALM συνεχίζει με 10^{-3} . Τα αποτελέσματα που προκύπτουν (λαμβάνοντας το συνδυασμό παραμέτρων που αναφέρεται στην αρχή της δεδομένης παραγράφου) παρουσιάζονται στον πίνακα 5.5.

Ανοχή Newton	x_1	x_2	f_1	f_2	Αξιολογήσεις
1,00E-03	0,609371	2,994609	0,371333	15,000874	173
1,00E-02	0,609371	2,994609	0,371333	15,000874	173
1,00E-01	0,609371	2,994608	0,371333	15,000870	165
1,00E+00	0,609410	2,993906	0,371381	14,999908	161

Πίνακας 5.5: Πρώτη μη-κυριαρχούμενη λύση και συνολικές αξιολογήσεις που πραγματοποιούνται για την εύρεση του μετώπου Pareto για διαφορετικές τιμές της ανοχής με την οποία ικανοποιείται το κριτήριο σύγκλισης της μεθόδου Newton όταν η ανοχή για την ALM παίρνει την τιμή 10^{-3} , $\hat{f}_2^{up} = 15$, $\vec{x}^0 = [10 \ 5]^T$, $\hat{\lambda}_2^0 = 0$, $\omega_p^0 = 0,3$, $\gamma = 1,5$, $\omega_p^{max} = 500$ και $n_{elit} = 30$.

Αναφέρεται ότι με βάση τα αποτελέσματα του πίνακα 5.5 προκύπτει ότι η μείωση της τιμής της επιθυμητής ανοχής με την οποία ικανοποιείται το κριτήριο σύγκλισης της μεθόδου Newton, οδηγεί μεν σε μείωση των αξιολογήσεων αλλά η διαφορά δεν είναι τόσο μεγάλη όσο στην περίπτωση χρήσης της μεθόδου της απότομης καθόδου. Τα αποτελέσματα από άποψη ακρίβειας είναι εξαιρετικά ικανοποιητικά και για τις τέσσερις περιπτώσεις. Αξίζει να αναφερθεί ότι λ.χ. στην τελευταία περίπτωση που φαίνεται στον παραπάνω πίνακα, για να συγκλίνει η ALM πραγματοποιήθηκαν 4 κύκλοι, ενώ η μέθοδος Newton 'έδρασε' 11 φορές σε αυτούς τους 4 κύκλους. Σημειώνεται ότι η ALM, στην εν λόγω περίπτωση, πραγματοποιεί $1 + 11(1 + n + 1) = 45$ αξιολογήσεις για την εύρεση του πρώτου σημείου του Pareto, καθώς, όπως αναφέρθηκε παραπάνω, οι 116 απαιτούνται για την εύρεση των υπόλοιπων 29 σημείων που ζητήθηκαν. Επίσης, πρέπει να ειπωθεί ότι

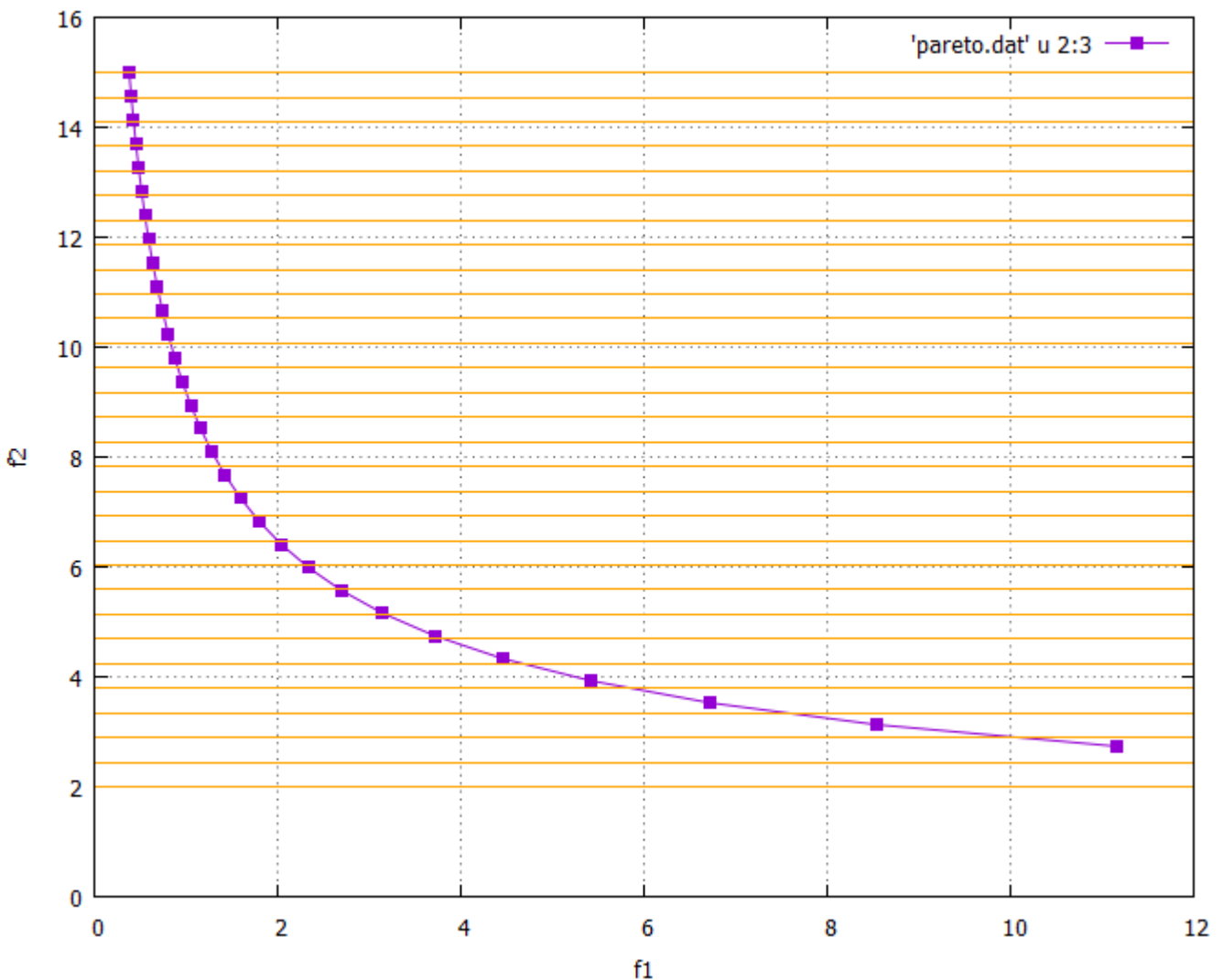
η ALM συνέκλινε για $\hat{\lambda}_2 = 0,049767$ και $\omega_p = 1,012500$. Στο σχήμα 5.4 παριστάνεται η πορεία που ακολούθησαν οι μεταβλητές σχεδιασμού κατά τη διάρκεια της εκτέλεσης των 11 επαναλήψεων της μεθόδου Newton ‘μέσα’ στην ALM.



Σχήμα 5.4: Διάγραμμα των μεταβλητών σχεδιασμού x_1 (κόκκινη καμπύλη) και x_2 (μπλε καμπύλη) συναρτήσει των επαναλήψεων που εκτελέστηκαν από τη μέθοδο Newton ‘μέσα’ στην ALM μέχρι την εύρεση του πρώτου σημείου του μετώπου Pareto όταν οι ανοχές με τις οποίες ικανοποιούνται τα κριτήρια σύγκλισης της μεθόδου Newton και της ALM είναι 1 και 10^{-3} αντίστοιχα, $\hat{f}_2^{up} = 15$, $\vec{x}^0 = [10 \ 5]^T$, $\hat{\lambda}_2^0 = 0$, $\omega_p^0 = 0,3$, $\gamma = 1,5$ και $\omega_p^{max} = 500$.

Στο σημείο αυτό, επιβάλλεται να επισημανθεί ότι αν μελετηθούν οι πίνακες 5.1 έως 5.5, θα διαπιστωθεί ότι με τη χρήση της μεθόδου Newton πραγματοποιούνται λιγότερες αξιολογήσεις και άρα είναι περισσότερο συμφέρουσα από άποψη υπολογιστικού κόστους. Πλεονέκτημα της μεθόδου Newton αποτελεί, επιπλέον, το γεγονός της σχετικής ανεξαρτησίας της από τη μεταβολή των παραμέτρων της ALM, όπως σχολιάσθηκε παραπάνω. Πρέπει, βέβαια, να τονιστεί ότι η μέθοδος Newton χρησιμοποιεί την πληροφορία των δεύτερων παραγώγων, ενώ η μέθοδος της απότομης καθόδου μόνο αυτή των πρώτων παραγώγων. Στη μέθοδο που διαχειρίζεται η παρούσα εργασία, όμως, η γνώση των δεύτερων παραγώγων είναι ούτως ή άλλως απαραίτητη για το τμήμα όπου πραγματοποιείται η συνέχιση του μετώπου Pareto άρα το εν λόγω μειονέκτημα της μεθόδου Newton ‘εξαλείφεται’.

Μετά την απόκτηση του πρώτου σημείου του μετώπου Pareto, το πρόγραμμα προβαίνει στην εύρεση και των υπολοίπων μη-κυριαρχούμενων λύσεων με βάση τις παραμέτρους n_{elit} και \hat{f}_2^{down} που ορίστηκαν παραπάνω (οι τιμές που δόθηκαν ήταν 30 και 2 αντίστοιχα). Το μέτωπο Pareto που προκύπτει, στη συγκεκριμένη περίπτωση, απεικονίζεται στο σχήμα 5.5. Στο εν λόγω σχήμα έχουν σχεδιαστεί, επίσης, και n_{elit} οριζόντιες γραμμές οι οποίες δηλώνουν την τιμή του \hat{f}_2 για καθένα από τα n_{elit} σημεία των οποίων ζητήθηκε η εύρεση. Αυτό έγινε για να δύναται να διαπιστωθεί αν υπάρχει διαφορά (και πόση είναι αυτή) μεταξύ της τιμής του \hat{f}_2 και του f_2 για κάθε σημείο του Pareto που προέκυψε από το πρόγραμμα.



Σχήμα 5.5: Μέτωπο Pareto (μωβ καμπύλη) για $\hat{f}_2^{up} = 15$, $\hat{f}_2^{down} = 2$ και $n_{elit} = 30$. Οι οριζόντιες γραμμές χρώματος πορτοκαλί αντιπροσωπεύουν τις τιμές των \hat{f}_2 για καθένα από τα n_{elit} σημεία που αναζητήθηκαν.

Από την παρατήρηση του σχήματος 5.5, εξάγεται το συμπέρασμα ότι στα πρώτα σημεία που αποκτήθηκαν από το πρόγραμμα οι διαφορές μεταξύ των τιμών του \hat{f}_2 και του f_2 είναι πολύ μικρές (πρακτικά

μπορούν να θεωρηθούν αμελητέες), ενώ καθώς ο αριθμός των υπολογισθέντων σημείων μεγαλώνει οι διαφορές αυτές αυξάνονται και γίνονται όλο και πιο σημαντικές. Αξίζει να αναφερθεί ότι στο δεύτερο αποκτηθέν σημείο ισχύει $f_2 - \hat{f}_2 \cong 0,0129$, στο δέκατο $f_2 - \hat{f}_2 \cong 0,1327$, ενώ στο τελευταίο είναι $f_2 - \hat{f}_2 \cong 0,7359$.

Για να κατασταθεί δυνατή η κατανόηση των αποτελεσμάτων που προέκυψαν, θα πρέπει κανείς να εισχωρήσει στη φιλοσοφία της μεθόδου της συνέχισης του μετώπου Pareto που χρησιμοποιήθηκε. Όπως παρουσιάστηκε στο κεφάλαιο 4, η ανανέωση των τιμών του διανύσματος \vec{z} βασίζεται στο θεώρημα πεπλεγμένης συνάρτησης. Για την ισχύ του εν λόγω θεωρήματος που επιδεικνύεται στην ενότητα 4.2.4 απαιτείται η ικανοποίηση των συνθηκών KKT. Αν αυτή η προϋπόθεση πληρείται, ισχύουν, όπως δηλώθηκε στην προαναφερθείσα ενότητα, τα εξής:

$$\vec{z} = \vec{h}(\hat{f}_2)$$

$$\frac{\partial \vec{h}}{\partial \hat{f}_2}(\hat{f}_2) = - \left\{ \frac{\partial \vec{H}}{\partial \vec{z}}[\vec{h}(\hat{f}_2), \hat{f}_2] \right\}^{-1} \frac{\partial \vec{H}}{\partial \hat{f}_2}[\vec{h}(\hat{f}_2), \hat{f}_2]$$

Ακολουθώντας, με την υπόθεση ότι το διάνυσμα $\vec{z}^k = \vec{h}(\hat{f}_2^k)$ είναι το k -οστό σημείο πάνω στο μέτωπο Pareto, λήφθηκε μία εφαπτομενική προσέγγιση \vec{z}^{k+1} του επόμενου σημείου \vec{z}^{k+1} πάνω στο μέτωπο Pareto μέσω της παρακάτω σχέσης:

$$\vec{z}^{k+1} = \vec{z}^k + \frac{\partial \vec{h}}{\partial \hat{f}_2}(\hat{f}_2^k) (\hat{f}_2^{k+1} - \hat{f}_2^k) \quad (5.1)$$

Στην ανάλυση, όμως, που παρουσιάστηκε στην ενότητα 5.3.3.1 θεωρήθηκε $\vec{z}^{k+1} = \vec{z}^{k+1}$, καθώς δεν πραγματοποιήθηκε διόρθωση.

Το πρόγραμμα, λοιπόν, που συνετέθη, έχοντας ήδη προχωρήσει στην εύρεση της πρώτης μη-κυριαρχούμενης λύσης (άρα ικανοποιούνται οι συνθήκες KKT και κατά συνέπεια ισχύει το θεώρημα πεπλεγμένης συνάρτησης), επιδιώκει την εύρεση του δεύτερου σημείου του μετώπου Pareto και για να την πετύχει κάνει χρήση της σχέσης (5.1).

Πρέπει, όμως, να τονιστεί ότι η συνάρτηση $\vec{z} = \vec{h}(\hat{f}_2)$, γενικά, είναι μη-γραμμική (δηλαδή η γραφική παράσταση $\hat{f}_2 - \vec{z}$ είναι καμπύλη), ενώ το πρόγραμμα μέσω της αναφερθείσας σχέσης επιχειρεί να κινηθεί γραμμικά για την εύρεση των νέων τιμών του διανύσματος \vec{z} , δηλαδή με άλλα λόγια λαμβάνει μία προσέγγιση πρώτης τάξης. Είναι, λοιπόν, φανερό το γεγονός ότι αυτή η πράξη ενέχει ένα σφάλμα, του οποίου το μέγεθος εξαρτάται άμεσα από την επιλογή του $\delta \hat{f}_2$ (είναι λογικό ότι όσο μεγαλύτερο $\delta \hat{f}_2$ επιλεγεί τόσο μεγαλύτερο είναι το σφάλμα). Πρέπει, βέβαια, να προστεθεί ότι το σφάλμα σε κάθε σημείο εξαρτάται και από το πόσο μεγάλη ή μικρή είναι η καμπυλότητα στο σημείο

αυτό, δηλαδή από το κατά πόσο η καμπύλη $\hat{f}_2 - \bar{z}$ διαφέρει από την ευθεία. Όταν υπάρχει μεγαλύτερη καμπυλότητα το σφάλμα είναι μεγαλύτερο αφού, όπως αναφέρθηκε, η προσέγγιση που γίνεται είναι γραμμική. Αναφέρεται ότι, στη συγκεκριμένη περίπτωση, είναι $\delta\hat{f}_2 \cong 0,4483$, μία τιμή που είναι σχετικά μικρή και, σε συνδυασμό με τη δεδομένη τιμή της καμπυλότητας στο πρώτο σημείο, οδηγεί σε πολύ ικανοποιητικά αποτελέσματα σε σχέση με τη δεύτερη μη-κυριαρχούμενη λύση αλλά όχι τόσο ώστε το σφάλμα να αποφεύγεται πλήρως. Καθώς το πρόγραμμα δεν προβαίνει σε διόρθωση του σημείου που μόλις υπολόγισε αλλά το δέχεται ως το δεύτερο σημείο του μετώπου Pareto, το προαναφερθέν σφάλμα δεν εξαλείφεται αλλά παραμένει. Για το δεύτερο σημείο, λοιπόν, οι συνθήκες ΚΚΤ δεν θα ικανοποιούνται 'πλήρως' αφού το f_2 θα έχει μία μικρή διαφορά με το \hat{f}_2 . Κατά συνέπεια, το θεώρημα πεπλεγμένης συνάρτησης, επίσης, δεν θα ισχύει απόλυτα. Άρα, όταν το πρόγραμμα κληθεί να υπολογίσει την τρίτη μη-κυριαρχούμενη λύση, το σφάλμα θα είναι ακόμα μεγαλύτερο από ότι στην περίπτωση του δεύτερου σημείου. Κατ' αυτόν τον τρόπο το σφάλμα 'μεταδίδεται' από το ένα σημείο στο επόμενο αυξανόμενο (συσσώρευση και επαύξηση σφάλματος) καθώς κάθε φορά το f_2 θα έχει όλο και μεγαλύτερη διαφορά με το \hat{f}_2 . Για το λόγο αυτό, είναι απόλυτα δικαιολογημένο το γεγονός της μεγαλύτερης απόκλισης του f_2 από το \hat{f}_2 όσο ο αριθμός των υπολογισθέντων σημείων αυξάνεται, όπως αυτό διαφάνηκε στο σχήμα 5.5.

Μετά τα ανωτέρω σχόλια, το επόμενο ζήτημα που είναι απαραίτητο να τελέσει υπό διαχείριση είναι το κατά πόσο το μέτωπο Pareto που απεικονίζεται στο σχήμα 5.5 ανταποκρίνεται στην πραγματικότητα. Για να πραγματοποιηθεί μία τέτοια διαπίστωση είναι αναγκαία η επίλυση του προβλήματος και με μία ακόμα μέθοδο. Λαμβάνεται, λοιπόν, η απόφαση το πρόβλημα να επιλυθεί και με τη χρήση των εξελικτικών αλγορίθμων κάτι που, όπως αναφέρθηκε στην ενότητα 5.2, πραγματοποιείται μέσω του λογισμικού βελτιστοποίησης EASY. Στη συνέχεια, ακολουθεί μία σύντομη περιγραφή της λειτουργίας του EASY αλλά και μία αναφορά των παραμέτρων που επιλέχθηκαν. Για κάθε υποψήφια λύση που πρέπει να αξιολογηθεί (σε οποιαδήποτε γενιά) το λογισμικό EASY δημιουργεί το αρχείο task.dat. Στην πρώτη γραμμή του εν λόγω αρχείου περιέχεται ένας ακέραιος ο οποίος δηλώνει το πλήθος των μεταβλητών σχεδιασμού, ενώ ακολουθούν τόσες γραμμές όσες ο ακέραιος αυτός, καθεμία με την τιμή της αντίστοιχης μεταβλητής σχεδιασμού με τη σειρά που αυτές δίνονται στο αρχείο δεδομένων του EASY. Στη συνέχεια, το λογισμικό EASY τρέχει ένα αρχείο εντολών που απαιτείται να έχει δημιουργηθεί από το χρήστη και ονομάζεται task.bat. Το αρχείο αυτό έχει την ακόλουθη μορφή:

```
@echo off
objective_case1.exe >nul
```

Η πρώτη γραμμή χρησιμεύει στο να μην εμφανίζονται στην οθόνη τα αποτελέσματα των κλήσεων που ακολουθούν. Η σύνταξη κώδικας.exe >nul χρησιμοποιείται ώστε το πρόγραμμα να έχει τυπική έξοδο το προσωρινό αρχείο nul, το οποίο σβήνεται με το τέλος της εντολής, ενώ κατά την εκτέλεσή της δεν εμφανίζονται αποτελέσματα στην οθόνη.

Το εκτελέσιμο αρχείο objective_case1.exe προκύπτει μετά από μετάφραση (compilation) του κώδικα objective_case1.for ο οποίος αποτελεί το λογισμικό αξιολόγησης (δηλαδή συντέθηκε ώστε να είναι ικανός να υπολογίζει τις τιμές των συναρτήσεων στόχου για κάθε υποψήφια λύση). Το εν λόγω πρόγραμμα λαμβάνει ως είσοδό του το αρχείο task.dat που δημιουργεί το λογισμικό EASY για κάθε υποψήφια λύση που πρέπει να αξιολογηθεί. Τα αποτελέσματά του (έξοδος) γράφονται σε ένα αρχείο που ονομάζεται task.res. Πρόκειται για ένα αρχείο το οποίο το λογισμικό EASY περιμένει να του επιστραφεί από το αρχείο εντολών task.bat και αποτελείται από τόσες γραμμές όσες είναι οι συναρτήσεις στόχου, με την κάθε γραμμή να περιέχει την τιμή της αντίστοιχης συνάρτησης στόχου.

Το επόμενο βήμα είναι η ρύθμιση των παραμέτρων του EASY. Παρακάτω αναφέρονται οι σημαντικότερες από αυτές, κατηγοριοποιημένες σύμφωνα με το πώς απαντώνται στο περιβάλλον του EASY:

Case Organization (Οργάνωση του προβλήματος)

Log file: EALog (στο αρχείο αυτό υπάρχει όλη η εξέλιξη του αλγορίθμου)

Solution file: out (στο αρχείο αυτό αποθηκεύονται οι τιμές των συναρτήσεων στόχου και των μεταβλητών σχεδιασμού για τα μέλη του Pareto).

Number of objectives: 2

Design Variables and Constraints (Μεταβλητές σχεδιασμού και περιορισμοί)

Στο πεδίο αυτό ορίζονται τα όρια των μεταβλητών σχεδιασμού ακριβώς όπως δόθηκαν στην ενότητα 5.3.1. Επιπλέον, εδώ ορίζεται ο αριθμός των bits που αντιστοιχούν σε κάθε μεταβλητή στην περίπτωση που χρησιμοποιείται η δυαδική κωδικοποίησή τους. Για αμφότερες τις μεταβλητές επιλέχθηκαν 10 bits.

Το συγκεκριμένο πρόβλημα δεν περιλαμβάνει περιορισμούς.

Search Engines (Μηχανές αναζήτησης)

Search Engine: Evolutionary Algorithm

General

Solution store step: 1 (δηλαδή σώζονται για την κάθε γενιά που τελειώνει οι τιμές των συναρτήσεων στόχου και των μεταβλητών σχεδιασμού για τα 'καλύτερα' μέλη στο αρχείο αποτελεσμάτων out

ώστε να μπορεί να διαπιστωθεί πότε το λογισμικό EASY φτάνει στο μέτωπο Pareto)

Multiobjective mode (fitness assignment technique): SPEA 2

Convergence (Σύγκλιση)

Maximum evaluations: 3000 (επιλέγεται ένας σχετικά μεγάλος αριθμός αξιολογήσεων ώστε να υπάρχει η σιγουριά ότι ο αλγόριθμος θα συγκλίνει στο μέτωπο Pareto)

Maximum idle generations (μέγιστος αριθμός γενιών χωρίς να υπάρχει αλλαγή στα σημεία του Pareto): 100

Population

Parent Population size (πληθυσμός γονέων): 15

Offspring Population size (πληθυσμός απογόνων): 60

Parents of one offspring (πόσοι γονείς είναι υπεύθυνοι για τη δημιουργία ενός απογόνου): 2

Elite archive size: 50 (ζητούνται πολλά σημεία πάνω στο μέτωπο Pareto ώστε να προκύψει με μεγαλύτερη ακρίβεια η καμπύλη των μη-κυριαρχούμενων λύσεων)

Elite individuals to force as new offspring (ο αριθμός των elite σημείων που επιβάλλονται ως νέοι απόγονοι): 1

Probability to select an elite (πιθανότητα να επιλεγεί ένα elite άτομο ως γονέας ενός απογόνου): 0,05

Tournament size (ο αριθμός των ατόμων που συμμετέχουν στο τουρνουά για την επιλογή γονέων): 4

Tournament probability (πιθανότητα να επιλεγεί το καλύτερο εκ των ατόμων που συμμετέχουν στο τουρνουά): 0,9

Operators

Coding (Κωδικοποίηση): Binary-Gray

Probability crossover (πιθανότητα διασταύρωσης): 0,95

Mode crossover (τρόπος διασταύρωσης μεταβλητών σχεδιασμού): two points/variable (δύο σημείων ανά μεταβλητή)

Mutation probability (πιθανότητα μετάλλαξης): 0,05

Multiplier: 0,8

Idle generations: 10 (αριθμός γενιών χωρίς να υπάρξει ανανέωση στα elite, αν ξεπεραστεί η πιθανότητα μετάλλαξης αυξάνεται σύμφωνα με τον multiplier)

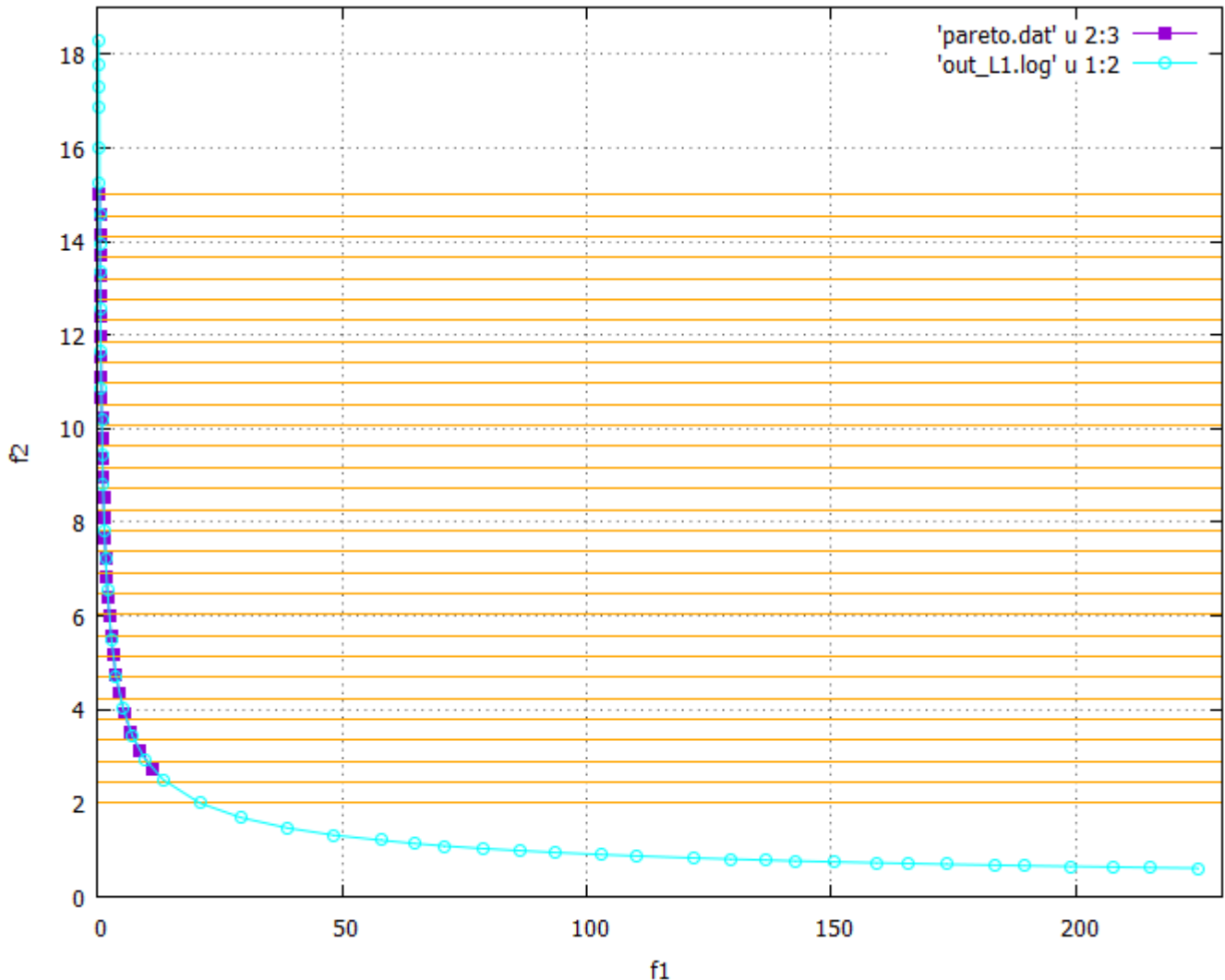
Evaluation scripts:

Script filename: task.bat

Maximum evaluations: 3000

Μετά τα ανωτέρω το λογισμικό EASY είναι έτοιμο να τρέξει και να παραδώσει τα ζητηθέντα αποτελέσματα. Μέσω των εν λόγω αποτελεσμάτων, καθίσταται δυνατό να διαπιστωθεί κατά πόσο το μέτωπο Pareto που απεικονίζεται στο σχήμα 5.5 ανταποκρίνεται στην

πραγματικότητα. Για το σκοπό αυτό, σχεδιάζονται στο ίδιο σχήμα το μέτωπο Pareto που προέκυψε από τη μέθοδο που βασίζεται στη συνέχιση της καμπύλης των μη-κυριαρχούμενων λύσεων και αυτό που προέκυψε από το λογισμικό των εξελικτικών αλγορίθμων (σχήμα 5.6).

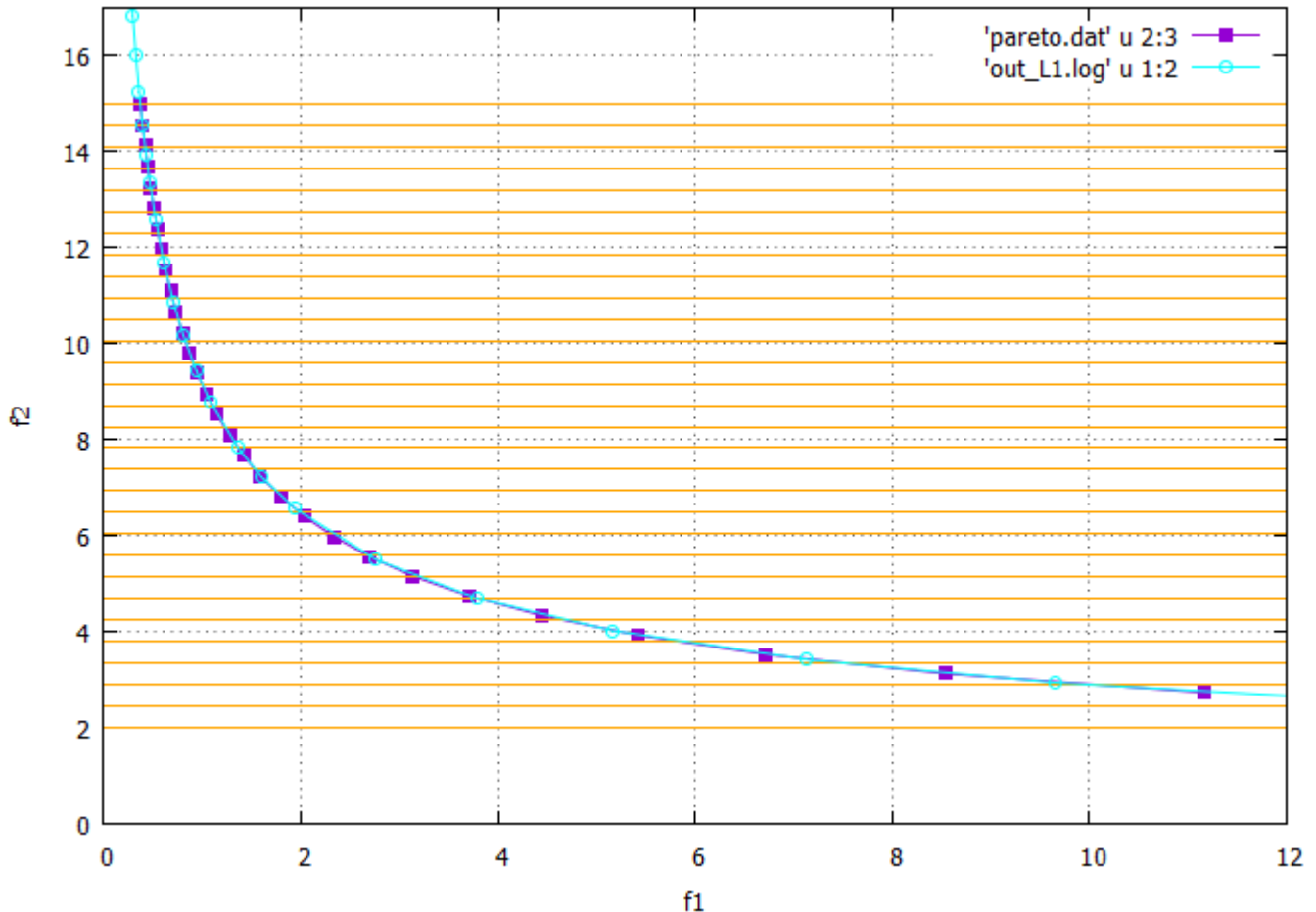


Σχήμα 5.6: Μέτωπο Pareto που προέκυψε από τη μέθοδο που βασίζεται στη συνέχιση της καμπύλης των μη-κυριαρχούμενων λύσεων για $\hat{f}_2^{up} = 15$, $\hat{f}_2^{down} = 2$ και $n_{elit} = 30$ (μωβ καμπύλη) και μέτωπο Pareto που προέκυψε από το λογισμικό των εξελικτικών αλγορίθμων (γαλάζια καμπύλη). Οι οριζόντιες γραμμές χρώματος πορτοκαλί αντιπροσωπεύουν τις τιμές των \hat{f}_2 για καθένα από τα n_{elit} σημεία που αναζητήθηκαν.

Η πρώτη διαπίστωση που θα μπορούσε να γίνει μετά την παρατήρηση του σχήματος 5.6 είναι ότι το μέτωπο Pareto που παρουσιάστηκε στο σχήμα 5.5 αποτελεί μόνο ένα τμήμα του πραγματικού μετώπου των μη-κυριαρχούμενων λύσεων που αποκτήθηκε μέσω του λογισμικού των εξελικτικών αλγορίθμων. Με άλλα λόγια, παρατηρείται ότι υπάρχουν σημεία του μετώπου Pareto για τα οποία η συνάρτηση στόχου f_2 λαμβάνει τιμές και μεγαλύτερες του $\hat{f}_2^{up} = 15$ και

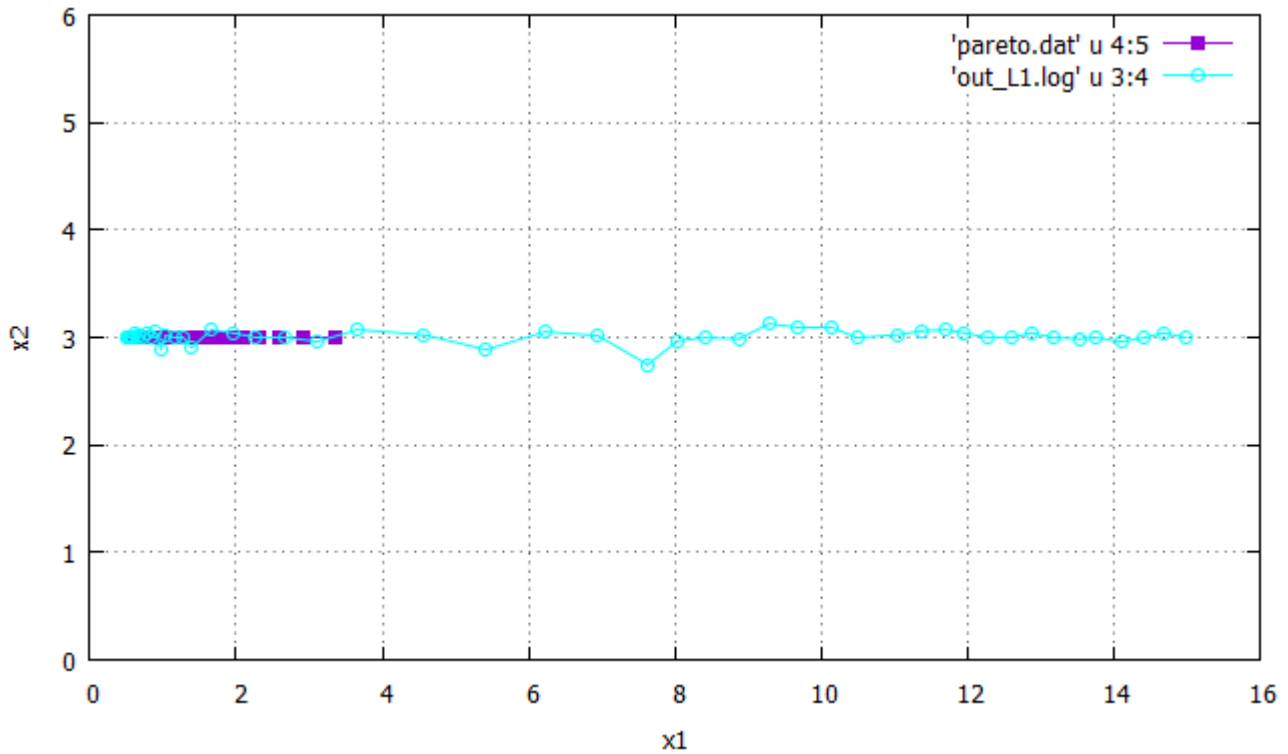
μικρότερες του $\hat{f}_2^{down} = 2$ (βέβαια, όπως σχολιάσθηκε παραπάνω, στην πραγματικότητα το τελευταίο σημείο που αποκτήθηκε είχε $f_2 = 2,735947$). Στο δεδομένο σημείο, πρέπει να ειπωθεί ότι η επιλογή του τμήματος του μετώπου Pareto που ενδιαφέρει να αποκτηθεί σε ένα (πραγματικό) πρόβλημα γίνεται από τον κάθε επιστήμονα σύμφωνα με τις ανάγκες του. Συνήθως, όμως, η τελική επιλογή την οποία καλείται να κάνει ο υπεύθυνος λήψης αποφάσεων σχετικά με το ποιος σχεδιασμός θα κατασκευαστεί γίνεται από το τμήμα του μετώπου Pareto όπου συναντάται το 'γόνατο', δηλαδή από το μεσαίο τμήμα, καθώς εκεί υπάρχει 'καλύτερος συμβιβασμός' μεταξύ των δύο στόχων. Συνεπώς, γενικά, αυτό είναι ένα τμήμα που επιθυμείται να αποκτηθεί. Στη συγκεκριμένη περίπτωση, όπως φαίνεται στο σχήμα 5.6, στο εν λόγω τμήμα αντιστοιχούν περίπου $\hat{f}_2^{up} = 6$ και $\hat{f}_2^{down} = 1$. Είναι σαφές ότι το τμήμα αυτό δεν έχει αποκτηθεί. Είναι αναγκαίο να σημειωθεί ότι ο μηχανικός πρέπει είτε να γνωρίζει τα όρια \hat{f}_2^{up} και \hat{f}_2^{down} που θα του προμηθεύσουν το τμήμα του μετώπου μη-κυριαρχούμενων λύσεων που επιθυμεί είτε να κάνει κάποιες δοκιμές με διαφορετικές τιμές των ορίων αυτών (συνοδευόμενες από το αντίστοιχο υπολογιστικό κόστος) ώστε να δει τι αποτελέσματα προκύπτουν.

Επιπλέον, αν παρατηρηθεί το σχήμα 5.6, εξάγεται το συμπέρασμα ότι η καμπύλη που προέκυψε από τη μέθοδο που βασίζεται στη συνέχιση της καμπύλης των μη-κυριαρχούμενων λύσεων για $\hat{f}_2^{up} = 15$, $\hat{f}_2^{down} = 2$ και $n_{elit} = 30$, πρακτικά, πέφτει πάνω σε αυτή που προέκυψε από το λογισμικό των εξελικτικών αλγορίθμων. Για να φανεί αυτό με μεγαλύτερη ευκρίνεια δημιουργείται το σχήμα 5.7. Το εν λόγω σχήμα αποτελεί ένα διάγραμμα το οποίο επικεντρώνεται στο κοινό τμήμα των δύο καμπυλών.

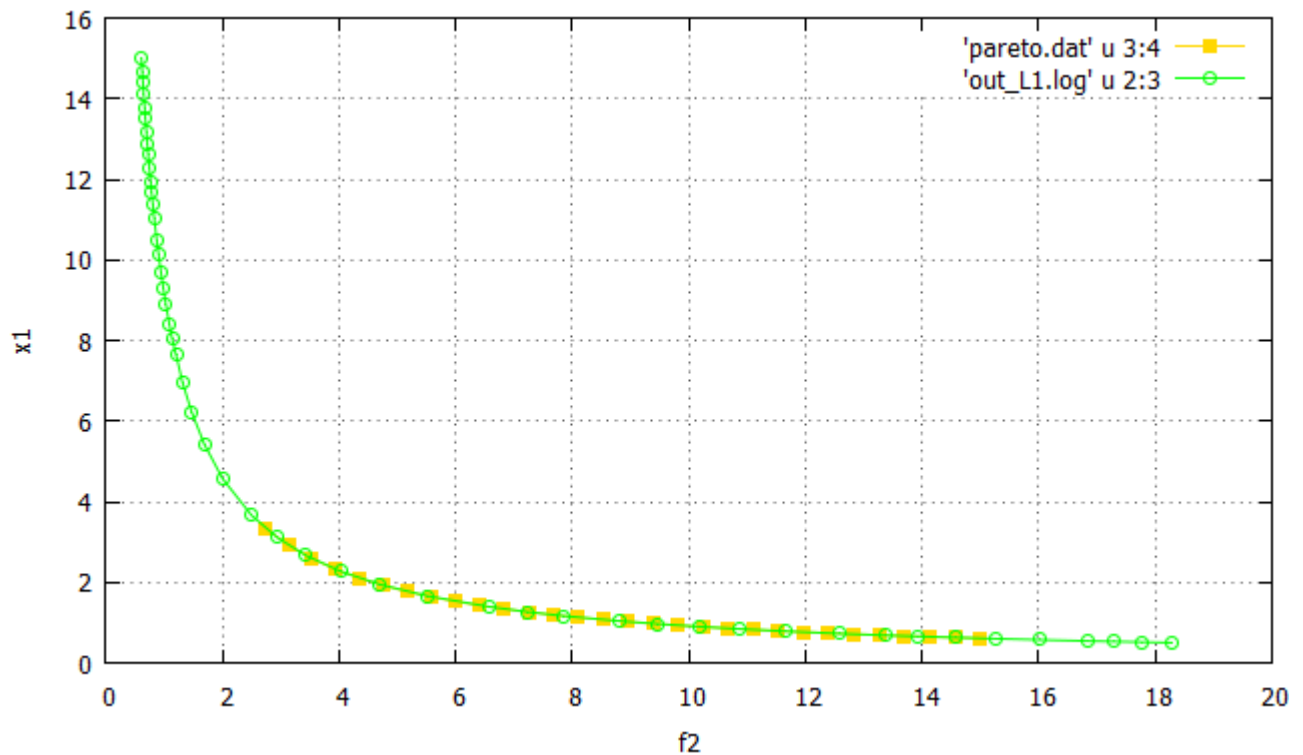


Σχήμα 5.7: Μέτωπο Pareto που προέκυψε από τη μέθοδο που βασίζεται στη συνέχιση της καμπύλης των μη-κυριαρχούμενων λύσεων για $\hat{f}_2^{up} = 15$, $\hat{f}_2^{down} = 2$ και $n_{elit} = 30$ (μωβ καμπύλη) και τμήμα του μετώπου Pareto που προέκυψε από το λογισμικό των εξελικτικών αλγορίθμων (γαλάζια καμπύλη). Οι οριζόντιες γραμμές χρώματος πορτοκαλί αντιπροσωπεύουν τις τιμές των \hat{f}_2 για καθένα από τα n_{elit} σημεία που αναζητήθηκαν.

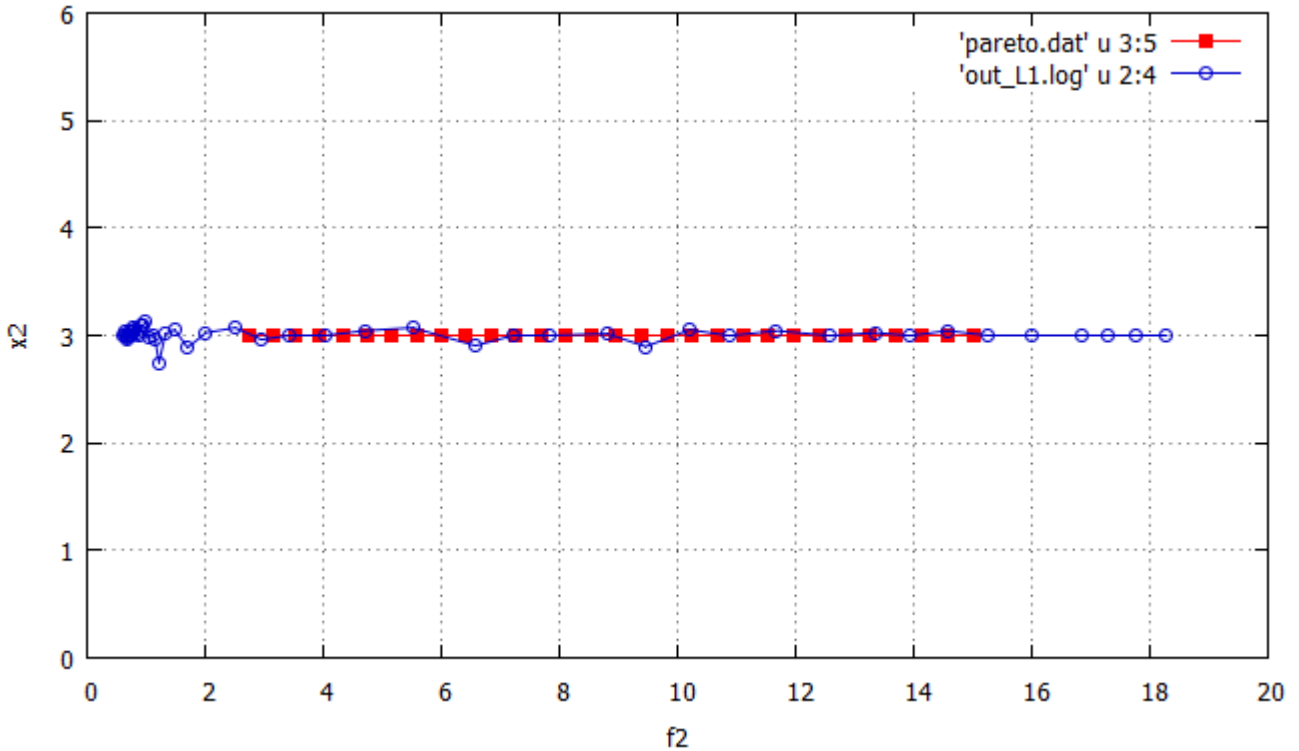
Στο σχήμα 5.7, λοιπόν, φαίνεται καθαρά ότι η καμπύλη που προέκυψε με βάση τη συνέχιση του μετώπου των μη-κυριαρχούμενων λύσεων για τις παραμέτρους που αναφέρθηκαν ανωτέρω, πρακτικά, συμπίπτει με το αντίστοιχο τμήμα του μετώπου Pareto που προέκυψε από το λογισμικό των εξελικτικών αλγορίθμων. Αυτό σημαίνει ότι έχει επιτευχθεί η εύρεση ενός τμήματος του μετώπου Pareto. Κρίνεται απαραίτητο να τονιστεί ότι αυτό ήταν δυνατό να συμβεί παρά το γεγονός ότι, όπως σχολιάσθηκε παραπάνω, η τιμή της διαφοράς $f_2 - \hat{f}_2$ συνεχώς αυξάνεται καθώς ο αλγόριθμος υπολογίζει όλο και περισσότερα σημεία. Για την κατανόηση της εν λόγω συμπεριφοράς είναι απαραίτητη η κατασκευή των διαγραμμάτων $x_1 - x_2$, $f_2 - x_1$ και $f_2 - x_2$ που προέκυψαν από τις δύο μεθόδους. Αυτά φαίνονται αντίστοιχα στα σχήματα 5.8, 5.9 και 5.10.



Σχήμα 5.8: Διάγραμμα των καμπυλών $x_1 - x_2$ που προέκυψαν από τη μέθοδο που βασίζεται στη συνέχιση της καμπύλης των μη-κυριαρχούμενων λύσεων για $\hat{f}_2^{up} = 15$, $\hat{f}_2^{down} = 2$ και $n_{elit} = 30$ (μωβ καμπύλη) και από το λογισμικό των εξελικτικών αλγορίθμων (γαλάζια καμπύλη).



Σχήμα 5.9: Διάγραμμα των καμπυλών $f_2 - x_1$ που προέκυψαν από τη μέθοδο που βασίζεται στη συνέχιση της καμπύλης των μη-κυριαρχούμενων λύσεων για $\hat{f}_2^{up} = 15$, $\hat{f}_2^{down} = 2$ και $n_{elit} = 30$ (κίτρινη καμπύλη) και από το λογισμικό των εξελικτικών αλγορίθμων (πράσινη καμπύλη).

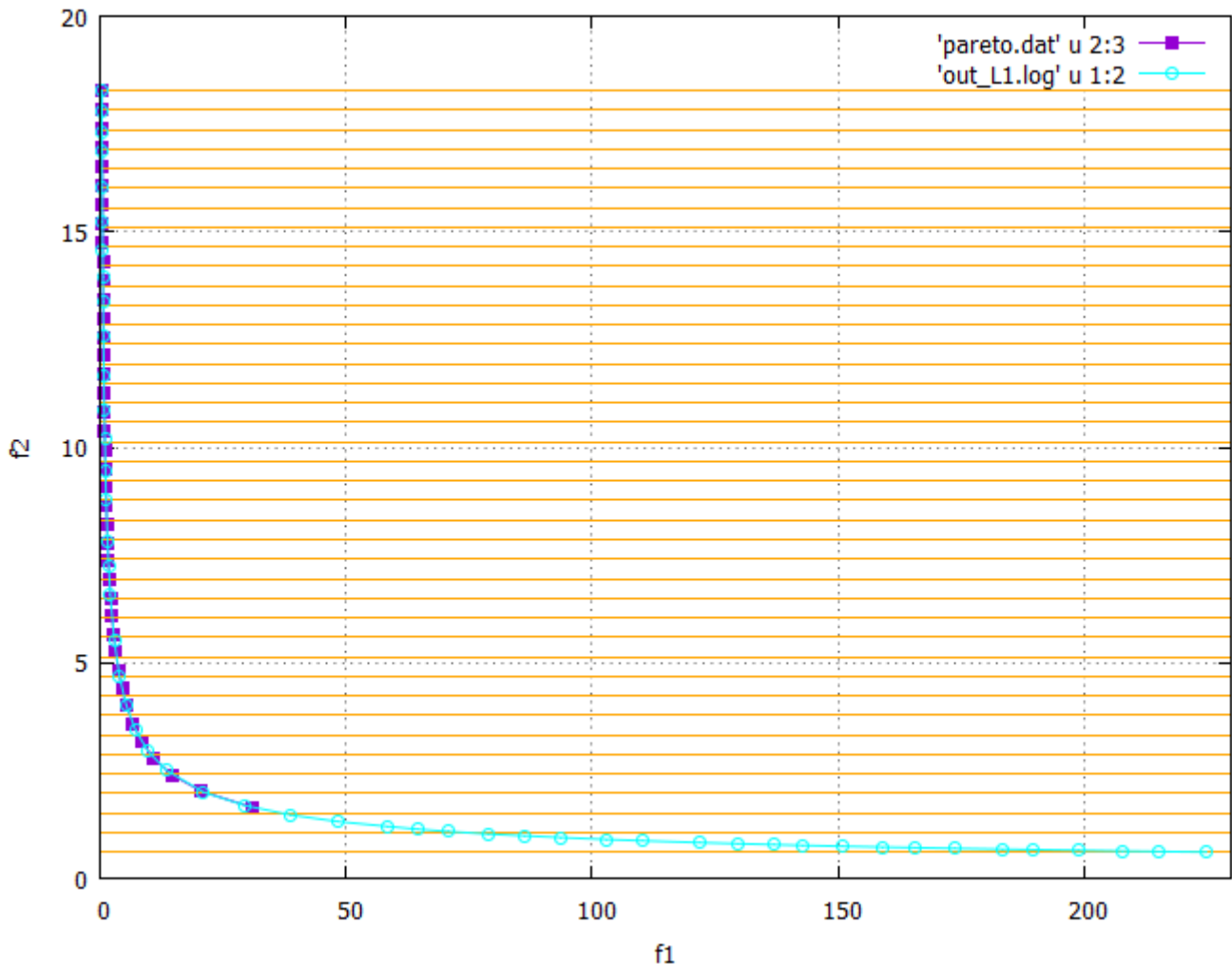


Σχήμα 5.10: Διάγραμμα των καμπυλών $f_2 - x_2$ που προέκυψαν από τη μέθοδο που βασίζεται στη συνέχιση της καμπύλης των μη-κυριαρχούμενων λύσεων για $f_2^{up} = 15$, $f_2^{down} = 2$ και $n_{elit} = 30$ (κόκκινη καμπύλη) και από το λογισμικό των εξελικτικών αλγορίθμων (μπλε καμπύλη).

Παρατηρείται ότι οι καμπύλες $x_1 - x_2$ (σχήμα 5.8) και $f_2 - x_2$ (σχήμα 5.10) που προέκυψαν από τις δύο μεθόδους είναι πρακτικά οριζόντιες ευθείες (ισχύει $x_2 \cong 3$ για όλες τις μη-κυριαρχούμενες λύσεις). Δεδομένου ότι η τιμή του x_2 είναι πρακτικά ίδια για όλα τα σημεία του μετώπου Pareto, ισχύει $f_2 = \alpha/x_1$, $\alpha = \text{σταθερό}$, οπότε είναι εύλογο το αποτέλεσμα που φαίνεται στο σχήμα 5.9. Η ανανέωση των τιμών των μεταβλητών σχεδιασμού, όπως δηλώθηκε παραπάνω, γίνεται με βάση τη σχέση 5.1 η οποία αποτελεί μία προσέγγιση πρώτης τάξης στο χώρο $\hat{f}_2 - \vec{z}$. Στην περίπτωση, λοιπόν, όπου η καμπύλη $\hat{f}_2 - x_2$ είναι ουσιαστικά μία οριζόντια ευθεία, η εφαπτομένη σε κάθε σημείο της είναι επίσης η ίδια ευθεία και έτσι το x_2 θα παραμένει ίδιο για όλα τα σημεία του μετώπου Pareto που υπολογίζει το πρόγραμμα. Συνεπώς, η τιμή του f_2 καθορίζεται από την τιμή που λαμβάνει το x_1 . Όμως, η καμπύλη $\hat{f}_2 - x_1$ είναι πρακτικά μία υπερβολή και άρα η εφαπτομενική προσέγγιση που γίνεται ενέχει σφάλμα, όπως περιγράφηκε λεπτομερώς σε ανωτέρω τμήμα της παρούσας ενότητας. Αυτό σημαίνει ότι οι τιμές των συναρτήσεων στόχου που έχουν τα αποκτηθέντα σημεία από το πρόγραμμα περικλείουν το σφάλμα που έχει γίνει στον υπολογισμό του x_1 (το οποίο μεγαλώνει σημείο με το σημείο, όπως εξηγήθηκε). Συγκεκριμένα, το x_1 που υπολογίζεται είναι μικρότερο του πραγματικού. Έτσι, το σφάλμα που αφορά στο x_1 οδηγεί σε

μεγαλύτερες τιμές του f_2 σε σχέση με το \hat{f}_2 που δόθηκε και μικρότερες τιμές του f_1 . Όμως, επειδή το x_2 διατηρείται σταθερό, τα προκύπτοντα σημεία θα είναι πάνω στο μέτωπο Pareto, απλά θα βρίσκονται πιο αριστερά από ότι ορίστηκε μέσω του \hat{f}_2 . Στο συγκεκριμένο, λοιπόν, πρόβλημα ουσιαστικά δεν απασχολεί το γεγονός της απόκλισης των τιμών των f_2 και \hat{f}_2 , καθώς τα σημεία που λαμβάνονται θα είναι πάντα πάνω στο μέτωπο Pareto.

Στη συνέχεια, κρίνεται σκόπιμο να παρουσιαστούν τα αποτελέσματα του αναπτυχθέντος προγράμματος αν ως \hat{f}_2^{up} και \hat{f}_2^{down} ληφθούν το άνω και κάτω όριο αντίστοιχα του f_2 που προέκυψαν από το λογισμικό των εξελικτικών αλγορίθμων (δηλαδή λαμβάνονται $\hat{f}_2^{up} = 18,282208$, $\hat{f}_2^{down} = 0,609407$). Στην εν λόγω περίπτωση, επειδή το εύρος των τιμών του \hat{f}_2 αυξάνεται, τίθεται $n_{elit} = 40$ ώστε να μην προκύψει υπερβολικά μεγάλο βήμα $\delta\hat{f}_2$. Επιπλέον, έστω ότι ως μέθοδος βελτιστοποίησης προβλημάτων ενός στόχου χωρίς περιορισμούς επιλέγεται η μέθοδος Newton και ισχύουν $\hat{\lambda}_2^0 = 0$, $\omega_p^0 = 0,2$, $\gamma = 1,5$ και $\vec{x}^0 = [10 \ 5]^T$. Τέλος, οι επιθυμητές ανοχές για το κριτήριο σύγκλισης της μεθόδου Newton και για αυτό της ALM τίθενται 10^{-3} και 1 αντίστοιχα. Για το πρώτο σημείο που λαμβάνεται, μετά την εκτέλεση του προγράμματος, ισχύουν $x_1 = 0,5$, $x_2 = 2,974022$, $f_1 = 0,25$ και $f_2 = 18,282969$. Για να συγκλίνει η ALM στο εν λόγω σημείο πραγματοποιήθηκαν 2 κύκλοι, ενώ η μέθοδος Newton 'έδρασε' 8 φορές σε αυτούς τους 2 κύκλους. Σημειώνεται ότι η ALM, στην εν λόγω περίπτωση, πραγματοποιεί $1 + 8(1 + n + 1) = 33$ αξιολογήσεις για την εύρεση του πρώτου σημείου του Pareto, ενώ για την εύρεση των υπόλοιπων 39 σημείων που ζητήθηκαν πραγματοποιούνται $(n_{elit} - 1)(1 + n + 1) = 156$ (δηλαδή συνολικά 189). Στο σχήμα 5.11 φαίνεται το μέτωπο Pareto που προέκυψε συνοδευόμενο από αυτό του εξελικτικού αλγορίθμου.

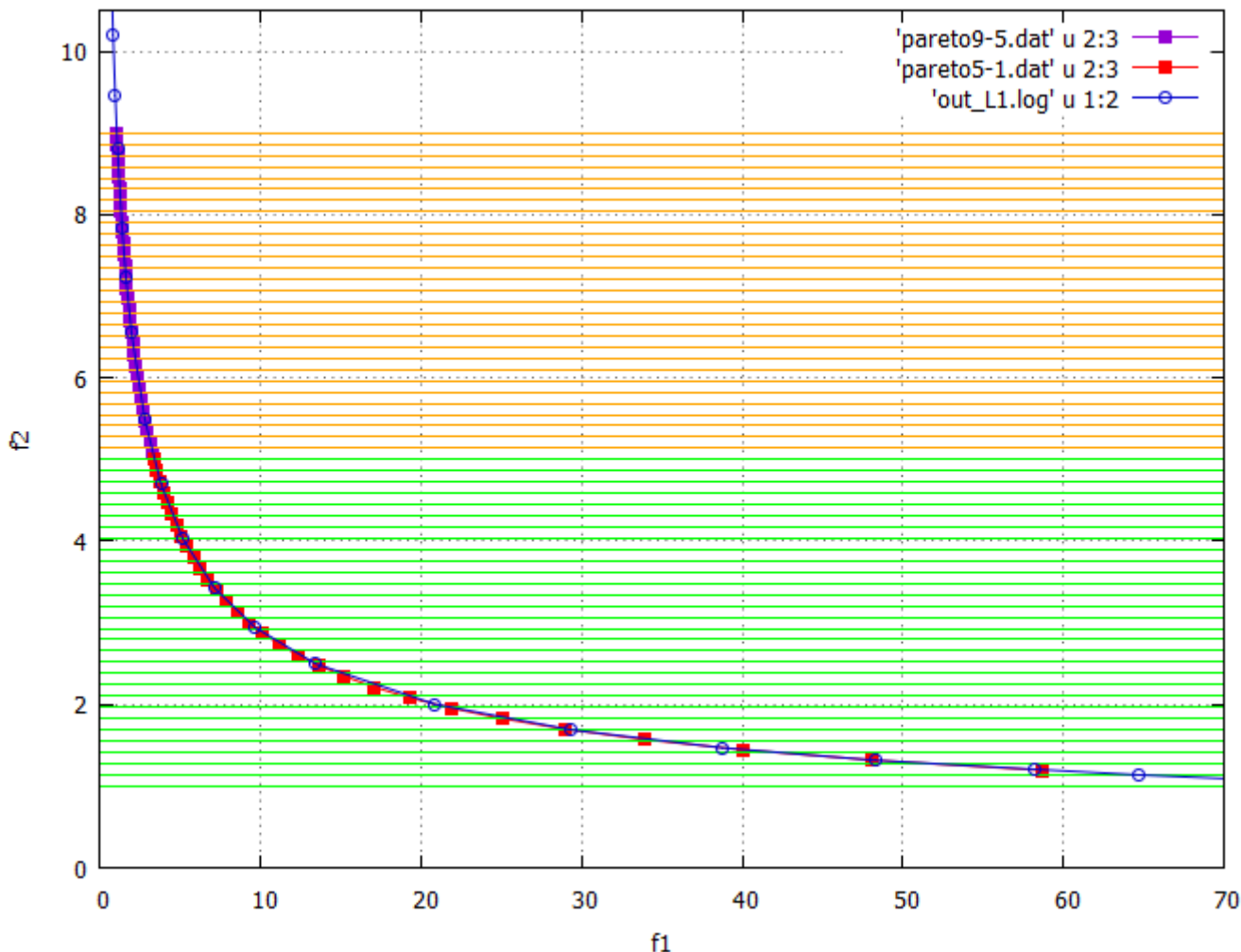


Σχήμα 5.11: Μέτωπο Pareto που προέκυψε από τη μέθοδο που βασίζεται στη συνέχιση της καμπύλης των μη-κυριαρχούμενων λύσεων για $f_2^{up} = 18,282208$, $f_2^{down} = 0,609407$ και $n_{elit} = 40$ (μωβ καμπύλη) και μέτωπο Pareto που προέκυψε από το λογισμικό των εξελικτικών αλγορίθμων (γαλάζια καμπύλη). Οι οριζόντιες γραμμές χρώματος πορτοκαλί αντιπροσωπεύουν τις τιμές των \hat{f}_2 για καθένα από τα n_{elit} σημεία που αναζητήθηκαν.

Παρατηρείται ότι το μέτωπο Pareto που αποκτήθηκε σε αυτή την περίπτωση καλύπτει μεγαλύτερο μέρος του συνολικού που δίνεται από το λογισμικό των εξελικτικών αλγορίθμων σε σχέση με προηγουμένως αλλά και πάλι η μέθοδος δεν δύναται να δώσει το τμήμα που αντιστοιχεί σε μεγαλύτερες τιμές του f_1 λόγω των σφαλμάτων που σχολιάσθηκαν παραπάνω. Είναι, όμως, σημαντικό το γεγονός ότι πλέον αποκτήθηκαν σημεία που καλύπτουν το 'γόνατο' που παρουσιάζει η καμπύλη σε αντίθεση με προηγουμένως. Πρέπει, επίσης, να αναφερθεί ότι στην εν λόγω περίπτωση, η διαφορά $f_2 - \hat{f}_2$ για το τελευταίο σημείο που αποκτήθηκε είναι μεγαλύτερη από ότι προηγουμένως (1,0424 και 0,7359 αντίστοιχα), κάτι που ήταν αναμενόμενο.

Ακολουθως, παρά το γεγονός ότι τα σημεία που λαμβάνονται στο συγκεκριμένο πρόβλημα είναι ούτως ή άλλως πάνω στο μέτωπο Pareto, κρίνεται απαραίτητο να ειπωθεί ότι για τη μείωση των διαφορών $f_2 - \hat{f}_2$ στα σημεία που υπολογίζονται, μπορεί κάποιος να προχωρήσει στην εύρεση του μετώπου μη-κυριαρχούμενων λύσεων ανά τμήματα, καθώς έτσι το εύρος των τιμών του \hat{f}_2 θα είναι μικρότερο. Στην περίπτωση αυτή, βέβαια, το υπολογιστικό κόστος ανεβαίνει. Όμως, υπάρχει και το ενδεχόμενο ο κληθείς να επιλύσει το πρόβλημα να χρειάζεται ένα συγκεκριμένο τμήμα του μετώπου Pareto και όχι ολόκληρο το μέτωπο (χαρακτηριστική είναι, όπως προαναφέρθηκε, η περίπτωση επιλογής του τμήματος όπου συναντάται το 'γόνατο'). Παρακάτω παρουσιάζονται δύο τέτοια τμήματα που επιλέχθηκαν.

Για το πρώτο τμήμα που επιλέχθηκε λαμβάνονται $\hat{f}_2^{up} = 9$, $\hat{f}_2^{down} = 5$ και $n_{elit} = 30$. Επιπλέον, έστω ότι ως μέθοδος βελτιστοποίησης προβλημάτων ενός στόχου χωρίς περιορισμούς επιλέγεται η μέθοδος Newton και ισχύουν $\hat{\lambda}_2^0 = 0$, $\omega_p^0 = 0,3$, $\gamma = 1,5$ και $\vec{x}^0 = [10 \ 5]^T$. Τέλος, οι επιθυμητές ανοχές για το κριτήριο σύγκλισης της μεθόδου Newton και για αυτό της ALM τίθενται 10^{-3} και 1 αντίστοιχα. Για το πρώτο σημείο που λαμβάνεται, μετά την εκτέλεση του προγράμματος, ισχύουν $x_1 = 1,015694$, $x_2 = 2,994466$, $f_1 = 1,031634$ και $f_2 = 8,999851$. Για να συγκλίνει η ALM στο εν λόγω σημείο πραγματοποιήθηκαν 3 κύκλοι, ενώ η μέθοδος Newton 'έδρασε' 10 φορές σε αυτούς τους 3 κύκλους. Σημειώνεται ότι η ALM, στην εν λόγω περίπτωση, πραγματοποιεί $1 + 10(1 + n + 1) = 41$ αξιολογήσεις για την εύρεση του πρώτου σημείου του Pareto, ενώ για την εύρεση των υπόλοιπων 29 σημείων που ζητήθηκαν πραγματοποιούνται $(n_{elit} - 1)(1 + n + 1) = 116$ (δηλαδή συνολικά 157). Για το δεύτερο επιλεχθέν τμήμα λαμβάνονται $\hat{f}_2^{up} = 5$ και $\hat{f}_2^{down} = 1$. Με εξαίρεση το ω_p^0 που λαμβάνει την τιμή 3, όλες οι άλλες παράμετροι ισχύουν ως ανωτέρω. Για το πρώτο σημείο που λαμβάνεται, σε αυτή την περίπτωση, ισχύουν $x_1 = 1,827941$, $x_2 = 2,994579$, $f_1 = 3,341368$ και $f_2 = 5,000760$. Για να συγκλίνει η ALM στο εν λόγω σημείο πραγματοποιήθηκαν 3 κύκλοι, ενώ η μέθοδος Newton 'έδρασε' 13 φορές σε αυτούς τους 3 κύκλους. Σημειώνεται ότι η ALM, στην εν λόγω περίπτωση, πραγματοποιεί $1 + 13(1 + n + 1) = 53$ αξιολογήσεις για την εύρεση του πρώτου σημείου του Pareto, ενώ για την εύρεση των υπόλοιπων 29 σημείων πραγματοποιούνται, όπως και για το προηγούμενο τμήμα, $(n_{elit} - 1)(1 + n + 1) = 116$ (δηλαδή συνολικά 169). Στο σχήμα 5.12 φαίνονται τα δύο αυτά τμήματα του μετώπου Pareto συνοδευόμενα από το μέτωπο του εξελικτικού αλγορίθμου.



Σχήμα 5.12: Τμήματα μετώπου Pareto που προέκυψαν από τη μέθοδο που βασίζεται στη συνέχιση της καμπύλης των μη-κυριαρχούμενων λύσεων με τη μωβ καμπύλη να αντιστοιχεί σε $\hat{f}_2^{up} = 9$ και $\hat{f}_2^{down} = 5$ και την κόκκινη καμπύλη σε $\hat{f}_2^{up} = 5$ και $\hat{f}_2^{down} = 1$. Η μπλε καμπύλη αποτελεί το αντίστοιχο τμήμα του μετώπου Pareto που προέκυψε από το λογισμικό των εξελικτικών αλγορίθμων και οι οριζόντιες πορτοκαλί και πράσινες γραμμές αντιπροσωπεύουν τις τιμές των \hat{f}_2 για καθένα από τα n_{elit} σημεία που αναζητήθηκαν στο πρώτο και το δεύτερο τμήμα αντίστοιχα.

Μέσω του σχήματος 5.12, γίνεται ξεκάθαρο το ότι οι διαφορές $f_2 - \hat{f}_2$ για τα αποκτηθέντα σημεία έχουν μειωθεί σημαντικά σε σχέση με τις δύο ανωτέρω παρουσιασθέντες περιπτώσεις. Αναφέρεται χαρακτηριστικά ότι για το τελευταίο σημείο του πρώτου τμήματος ($\hat{f}_2^{up} = 9$ και $\hat{f}_2^{down} = 5$) ισχύει $f_2 - \hat{f}_2 \cong 0,0780$, ενώ για αυτό του δεύτερου τμήματος ($\hat{f}_2^{up} = 5$ και $\hat{f}_2^{down} = 1$) είναι $f_2 - \hat{f}_2 \cong 0,1929$ (υπενθυμίζεται ότι στις δύο ανωτέρω περιπτώσεις ήταν 0,7359 και 1,0424 αντίστοιχα). Παρατηρείται ότι, στο δεύτερο τμήμα, η διαφορά είναι μεγαλύτερη από ότι στο πρώτο τμήμα. Αυτό συμβαίνει διότι, στη δεύτερη περίπτωση, συναντάται μεγαλύτερη καμπυλότητα και άρα μεγαλύτερη απόκλιση από τη γραμμική προσέγγιση που χρησιμοποιήθηκε για την ανανέωση των τιμών των μεταβλητών

σχεδιασμού. Επιπλέον, πρέπει να ειπωθεί ότι, λόγω των μικρότερων σφαλμάτων-αποκλίσεων, έχει επιτευχθεί μέσω του δεύτερου τμήματος η εύρεση σημείων του μετώπου Pareto που παρουσιάζουν μεγαλύτερες τιμές f_1 από ότι στις προηγούμενες περιπτώσεις.

5.3.4.2 Μέθοδος πρόβλεψης-διόρθωσης

Στην παρούσα ενότητα, κρίνεται σκόπιμο να επιχειρηθεί η εύρεση του μετώπου Pareto με την προσθήκη του σταδίου της διόρθωσης μετά από κάθε σημείο το οποίο 'προβλέπεται'. Λαμβάνεται η απόφαση να γίνει χρήση της μεθόδου Newton ($method = 2$) αντί της μεθόδου της απότομης καθόδου στο τμήμα της ALM όπου πραγματοποιείται η βελτιστοποίηση ενός μονοκριτηριακού προβλήματος χωρίς περιορισμούς, καθώς στην ενότητα που προηγήθηκε επιδείχθηκε ότι παρουσιάζει καλύτερα αποτελέσματα από άποψη υπολογιστικού κόστους. Σε πρώτη φάση, επιλέγονται $\hat{f}_2^{up} = 15$, $n_{elit} = 10$ και $\hat{f}_2^{down} = 1$. Επίσης, επιλέγονται ως επιθυμητές ανοχές για τα κριτήρια σύγκλισης της μεθόδου Newton και της ALM οι τιμές 1 και 10^{-3} αντίστοιχα (στην προηγούμενη ενότητα επιδείχθηκε ότι δίνουν πολύ ικανοποιητικά αποτελέσματα). Όπως έγινε και στην ενότητα που προηγήθηκε, έλαβε χώρα διερεύνηση με στόχο να μελετηθεί η συμπεριφορά της μεθόδου πρόβλεψης-διόρθωσης όσον αφορά το υπολογιστικό κόστος με τη μεταβολή των παραμέτρων της ALM αλλά και του σημείου εκκίνησης \vec{x}^0 . Στον πίνακα 5.6 φαίνονται οι συνολικές αξιολογήσεις που εκτελεί το πρόγραμμα για την απόκτηση του μετώπου Pareto για διάφορες τιμές των παραμέτρων $\hat{\lambda}_2^0$, ω_p^0 , γ και ω_p^{max} ενώ γίνεται επιλογή και διαφορετικών σημείων εκκίνησης \vec{x}^0 .

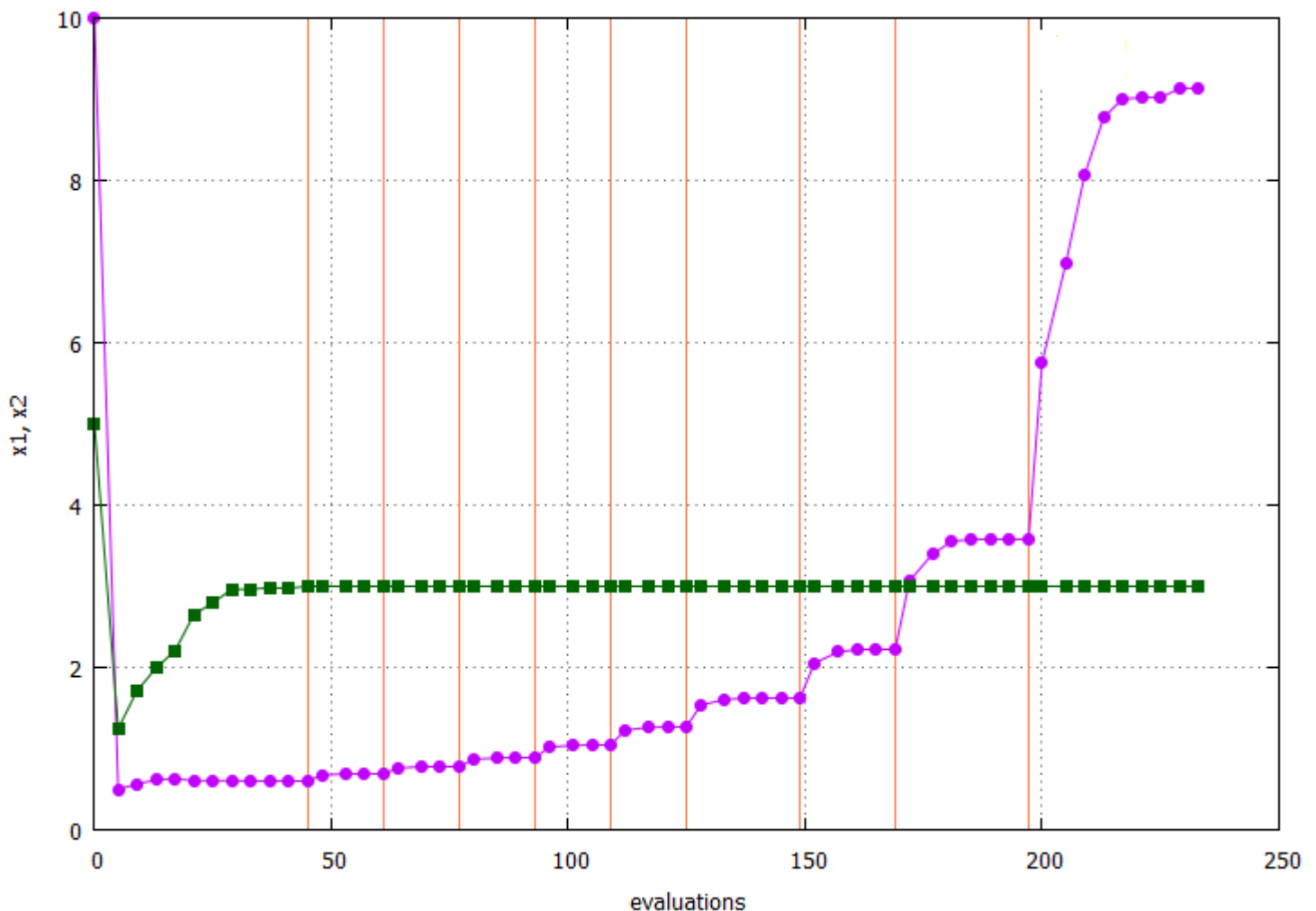
$\hat{\lambda}_2^0$	ω_p^0	γ	ω_p^{max}	\vec{x}^0	Αξιολογήσεις
0	0,001	1,5	500	[10 5]	277
0	0,006	1,5	500	[10 5]	269
0	0,01	1,5	500	[10 5]	293
0	0,1	1,5	500	[10 5]	305
0	0,2	1,5	500	[10 5]	273
0	0,6	1,5	500	[10 5]	277
0	2	1,5	500	[10 5]	329
1	0,2	1,5	500	[10 5]	281
5	0,2	1,5	500	[10 5]	293
10	0,2	1,5	500	[10 5]	317
0	0,2	1,1	500	[10 5]	449
0	0,2	2	500	[10 5]	261
0	0,2	5	500	[10 5]	257
0	0,2	20	500	[10 5]	249
0	0,2	30	500	[10 5]	253
0	0,2	20	5000	[10 5]	233
0	0,2	20	50000	[10 5]	233
0	0,2	20	500000	[10 5]	237
0	0,2	20	5000	[2 4]	233
0	0,2	20	5000	[1 1]	229
0	0,2	20	5000	[0,5 0,5]	233

Πίνακας 5.6: Συνολικές αξιολογήσεις που πραγματοποιούνται για την εύρεση του μετώπου Pareto για διαφορετικές τιμές των παραμέτρων της ALM όταν οι ανοχές για τα κριτήρια σύγκλισης της μεθόδου Newton και της ALM λαμβάνουν τις τιμές 1 και 10^{-3} αντίστοιχα, $\hat{f}_2^{up} = 15$ και $n_{elit} = 10$.

Η διερεύνηση που απεικονίζεται στον πίνακα 5.6 ξεκίνησε με σταδιακή αύξηση του ω_p^0 , ενώ οι τιμές $\hat{\lambda}_2^0$, γ και ω_p^{max} διατηρήθηκαν σταθερές με το σημείο εκκίνησης να παραμένει επίσης ίδιο. Διαπιστώθηκε ότι τα ταχύτερα αποτελέσματα επιτυγχάνονται για $\omega_p^0 = 0,2$. Επιχειρήθηκε, έπειτα, η αύξηση του $\hat{\lambda}_2^0$. Οποιαδήποτε αύξηση του $\hat{\lambda}_2^0$ οδηγεί σε αύξηση των αξιολογήσεων που απαιτούνται. Στη συνέχεια, πραγματοποιήθηκε μείωση και αύξηση του γ . Παρατηρήθηκε ότι η μείωση του γ οδηγεί σε σχετικά μεγάλη αύξηση των αξιολογήσεων, ενώ η αύξηση του γ μέχρι ένα βαθμό μειώνει αρκετά τις αξιολογήσεις. Ακολούθως, έγινε απόπειρα αύξησης του ω_p^{max} και το αποτέλεσμα ήταν η μείωση των απαιτούμενων αξιολογήσεων. Τέλος, εξετάστηκε και η περίπτωση εκκίνησης από διαφορετικό σημείο με τα αποτελέσματα από άποψη υπολογιστικού κόστους να διαφέρουν ελάχιστα. Στο σημείο αυτό κρίνεται απαραίτητο να σχολιαστεί ότι, σε σχέση με την περίπτωση παράλειψης του σταδίου της διόρθωσης, οι αξιολογήσεις που πραγματοποιούνται από το πρόγραμμα είναι περισσότερες (για $\hat{f}_2^{up} = 15$ ακόμα και αν είχε επιλεγεί $n_{elit} = 30$ οι αξιολογήσεις ήταν 161 σε σχέση με τις 233 που

συναντώνται εδώ, αν συγκριθούν οι καλύτεροι συνδυασμοί που βρέθηκαν στις δύο περιπτώσεις για το ίδιο σημείο εκκίνησης), όπως ήταν αναμενόμενο. Στη συγκεκριμένη περίπτωση, βέβαια, δεν συναντώνται αποκλίσεις του f_2 από το \hat{f}_2 και άρα επιτυγχάνεται μεγαλύτερη ακρίβεια στην απόκτηση των σημείων του Pareto (αν και τονίζεται ότι στο συγκεκριμένο πρόβλημα, όπως αναλύθηκε παραπάνω, αυτές δεν απασχολούν καθώς τα σημεία ούτως ή άλλως ήταν πάνω στο μέτωπο Pareto, κάτι που βέβαια δεν είναι δυνατό να το γνωρίζει πάντοτε ο χρήστης εκ των προτέρων).

Στο σημείο αυτό, έστω ότι από τους συνδυασμούς που συναντώνται στον πίνακα 5.6 επιλέγεται αυτός για τον οποίο ισχύουν $\hat{\lambda}_2^0 = 0$, $\omega_p^0 = 0,2$, $\gamma = 20$, $\omega_p^{max} = 5000$ και $\vec{x}^0 = [10 \ 5]^T$. Για τον εν λόγω συνδυασμό παραμέτρων, κρίνεται σκόπιμο να παρατεθεί ένα διάγραμμα στο οποίο παριστάνεται η πορεία που ακολούθησαν οι μεταβλητές σχεδιασμού κατά τη διάρκεια της εκτέλεσης του προγράμματος. Το διάγραμμα αυτό αποτελεί το σχήμα 5.13.



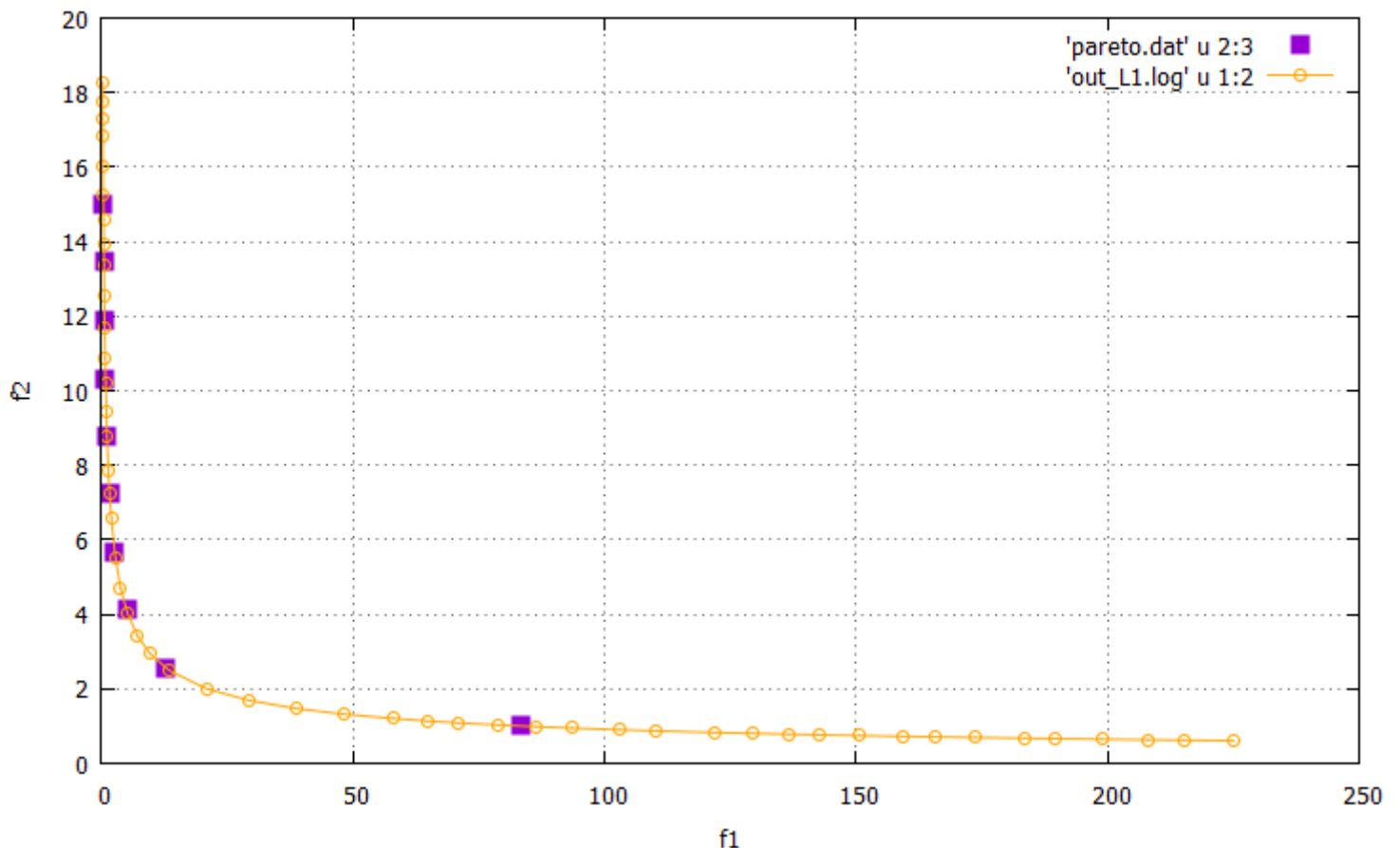
Σχήμα 5.13: Διάγραμμα των μεταβλητών σχεδιασμού x_1 (μωβ καμπύλη) και x_2 (πράσινη καμπύλη) συναρτήσει των επαναλήψεων που εκτελούνται από το πρόγραμμα όταν οι ανοχές με τις οποίες ικανοποιούνται τα κριτήρια

σύγκλισης της μεθόδου Newton και της ALM είναι 1 και 10^{-3} αντίστοιχα, $\hat{f}_2^{up} = 15$, $\bar{x}^0 = [10 \ 5]^T$, $\hat{\lambda}_2^0 = 0$, $\omega_p^0 = 0,2$, $\gamma = 20$ και $\omega_p^{max} = 5000$, $\hat{f}_2^{down} = 1$ και $n_{elit} = 10$. Οι κατακόρυφες γραμμές χρώματος πορτοκαλί χωρίζουν τα 10 τμήματα που αντιστοιχούν στα 10 σημεία του μετώπου Pareto που ευρέθησαν.

Το διάγραμμα 5.13 αποτελεί την απεικόνιση των ‘διεργασιών’ που εκτελούνται από το πρόγραμμα στο σύνολό του. Στο πρώτο τμήμα του εν λόγω διαγράμματος (το οποίο οριοθετείται από την πρώτη πορτοκαλί κατακόρυφη γραμμή) φαίνεται η πορεία που ακολουθεί το πρόγραμμα για να επιτύχει την εύρεση της πρώτης μη-κυριαρχούμενης λύσης. Ξεκινάει από το επιλεχθέν \bar{x}^0 και μέσω της ALM βρίσκει το πρώτο σημείο του Pareto (οι τιμές των μεταβλητών σχεδιασμού για το σημείο αυτό βρίσκονται πάνω στην πρώτη πορτοκαλί κατακόρυφη γραμμή και είναι $x_1 = 0,609391$ και $x_2 = 2,993293$). Για να συγκλίνει η ALM στην πρώτη μη-κυριαρχούμενη λύση πραγματοποιήθηκαν 3 κύκλοι, ενώ η μέθοδος Newton ‘έδρασε’ 11 φορές σε αυτούς τους 3 κύκλους. Σημειώνεται ότι η ALM, στην εν λόγω περίπτωση, πραγματοποιεί $1 + 11(1 + n + 1) = 45$ αξιολογήσεις για την εύρεση του πρώτου σημείου του Pareto, κάτι που φαίνεται και στο σχήμα 5.13. Επίσης, σημειώνεται, ενδεικτικά, ότι η ALM συνέκλινε για $\hat{\lambda}_2 = 0,053018$ και $\omega_p = 80$. Το καθένα από τα επόμενα τμήματα του σχήματος 5.13 αντιστοιχεί σε καθεμία από τις υπολειπόμενες μη-κυριαρχούμενες λύσεις των οποίων η εύρεση έχει απαιτηθεί. Τα πρώτα σημεία των δύο καμπυλών σε κάθε τέτοιο τμήμα αντιπροσωπεύουν την πρόβλεψη της μη-κυριαρχούμενης λύσης στην οποία αντιστοιχεί το εν λόγω τμήμα ενώ τα υπόλοιπα αντιπροσωπεύουν τη διόρθωσή της που γίνεται μέσω της ALM. Τα τελευταία σημεία του κάθε τμήματος αποτελούν το σημείο Pareto στο οποίο αντιστοιχεί το τμήμα αυτό και βρίσκονται πάνω στην αντίστοιχη πορτοκαλί κατακόρυφη γραμμή. Παρατηρείται ότι για την εύρεση καθεμίας από τις επόμενες μη-κυριαρχούμενες λύσεις απαιτούνται λιγότερες αξιολογήσεις από αυτές που πραγματοποιούνται για την απόκτηση της πρώτης, γεγονός που επιβεβαιώνει το σκεπτικό πάνω στο οποίο ‘χτίστηκε’ η μέθοδος πρόβλεψης-διόρθωσης που περιγράφηκε στο κεφάλαιο 4. Βέβαια, πρέπει να αναφερθεί ότι για τις πέντε μη-κυριαρχούμενες λύσεις που έπονται της πρώτης πραγματοποιούνται μόλις 16 αξιολογήσεις ($1 + n = 3$ για το στάδιο της πρόβλεψης και $1 + 3(1 + n + 1) = 13$ για το στάδιο της διόρθωσης που καλεί 3 φορές τη μέθοδο Newton στον έναν κύκλο της ALM που απαιτείται), ενώ για τις υπόλοιπες οι απαιτούμενες αξιολογήσεις αρχίζουν να αυξάνονται (είναι αντίστοιχα 24, 20, 28 και 36). Αυτό συμβαίνει διότι, όπως εξηγήθηκε στην προηγούμενη ενότητα, η καμπυλότητα (στο χώρο όπου πραγματοποιείται η διόρθωση) για τα τελευταία σημεία είναι μεγαλύτερη και άρα η προσέγγιση που γίνεται ενέχει μεγαλύτερο σφάλμα. Πρέπει, όμως, να τονιστεί ότι οι αξιολογήσεις που απαιτούνται για την εύρεση καθεμίας από τις

υπόλοιπες ζητηθείσες μη-κυριαρχούμενες λύσεις παραμένουν αρκετά λιγότερες από αυτές που πραγματοποιούνται για την εύρεση της πρώτης. Τέλος, παρατηρείται ότι όσο μειώνεται η τιμή του \hat{f}_2 αυξάνεται η τιμή του x_1 με την τιμή του x_2 να παραμένει πρακτικά σταθερή, κάτι που αναλύθηκε στην προηγούμενη ενότητα.

Το σημεία του μετώπου Pareto που προκύπτουν, στη συγκεκριμένη περίπτωση, απεικονίζονται στο σχήμα 5.14. Στο εν λόγω σχήμα έχει σχεδιαστεί, επίσης, και το μέτωπο μη-κυριαρχούμενων λύσεων που αποκτήθηκε μέσω του λογισμικού των εξελικτικών αλγορίθμων, όπως περιγράφηκε στην ενότητα που προηγήθηκε.

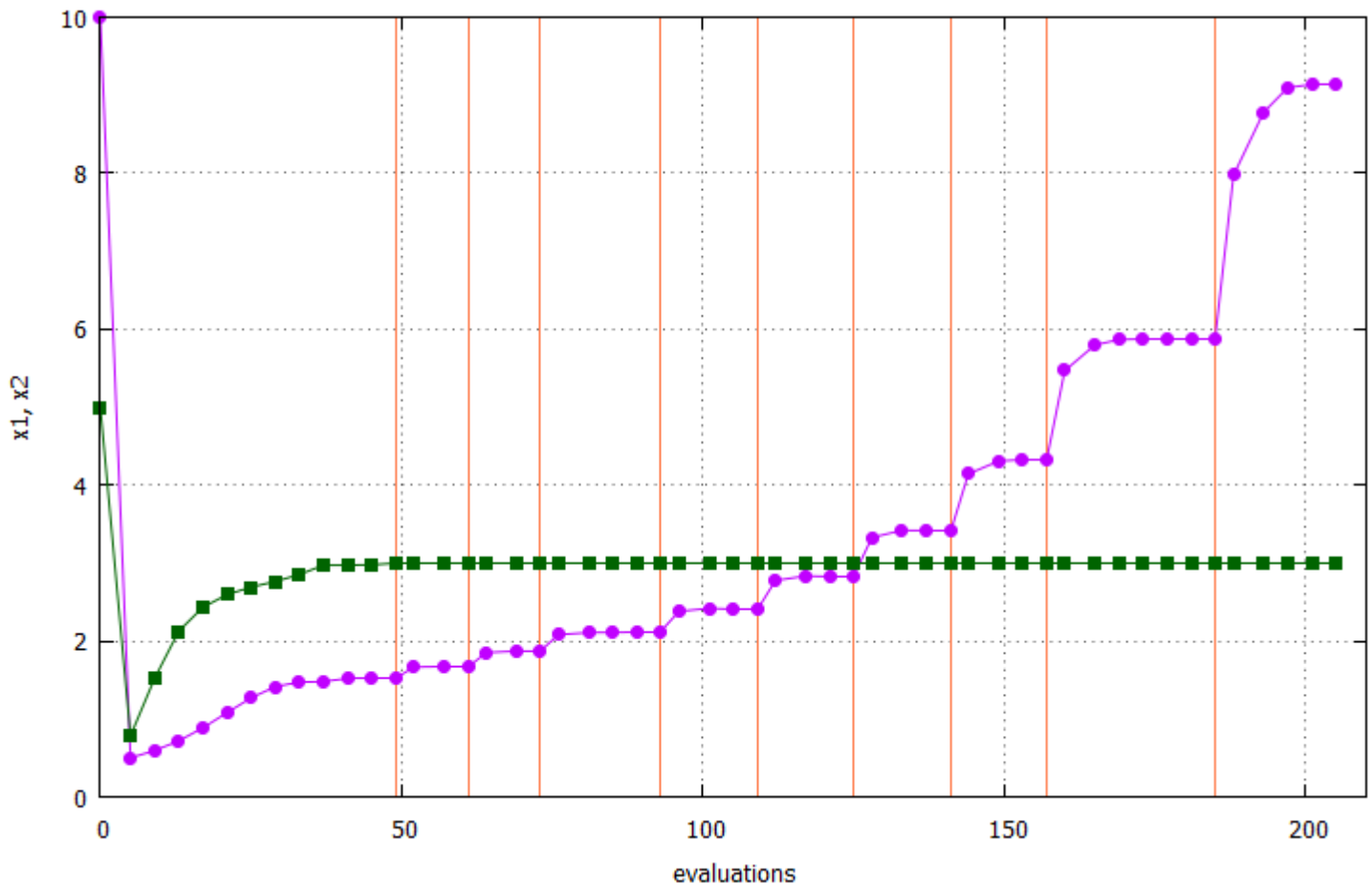


Σχήμα 5.14: Σημεία μετώπου Pareto που προέκυψαν από τη μέθοδο πρόβλεψης-διόρθωσης για $\hat{f}_2^{up} = 15$, $\hat{f}_2^{down} = 1$ και $n_{elit} = 10$ (μεγάλα μωβ σημεία) και μέτωπο Pareto που προέκυψε από το λογισμικό των εξελικτικών αλγορίθμων (πορτοκαλί καμπύλη).

Παρατηρείται ότι τα εννέα πρώτα αποκτηθέντα σημεία βρίσκονται αρκετά ‘κοντά’ μεταξύ τους ενώ το δέκατο σε μεγάλη απόσταση. Αυτό συμβαίνει διότι το μεγάλο εύρος τιμών \hat{f}_2 που δόθηκε αλλά και ο μικρός αριθμός n_{elit} οδήγησαν σε αρκετά μεγάλο $\delta\hat{f}_2$ με αποτέλεσμα όταν το πρόγραμμα φτάνει στο τμήμα του μετώπου Pareto που τείνει να οριζοντιωθεί το προκύπτον $\delta\hat{f}_2$ να μην ‘επαρκεί’ για να δώσει περισσότερα σημεία σε αυτό το τμήμα. Αν ενδιαφέρει να αποκτηθούν

περισσότερα σημεία ενδιάμεσο στο ένατο και το δέκατο που αποκτήθηκαν σε αυτή την περίπτωση θα πρέπει είτε να αυξηθεί το n_{elit} (με ανάλογη αύξηση του υπολογιστικού κόστους) είτε να μειωθεί η διαφορά $\hat{f}_2^{up} - \hat{f}_2^{down}$.

Έστω, λοιπόν, ότι γίνεται το δεύτερο και επιλέγονται $\hat{f}_2^{up} = 6$, $n_{elit} = 10$ και $\hat{f}_2^{down} = 1$. Επιπλέον, επιλέγονται $\hat{\lambda}_2^0 = 0$, $\omega_p^0 = 2$, $\gamma = 50$, $\omega_p^{max} = 50000$ και $\vec{x}^0 = [10 \ 5]^T$. Στη συγκεκριμένη περίπτωση, απαιτείται να πραγματοποιηθούν 205 αξιολογήσεις από το πρόγραμμα. Για τον εν λόγω συνδυασμό παραμέτρων το διάγραμμα στο οποίο παριστάνεται η πορεία που ακολούθησαν οι μεταβλητές σχεδιασμού κατά τη διάρκεια της εκτέλεσης του προγράμματος φαίνεται στο σχήμα 5.15.

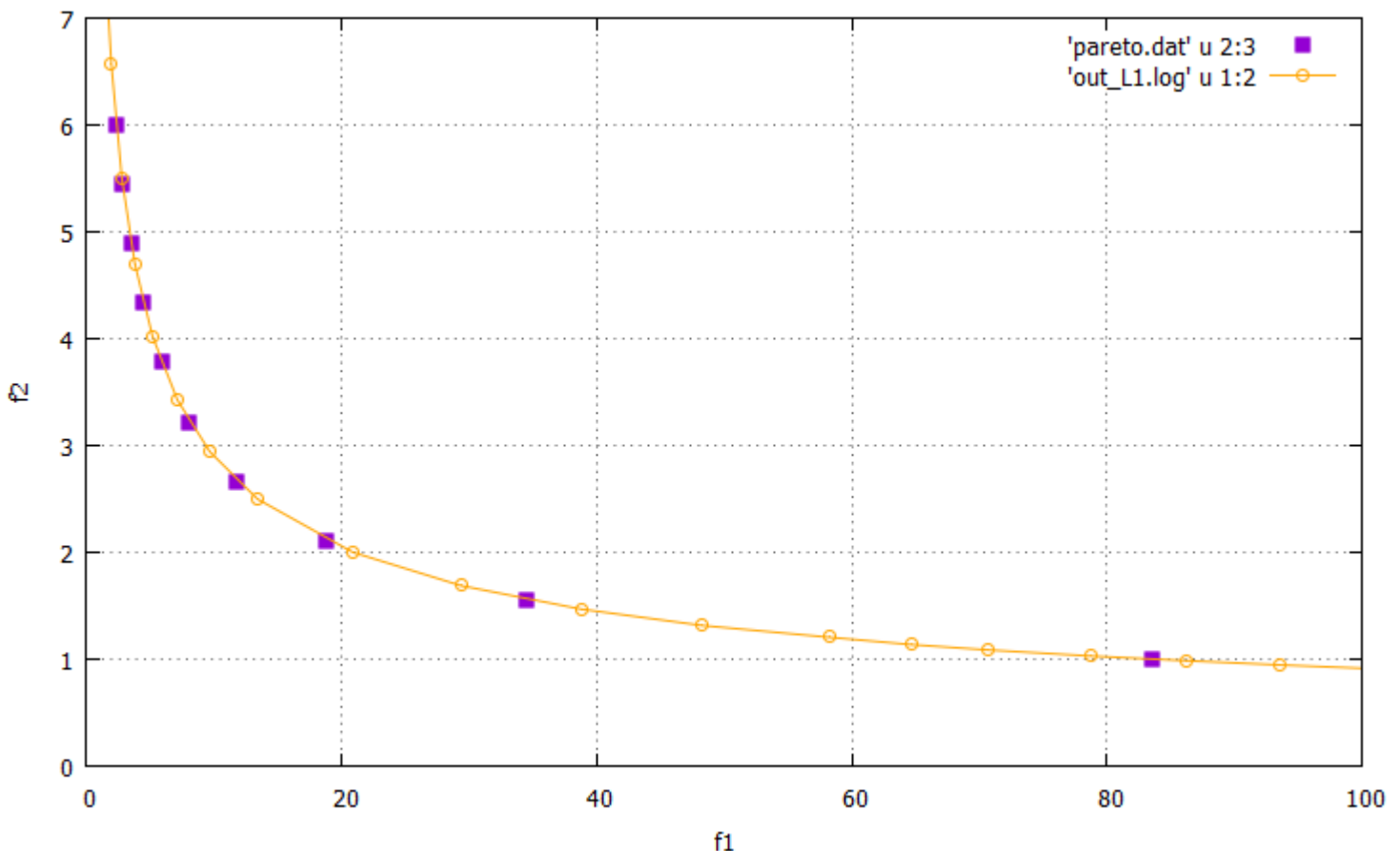


Σχήμα 5.15: Διάγραμμα των μεταβλητών σχεδιασμού x_1 (μωβ καμπύλη) και x_2 (πράσινη καμπύλη) συναρτήσει των επαναλήψεων που εκτελούνται από το πρόγραμμα όταν οι ανοχές με τις οποίες ικανοποιούνται τα κριτήρια σύγκλισης της μεθόδου Newton και της ALM είναι 1 και 10^{-3} αντίστοιχα, $\hat{f}_2^{up} = 6$, $\vec{x}^0 = [10 \ 5]^T$, $\hat{\lambda}_2^0 = 0$, $\omega_p^0 = 2$, $\gamma = 50$ και $\omega_p^{max} = 50000$, $\hat{f}_2^{down} = 1$ και $n_{elit} = 10$. Οι κατακόρυφες γραμμές χρώματος πορτοκαλί χωρίζουν τα 10 τμήματα που αντιστοιχούν στα 10 σημεία του μετώπου Pareto που ευρέθησαν.

Για να συγκλίνει η ALM στην πρώτη μη-κυριαρχούμενη λύση πραγματοποιήθηκαν 49 αξιολογήσεις, ενώ για τα υπόλοιπα σημεία του Pareto οι απαιτούμενες αξιολογήσεις είναι αντίστοιχα 12, 12, 20,

16, 16, 16, 16, 28 και 20. Διαπιστώνεται ότι σε σχέση με την προηγούμενη περίπτωση οι αξιολογήσεις που απαιτούνται για την απόκτηση καθενός από τα σημεία του Pareto, πέραν του πρώτου, είναι λιγότερες λόγω του μικρότερου $\delta \hat{f}_2$ (και άρα λιγότερες είναι και οι συνολικές αξιολογήσεις που γίνονται από το πρόγραμμα).

Το σημεία του μετώπου Pareto που προκύπτουν, στη συγκεκριμένη περίπτωση, απεικονίζονται στο σχήμα 5.16. Στο εν λόγω σχήμα έχει σχεδιαστεί, επίσης, και το αντίστοιχο τμήμα μετώπου μη-κυριαρχούμενων λύσεων που αποκτήθηκε μέσω του λογισμικού των εξελικτικών αλγορίθμων.



Σχήμα 5.16: Σημεία μετώπου Pareto που προέκυψαν από τη μέθοδο πρόβλεψης-διόρθωσης για $\hat{f}_2^{up} = 6$, $\hat{f}_2^{down} = 1$ και $n_{elit} = 10$ (μεγάλα μωβ σημεία) και αντίστοιχο τμήμα μετώπου Pareto που προέκυψε από το λογισμικό των εξελικτικών αλγορίθμων (πορτοκαλί καμπύλη).

Όπως φαίνεται στο σχήμα 5.16, στη συγκεκριμένη περίπτωση, έχουν ληφθεί περισσότερα σημεία στο τμήμα όπου συναντάται το 'γόνατο' της καμπύλης σε σχέση με προηγουμένως.

Τέλος, κρίνεται σκόπιμο να αναφερθεί σε πόσες περίπου αξιολογήσεις δύναται το λογισμικό των εξελικτικών αλγορίθμων να φτάσει στο μέτωπο Pareto. Για να μπορέσει το εν λόγω λογισμικό να δώσει το μέτωπο Pareto με λιγότερες αξιολογήσεις γίνεται χρήση της τεχνικής

IPE (Inexact Pre-Evaluation). Όλες οι παράμετροι που είχαν οριστεί στην προηγούμενη ενότητα παραμένουν ίδιες. Το μόνο πράγμα που αλλάζει είναι ότι τώρα δεν θα γίνονται μόνο ακριβείς αξιολογήσεις των συναρτήσεων στόχου για μία υποψήφια λύση μέσω του λογισμικού αξιολόγησης αλλά θα χρησιμοποιούνται μεταπρότυπα, [19]. Αυτό σημαίνει ότι οι τιμές των συναρτήσεων στόχου για μία υποψήφια λύση υπολογίζονται προσεγγιστικά από τις ‘γειτονικές’ της. Αυτό προϋποθέτει ότι αρχικά θα έχει γίνει ικανοποιητικός αριθμός αξιολογήσεων ώστε να μπορεί το μεταπρότυπο να βγάλει ένα σχετικά αξιόπιστο αποτέλεσμα από τα ‘γειτονικά’ άτομα. Η χρήση μεταπρωτύπων επιτρέπει την αξιολόγηση μόνο των ατόμων που ‘αξίζουν’ στις μεταγενέστερες γενιές και έτσι γίνεται εξοικονόμηση από άποψη υπολογιστικού κόστους.

Η προαναφερθείσα τεχνική IPE συναντάται στο τμήμα Search Engines που αναφέρθηκε στην προηγούμενη ενότητα και οι σημαντικότερες παράμετροι που τέθηκαν φαίνονται παρακάτω.

Metamodel type: RBF

Exact evaluations min (ελάχιστος αριθμός ακριβών αξιολογήσεων που πραγματοποιούνται σε κάθε γενιά): 1

Exact evaluations max (μέγιστος αριθμός ακριβών αξιολογήσεων που πραγματοποιούνται σε κάθε γενιά): 5

IPE pause generations: 10 (ο μέγιστος αριθμός γενιών χωρίς να ανανεώνεται το αρχείο με τα elite, αν ξεπεραστούν η IPE σταματά)

Min DB entries (ελάχιστες ακριβείς αξιολογήσεις πριν χρησιμοποιηθούν μεταμοντέλα): 100

Διαπιστώνεται ότι μέσω της εν λόγω επιλογής παραμέτρων το λογισμικό των εξελικτικών αλγορίθμων δύναται να δώσει το μέτωπο Pareto μετά από περίπου 300 αξιολογήσεις. Σύμφωνα με τα όσα παρουσιάστηκαν στην εν λόγω ενότητα αλλά και την προηγούμενη η μέθοδος που βασίζεται στη συνέχιση του μετώπου Pareto (είτε περιλαμβάνει το στάδιο της διόρθωσης είτε όχι) φαίνεται να επιτυγχάνει την εύρεση του μετώπου μη-κυριαρχούμενων λύσεων αρκετά γρηγορότερα (στην περίπτωση αφαίρεσης του σταδίου της διόρθωσης για $n_{elit} = 30$ οι αξιολογήσεις ήταν 161 και για τη μέθοδο πρόβλεψης-διόρθωσης για $n_{elit} = 10$ ήταν 233). Αυτό, βέβαια, ισχύει για το δεδομένο πρόβλημα αλλά και για το σετ παραμέτρων που επιλέχθηκε στο λογισμικό των εξελικτικών αλγορίθμων (δεν είναι απαραίτητο το σετ που χρησιμοποιήθηκε να δίνει το ‘ταχύτερο’ αποτέλεσμα). Τονίζεται ότι στόχο της παρούσας εργασίας δεν αποτελεί η σύγκριση των δύο μεθόδων. Ο αριθμός των αξιολογήσεων που απαιτεί το λογισμικό των εξελικτικών αλγορίθμων δίνεται απλά για να αποκτηθεί μία ιδέα σχετικά με το αν οι αξιολογήσεις που πραγματοποιεί το αναπτυχθέν πρόγραμμα είναι πολλές ή λίγες (ενδιαφέρει δηλαδή μία τάξη μεγέθους).

5.4 Δεύτερο μαθηματικό πρόβλημα

5.4.1 Διατύπωση προβλήματος

Στην ενότητα αυτή, παρουσιάζεται το δεύτερο μαθηματικό πρόβλημα που επιλέχθηκε να επιλυθεί με βάση τη θεωρία της συνέχισης του μετώπου Pareto. Πρόκειται για ένα πρόβλημα ελαχιστοποίησης δύο στόχων (f_1 και f_2) με δύο μεταβλητές σχεδιασμού (x_1 και x_2). Το πρόβλημα επιβάλλει τόσο άνω όσο και κάτω όρια στις δύο μεταβλητές σχεδιασμού. Αν εξαιρεθούν οι περιορισμοί που υποχρεώνουν τις δύο μεταβλητές σχεδιασμού να κινούνται ανάμεσα στο κάτω και το άνω όριό τους, το πρόβλημα δεν υπόκειται σε κανέναν άλλο περιορισμό, ισότητας ή ανισότητας. Οι δύο συναρτήσεις στόχου που συναντώνται στο εν λόγω πρόβλημα είναι οι εξής:

$$f_1(x_1, x_2) = [g_1(x_1, x_2) - a_1]^2 + [g_2(x_1, x_2) - b_1]^2$$

$$f_2(x_1, x_2) = [g_1(x_1, x_2) - a_2]^2 + [g_2(x_1, x_2) - b_2]^2$$

όπου

$$a_1 = 0,82, \quad a_2 = 1,02, \quad b_1 = 0,67, \quad b_2 = 0,194$$

$$g_1(x_1, x_2) = \frac{10}{r(x_1, x_2)}, \quad g_2(x_1, x_2) = \frac{x_2}{r(x_1, x_2)}, \quad r(x_1, x_2) = 10x_1 \ln\left(1 + \frac{1}{15}x_2\right)$$

Έτσι, το πρόβλημα βελτιστοποίησης που αποτελεί αντικείμενο της δεδομένης ενότητας διατυπώνεται ως εξής:

$$\min \vec{f}(\vec{x}) = \min \begin{bmatrix} f_1(\vec{x}) \\ f_2(\vec{x}) \end{bmatrix}$$

που υπόκειται σε (subject to – s.t.)

$$1 \leq x_1 \leq 10$$

$$1 \leq x_2 \leq 10$$

5.4.2 Εφαρμογή της θεωρίας της συνέχισης του μετώπου Pareto στο πρόβλημα

Η εν λόγω ανάλυση πραγματοποιήθηκε στην ενότητα 5.3.2 και, καθώς δεν παρουσιάζονται διαφορές, δεν θα επαναληφθεί εδώ.

5.4.3 Περιγραφή του αλγορίθμου που αναπτύχθηκε για την επίλυση του προβλήματος

Οι αλγόριθμοι που αναπτύχθηκαν για την επίλυση του παρουσιασθέντος στην ενότητα 5.4.1 μαθηματικού προβλήματος διαφέρουν σε σχέση με αυτούς που περιγράφηκαν διεξοδικώς στις ενότητες 5.3.3.1 και 5.3.3.2 μόνο ως προς το τμήμα υπολογισμού των παραγώγων.

Οι πρώτες και δεύτερες παράγωγοι των συναρτήσεων στόχου ως προς τις μεταβλητές σχεδιασμού για το δεδομένο μαθηματικό πρόβλημα υπολογίζονται, όπως και για το πρώτο πρόβλημα που παρουσιάστηκε στην ενότητα 5.3, αναλυτικά ως εξής:

$$\frac{\partial f_i}{\partial x_j} = 2 \left[(g_1 - a_i) \frac{\partial g_1}{\partial x_j} + (g_2 - b_i) \frac{\partial g_2}{\partial x_j} \right], \quad i, j = 1, 2$$

$$\frac{\partial^2 f_i}{\partial x_j \partial x_k} = 2 \left[\frac{\partial g_1}{\partial x_j} \frac{\partial g_1}{\partial x_k} + (g_1 - a_i) \frac{\partial^2 g_1}{\partial x_j \partial x_k} + \frac{\partial g_2}{\partial x_j} \frac{\partial g_2}{\partial x_k} + (g_2 - b_i) \frac{\partial^2 g_2}{\partial x_j \partial x_k} \right],$$

$$i, j, k = 1, 2$$

όπου

$$\frac{\partial g_1}{\partial x_1} = -\frac{1}{x_1^2 \ln\left(1 + \frac{1}{15}x_2\right)}, \quad \frac{\partial g_1}{\partial x_2} = -\frac{\frac{1}{15}}{x_1 \left(1 + \frac{1}{15}x_2\right) \left[\ln\left(1 + \frac{1}{15}x_2\right)\right]^2}$$

$$\frac{\partial g_2}{\partial x_1} = -\frac{x_2}{10x_1^2 \ln\left(1 + \frac{1}{15}x_2\right)}$$

$$\frac{\partial g_2}{\partial x_2} = \frac{1}{10x_1} \left\{ \frac{1}{\ln\left(1 + \frac{1}{15}x_2\right)} - \frac{\frac{1}{15}x_2}{\left(1 + \frac{1}{15}x_2\right) \left[\ln\left(1 + \frac{1}{15}x_2\right)\right]^2} \right\}$$

$$\frac{\partial^2 g_1}{\partial x_1^2} = \frac{2}{x_1^3 \ln\left(1 + \frac{1}{15}x_2\right)}, \quad \frac{\partial^2 g_1}{\partial x_1 \partial x_2} = \frac{\frac{1}{15}}{x_1^2 \left(1 + \frac{1}{15}x_2\right) \left[\ln\left(1 + \frac{1}{15}x_2\right)\right]^2}$$

$$\frac{\partial^2 g_1}{\partial x_1 \partial x_2} = \frac{\partial^2 g_1}{\partial x_2 \partial x_1}, \quad \frac{\partial^2 g_1}{\partial x_2^2} = \frac{1}{x_1} \left(\frac{\frac{1}{15}}{1 + \frac{1}{15}x_2} \right)^2 \frac{\ln\left(1 + \frac{1}{15}x_2\right) + 2}{\left[\ln\left(1 + \frac{1}{15}x_2\right)\right]^3}$$

$$\frac{\partial^2 g_2}{\partial x_1^2} = \frac{2x_2}{10x_1^3 \ln\left(1 + \frac{1}{15}x_2\right)}$$

$$\frac{\partial^2 g_2}{\partial x_1 \partial x_2} = -\frac{1}{10x_1^2} \left\{ \frac{1}{\ln\left(1 + \frac{1}{15}x_2\right)} - \frac{\frac{1}{15}x_2}{\left(1 + \frac{1}{15}x_2\right) \left[\ln\left(1 + \frac{1}{15}x_2\right)\right]^2} \right\} = \frac{\partial^2 g_2}{\partial x_2 \partial x_1}$$

$$\frac{\partial^2 g_2}{\partial x_2^2} =$$

$$= \frac{1}{10x_1} \frac{\left(\frac{1}{15}\right)^2 x_2 \ln\left(1 + \frac{1}{15}x_2\right) - \left[\ln\left(1 + \frac{1}{15}x_2\right) - \frac{\frac{1}{15}x_2}{1 + \frac{1}{15}x_2}\right] 2 \frac{\frac{1}{15}x_2}{1 + \frac{1}{15}x_2}}{\left[\ln\left(1 + \frac{1}{15}x_2\right)\right]^3}$$

5.4.4 Παρουσίαση αποτελεσμάτων

5.4.4.1 Περίπτωση αφαίρεσης σταδίου διόρθωσης

Έχοντας καλή γνώση του προβλήματος προς επίλυση, λαμβάνεται η απόφαση να τροφοδοτηθεί στο πρόγραμμα ως αρχική εκτίμηση για το \hat{f}_2 η τιμή 0,25. Μία αρχική επιλογή των παραμέτρων που σχετίζονται με την ALM δύναται να είναι η εξής: $method = 1$ (μέθοδος απότομης καθόδου), $\vec{x}^0 = [5 \ 5]^T$, $\hat{\lambda}_2^0 = 0$, $\omega_p^0 = 0,1$, $\gamma = 1,5$, $\omega_p^{max} = 500$ και $\eta = 0,1$. Επιπλέον, έστω ότι στην αρχική αυτή φάση επιλέγεται $n_{elit} = 30$ και $\hat{f}_2^{down} = 0,01$. Στο παρόν σημείο, κρίνεται σκόπιμο να παρατεθούν τα αποτελέσματα του προγράμματος όσον αφορά το πρώτο σημείο του μετώπου Pareto που λαμβάνεται μέσω της ALM (με δεδομένες τις προαναφερθείσες τιμές των παραμέτρων) για διαφορετικές τιμές των ανοχών με τις οποίες ικανοποιούνται τα κριτήρια σύγκλισης της μεθόδου της απότομης καθόδου και της ALM. Τα εν λόγω αποτελέσματα συναντώνται στον πίνακα 5.7, συνοδευόμενα από τις συνολικές αξιολογήσεις που πραγματοποιούνται από το πρόγραμμα (δηλαδή συνυπολογίζονται και οι αξιολογήσεις που γίνονται κατά τη διάρκεια της εύρεσης των υπολοίπων σημείων του Pareto) σε κάθε περίπτωση.

Ανοχή ALM	Ανοχή SD	x_1	x_2	f_1	f_2	Αξιολογήσεις
1,00E-06	1,00E-06	2,852645	7,926304	0,265999E-03	0,250001	98927
1,00E-06	1,00E-05	2,852681	7,926379	0,266000E-03	0,250001	63237
1,00E-06	1,00E-04	2,853017	7,927093	0,266032E-03	0,250001	30181
1,00E-06	1,00E-03	2,856024	7,933390	0,269048E-03	0,250000	9301
1,00E-06	1,00E-02	2,842669	7,904647	0,293328E-03	0,249999	4433
1,00E-05	1,00E-04	2,853007	7,927174	0,265876E-03	0,250005	30165
1,00E-04	1,00E-04	2,852933	7,927585	0,265011E-03	0,250032	30089
1,00E-03	1,00E-04	2,851358	7,935274	0,248411E-03	0,250550	29077

Πίνακας 5.7: Πρώτη μη-κυριαρχούμενη λύση και συνολικές αξιολογήσεις που πραγματοποιούνται για την εύρεση του μετώπου Pareto για διαφορετικές τιμές των ανοχών με τις οποίες ικανοποιούνται τα κριτήρια σύγκλισης της μεθόδου της απότομης καθόδου και της ALM όταν $\hat{f}_2^{up} = 0,25$, $\bar{x}^0 = [5 \ 5]^T$, $\hat{\lambda}_2^0 = 0$, $\omega_p^0 = 0,1$, $\gamma = 1,5$, $\omega_p^{max} = 500$, $\eta = 0,1$ και $n_{elit} = 30$.

Στη δεδομένη περίπτωση, λαμβάνοντας υπόψη τα δεδομένα του παραπάνω πίνακα, έστω ότι επιβάλλονται ως ικανοποιητικές ανοχές για τα κριτήρια σύγκλισης της μεθόδου της απότομης καθόδου και της ALM οι τιμές 10^{-4} και 10^{-5} αντίστοιχα. Για τις επιλεχθείσες ανοχές αλλά και για τις ανωτέρω αναφερθείσες παραμέτρους, όπως φαίνεται στον πίνακα 5.7, οι συνολικές αξιολογήσεις που χρειάζονται για την απόκτηση του μετώπου Pareto είναι 30165. Στο σημείο αυτό, λαμβάνει χώρα διερεύνηση με στόχο να επιδειχθεί το πώς συμπεριφέρεται το πρόγραμμα όσον αφορά το υπολογιστικό κόστος με τη μεταβολή των παραμέτρων της ALM. Τα αποτελέσματα της αναφερθείσας διερεύνησης επιδεικνύονται στον πίνακα 5.8 (σημειώνεται ότι οι αναγραφόμενες αξιολογήσεις είναι οι συνολικές, ενώ αν επιθυμείται να βρεθούν αυτές που αντιστοιχούν στον υπολογισμό του πρώτου σημείου του μετώπου Pareto πρέπει να αφαιρεθούν οι αξιολογήσεις του τμήματος που υπολογίζει τα υπόλοιπα 29 σημεία που είναι σταθερές σε όλη τη διαδικασία και ίσες με $(n_{elit} - 1)(1 + n + 1) = 116$).

$\hat{\lambda}_2^0$	ω_p^0	γ	η	$(\bar{x}^0)^T$	Αξιολογήσεις
0	0,1	1,5	0,1	[5 5]	30165
0	0,1	1,5	0,5	[5 5]	6131
0	0,1	1,5	1	[5 5]	3129
0	0,1	1,5	2	[5 5]	1627
0	0,1	1,5	4	[5 5]	883
0	0,1	1,5	4,9	[5 5]	759
0	0,1	1,5	4,9	[1 1]	611
0	0,1	1,5	5	[5 5]	δεν συγκλίνει
0	0,2	1,5	4,9	[5 5]	659
0	0,5	1,5	4,9	[5 5]	δεν συγκλίνει
0	0,5	1,5	4,4	[5 5]	565
0	1	1,5	4,4	[5 5]	443
0	1,1	1,5	4,4	[5 5]	427
0	1,2	1,5	4,4	[5 5]	δεν συγκλίνει
0	1,2	1,5	4,1	[5 5]	437
0	5	1,1	4,4	[5 5]	261
0	6	1,01	4,4	[5 5]	255
0,1	6	1,01	4,4	[5 5]	307
0	6	1,01	4,4	[10 10]	255
0	6	1,01	4,4	[5 10]	237

Πίνακας 5.8: Συνολικές αξιολογήσεις που πραγματοποιούνται για την εύρεση του μετώπου Pareto για διαφορετικές τιμές των παραμέτρων της ALM όταν οι ανοχές για τα κριτήρια σύγκλισης της μεθόδου της απότομης καθόδου και της ALM λαμβάνουν τις τιμές 10^{-4} και 10^{-5} αντίστοιχα, $\hat{f}_2^{up} = 0,25$ και $n_{elit} = 30$.

Η προσπάθεια για ελάττωση των αξιολογήσεων που απαιτούνται για τη σύγκλιση της ALM (άρα κατά συνέπεια και του προγράμματος στο σύνολό του), όπως αυτή απεικονίζεται στον παραπάνω πίνακα, ξεκίνησε με σταδιακή αύξηση του η , ενώ οι τιμές $\hat{\lambda}_2^0$, ω_p^0 και γ διατηρήθηκαν σταθερές. Μέσω του πίνακα 5.8, γίνεται αντιληπτό ότι με την αύξηση του η επιτυγχάνεται δραματική μείωση των αξιολογήσεων. Επίσης, είναι φανερό η δυνατότητα αύξησης του η σε μεγάλο βαθμό (μέχρι 4,9) πριν η μέθοδος φτάσει στο σημείο να μην συγκλίνει. Έπειτα, έγινε σταδιακή αύξηση του ω_p^0 με ταυτόχρονη αναγκαστική μείωση του η (λόγω μη-σύγκλισης) και παρατηρήθηκε ότι οι αξιολογήσεις βαίνουν μειούμενες μέχρι το $\omega_p^0 = 1,1$, ενώ στη συνέχεια αυξάνονται. Στη συνέχεια, πραγματοποιήθηκε μείωση του γ , γεγονός που έδωσε δυνατότητα για σημαντική αύξηση του ω_p^0 με αποτέλεσμα τη μείωση των αξιολογήσεων. Τέλος, παρατηρήθηκε ότι ακόμα και μία μικρή αύξηση του $\hat{\lambda}_2^0$ οδηγεί σε αισθητή αύξηση των αξιολογήσεων. Ο καλύτερος συνδυασμός παραμέτρων που επιτεύχθηκε για $\vec{x}^0 = [5 \ 5]^T$, όπως φαίνεται από τον ανωτέρω πίνακα, είναι $\hat{\lambda}_2^0 = 0$, $\omega_p^0 = 6$, $\gamma = 1,01$ και $\eta = 4,4$. Ο συνδυασμός αυτός απαιτεί την πραγματοποίηση 255 συνολικών αξιολογήσεων (ενώ αν, λ.χ. επιλεγεί $\vec{x}^0 = [5 \ 10]^T$ για τις ίδιες παραμέτρους θα απαιτούνταν 237).

Ακολούθως, κρίνεται σκόπιμο να επιχειρηθεί η εύρεση του πρώτου σημείου του μετώπου Pareto με τη χρήση της μεθόδου Newton αντί της μεθόδου της απότομης καθόδου στο τμήμα της ALM όπου πραγματοποιείται η βελτιστοποίηση ενός μονοκριτηριακού προβλήματος χωρίς περιορισμούς. Όπως και στην περίπτωση χρήσης της μεθόδου της απότομης καθόδου, επιλέγονται $\hat{f}_2^{up} = 0,25$, $\omega_p^{max} = 500$, $n_{elit} = 30$ και $\hat{f}_2^{down} = 0,01$. Στην περίπτωση χρήσης της μεθόδου Newton παρατηρήθηκε ότι τα αποτελέσματα ήταν εξαιρετικά καλά ακόμα και αν η ανοχή με την οποία ικανοποιείται το κριτήριο σύγκλισης της μειωθεί σε 10^{-2} (οπότε και υιοθετείται), ενώ η ανοχή για το κριτήριο σύγκλισης της ALM διατηρείται σε 10^{-5} . Τα αποτελέσματα της διερεύνησης που έγινε για τη μέθοδο Newton επιδεικνύονται στον πίνακα 5.9 (σημειώνεται ότι οι αναγραφόμενες αξιολογήσεις είναι οι συνολικές, ενώ αν επιθυμείται να βρεθούν αυτές που αντιστοιχούν στον υπολογισμό του πρώτου σημείου του μετώπου Pareto πρέπει να αφαιρεθούν οι αξιολογήσεις του τμήματος που υπολογίζει τα υπόλοιπα 29 σημεία που είναι σταθερές σε όλη τη διαδικασία και ίσες με $(n_{elit} - 1)(1 + n + 1) = 116$).

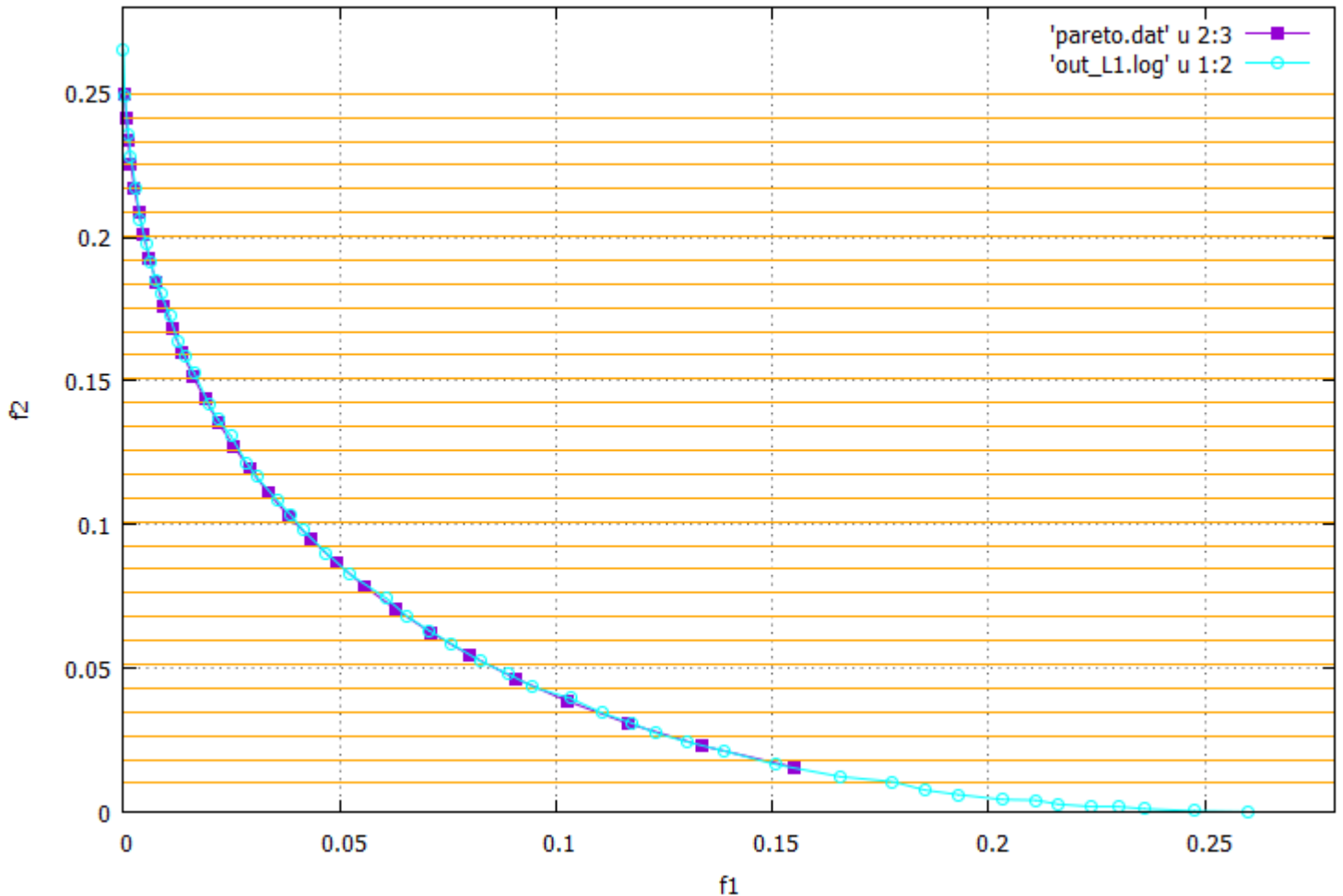
$\hat{\lambda}_2^0$	ω_p^0	γ	$(\bar{x}^0)^T$	Αξιολογήσεις
0	0,1	1,5	[5 5]	201
0	0,5	1,5	[5 5]	185
0	1	1,5	[5 5]	185
0	2	1,5	[5 5]	189
0	5	1,5	[5 5]	201
0	5	1,1	[5 5]	201
0	6	1,1	[5 5]	197
0	6	1,01	[5 5]	197
0	0,5	2	[5 5]	177
0	0,5	3	[5 5]	173
0	0,5	4	[5 5]	169
0,1	0,5	4	[5 5]	189
1	0,5	4	[5 5]	225
0	0,5	4	[2 2]	173
0	0,5	4	[1 10]	165

Πίνακας 5.9: Συνολικές αξιολογήσεις που πραγματοποιούνται για την εύρεση του μετώπου Pareto για διαφορετικές τιμές των παραμέτρων της ALM όταν οι ανοχές για τα κριτήρια σύγκλισης της μεθόδου Newton και της ALM λαμβάνουν τις τιμές 10^{-2} και 10^{-5} αντίστοιχα, $\hat{f}_2^{up} = 0,25$ και $n_{elit} = 30$.

Παρατηρείται ότι στην περίπτωση χρήσης της μεθόδου Newton, όπως δηλώνουν τα αποτελέσματα που φαίνονται στον πίνακα 5.9, μία αλλαγή σε οιαδήποτε από τις παραμέτρους που αναφέρθηκαν δεν επιφέρει μεγάλη αυξομείωση στον αριθμό των αξιολογήσεων που πραγματοποιούνται, σε αντίθεση με την περίπτωση επιλογής της μεθόδου της απότομης καθόδου όπου παρατηρήθηκαν σημαντικές αυξομειώσεις (κάτι που διαπιστώθηκε και στην περίπτωση του πρώτου μαθηματικού προβλήματος). Επίσης, είναι φανερό ότι στην εν λόγω περίπτωση επιτυγχάνονται αποτελέσματα χαμηλότερου υπολογιστικού κόστους σε σχέση με προηγούμενως.

Μετά την απόκτηση του πρώτου σημείου του μετώπου Pareto, το πρόγραμμα προβαίνει στην εύρεση και των υπολοίπων μη-κυριαρχούμενων λύσεων με βάση τις παραμέτρους n_{elit} και \hat{f}_2^{down} που ορίστηκαν παραπάνω (οι τιμές που δόθηκαν ήταν 30 και 0,01 αντίστοιχα). Το μέτωπο Pareto που προκύπτει, στη συγκεκριμένη περίπτωση, απεικονίζεται στο σχήμα 5.17. Στο ίδιο σχήμα φαίνεται και το μέτωπο Pareto που προέκυψε από το λογισμικό των εξελικτικών αλγορίθμων (οι τιμές των παραμέτρων που χρησιμοποιήθηκαν είναι ακριβώς οι ίδιες που παρατέθηκαν για το πρώτο μαθηματικό πρόβλημα στην ενότητα 5.3.4). Στο σχήμα 5.17 έχουν σχεδιαστεί, επίσης, και n_{elit} οριζόντιες γραμμές οι οποίες δηλώνουν την τιμή του \hat{f}_2 για καθένα από τα n_{elit} σημεία των οποίων ζητήθηκε η εύρεση. Αυτό έγινε για να δύναται να διαπιστωθεί αν

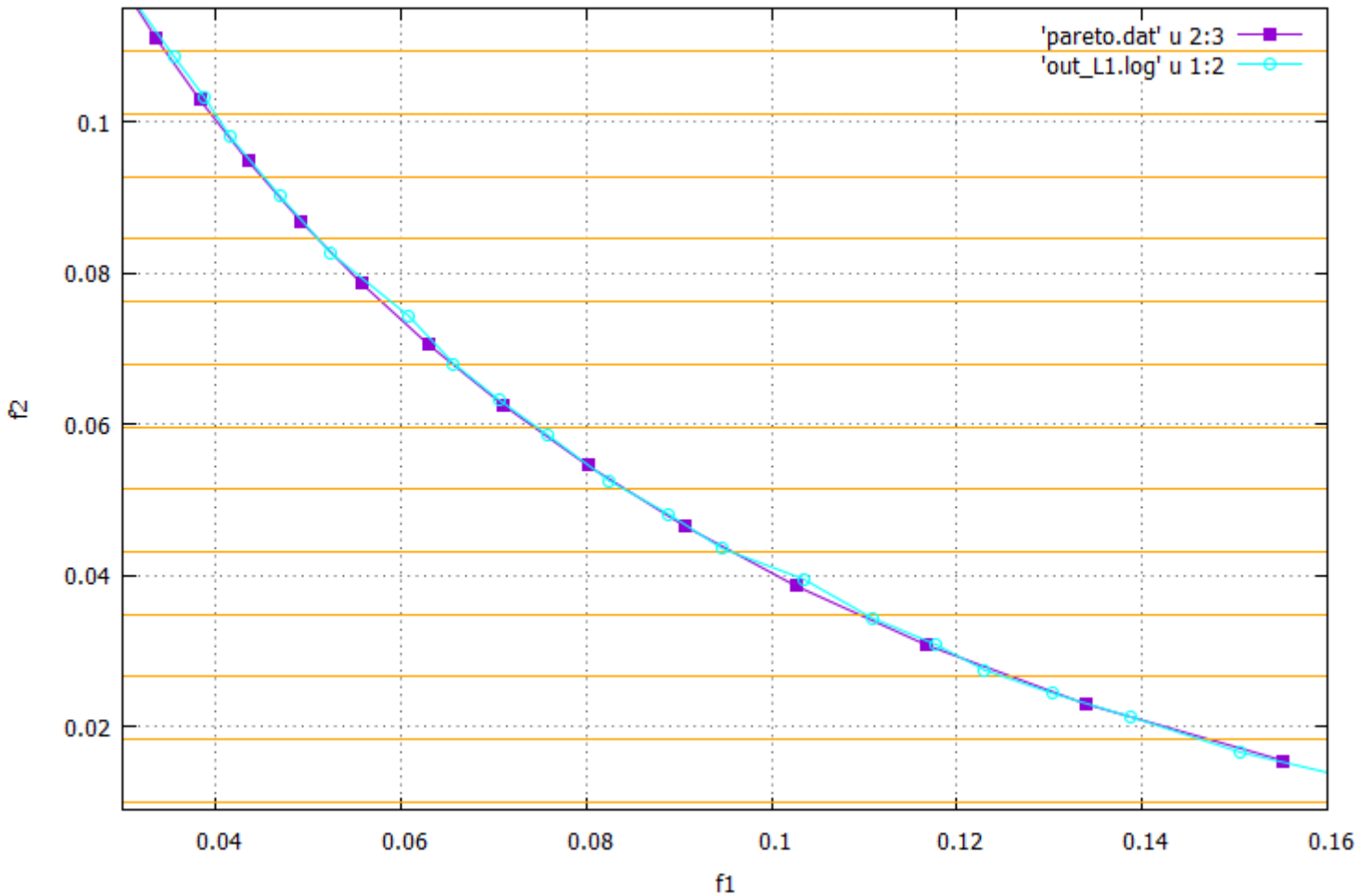
υπάρχει διαφορά (και πόση είναι αυτή) μεταξύ της τιμής του \hat{f}_2 και του f_2 για κάθε σημείο του Pareto που προέκυψε από το πρόγραμμα.



Σχήμα 5.17: Μέτωπο Pareto που προέκυψε από τη μέθοδο που βασίζεται στη συνέχιση της καμπύλης των μη-κυριαρχούμενων λύσεων για $\hat{f}_2^{up} = 0,25$, $\hat{f}_2^{down} = 0,01$ και $n_{elit} = 30$ (μωβ καμπύλη) και μέτωπο Pareto που προέκυψε από το λογισμικό των εξελικτικών αλγορίθμων (γαλάζια καμπύλη). Οι οριζόντιες γραμμές χρώματος πορτοκαλί αντιπροσωπεύουν τις τιμές των \hat{f}_2 για καθένα από τα n_{elit} σημεία που αναζητήθηκαν.

Από την παρατήρηση του σχήματος 5.17, κατά πρώτο λόγο, καθίσταται φανερό (όπως και στην περίπτωση του πρώτου μαθηματικού προβλήματος) το γεγονός ότι στα πρώτα σημεία που αποκτήθηκαν από το πρόγραμμα οι διαφορές μεταξύ των τιμών του \hat{f}_2 και του f_2 είναι πολύ μικρές (πρακτικά μπορούν να θεωρηθούν αμελητέες), ενώ καθώς ο αριθμός των υπολογισθέντων σημείων μεγαλώνει οι διαφορές αυτές αυξάνονται και γίνονται όλο και πιο αισθητές. Αξίζει να αναφερθεί ότι στο δεύτερο αποκτηθέν σημείο ισχύει $f_2 - \hat{f}_2 \cong 0,000081$, στο δέκατο $f_2 - \hat{f}_2 \cong 0,000789$, ενώ στο τελευταίο είναι $f_2 - \hat{f}_2 \cong 0,005478$. Οι αιτίες εμφάνισης των εν λόγω διαφορών εξηγήθηκαν λεπτομερώς στην ενότητα 5.3.4 και δεν κρίνεται σκόπιμο να επαναληφθούν εδώ.

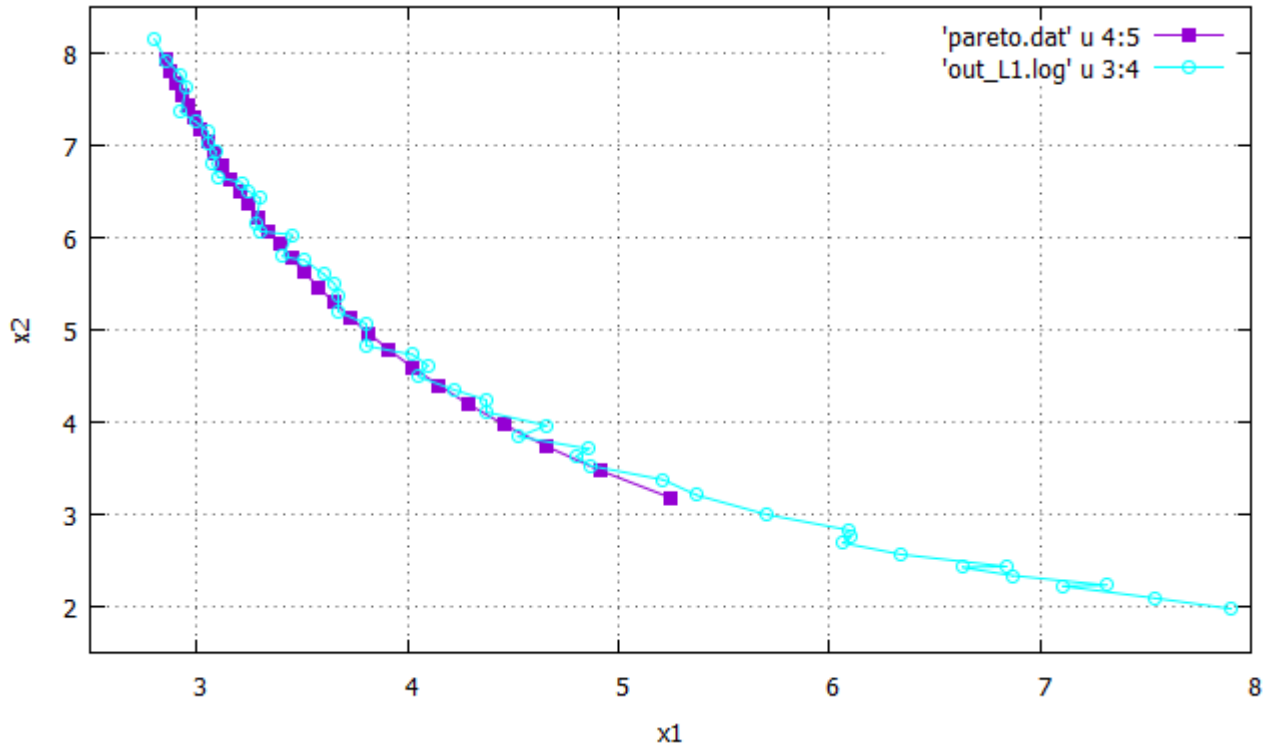
Μία δεύτερη διαπίστωση που θα μπορούσε να γίνει μετά την παρατήρηση του σχήματος 5.17 είναι ότι το μέτωπο Pareto που αποκτήθηκε με βάση τη μέθοδο της συνέχισης της καμπύλης των μη-κυριαρχούμενων λύσεων αποτελεί ένα τμήμα του πραγματικού μετώπου που αποκτήθηκε μέσω του λογισμικού των εξελικτικών αλγορίθμων, καθώς η γαλάζια καμπύλη εκτείνεται προς αμφότερες τις πλευρές της μωβ (κυρίως προς την πλευρά των μικρότερων f_2). Πρέπει, όμως, να ειπωθεί ότι το αποκτηθέν τμήμα καλύπτει πολύ μεγάλο κομμάτι του μετώπου Pareto που δόθηκε από το λογισμικό των εξελικτικών αλγορίθμων και μάλιστα τα σημεία που ως επί το πλείστον 'ενδιαφέρουν' τον κληθέντα να επιλύσει το πρόβλημα δεν περιλαμβάνονται στην περιοχή που δεν αποκτήθηκε, καθώς η τελική επιλογή προς κατασκευή γενικά δεν πραγματοποιείται από τα ακραία σημεία του μετώπου. Επαναλαμβάνεται ότι τα όρια \hat{f}_2^{up} και \hat{f}_2^{down} επιλέγονται από τον κάθε επιστήμονα σύμφωνα με τις ανάγκες του. Επιπλέον, το σχήμα 5.17 οδηγεί στο συμπέρασμα ότι η καμπύλη που προέκυψε από τη μέθοδο που βασίζεται στη συνέχιση της καμπύλης των μη-κυριαρχούμενων λύσεων για $\hat{f}_2^{up} = 0,25$, $\hat{f}_2^{down} = 0,01$ και $n_{elit} = 30$, πρακτικά, πέφτει πάνω σε αυτή που προέκυψε από το λογισμικό των εξελικτικών αλγορίθμων. Για να φανεί αυτό με μεγαλύτερη ευκρίνεια δημιουργείται το σχήμα 5.18 στο οποίο απεικονίζονται μόνο τα 13 τελευταία σημεία που προέκυψαν από το πρόγραμμα (επιλέχθηκαν τα τελευταία σημεία καθώς αυτά παρουσιάζουν τις μεγαλύτερες διαφορές $f_2 - \hat{f}_2$).



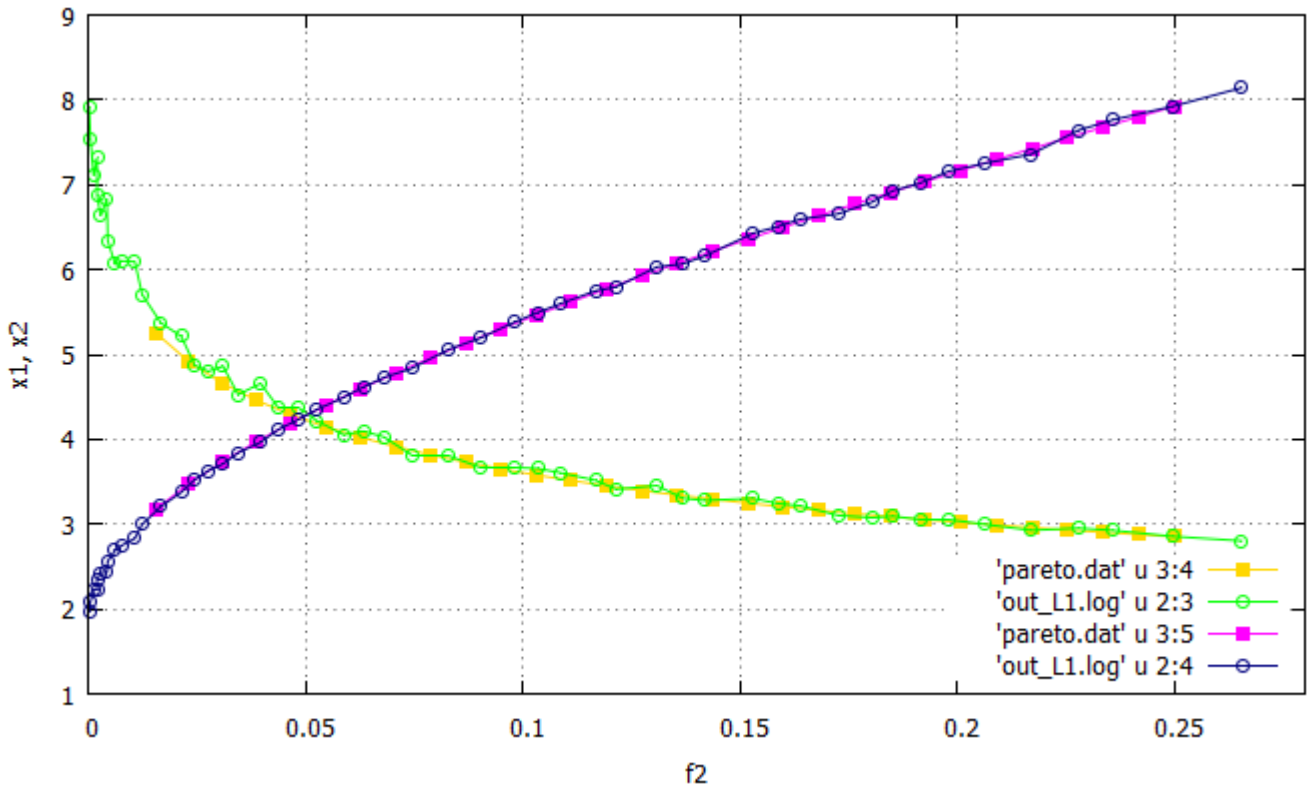
Σχήμα 5.18: Τελευταίες 13 μη-κυριαρχούμενες λύσεις που προέκυψαν από τη μέθοδο που βασίζεται στη συνέχιση της καμπύλης των μη-κυριαρχούμενων λύσεων για $\hat{f}_2^{up} = 0,25$, $\hat{f}_2^{down} = 0,01$ και $n_{elit} = 30$ (μωβ καμπύλη) και αντίστοιχο τμήμα μετώπου Pareto που προέκυψε από το λογισμικό των εξελικτικών αλγορίθμων (γαλάζια καμπύλη). Οι οριζόντιες γραμμές χρώματος πορτοκαλί αντιπροσωπεύουν τις τιμές των \hat{f}_2 για καθένα από τα 13 τελευταία σημεία που αναζητήθηκαν.

Με τη βοήθεια του σχήματος 5.18 καθίσταται φανερό το γεγονός ότι η καμπύλη που προέκυψε με βάση τη συνέχιση του μετώπου των μη-κυριαρχούμενων λύσεων για τις παραμέτρους που αναφέρθηκαν παραπάνω πράγματι συμπίπτει, πρακτικά, με το αντίστοιχο τμήμα του μετώπου Pareto που προέκυψε από το λογισμικό των εξελικτικών αλγορίθμων. Πρέπει, επίσης, να αναφερθεί ότι η καμπύλη που προέκυψε από το αναπτυχθέν πρόγραμμα είναι ομαλή (όπως ήταν αναμενόμενο λόγω του ορισμού της μεθόδου) σε αντίθεση με αυτή που αποκτήθηκε μέσω του εξελικτικού αλγορίθμου (λόγω του τρόπου με τον οποίο δημιουργούνται τα νέα άτομα από την εξελικτική διαδικασία), γεγονός που οδηγεί την πρώτη καμπύλη να βρίσκεται κάτω από τη δεύτερη σε κάποια μικρά τμήματα (δηλαδή έχει επιτευχθεί η εύρεση ‘καλύτερων’ σημείων στα τμήματα αυτά). Αυτό είναι ένα πλεονέκτημα της μεθόδου που πραγματεύεται η παρούσα εργασία. Επιπροσθέτως, τονίζεται ότι η αναφερθείσα πρακτική ταύτιση

των καμπυλών ήταν δυνατή παρά το γεγονός ότι, όπως σχολιάστηκε παραπάνω, η τιμή της διαφοράς $f_2 - \hat{f}_2$ συνεχώς αυξάνεται καθώς ο αλγόριθμος υπολογίζει όλο και περισσότερα σημεία. Στο σημείο αυτό κρίνεται απαραίτητη η κατασκευή των διαγραμμάτων $x_1 - x_2$ (σχήμα 5.19), $f_2 - x_1$ και $f_2 - x_2$ (σχήμα 5.20) που προέκυψαν από τις δύο μεθόδους.



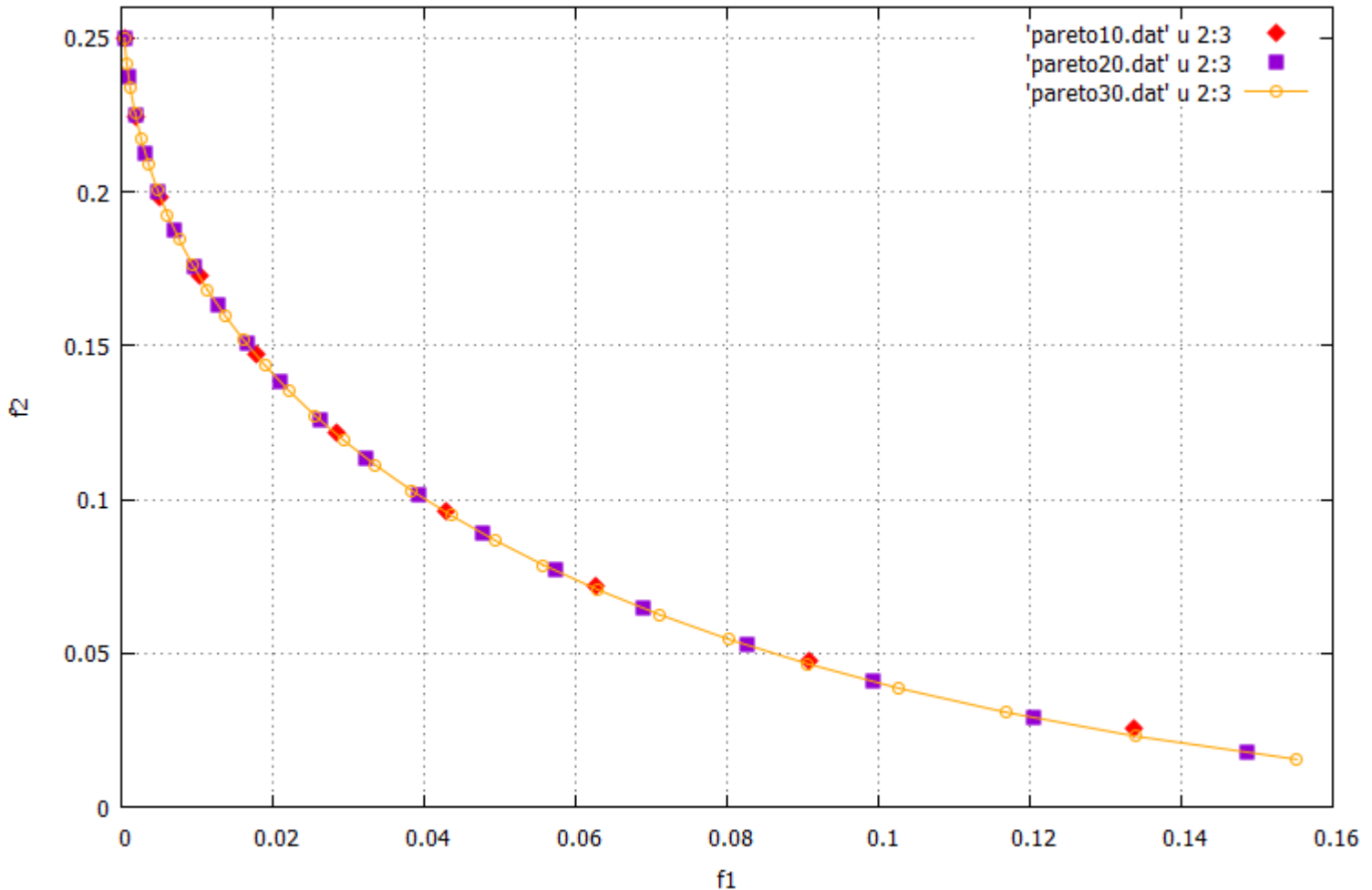
Σχήμα 5.19: Διάγραμμα των καμπυλών $x_1 - x_2$ που προέκυψαν από τη μέθοδο που βασίζεται στη συνέχιση της καμπύλης των μη-κυριαρχούμενων λύσεων για $\hat{f}_2^{up} = 0,25$, $\hat{f}_2^{down} = 0,01$ και $n_{elit} = 30$ (μωβ καμπύλη) και από το λογισμικό των εξελικτικών αλγορίθμων (γαλάζια καμπύλη).



Σχήμα 5.20: Καμπύλες $f_2 - x_1$ και $f_2 - x_2$ που προέκυψαν από τη μέθοδο που βασίζεται στη συνέχιση της καμπύλης των μη-κυριαρχούμενων λύσεων για $\hat{f}_2^{up} = 0,25$, $\hat{f}_2^{down} = 0,01$ και $n_{elit} = 30$ (κίτρινη και φούξια καμπύλη αντίστοιχα) και από το λογισμικό των εξελικτικών αλγορίθμων (πράσινη και μπλε καμπύλη αντίστοιχα).

Παρατηρείται ότι στα διαγράμματα $x_1 - x_2$ (σχήμα 5.19), $f_2 - x_1$ και $f_2 - x_2$ (σχήμα 5.20) απεικονίζονται καμπύλες γραμμές, σε αντίθεση με την περίπτωση του πρώτου μαθηματικού προβλήματος όπου στα διαγράμματα $x_1 - x_2$ (σχήμα 5.8) και $f_2 - x_2$ (σχήμα 5.10) πρακτικά απεικονίζονταν οριζόντιες ευθείες. Η ανανέωση των τιμών των μεταβλητών σχεδιασμού, όπως δηλώθηκε παραπάνω, γίνεται με βάση τη σχέση 5.1 η οποία αποτελεί μία προσέγγιση πρώτης τάξης στο χώρο $\hat{f}_2 - \vec{z}$ και δεν λαμβάνει υπόψη την καμπυλότητα. Δεδομένου, λοιπόν, ότι οι γραφικές παραστάσεις $f_2 - x_1$ και $f_2 - x_2$ είναι καμπύλες γραμμές, η προσέγγιση που γίνεται ενέχει σφάλμα και για τις δύο μεταβλητές σχεδιασμού (σε αντίθεση με το πρώτο μαθηματικό πρόβλημα όπου το σφάλμα υπήρχε μόνο στον υπολογισμό της πρώτης μεταβλητής). Το σφάλμα που αφορά στις δύο μεταβλητές σχεδιασμού σαφέστατα 'μεταφέρεται' και στον υπολογισμό των συναρτήσεων στόχου. Η παραπάνω διαπιστωθείσα ολοένα αυξανόμενη τιμή της διαφοράς $f_2 - \hat{f}_2$, λοιπόν, οφείλεται σε 'μεταφερόμενα' σφάλματα στον υπολογισμό και των δύο μεταβλητών σχεδιασμού. Αυτό σημαίνει ότι σε αντιπαράθεση με το πρώτο μαθηματικό πρόβλημα, τα προκύπτοντα από το πρόγραμμα σημεία δεν είναι απαραίτητο να βρίσκονται πάνω στο μέτωπο Pareto. Παρόλα αυτά, παρατηρήθηκε μέσω των σχημάτων 5.17 και 5.18 ότι κάτι τέτοιο

επετεύχθη. Η απόκτηση σημείων του πραγματικού μετώπου Pareto (τα οποία έχουν λίγο μικρότερο f_2 από αυτό που στοχεύθηκε) ήταν δυνατή λόγω της επιλογής κατάλληλου-ικανοποιητικά μικρού $\delta\hat{f}_2$ (μέσω της επιλογής n_{elit}). Στο σημείο αυτό κρίνεται ενδιαφέρον να παρουσιαστεί η συμπεριφορά του προγράμματος στην περίπτωση μικρότερου n_{elit} (άρα μεγαλύτερου $\delta\hat{f}_2$). Στο σχήμα 5.21 φαίνονται τα σημεία που δίνει το πρόγραμμα για τις περιπτώσεις όπου το n_{elit} παίρνει τις τιμές 30, 20 και 10 (ισχύουν, όπως προηγουμένως, $\hat{f}_2^{up} = 0,25$ και $\hat{f}_2^{down} = 0,01$).



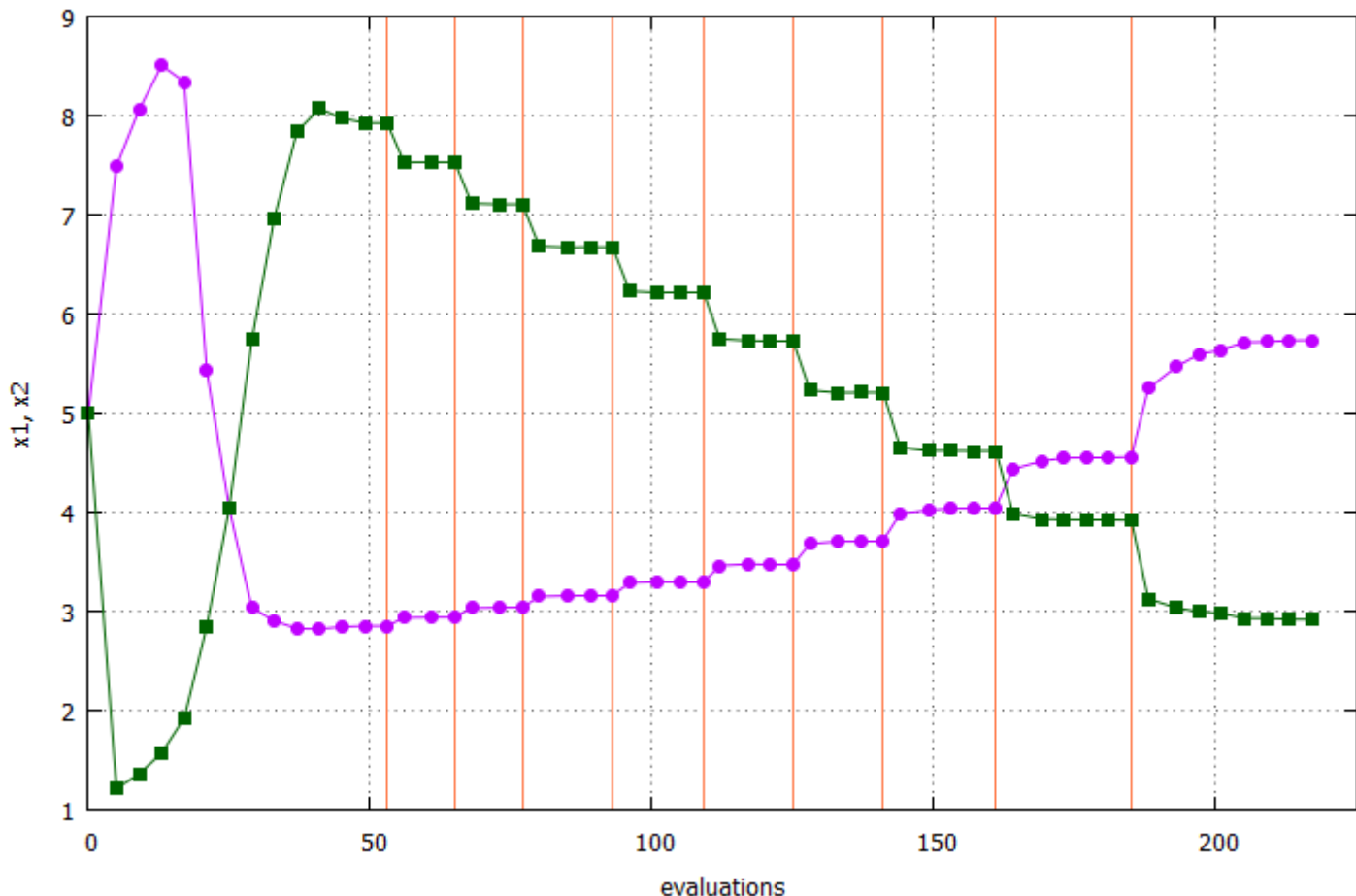
Σχήμα 5.21: Μη-κυριαρχούμενες λύσεις που προέκυψαν από τη μέθοδο που βασίζεται στη συνέχιση του μετώπου Pareto για $n_{elit} = 30$ (πορτοκαλί χρώμα), για $n_{elit} = 20$ (μωβ χρώμα) και για $n_{elit} = 10$ (κόκκινο χρώμα). Σημειώνεται ότι σε όλες τις περιπτώσεις ισχύουν $\hat{f}_2^{up} = 0,25$ και $\hat{f}_2^{down} = 0,01$.

Αν παρατηρηθεί το σχήμα 5.21, καθίσταται φανερό το γεγονός ότι όσο το n_{elit} μειώνεται τόσο μεγαλύτερη είναι η διαφορά $f_2 - \hat{f}_2$ για τα υπολογισθέντα σημεία (χαρακτηριστικά το τελευταίο σημείο βρίσκεται όλο και υψηλότερα στον άξονα των f_2). Ωστόσο, φαίνεται ότι στην περίπτωση όπου $n_{elit} = 20$ τα υπολογισθέντα σημεία παραμένουν πάνω στο Pareto, ενώ οι συνολικές αξιολογήσεις που απαιτούνται μειώνονται κατά 40 (αφού για κάθε σημείο είναι απαραίτητες

$1 + n + 1 = 4$ αξιολογήσεις) σε σχέση με την περίπτωση όπου $n_{elit} = 30$ (αν ληφθούν $method = 2$, $\vec{x}^0 = [5 \ 5]^T$, $\hat{\lambda}_2^0 = 0$, $\omega_p^0 = 0,5$, $\gamma = 4$ και $\omega_p^{max} = 500$, το αποτέλεσμα επιτυγχάνεται μόνο με 129 αξιολογήσεις). Όσον αφορά την περίπτωση για την οποία ισχύει $n_{elit} = 10$, διαπιστώνεται ότι τα αρχικά σημεία παραμένουν πάνω στο μέτωπο Pareto αλλά τα τελευταία απομακρύνονται από αυτό (πρέπει να σημειωθεί ότι τα σημεία που βρίσκονται στα τμήματα των καμπυλών $f_2 - x_1$ και $f_2 - x_2$ που αντιστοιχούν σε μεγαλύτερες τιμές f_2 , όπως φαίνεται στο σχήμα 5.20, έχουν μικρότερη καμπυλότητα σε σχέση με αυτά που βρίσκονται στα τμήματα μικρότερων f_2 , κάτι που σημαίνει ότι το σφάλμα αυξάνεται για τα δεύτερα). Συγκεκριμένα, τα 3 τελευταία σημεία που υπολογίσθηκαν είναι αισθητά πάνω από το μέτωπο Pareto, χωρίς οι διαφορές να είναι πολύ μεγάλες. Βέβαια, στην εν λόγω περίπτωση οι πραγματοποιηθείσες αξιολογήσεις μειώνονται κατά 80 σε σχέση με την περίπτωση όπου $n_{elit} = 30$ (για τον ανωτέρω συνδυασμό παραμέτρων το αποτέλεσμα επιτυγχάνεται μόνο με 89 αξιολογήσεις). Επιπλέον, όπως αναφέρθηκε στην ενότητα 5.3.4, είναι δυνατό ο χρήστης να επιθυμεί την απόκτηση μικρότερου τμήματος του μετώπου Pareto οπότε μειώνει τη διαφορά $\hat{f}_2^{up} - \hat{f}_2^{down}$ και έτσι η περίπτωση για την οποία ισχύει $n_{elit} = 10$ θα δύναται να αποδώσει ακριβέστερα αποτελέσματα. Εν τέλει, όλα εξαρτώνται από τις ανάγκες του χρήστη σχετικά με το πόσα σημεία θα αποκτηθούν πάνω στο μέτωπο Pareto, την περιοχή του μετώπου που αυτά θα καλύπτουν, την ακρίβεια και το υπολογιστικό κόστος.

5.4.4.2 Μέθοδος πρόβλεψης-διόρθωσης

Στην παρούσα ενότητα, επιχειρείται η εύρεση του μετώπου Pareto με την προσθήκη του σταδίου της διόρθωσης μετά από κάθε σημείο το οποίο 'προβλέπεται'. Μετά από παρατήρηση των πινάκων που παρατέθηκαν στην ενότητα που προηγήθηκε, λαμβάνεται η απόφαση να επιλεχθεί ο ακόλουθος συνδυασμός των δεδομένων που εισάγονται στο πρόγραμμα: $\hat{f}_2^{up} = 0,25$, $\hat{\lambda}_2^0 = 0$, $\omega_p^0 = 0,5$, $\gamma = 5$, $\omega_p^{max} = 500$, $method = 2$ (μέθοδος Newton), $n_{elit} = 10$, $\hat{f}_2^{down} = 0,01$ και $\vec{x}^0 = [5 \ 5]^T$. Επίσης, επιλέγονται ως επιθυμητές ανοχές για τα κριτήρια σύγκλισης της μεθόδου Newton και της ALM οι τιμές 10^{-2} και 10^{-5} αντίστοιχα. Στην περίπτωση αυτή, το πρόγραμμα απαιτεί την πραγματοποίηση 217 αξιολογήσεων για να δώσει τις 10 ζητηθείσες μη-κυριαρχούμενες λύσεις. Για τον εν λόγω συνδυασμό παραμέτρων, κρίνεται σκόπιμο να παρατεθεί ένα διάγραμμα στο οποίο παριστάνεται η πορεία που ακολούθησαν οι μεταβλητές σχεδιασμού κατά τη διάρκεια της εκτέλεσης του προγράμματος. Το διάγραμμα αυτό αποτελεί το σχήμα 5.22.

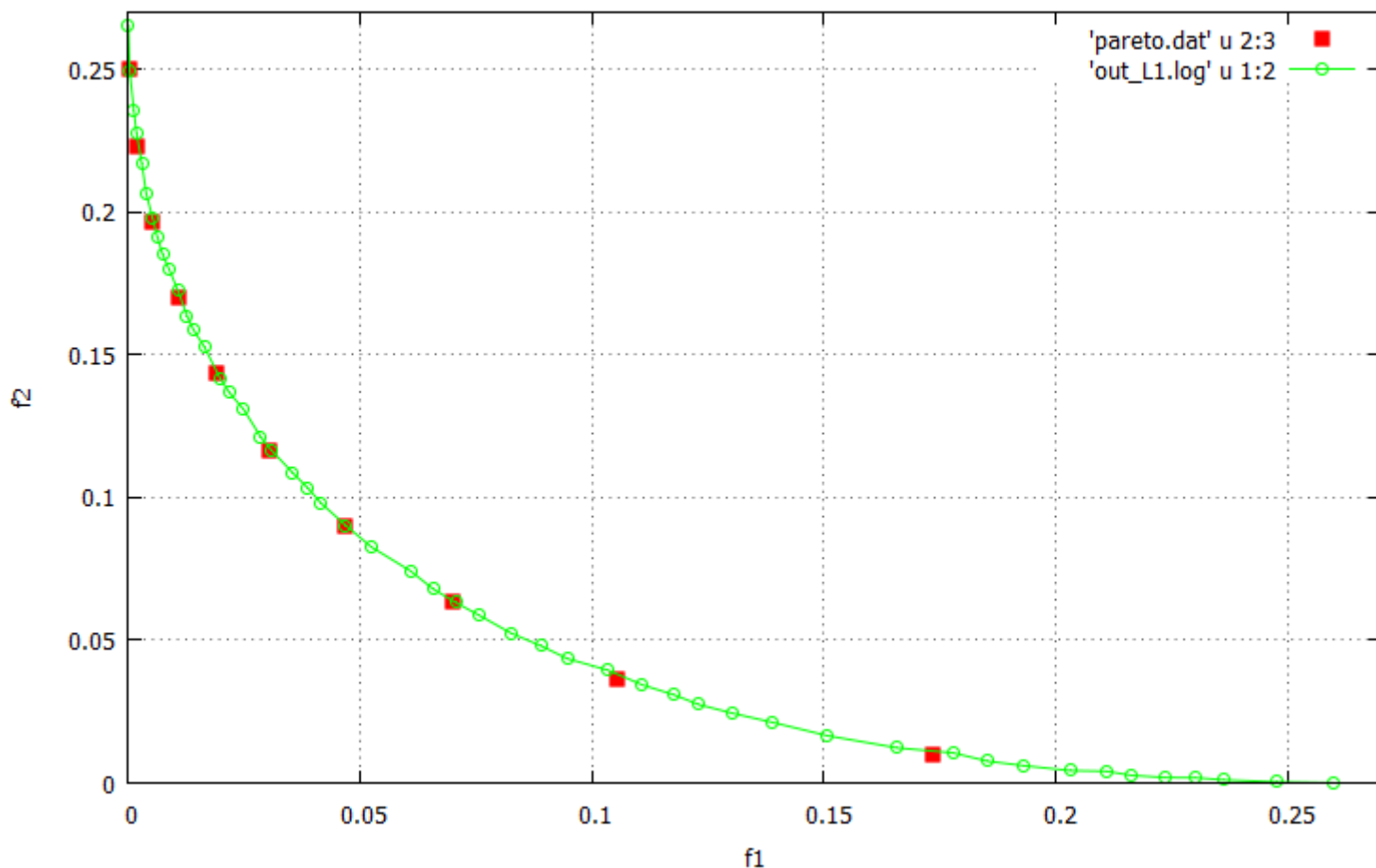


Σχήμα 5.22: Διάγραμμα των μεταβλητών σχεδιασμού x_1 (μωβ καμπύλη) και x_2 (πράσινη καμπύλη) συναρτήσει των επαναλήψεων που εκτελούνται από το πρόγραμμα για $\hat{f}_2^{up} = 0,25$, $\hat{f}_2^{down} = 0,01$ και $n_{elit} = 10$. Οι κατακόρυφες γραμμές χρώματος πορτοκαλί χωρίζουν τα 10 τμήματα που αντιστοιχούν στα 10 σημεία του μετώπου Pareto που ευρέθησαν.

Μέσω του σχήματος 5.22, παρατηρείται ότι η απόκτηση της πρώτης μη-κυριαρχούμενης λύσης απαιτεί την πραγματοποίηση αισθητά περισσότερων αξιολογήσεων από ότι η εύρεση οιασδήποτε από τις υπόλοιπες, γεγονός που επιβεβαιώνει το σκεπτικό πάνω στο οποίο ‘χτίστηκε’ η μέθοδος πρόβλεψης-διόρθωσης που περιγράφηκε στο κεφάλαιο 4 (όπως επιβεβαιώθηκε και στην περίπτωση του πρώτου μαθηματικού προβλήματος). Βέβαια, πρέπει να αναφερθεί ότι οι απαιτούμενες αξιολογήσεις για την εύρεση των τελευταίων σημείων είναι περισσότερες από αυτές που αντιστοιχούν στα πρώτα σημεία (εξαιρουμένου, βέβαια, του πρώτου). Αυτό συμβαίνει διότι, όπως επιδείχθηκε στην προηγούμενη ενότητα (σχήμα 5.20), η καμπυλότητα (στο χώρο όπου πραγματοποιείται η διόρθωση) για τα τελευταία σημεία είναι μεγαλύτερη και άρα η προσέγγιση πρώτης τάξης που γίνεται (σχέση 5.1) ενέχει μεγαλύτερο σφάλμα.

Το σημεία του μετώπου Pareto που προκύπτουν, στη συγκεκριμένη περίπτωση, απεικονίζονται στο σχήμα 5.23. Στο εν λόγω σχήμα έχει

σχεδιαστεί, επίσης, και το μέτωπο μη-κυριαρχούμενων λύσεων που αποκτήθηκε μέσω του λογισμικού των εξελικτικών αλγορίθμων.



Σχήμα 5.23: Σημεία μετώπου Pareto που προέκυψαν από τη μέθοδο πρόβλεψης-διόρθωσης για $\hat{f}_2^{up} = 0,25$, $\hat{f}_2^{down} = 0,01$ και $n_{elit} = 10$ (μεγάλα κόκκινα σημεία) και μέτωπο Pareto που προέκυψε από το λογισμικό των εξελικτικών αλγορίθμων (πράσινη καμπύλη).

Αν παρατηρηθεί το σχήμα 5.23, γίνεται αντιληπτό ότι σε κάποιες περιπτώσεις τα αποκτηθέντα σημεία από το πρόγραμμα βρίσκονται λίγο κάτω από το μέτωπο Pareto που έδωσε το λογισμικό των εξελικτικών αλγορίθμων (άρα είναι λίγο ‘καλύτερα’). Αυτό οφείλεται στο ότι το εν λόγω λογισμικό δεν παρέχει μία ομαλή καμπύλη (λόγω του τρόπου με τον οποίο δημιουργούνται τα νέα άτομα από την εξελικτική διαδικασία), όπως αναφέρθηκε και στην προηγούμενη ενότητα.

Στο σημείο αυτό, κρίνεται απαραίτητο να δηλωθεί ότι για την επίλυση του συγκεκριμένου μαθηματικού προβλήματος αξίζει κανείς να προχωρήσει στη χρήση της μεθόδου χωρίς το στάδιο της διόρθωσης, καθώς στο τέλος της προηγούμενης ενότητας επιδείχθηκε ότι με αυτόν τον τρόπο δύναται να αποκτηθούν 20 μη-κυριαρχούμενες λύσεις με μόλις 129 αξιολογήσεις και μάλιστα, όπως δείχνει το σχήμα 5.21, με αρκετά μεγάλη ακρίβεια, ενώ στην περίπτωση που εξετάστηκε εδώ για

10 σημεία του μετώπου Pareto πραγματοποιήθηκαν 217 αξιολογήσεις (σαφέστατα η σύγκριση γίνεται για τις ίδιες τιμές παραμέτρων).

Τέλος, κρίνεται σκόπιμο να αναφερθεί σε πόσες περίπου αξιολογήσεις δύναται το λογισμικό EASY να φτάσει στο μέτωπο Pareto. Για να επιτευχθεί γρήγορη σύγκλιση γίνεται ταυτόχρονη χρήση κατανεμημένης ανίχνευσης και μεταπροτύπων (DMAEA – Distributed Metamodel Assisted Evolutionary Algorithm, [19]). Η ιδέα της κατανεμημένης ανίχνευσης είναι αντί για έναν πληθυσμό, να υπάρχουν υποπληθυσμοί, οι οποίοι ονομάζονται δήμοι (demes) ή νησιά (islands). Αυτοί οι δήμοι δρουν αυτόνομα και μετά από έναν συγκεκριμένο αριθμό γενιών (που ρυθμίζουμε σαν παράμετρο) επικοινωνούν ο ένας με τον άλλο (με τρόπο που επίσης ρυθμίζουμε) και ανταλλάσσουν πληροφορίες. Αυτή η ιδέα μας βοηθά να έχουμε γρηγορότερη σύγκλιση. Στη συγκεκριμένη περίπτωση έχουμε DMAEA, δηλαδή κάθε δήμος ξεχωριστά χρησιμοποιεί μεταπρότυπα. Ακολούθως, γίνεται αναφορά στις σημαντικότερες από τις παραμέτρους που επιλέχθηκαν, με την εξαίρεση όσων μένουν αμετάβλητες σε σχέση με τις αναφερθείσες στην ενότητα 5.3.4.

Search Engines (Μηχανές αναζήτησης)

General

Demes (δήμοι): 4

Distributed

Migration frequency: 1 (συχνότητα σε γενιές με την οποία θα γίνει μετανάστευση μεταξύ των δήμων-ενδοεπικοινωνία δήμων)

Maximum migrations: -1 (για να γίνουν άπειρες μεταναστεύσεις)

Emigrants (best): 2 (είναι τα καλύτερα συμπεριφερόμενα άτομα τα οποία θα μεταναστεύσουν από ένα δήμο)

Emigrants (random): 2 (είναι οι τυχαίοι απόγονοι που θα μεταναστεύσουν από ένα δήμο)

Immigrants: 2 (ο αριθμός των ατόμων που έρχονται σε ένα δήμο)

Migration mode: tournament (αντικαθιστά το χειρότερο απόγονο)

Migration graph: round (θέτει τους δήμους στην περιφέρεια ενός δακτυλίου και επιτρέπει την επικοινωνία ενός δήμου μόνο με τους διπλανούς του)

Population

Parent Population size: 10 (εδώ πληθυσμός γονέων είναι ο πληθυσμός του κάθε δήμου)

Offspring Population size: 15 (ομοίως πληθυσμός απογόνων του κάθε δήμου)

Διαπιστώνεται ότι μέσω της εν λόγω επιλογής παραμέτρων το λογισμικό των εξελικτικών αλγορίθμων δύναται να δώσει το μέτωπο Pareto μετά από περίπου 500 αξιολογήσεις. Σύμφωνα με τα όσα

παρουσιάσθηκαν στην εν λόγω ενότητα αλλά και την προηγούμενη η μέθοδος που βασίζεται στη συνέχιση του μετώπου Pareto (είτε περιλαμβάνει το στάδιο της διόρθωσης είτε όχι) φαίνεται να επιτυγχάνει την εύρεση του μετώπου μη-κυριαρχούμενων λύσεων γρηγορότερα. Συγκεκριμένα, η μέθοδος πρόβλεψης-διόρθωσης, όπως δηλώθηκε παραπάνω, παρέχει 10 μη-κυριαρχούμενες λύσεις με την πραγματοποίηση 217 αξιολογήσεων που είναι λιγότερες από τις μισές που απαιτούνται από το λογισμικό των εξελικτικών αλγορίθμων. Όταν δε αφαιρείται το στάδιο της διόρθωσης, δύναται να αποκτηθούν 20 σημεία του μετώπου Pareto με μόλις 129 αξιολογήσεις, σχεδόν 4 φορές λιγότερες από αυτές του εξελικτικού αλγορίθμου. Αυτό, βέβαια, ισχύει για το δεδομένο πρόβλημα αλλά και για το σετ παραμέτρων που επιλέχθηκε στο λογισμικό των εξελικτικών αλγορίθμων (δεν είναι απαραίτητο να έχει αποκτηθεί το 'ταχύτερο' αποτέλεσμα που μπορεί να δώσει το λογισμικό). Τονίζεται ότι η πραγματοποιηθείσα σύγκριση είναι ενδεικτική και δεν έχει στόχο τη σύγκριση των δύο μεθόδων.

5.5 Τρίτο μαθηματικό πρόβλημα

5.5.1 Διατύπωση προβλήματος

Στην ενότητα αυτή, παρουσιάζεται το τρίτο μαθηματικό πρόβλημα που επελέχθη να επιλυθεί με βάση τη θεωρία της συνέχισης του μετώπου Pareto. Το εν λόγω πρόβλημα αντλείται από τη βιβλιογραφία, Poloni κ.ά., [20]. Πρόκειται για ένα πρόβλημα ελαχιστοποίησης δύο στόχων (f_1 και f_2) με δύο μεταβλητές σχεδιασμού (x_1 και x_2). Το πρόβλημα επιβάλλει τόσο άνω όσο και κάτω όρια στις δύο μεταβλητές σχεδιασμού. Αν εξαιρεθούν οι περιορισμοί που υποχρεώνουν τις δύο μεταβλητές σχεδιασμού να κινούνται ανάμεσα στο κάτω και το άνω όριό τους, το πρόβλημα δεν υπόκειται σε κανέναν άλλο περιορισμό, ισότητας ή ανισότητας. Οι δύο συναρτήσεις στόχου είναι οι εξής:

$$f_1(x_1, x_2) = 1 + [A_1 - B_1(x_1, x_2)]^2 + [A_2 - B_2(x_1, x_2)]^2$$

$$f_2(x_1, x_2) = (x_1 + 3)^2 + (x_2 + 1)^2$$

όπου

$$A_1 = 0,5 \sin 1 - 2 \cos 1 + \sin 2 - 1,5 \cos 2$$

$$A_2 = 1,5 \sin 1 - \cos 1 + 2 \sin 2 - 0,5 \cos 2$$

$$B_1 = 0,5 \sin x_1 - 2 \cos x_1 + \sin x_2 - 1,5 \cos x_2$$

$$B_2 = 1,5 \sin x_1 - \cos x_1 + 2 \sin x_2 - 0,5 \cos x_2$$

Έτσι, το πρόβλημα βελτιστοποίησης που αποτελεί αντικείμενο της δεδομένης ενότητας διατυπώνεται ως εξής:

$$\min \vec{f}(\vec{x}) = \min \begin{bmatrix} f_1(\vec{x}) \\ f_2(\vec{x}) \end{bmatrix}$$

που υπόκειται σε (subject to – s.t.)

$$-\pi \leq x_1 \leq \pi$$

$$-\pi \leq x_2 \leq \pi$$

5.5.2 Εφαρμογή της θεωρίας της συνέχισης του μετώπου Pareto στο πρόβλημα

Η εν λόγω ανάλυση πραγματοποιήθηκε στην ενότητα 5.3.2 και, καθώς δεν παρουσιάζονται διαφορές, δεν επαναλαμβάνεται εδώ.

5.5.3 Περιγραφή του αλγορίθμου που αναπτύχθηκε για την επίλυση του προβλήματος

Για να επιτευχθεί η επίλυση του μαθηματικού προβλήματος βελτιστοποίησης που παρουσιάστηκε παραπάνω, λαμβάνεται η απόφαση να χρησιμοποιηθεί η μέθοδος που βασίζεται στη συνέχιση του μετώπου Pareto χωρίς το στάδιο της διόρθωσης. Ο αλγόριθμος που αναπτύχθηκε διαφέρει σε σχέση με αυτόν που περιγράφηκε διεξοδικώς στην ενότητα 5.3.3.1 μόνο ως προς το τμήμα υπολογισμού των πρώτων και δεύτερων παραγώγων των συναρτήσεων στόχου ως προς τις μεταβλητές σχεδιασμού για το τρέχον \vec{x} . Ο προαναφερθείς υπολογισμός για το δεδομένο μαθηματικό πρόβλημα πραγματοποιείται, όπως και για τα δύο προηγούμενα που παρουσιάστηκαν στις ενότητες 5.3 και 5.4, αναλυτικά ως εξής:

$$\frac{\partial f_1}{\partial x_1} = 2[(A_1 - B_1)(-0,5 \cos x_1 - 2 \sin x_1) + (A_2 - B_2)(-1,5 \cos x_1 - \sin x_1)]$$

$$\frac{\partial f_1}{\partial x_2} = 2[(A_1 - B_1)(-\cos x_2 - 1,5 \sin x_2) + (A_2 - B_2)(-2 \cos x_2 - 0,5 \sin x_2)]$$

$$\frac{\partial f_2}{\partial x_1} = 2(x_1 + 3), \quad \frac{\partial f_2}{\partial x_2} = 2(x_2 + 1)$$

$$\begin{aligned} \frac{\partial^2 f_1}{\partial x_1^2} = & 2[(0,5 \cos x_1 + 2 \sin x_1)^2 + (A_1 - B_1)(0,5 \sin x_1 - 2 \cos x_1) + \\ & + (1,5 \cos x_1 + \sin x_1)^2 + (A_2 - B_2)(1,5 \sin x_1 - \cos x_1)] \end{aligned}$$

$$\frac{\partial^2 f_1}{\partial x_1 \partial x_2} = 2[(\cos x_2 + 1,5 \sin x_2)(0,5 \cos x_1 + 2 \sin x_1) + (2 \cos x_2 + 0,5 \sin x_2)(1,5 \cos x_1 + \sin x_1)] = \frac{\partial^2 f_1}{\partial x_2 \partial x_1}$$

$$\frac{\partial^2 f_1}{\partial x_2^2} = 2[(\cos x_2 + 1,5 \sin x_2)^2 + (A_1 - B_1)(\sin x_2 - 1,5 \cos x_2) + (2 \cos x_2 + 0,5 \sin x_2)^2 + (A_2 - B_2)(2 \sin x_2 - 0,5 \cos x_2)]$$

$$\frac{\partial^2 f_2}{\partial x_1^2} = 2, \quad \frac{\partial^2 f_2}{\partial x_1 \partial x_2} = 0 = \frac{\partial^2 f_2}{\partial x_2 \partial x_1}, \quad \frac{\partial^2 f_2}{\partial x_2^2} = 2$$

5.5.4 Παρουσίαση αποτελεσμάτων

Τονίστηκε ήδη ότι για την εύρεση του μετώπου μη-κυριαρχούμενων λύσεων μέσω της μεθόδου που έχει ως βάση τη συνέχιση του μετώπου Pareto, απαιτείται καλή γνώση του προβλήματος προς επίλυση. Σύμφωνα με τα όσα παρουσιάστηκαν στην ενότητα 5.3.3, αυτό αντικατοπτρίζεται στην απόφαση που πρέπει να λάβει ο χρήστης σχετικά με τις τιμές των ορίων του μετώπου Pareto \hat{f}_2^{up} και \hat{f}_2^{down} που θα τροφοδοτηθούν στο πρόγραμμα. Στην περίπτωση του μαθηματικού προβλήματος που παρουσιάστηκε στην ενότητα 5.5.1, ο χρήστης οφείλει να γνωρίζει ότι το μέτωπο Pareto που αναμένεται να προκύψει 'αποτελείται από δύο τμήματα', δηλαδή είναι διακοπτόμενο ή ασυνεχές (discontinuous Pareto front). Ο χρήστης πρέπει, επιπλέον, να γνωρίζει που κυμαίνονται οι τιμές που αναμένεται να λάβουν τα όρια \hat{f}_2^{up} και \hat{f}_2^{down} για τα δύο τμήματα του μετώπου Pareto. Σε αντίθετη περίπτωση, ο χρήστης πρέπει να πραγματοποιήσει δοκιμές, οι οποίες, βέβαια, αυξάνουν το υπολογιστικό κόστος.

Έστω, λοιπόν, ότι σε πρώτη φάση επιθυμείται η εύρεση του τμήματος του μετώπου Pareto που αντιστοιχεί σε μεγαλύτερες τιμές του f_2 . Για το σκοπό αυτό, λαμβάνεται η απόφαση να δοθεί στο πρόγραμμα ως αρχική εκτίμηση για το \hat{f}_2 η τιμή 25. Με δεδομένη την τιμή του \hat{f}_2^{up} , το πρόγραμμα απαιτεί κατάλληλες τιμές των παραμέτρων που σχετίζονται με την ALM ώστε να συγκλίνει στο πρώτο σημείο του Pareto. Ως μέθοδος βελτιστοποίησης προβλημάτων ενός στόχου χωρίς περιορισμούς επιλέγεται η μέθοδος της απότομης καθόδου. Μία αρχική επιλογή των παραμέτρων που σχετίζονται με την ALM δύναται να είναι η εξής: $\vec{x}^0 = [0 \ 1]^T$, $\hat{\lambda}_2^0 = 0$, $\omega_p^0 = 0,01$, $\gamma = 1,1$, $\omega_p^{max} = 500$ και $\eta = 0,01$. Επιπλέον, έστω ότι στην αρχική αυτή φάση επιλέγονται $n_{elit} = 10$ και $\hat{f}_2^{down} = 21$. Στο παρόν σημείο, κρίνεται σκόπιμο να παρατεθούν τα αποτελέσματα του προγράμματος όσον αφορά το πρώτο σημείο του μετώπου Pareto που λαμβάνεται μέσω της ALM (με δεδομένες τις προαναφερθείσες τιμές των παραμέτρων) για διαφορετικές τιμές των ανοχών με τις οποίες ικανοποιούνται τα

κριτήρια σύγκλισης της μεθόδου της απότομης καθόδου και της ALM. Τα εν λόγω αποτελέσματα συναντώνται στον πίνακα 5.10, συνοδευόμενα από τις συνολικές αξιολογήσεις που πραγματοποιούνται από το πρόγραμμα (δηλαδή συνυπολογίζονται και οι αξιολογήσεις που γίνονται κατά τη διάρκεια της εύρεσης των υπολοίπων σημείων του Pareto) σε κάθε περίπτωση.

Ανοχή ALM	Ανοχή SD	x_1	x_2	f_1	f_2	Αξιολογήσεις
1,00E-06	1,00E-06	1,000000	2,000000	1,000000	24,999999	1437
1,00E-06	1,00E-05	1,000001	1,999998	1,000000	24,999999	1223
1,00E-06	1,00E-04	1,000008	1,999990	1,000000	25,000001	1007
1,00E-06	1,00E-03	1,000035	1,999953	1,000000	25,000001	805
1,00E-06	1,00E-02	1,000246	1,999671	1,000000	25,000000	579
1,00E-06	1,00E-01		δεν	συγκλίνει		
1,00E-05	1,00E-02	1,000346	1,999540	1,000001	25,000009	561
1,00E-05	1,00E-01	1,002871	1,996170	1,000042	25,000009	319
1,00E-04	1,00E-02	1,000797	1,998952	1,000003	25,000089	517
1,00E-04	1,00E-01	1,008662	1,988408	1,000377	24,999958	259

Πίνακας 5.10: Πρώτη μη-κυριαρχούμενη λύση και συνολικές αξιολογήσεις που πραγματοποιούνται για την εύρεση του πρώτου τμήματος του μετώπου Pareto για διαφορετικές τιμές των ανοχών με τις οποίες ικανοποιούνται τα κριτήρια σύγκλισης της μεθόδου της απότομης καθόδου και της ALM όταν $\hat{f}_2^{up} = 25$, $\bar{x}^0 = [0 \ 1]^T$, $\hat{\lambda}_2^0 = 0$, $\omega_p^0 = 0,01$, $\gamma = 1,1$, $\omega_p^{max} = 500$, $\eta = 0,01$ και $n_{elit} = 10$.

Λαμβάνοντας υπόψη τα δεδομένα του παραπάνω πίνακα, επιβάλλονται ως ικανοποιητικές ανοχές για τα κριτήρια σύγκλισης της μεθόδου της απότομης καθόδου και της ALM οι τιμές 10^{-2} και 10^{-4} αντίστοιχα. Για τις επιλεχθείσες ανοχές αλλά και για τις ανωτέρω αναφερθείσες παραμέτρους, όπως φαίνεται στον πίνακα 5.10, οι συνολικές αξιολογήσεις που χρειάζονται για την απόκτηση του μετώπου Pareto είναι 517. Στο σημείο αυτό, λαμβάνει χώρα μία διερεύνηση με στόχο να επιδειχθεί το πώς συμπεριφέρεται το πρόγραμμα όσον αφορά το υπολογιστικό κόστος με τη μεταβολή των παραμέτρων της ALM. Τα αποτελέσματα της αναφερθείσας διερεύνησης επιδεικνύονται στον πίνακα 5.11 (σημειώνεται ότι οι αναγραφόμενες αξιολογήσεις είναι οι συνολικές, ενώ αν επιθυμείται να βρεθούν αυτές που αντιστοιχούν στον υπολογισμό του πρώτου σημείου του μετώπου Pareto πρέπει να αφαιρεθούν οι αξιολογήσεις του τμήματος που υπολογίζει τα υπόλοιπα 9 σημεία που είναι σταθερές σε όλη τη διαδικασία και ίσες με $(n_{elit} - 1)(1 + n + 1) = 36$).

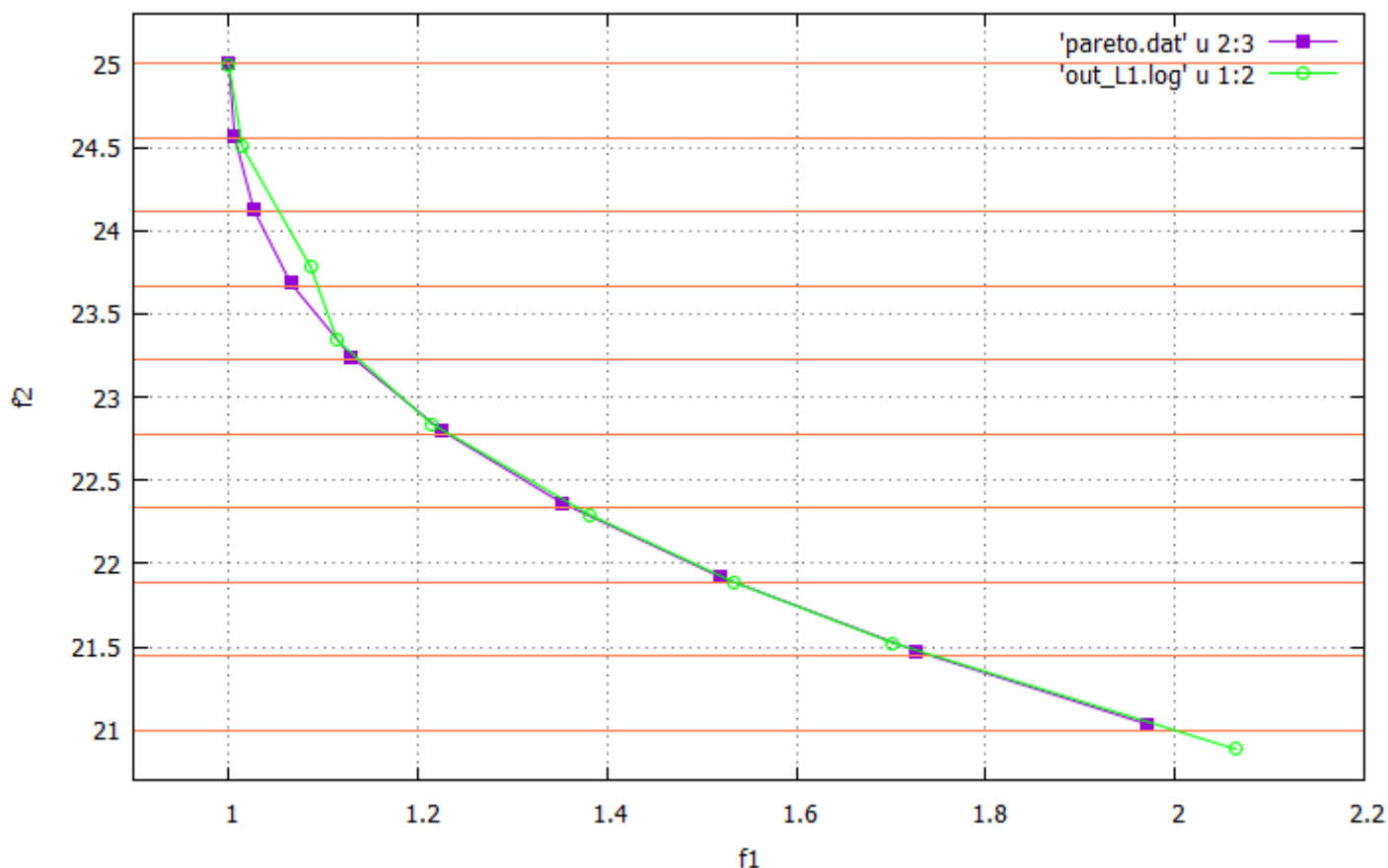
$\hat{\lambda}_2^0$	ω_p^0	γ	η	$(\bar{x}^0)^T$	Αξιολογήσεις
0	0,01	1,1	0,01	[0 1]	517
0	0,01	1,1	0,01	[0 3]	551
0	0,01	1,1	0,01	[-0,5 1]	465
0	0,01	1,1	0,01	[-0,1 1]	411
0	0,01	1,1	0,02	[-0,1 1]	223
0	0,01	1,1	0,03	[-0,1 1]	67
0	0,01	1,1	0,04	[-0,1 1]	115
0	0,01	1,1	0,06	[-0,1 1]	137
0	0,01	1,1	0,11	[-0,1 1]	83
0	0,01	1,1	0,12	[-0,1 1]	δεν συγκλίνει
0	0,005	1,1	0,11	[-0,1 1]	115
0	0,02	1,1	0,07	[-0,1 1]	119
0	0,02	1,1	0,08	[-0,1 1]	δεν συγκλίνει
0	0,03	1,1	0,07	[-0,1 1]	91
0	0,06	1,1	0,05	[-0,1 1]	109
0	0,01	1,05	0,11	[-0,1 1]	105
0	0,01	1,2	0,1	[-0,1 1]	87
0,1	0,01	1,1	0,03	[-0,1 1]	617

Πίνακας 5.11: Συνολικές αξιολογήσεις που πραγματοποιούνται για την εύρεση του πρώτου τμήματος του μετώπου Pareto για διαφορετικές τιμές των παραμέτρων της ALM όταν οι ανοχές για τα κριτήρια σύγκλισης της μεθόδου της απότομης καθόδου και της ALM λαμβάνουν τις τιμές 10^{-2} και 10^{-4} αντίστοιχα, $\hat{f}_2^{up} = 25$ και $n_{elit} = 10$.

Ο καλύτερος συνδυασμός παραμέτρων που επετεύχθη, όπως φαίνεται από τον πίνακα 5.11, είναι $\bar{x}^0 = [-0,1 \ 1]^T$, $\hat{\lambda}_2^0 = 0$, $\omega_p^0 = 0,01$ και $\eta = 0,03$ (σημειώνεται ότι ο πολλαπλασιαστής γ για αυτό το συνδυασμό παραμέτρων δεν παίζει ρόλο καθώς τυχαίνει η ALM να συγκλίνει από τον πρώτο κύκλο χωρίς να προχωράει σε δεύτερο). Ο συνδυασμός αυτός απαιτεί την πραγματοποίηση μόλις 67 αξιολογήσεων συνολικά.

Μετά την εύρεση της πρώτης, το πρόγραμμα προβαίνει στον υπολογισμό και των υπολοίπων μη-κυριαρχούμενων λύσεων που αντιστοιχούν στο πρώτο τμήμα του διακοπτόμενου μετώπου με βάση τις παραμέτρους n_{elit} και \hat{f}_2^{down} που ορίστηκαν παραπάνω. Το τμήμα του μετώπου Pareto που προκύπτει απεικονίζεται στο σχήμα 5.24. Στο ίδιο σχήμα φαίνεται και το αντίστοιχο τμήμα του μετώπου Pareto που προέκυψε από το λογισμικό των εξελικτικών αλγορίθμων (οι τιμές των παραμέτρων που χρησιμοποιήθηκαν είναι ακριβώς οι ίδιες με αυτές που παρατέθηκαν στην ενότητα 5.3.4 με τη μοναδική εξαίρεση ότι εδώ πραγματοποιήθηκαν 10000 αξιολογήσεις ώστε να επιτευχθεί καλύτερη σύγκλιση). Στο σχήμα 5.24 έχουν σχεδιαστεί, επίσης, και n_{elit} οριζόντιες γραμμές οι οποίες δηλώνουν την τιμή του \hat{f}_2 για καθένα από τα n_{elit} σημεία των οποίων ζητήθηκε η εύρεση (ώστε να δύναται να διαπιστωθεί αν υπάρχει διαφορά μεταξύ της τιμής του \hat{f}_2

και του f_2 για κάθε σημείο του πρώτου τμήματος του μετώπου Pareto που προέκυψε από το πρόγραμμα).



Σχήμα 5.24: Άνω τμήμα ασυνεχούς μετώπου Pareto που προέκυψε από τη μέθοδο που βασίζεται στη συνέχιση της καμπύλης των μη-κυριαρχούμενων λύσεων για $\hat{f}_2^{up} = 25$, $\hat{f}_2^{down} = 21$ και $n_{elit} = 10$ (μωβ καμπύλη) και αντίστοιχο τμήμα μετώπου Pareto που προέκυψε από το λογισμικό των εξελικτικών αλγορίθμων (πράσινη καμπύλη). Οι οριζόντιες γραμμές χρώματος πορτοκαλί αντιπροσωπεύουν τις τιμές των \hat{f}_2 για καθένα από τα n_{elit} σημεία που αναζητήθηκαν.

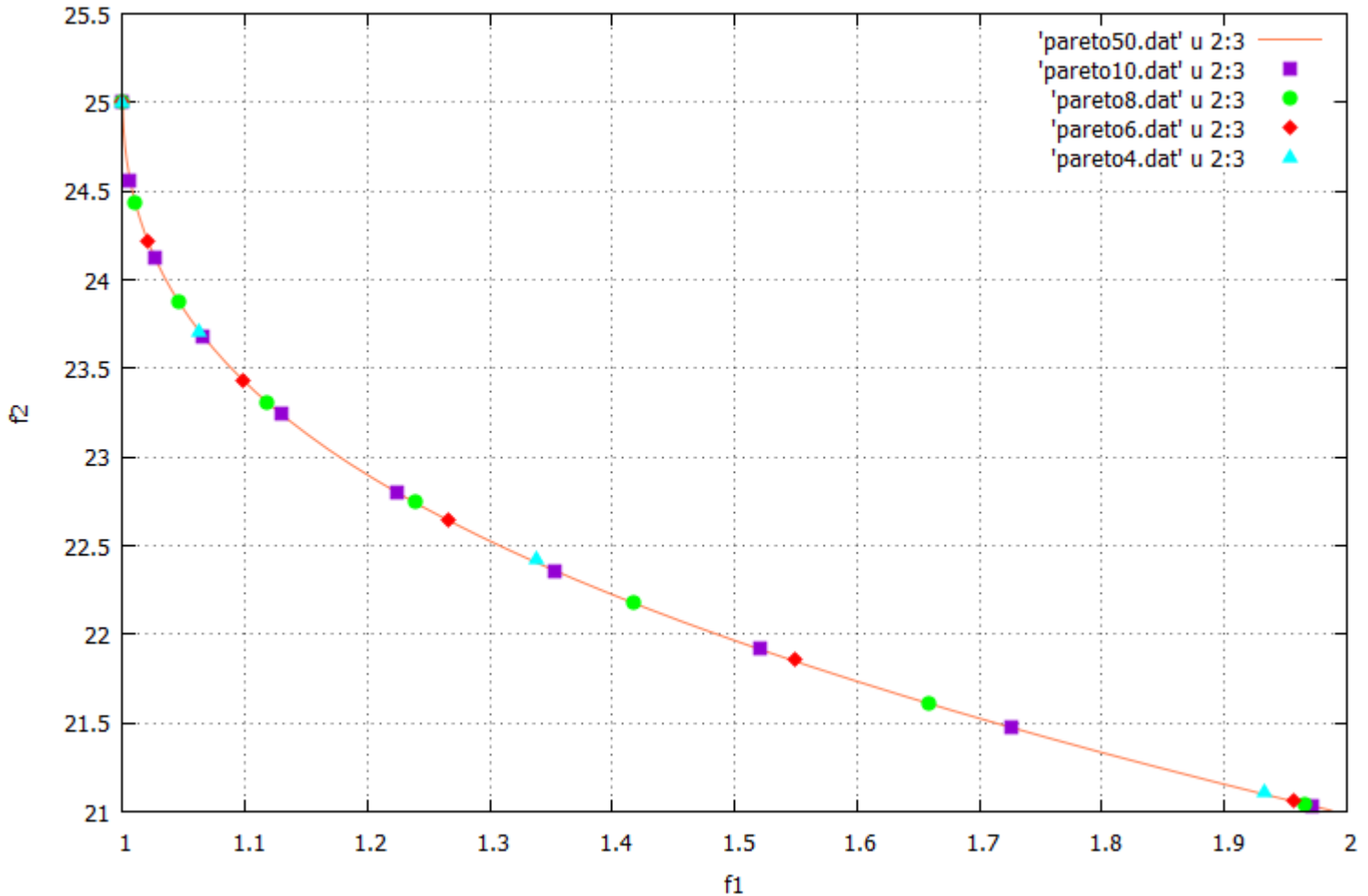
Από την παρατήρηση του σχήματος 5.24, κατά πρώτο λόγο, καθίσταται φανερό (όπως και στην περίπτωση των δύο μαθηματικών προβλημάτων που προηγήθηκαν) το γεγονός ότι στα πρώτα σημεία που αποκτήθηκαν από το πρόγραμμα οι διαφορές μεταξύ των τιμών του \hat{f}_2 και του f_2 είναι πολύ μικρές (πρακτικά αμελητέες), ενώ καθώς ο αριθμός των υπολογισθέντων σημείων μεγαλώνει οι διαφορές αυτές αυξάνονται και γίνονται όλο και πιο αισθητές. Σε κάθε περίπτωση, όμως, οι εν λόγω διαφορές παραμένουν σχετικά μικρές. Αξίζει να αναφερθεί ότι στο δεύτερο αποκτηθέν σημείο ισχύει $f_2 - \hat{f}_2 \cong 0,005048$, ενώ στο τελευταίο είναι $f_2 - \hat{f}_2 \cong 0,034393$. Οι αιτίες εμφάνισης των εν λόγω διαφορών εξηγήθηκαν λεπτομερώς στην ενότητα 5.3.4 και δεν κρίνεται σκόπιμο να επαναληφθούν εδώ.

Μία δεύτερη διαπίστωση που θα μπορούσε να γίνει μετά την παρατήρηση του σχήματος 5.24 είναι ότι η καμπύλη που

αποκτήθηκε με βάση τη μέθοδο της συνέχισης της καμπύλης των μη-κυριαρχούμενων λύσεων αποτελεί το μεγαλύτερο μέρος του πραγματικού πρώτου τμήματος του μετώπου που αποκτήθηκε μέσω του λογισμικού των εξελικτικών αλγορίθμων αλλά όχι ολόκληρο, καθώς η πράσινη καμπύλη εκτείνεται ελαφρώς προς τα δεξιά της μωβ. Πρέπει, όμως, να ειπωθεί ότι το αποκτηθέν τμήμα καλύπτει πολύ μεγάλο κομμάτι του πρώτου τμήματος του μετώπου Pareto που δόθηκε μέσω του εξελικτικού αλγορίθμου οπότε, γενικά, θεωρείται ότι καλύπτει τις ανάγκες του χρήστη. Επαναλαμβάνεται, άλλωστε, ότι τα όρια \hat{f}_2^{up} και \hat{f}_2^{down} επιλέγονται από το χρήστη σύμφωνα με τις ανάγκες του.

Επιπλέον, το σχήμα 5.24 οδηγεί στο συμπέρασμα ότι η καμπύλη που προέκυψε από τη μέθοδο που βασίζεται στη συνέχιση της καμπύλης των μη-κυριαρχούμενων λύσεων για $\hat{f}_2^{up} = 25$, $\hat{f}_2^{down} = 21$ και $n_{elit} = 10$, πρακτικά, πέφτει πάνω σε αυτή που προέκυψε από το λογισμικό των εξελικτικών αλγορίθμων. Πρέπει, επίσης, να αναφερθεί ότι η καμπύλη που προέκυψε από το αναπτυσσόμενο πρόγραμμα είναι ομαλή (όπως ήταν αναμενόμενο λόγω του ορισμού της μεθόδου) σε αντίθεση με αυτή που αποκτήθηκε μέσω του εξελικτικού αλγορίθμου (λόγω του τρόπου με τον οποίο δημιουργούνται τα νέα άτομα από την εξελικτική διαδικασία), γεγονός που οδηγεί την πρώτη καμπύλη να βρίσκεται κάτω από τη δεύτερη σε κάποια μικρά τμήματα (δηλαδή έχει επιτευχθεί η εύρεση ‘καλύτερων’ σημείων στα τμήματα αυτά). Αυτό είναι ένα πλεονέκτημα της μεθόδου που πραγματεύεται η παρούσα εργασία, όπως διαπιστώθηκε και μέσω του δεύτερου μαθηματικού προβλήματος. Επιπροσθέτως, τονίζεται ότι η αναφερθείσα πρακτική ταύτιση των καμπυλών ήταν δυνατή παρά το γεγονός ότι, όπως σχολιάσθηκε παραπάνω, η τιμή της διαφοράς $f_2 - \hat{f}_2$ συνεχώς αυξάνεται καθώς ο αλγόριθμος υπολογίζει όλο και περισσότερα σημεία.

Στο σημείο αυτό κρίνεται ενδιαφέρον να παρουσιαστεί η συμπεριφορά του προγράμματος στην περίπτωση μικρότερου n_{elit} (άρα μεγαλύτερου $\delta\hat{f}_2$). Στο σχήμα 5.25 φαίνονται τα σημεία που δίνει το πρόγραμμα για τις περιπτώσεις όπου το n_{elit} παίρνει τις τιμές 10, 8, 6 και 4 (ισχύουν, όπως προηγουμένως, $\hat{f}_2^{up} = 25$ και $\hat{f}_2^{down} = 21$).

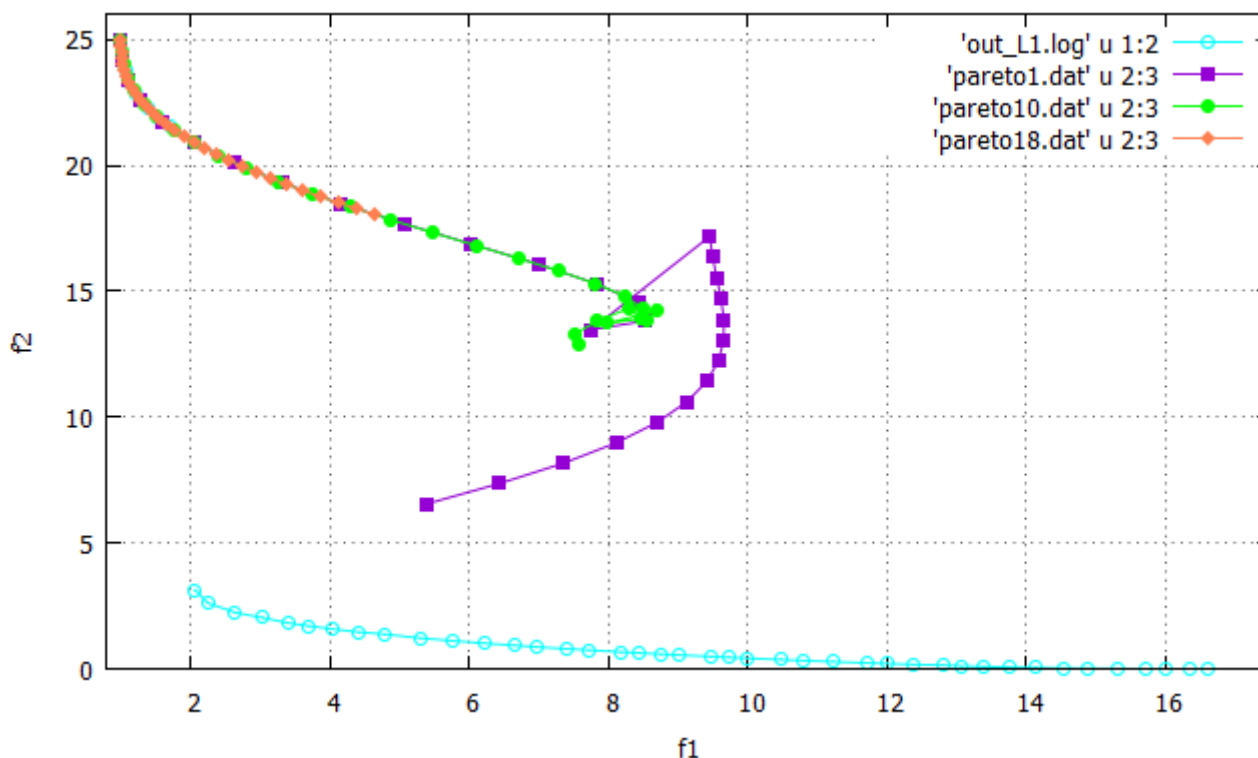


Σχήμα 5.25: Μη-κυριαρχούμενες λύσεις που προέκυψαν από τη μέθοδο που βασίζεται στη συνέχιση του μετώπου Pareto για $n_{elit} = 10$ (μωβ χρώμα), για $n_{elit} = 8$ (πράσινο χρώμα), για $n_{elit} = 6$ (κόκκινο χρώμα) και για $n_{elit} = 4$ (γαλάζιο χρώμα). Η καμπύλη χρώματος πορτοκαλί αντιστοιχεί σε $n_{elit} = 50$ και σχεδιάστηκε ως αναφορά για τις προαναφερθείσες περιπτώσεις, καθώς ουσιαστικά αποτελεί το πρώτο τμήμα του μετώπου Pareto. Σημειώνεται ότι σε όλες τις περιπτώσεις ισχύουν $\hat{f}_2^{up} = 25$ και $\hat{f}_2^{down} = 21$.

Αν παρατηρηθεί το σχήμα 5.25, καθίσταται φανερό το γεγονός ότι όσο το n_{elit} μειώνεται τόσο μεγαλύτερη είναι η διαφορά $f_2 - \hat{f}_2$ για τα υπολογισθέντα σημεία (χαρακτηριστικά το τελευταίο σημείο βρίσκεται όλο και υψηλότερα στον άξονα των f_2). Ωστόσο, φαίνεται ότι στην περίπτωση όπου $n_{elit} = 8$ τα υπολογισθέντα σημεία παραμένουν πάνω στο πρώτο τμήμα του μετώπου Pareto, ενώ οι συνολικές αξιολογήσεις που απαιτούνται μειώνονται κατά 8 (αφού για κάθε σημείο είναι απαραίτητες $1 + n + 1 = 4$ αξιολογήσεις) σε σχέση με την περίπτωση όπου $n_{elit} = 10$ (αν ληφθούν $\vec{x}^0 = [-0,1 \ 1]^T$, $\hat{\lambda}_2^0 = 0$, $\omega_p^0 = 0,01$, $\eta = 0,03$ και $\omega_p^{max} = 500$, το αποτέλεσμα επιτυγχάνεται μόνο με 59 αξιολογήσεις). Όσον αφορά την περίπτωση με $n_{elit} = 6$, τα αποτελέσματα είναι, επίσης, αρκετά καλά, καθώς τα σημεία φαίνεται να παραμένουν πάνω στο πρώτο τμήμα του μετώπου Pareto, ενώ για τον ανωτέρω συνδυασμό παραμέτρων το αποτέλεσμα επιτυγχάνεται

μόνο με 51 αξιολογήσεις. Τέλος, για τη περίπτωση $n_{elit} = 4$, το δεύτερο σημείο φαίνεται να παραμένει πάνω στο πρώτο τμήμα του μετώπου Pareto αλλά τα δύο τελευταία αποκτούν μία ‘αισθητή’ απόσταση από αυτό, χωρίς οι διαφορές να είναι πολύ μεγάλες. Βέβαια, στην εν λόγω περίπτωση οι πραγματοποιηθείσες αξιολογήσεις μειώνονται κατά 24 σε σχέση με την περίπτωση όπου $n_{elit} = 10$ (για τον ανωτέρω συνδυασμό παραμέτρων το αποτέλεσμα επιτυγχάνεται μόνο με 43 αξιολογήσεις). Όπως σημειώθηκε και στις περιπτώσεις των μαθηματικών προβλημάτων που προηγήθηκαν, το ποια περίπτωση θα επιλεγεί εξαρτάται από τις ανάγκες του χρήστη σχετικά με το πόσα σημεία θα αποκτηθούν πάνω στο πρώτο τμήμα του μετώπου Pareto (αν ο χρήστης ενδιαφέρεται να αποκτήσει σημεία σε αυτό το πρώτο τμήμα του μετώπου), την ακρίβεια και το υπολογιστικό κόστος.

Στο σημείο αυτό, παρουσιάζει ενδιαφέρον να ερευνηθεί η συμπεριφορά του προγράμματος αν επιχειρηθεί να δοθεί στο \hat{f}_2^{down} μία τιμή μικρότερη από το 21 με την τιμή του \hat{f}_2^{up} να παραμένει 25 (ενώ έστω ότι επιλέγεται $n_{elit} = 30$). Οι καμπύλες που προκύπτουν για τις περιπτώσεις όπου το \hat{f}_2^{down} παίρνει λ.χ. τις τιμές 18, 10 και 1 φαίνονται στο σχήμα 5.26 συνοδευόμενες από το μέτωπο Pareto που προκύπτει από το λογισμικό των εξελικτικών αλγορίθμων.



Σχήμα 5.26: Καμπύλες που προέκυψαν από τη μέθοδο που βασίζεται στη συνέχιση του μετώπου Pareto για $\hat{f}_2^{down} = 18$ (πορτοκαλί χρώμα), για $\hat{f}_2^{down} = 10$ (πράσινο χρώμα) και για $\hat{f}_2^{down} = 1$ (μωβ χρώμα). Σημειώνεται ότι και για τις τρεις περιπτώσεις ισχύουν $\hat{f}_2^{up} = 25$ και $n_{elit} = 30$. Το μέτωπο Pareto που προέκυψε από το λογισμικό των εξελικτικών αλγορίθμων αποτελεί η γαλάζια καμπύλη.

Μέσω της παρατήρησης του διακοπτόμενου μετώπου Pareto που έχει αποκτηθεί από το λογισμικό των εξελικτικών αλγορίθμων (γαλάζια καμπύλη σχήματος 5.26), διαπιστώνεται ότι μόνο στην τρίτη περίπτωση από αυτές που εξετάστηκαν, για την οποία ισχύει $\hat{f}_2^{down} = 1$ (μωβ καμπύλη σχήματος 5.26), 'περικλείονται' τιμές του \hat{f}_2 οι οποίες ανήκουν στο διάστημα των τιμών του f_2 που συναντάται στο δεύτερο τμήμα του μετώπου των μη-κυριαρχούμενων λύσεων. Το σχήμα 5.26, ωστόσο, καθιστά φανερό το γεγονός ότι η μέθοδος που βασίζεται στη συνέχιση του μετώπου Pareto δεν δύναται να αποδώσει σημεία πάνω στο δεύτερο τμήμα του μετώπου των μη-κυριαρχούμενων λύσεων αν ισχύει $\hat{f}_2^{up} = 25$ (δηλαδή αν η τιμή του \hat{f}_2^{up} οδηγεί σε σημείο πάνω στο πρώτο τμήμα του μετώπου Pareto), καθώς η μωβ καμπύλη, μπαίνοντας στο διάστημα των τιμών του f_2 που δεν απαντάται στο μέτωπο Pareto, αρχικά συνεχίζει μία 'ομαλή' πορεία εκτεινώντας λανθασμένα το πρώτο τμήμα του μετώπου και στη συνέχεια ακολουθεί μία εντελώς 'ανώμαλη' πορεία που δεν μπορεί να οδηγήσει σε καμία περίπτωση στο δεύτερο τμήμα του μετώπου. Αυτό, βέβαια, ήταν απολύτως αναμενόμενο, καθώς η μαθηματική βάση της μεθόδου την οποία πραγματεύεται η παρούσα εργασία, όπως αυτή παρουσιάστηκε αναλυτικά στο κεφάλαιο 4, δεν δίνει τη δυνατότητα για να υπάρξει αυτό το 'πήδημα' από το πρώτο τμήμα του μετώπου στο δεύτερο. Σε μία τέτοια, λοιπόν, περίπτωση, αν ο χρήστης ενδιαφέρεται να αποκτήσει σημεία και στο δεύτερο τμήμα του μετώπου Pareto, υποχρεούται να επαναλάβει την εκτέλεση του προγράμματος δίνοντάς του αυτή τη φορά μία τιμή για το \hat{f}_2^{up} που θα οδηγήσει σε σημείο του εν λόγω τμήματος του μετώπου. Το γεγονός αυτό είναι φανερό ότι αυξάνει το υπολογιστικό κόστος σε σχέση με μία ανάλογη περίπτωση συνεχούς μετώπου Pareto, καθώς εδώ ο χρήστης 'πληρώνει' το στάδιο του να φτάσει στο Pareto δύο φορές. Στο σημείο αυτό, υπενθυμίζεται ότι ο χρήστης πρέπει να έχει καλή γνώση του προβλήματος ώστε να είναι ικανός να δώσει στο πρόγραμμα δύο κατάλληλες τιμές για το \hat{f}_2^{up} που θα στοχεύουν σε δύο 'ξεχωριστά' τμήματα του μετώπου Pareto. Αξίζει να αναφερθεί ότι, αν ο χρήστης δεν γνωρίζει ότι πρόκειται για διακοπτόμενο Pareto και θέσει $\hat{f}_2^{up} = 25$ και $\hat{f}_2^{down} = 1$, η μορφή της μωβ καμπύλης θα τον βοηθήσει να αποκτήσει αυτή την πληροφορία με το ανάλογο υπολογιστικό κόστος, βέβαια, καθώς θα έχει 'πληρώσει' για σημεία που δεν αποτελούν μη-κυριαρχούμενες λύσεις. Το ίδιο θα συμβεί και στην περίπτωση που θέσει $\hat{f}_2^{up} = 25$ και $\hat{f}_2^{down} = 10$, καθώς και η πράσινη καμπύλη καταλήγει να ακολουθεί ανώμαλη πορεία. Αν, όμως, ο χρήστης θέσει $\hat{f}_2^{up} = 25$ και $\hat{f}_2^{down} = 18$, κινδυνεύει να οδηγηθεί σε λανθασμένη εντύπωση για το μέτωπο Pareto, καθώς η πορτοκαλί καμπύλη δεν προδίδει την ύπαρξη διακοπτόμενου μετώπου.

Στη συνέχεια, ακολουθεί η εύρεση του δεύτερου τμήματος του μετώπου Pareto το οποίο αντιστοιχεί σε μικρότερες τιμές του f_2 . Για να επιτευχθεί η απόκτηση του εν λόγω τμήματος, λαμβάνεται η απόφαση να τροφοδοτηθεί στο πρόγραμμα ως \hat{f}_2^{up} η τιμή 3. Η αρχική επιλογή των παραμέτρων της ALM είναι η εξής (αναφέρεται ότι, όπως και στην περίπτωση του πρώτου τμήματος του μετώπου, ως μέθοδος βελτιστοποίησης προβλημάτων ενός στόχου χωρίς περιορισμούς επιλέγεται η μέθοδος της απότομης καθόδου): $\vec{x}^0 = [-2 \ 1]^T$, $\hat{\lambda}_2^0 = 0$, $\omega_p^0 = 0,01$, $\gamma = 1,1$, $\omega_p^{max} = 500$ και $\eta = 0,01$. Επιπλέον, έστω ότι επιλέγονται $n_{elit} = 18$ και $\hat{f}_2^{down} = 0,01$. Ακολουθούν τα αποτελέσματα του προγράμματος όσον αφορά το πρώτο σημείο του μετώπου Pareto που λαμβάνεται μέσω της ALM (με δεδομένες τις προαναφερθείσες τιμές των παραμέτρων) για διαφορετικές τιμές των ανοχών με τις οποίες ικανοποιούνται τα κριτήρια σύγκλισης της μεθόδου της απότομης καθόδου και της ALM, συνοδευόμενα από τις συνολικές αξιολογήσεις που πραγματοποιούνται σε κάθε περίπτωση (πίνακας 5.12).

Ανοχή ALM	Ανοχή SD	x_1	x_2	f_1	f_2	Αξιολογήσεις
1,00E-06	1,00E-06	-3,141593	0,726254	2,078307	3,000000	4703
1,00E-06	1,00E-04	-3,141593	0,726254	2,078307	3,000000	2559
1,00E-06	1,00E-02	-3,141593	0,726254	2,078307	3,000001	763
1,00E-06	1,00E-01	-3,141593	0,726253	2,078307	2,999999	385
1,00E-06	1,00E-00	-3,141593	0,726254	2,078307	3,000001	335
1,00E-04	1,00E-02	-3,141593	0,726243	2,078313	2,999965	737
1,00E-04	1,00E-01	-3,141593	0,726265	2,078300	3,000040	367
1,00E-04	1,00E-00	-3,141593	0,726255	2,078306	3,000003	303
1,00E-03	1,00E-00	-3,141593	0,726006	2,078449	2,999144	281

Πίνακας 5.12: Πρώτη μη-κυριαρχούμενη λύση και συνολικές αξιολογήσεις που πραγματοποιούνται για την εύρεση του δεύτερου τμήματος του μετώπου Pareto για διαφορετικές τιμές των ανοχών με τις οποίες ικανοποιούνται τα κριτήρια σύγκλισης της μεθόδου της απότομης καθόδου και της ALM όταν $\hat{f}_2^{up} = 3$, $\vec{x}^0 = [-2 \ 1]^T$, $\hat{\lambda}_2^0 = 0$, $\omega_p^0 = 0,01$, $\gamma = 1,1$, $\omega_p^{max} = 500$, $\eta = 0,01$ και $n_{elit} = 18$.

Στη δεδομένη περίπτωση, λαμβάνοντας υπόψη τα δεδομένα του παραπάνω πίνακα, επιβάλλονται ως ικανοποιητικές ανοχές για τα κριτήρια σύγκλισης της μεθόδου της απότομης καθόδου και της ALM οι τιμές 1 και 10^{-4} αντίστοιχα. Για τις επιλεγθείσες ανοχές αλλά και για τις ανωτέρω αναφερθείσες παραμέτρους, όπως φαίνεται στον πίνακα 5.12, οι συνολικές αξιολογήσεις που χρειάζονται για την απόκτηση του μετώπου Pareto είναι 303. Στο σημείο αυτό, λαμβάνει χώρα μία διερεύνηση, όπως και στην περίπτωση εύρεσης του πρώτου τμήματος του μετώπου, με στόχο να επιδειχθεί το πώς συμπεριφέρεται το πρόγραμμα όσον αφορά το υπολογιστικό κόστος με τη μεταβολή των παραμέτρων της ALM. Τα αποτελέσματα της

αναφερθείσας διερεύνησης επιδεικνύονται στον πίνακα 5.13 (σημειώνεται ότι οι αναγραφόμενες αξιολογήσεις είναι οι συνολικές, ενώ αν επιθυμείται να βρεθούν αυτές που αντιστοιχούν στον υπολογισμό του πρώτου σημείου του μειώπου Pareto πρέπει να αφαιρεθούν οι αξιολογήσεις του τμήματος που υπολογίζει τα υπόλοιπα 17 σημεία που είναι σταθερές σε όλη τη διαδικασία και ίσες με $(n_{elit} - 1)(1 + n + 1) = 68$).

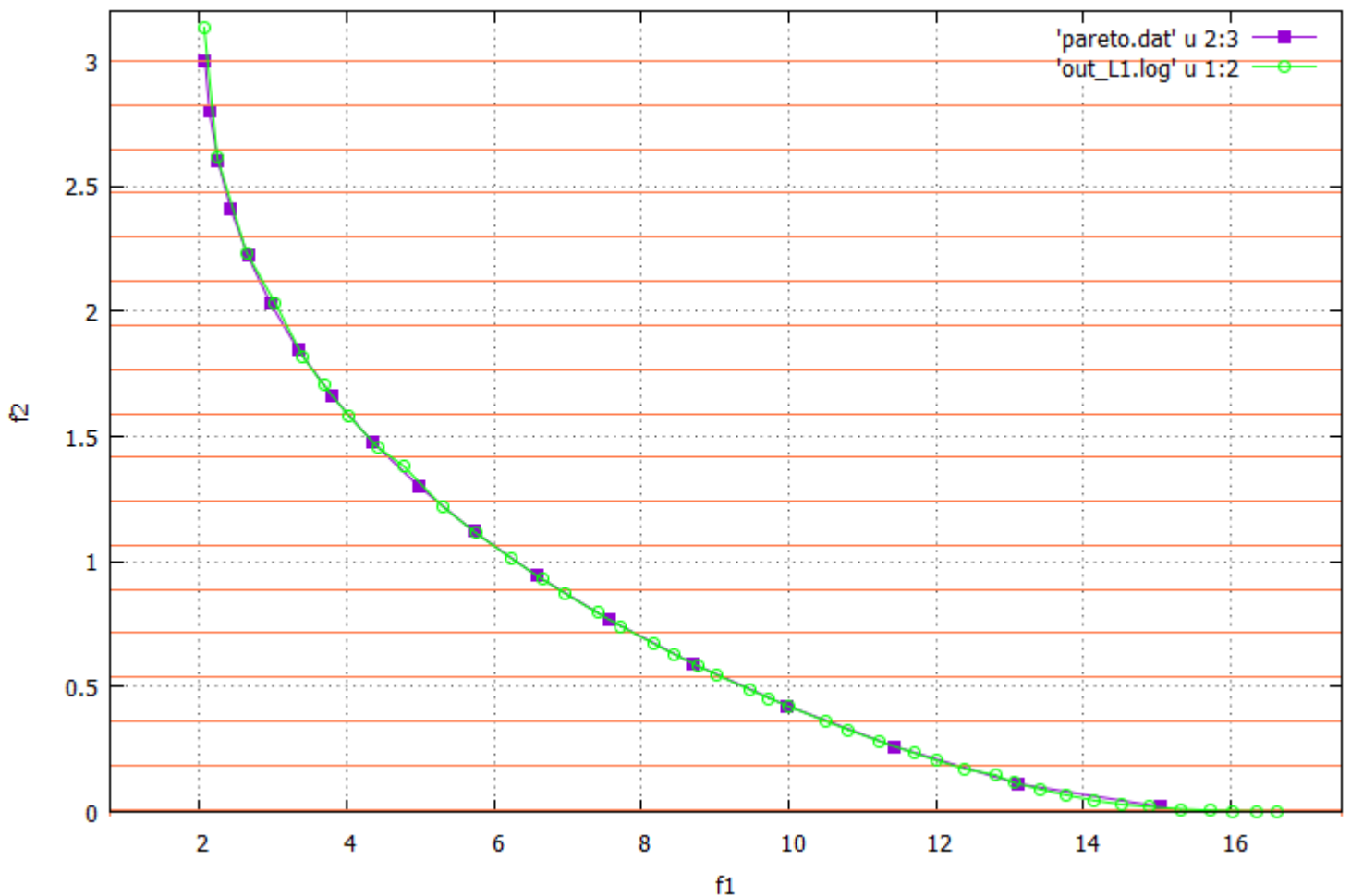
$\hat{\lambda}_2^0$	ω_p^0	γ	η	$(\bar{x}^0)^T$	Αξιολογήσεις
0	0,01	1,1	0,01	[-2 1]	303
0	0,01	1,1	0,01	[-2 2]	355
0	0,01	1,1	0,01	[-2 0]	217
0	0,01	1,1	0,01	[-2 -1]	213
0	0,01	1,1	0,01	[-1 -1]	209
0	0,01	1,1	0,01	[-1 0]	349
0	0,01	1,1	0,02	[-1 -1]	165
0	0,01	1,1	0,06	[-1 -1]	163
0	0,01	1,1	0,09	[-2 1]	171
0	0,02	1,1	0,06	[-1 -1]	149
0	0,04	1,1	0,06	[-1 -1]	137
0	0,04	1,1	0,08	[-1 -1]	135
0	0,06	1,1	0,06	[-1 -1]	129
0	0,1	1,1	0,06	[-1 -1]	121
0	0,14	1,1	0,06	[-1 -1]	115
0	0,18	1,1	0,06	[-1 -1]	111
0	0,28	1,1	0,05	[-1 -1]	105
0	0,36	1,1	0,04	[-1 -1]	103
0	0,81	1,1	0,03	[-1 -1]	97
0	0,85	1,1	0,02	[-1 -1]	95
0	0,85	1,1	0,02	[-2 1]	111
0	1,45	1,1	0,02	[-2 1]	95
0	0,85	1,2	0,02	[-1 -1]	99
0	0,85	1,05	0,02	[-1 -1]	103
0,1	0,85	1,1	0,02	[-1 -1]	101

Πίνακας 5.13: Συνολικές αξιολογήσεις που πραγματοποιούνται για την εύρεση του δεύτερου τμήματος του μειώπου Pareto για διαφορετικές τιμές των παραμέτρων της ALM όταν οι ανοχές για τα κριτήρια σύγκλισης της μεθόδου της απότομης καθόδου και της ALM λαμβάνουν τις τιμές 1 και 10^{-4} αντίστοιχα, $\hat{f}_2^{up} = 3$ και $n_{elit} = 18$.

Όπως φαίνεται από τον πίνακα 5.13, ο καλύτερος συνδυασμός παραμέτρων που επιτεύχθη για $\bar{x}^0 = [-1 \ -1]^T$ είναι $\hat{\lambda}_2^0 = 0$, $\omega_p^0 = 0,85$, $\gamma = 1,1$ και $\eta = 0,02$, ενώ για $\bar{x}^0 = [-2 \ 1]^T$ είναι $\hat{\lambda}_2^0 = 0$, $\omega_p^0 = 1,45$, $\gamma = 1,1$ και $\eta = 0,02$. Αμφότεροι οι προαναφερθέντες συνδυασμοί απαιτούν την πραγματοποίηση μόλις 95 συνολικών αξιολογήσεων (27

από τις οποίες αφορούν στην εύρεση του πρώτου σημείου του δεύτερου τμήματος του μετώπου Pareto).

Μετά την εύρεση της πρώτης, το πρόγραμμα προβαίνει στον υπολογισμό και των υπολοίπων μη-κυριαρχούμενων λύσεων που αντιστοιχούν στο δεύτερο τμήμα του διακοπόμενου μετώπου με βάση τις παραμέτρους n_{elit} και \hat{f}_2^{down} που ορίστηκαν παραπάνω. Το τμήμα του μετώπου Pareto που προκύπτει απεικονίζεται στο σχήμα 5.27. Στο ίδιο σχήμα φαίνεται και το αντίστοιχο τμήμα του μετώπου Pareto που προέκυψε από το λογισμικό των εξελικτικών αλγορίθμων. Στο σχήμα 5.27 έχουν σχεδιαστεί, επίσης, και n_{elit} οριζόντιες γραμμές οι οποίες δηλώνουν την τιμή του \hat{f}_2 για καθένα από τα n_{elit} σημεία των οποίων ζητήθηκε η εύρεση.

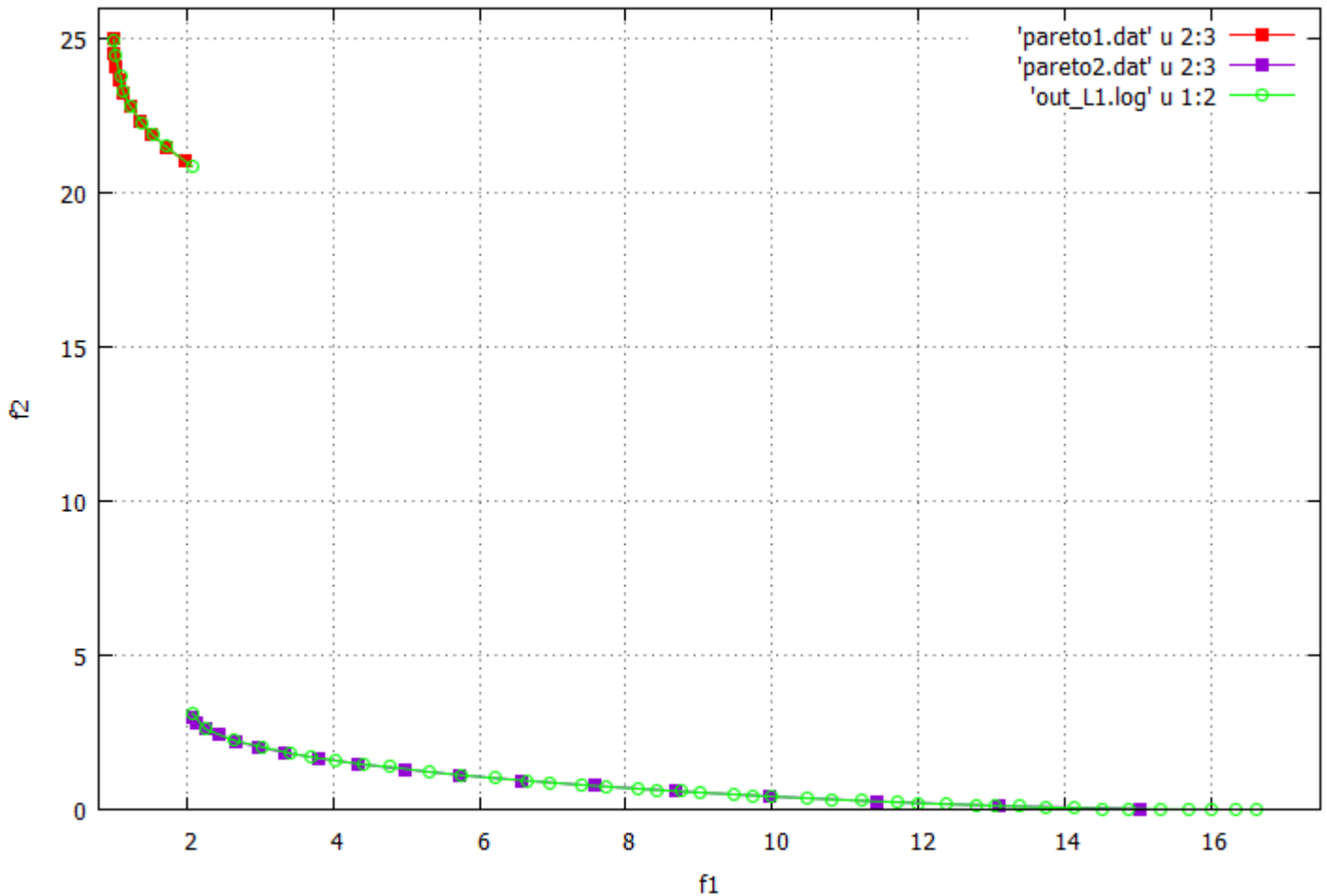


Σχήμα 5.27: Κάτω τμήμα ασυνεχούς μετώπου Pareto που προέκυψε από τη μέθοδο που βασίζεται στη συνέχιση της καμπύλης των μη-κυριαρχούμενων λύσεων για $\hat{f}_2^{up} = 3$, $\hat{f}_2^{down} = 0,01$ και $n_{elit} = 18$ (μωβ καμπύλη) και αντίστοιχο τμήμα μετώπου Pareto που προέκυψε από το λογισμικό των εξελικτικών αλγορίθμων (πράσινη καμπύλη). Οι οριζόντιες γραμμές χρώματος πορτοκαλί αντιπροσωπεύουν τις τιμές των \hat{f}_2 για καθένα από τα n_{elit} σημεία που αναζητήθηκαν.

Από την παρατήρηση του σχήματος 5.27, κατά πρώτο λόγο, καθίσταται φανερό το γεγονός ότι και εδώ υφίστανται διαφορές μεταξύ των τιμών του \hat{f}_2 και του f_2 για τα υπολογισθέντα σημεία από το πρόγραμμα. Στη συγκεκριμένη περίπτωση, οι εν λόγω διαφορές αρχικά είναι πολύ μικρές, έπειτα αυξάνονται και γίνονται όλο και πιο αισθητές και στα 6 τελευταία σημεία αρχίζουν και μειώνονται. Σε κάθε περίπτωση, όμως, μέσω του σχήματος, καθίσταται ξεκάθαρο ότι οι εν λόγω διαφορές παραμένουν σχετικά μικρές. Χαρακτηριστικά, στο δεύτερο αποκτηθέν σημείο ισχύει $f_2 - \hat{f}_2 \cong -0,024096$, στο δέκατο $f_2 - \hat{f}_2 \cong -0,116536$, στο προτελευταίο $f_2 - \hat{f}_2 \cong -0,074203$ ενώ στο τελευταίο είναι $f_2 - \hat{f}_2 \cong 0,010219$. Οι αιτίες εμφάνισης των εν λόγω διαφορών εξηγήθηκαν λεπτομερώς στην ενότητα 5.3.4 και δεν κρίνεται σκόπιμο να επαναληφθούν εδώ. Παρά, όμως, τις υφιστάμενες διαφορές $f_2 - \hat{f}_2$, παρατηρείται ότι η καμπύλη που προέκυψε από τη μέθοδο που βασίζεται στη συνέχιση της καμπύλης των μη-κυριαρχούμενων λύσεων για $\hat{f}_2^{up} = 3$, $\hat{f}_2^{down} = 0,01$ και $n_{elit} = 18$, πρακτικά, πέφτει πάνω σε αυτή που προέκυψε από το λογισμικό των εξελικτικών αλγορίθμων.

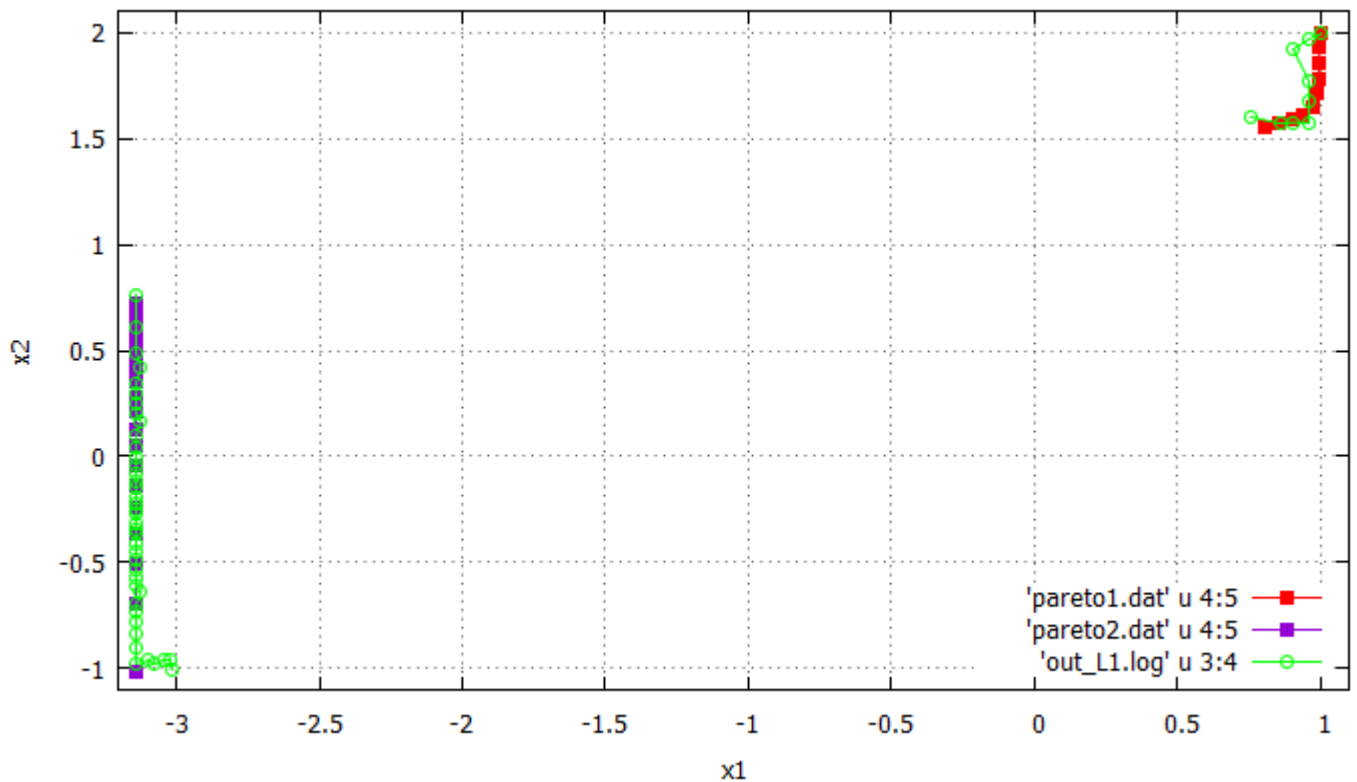
Επιπλέον, μέσω του σχήματος 5.27, διαπιστώνεται ότι η καμπύλη που αποκτήθηκε με τη μέθοδο της συνέχισης της καμπύλης των μη-κυριαρχούμενων λύσεων αποτελεί το μεγαλύτερο μέρος του πραγματικού δεύτερου τμήματος του μετώπου που αποκτήθηκε μέσω του εξελικτικού αλγορίθμου αλλά όχι ολόκληρο, καθώς η πράσινη καμπύλη εκτείνεται ελαφρώς και προς τις δύο πλευρές της μωβ. Πρέπει, όμως, να ειπωθεί ότι το αποκτηθέν τμήμα καλύπτει πολύ μεγάλο κομμάτι του δεύτερου τμήματος του μετώπου Pareto που δόθηκε από το λογισμικό των εξελικτικών αλγορίθμων οπότε, γενικά, θεωρείται ότι καλύπτει τις ανάγκες του χρήστη (υπενθυμίζεται ότι η επιλογή των \hat{f}_2^{up} και \hat{f}_2^{down} γίνεται από το χρήστη σύμφωνα με τις ανάγκες του).

Ακολουθεί η παρουσίαση των δύο τμημάτων του ασυνεχούς μετώπου Pareto που προέκυψαν από τις δύο μεθόδους σε κοινό διάγραμμα (σχήμα 5.28).

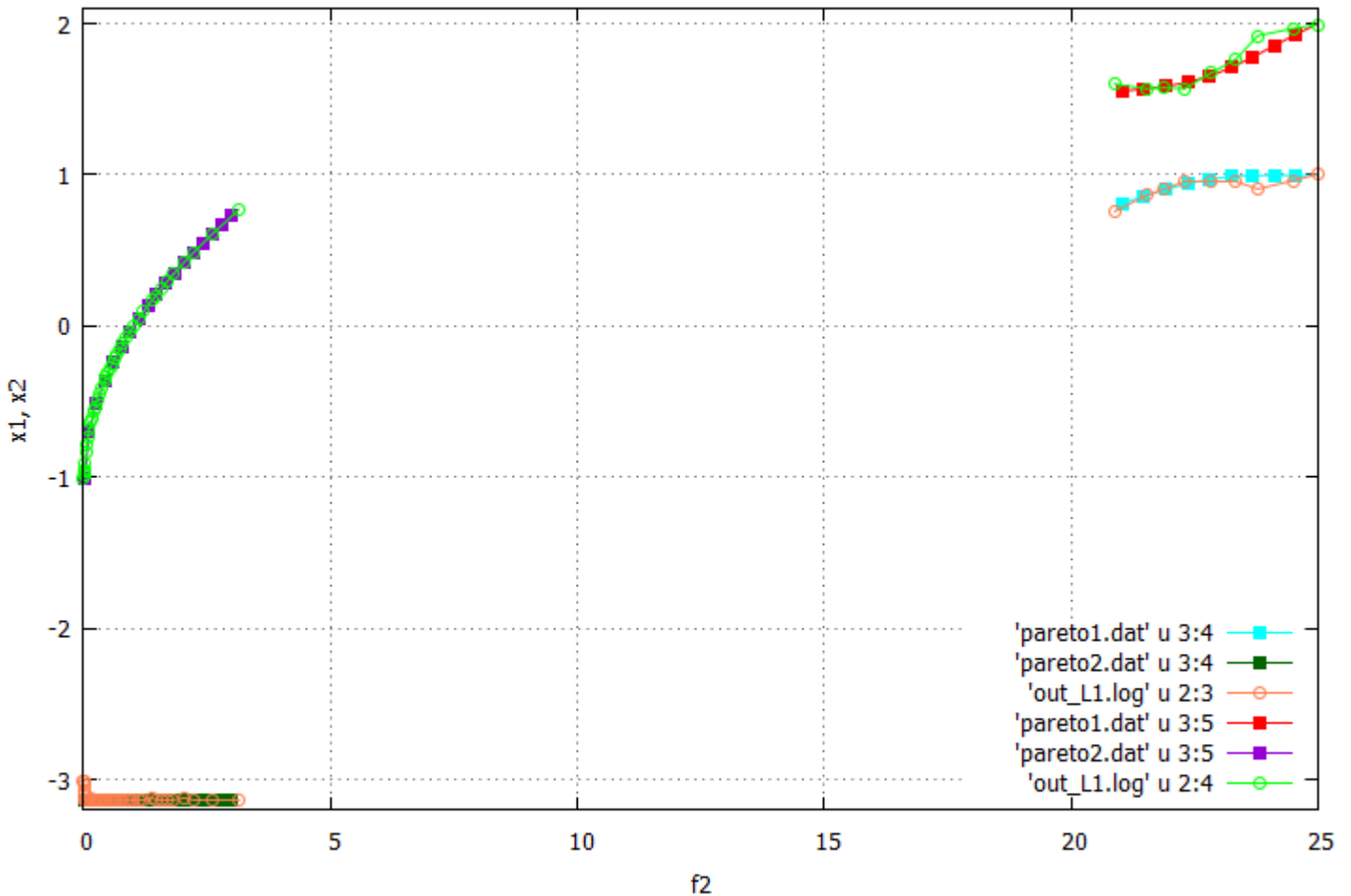


Σχήμα 5.28: Άνω και κάτω τμήματα ασυνεχούς μετώπου Pareto που προέκυψαν από τη μέθοδο που βασίζεται στη συνέχιση της καμπύλης των μη-κυριαρχούμενων λύσεων για $\hat{f}_2^{up} = 25$, $\hat{f}_2^{down} = 21$, $n_{elit} = 10$ (κόκκινη καμπύλη) και για $\hat{f}_2^{up} = 3$, $\hat{f}_2^{down} = 0,01$, $n_{elit} = 18$ (μωβ καμπύλη) αντίστοιχα και μέτωπο Pareto που προέκυψε από το λογισμικό των εξελικτικών αλγορίθμων (πράσινη καμπύλη).

Στη συνέχεια, παρατίθενται τα διαγράμματα $x_1 - x_2$ (σχήμα 5.29), $f_2 - x_1$ και $f_2 - x_2$ (σχήμα 5.30) που προέκυψαν από τις δύο μεθόδους.



Σχήμα 5.29: Διάγραμμα των καμπυλών $x_1 - x_2$ που προέκυψαν από τη μέθοδο που βασίζεται στη συνέχιση της καμπύλης των μη-κυριαρχούμενων λύσεων για $\hat{f}_2^{up} = 25$, $\hat{f}_2^{down} = 21$, $n_{elit} = 10$ (κόκκινη καμπύλη) και για $\hat{f}_2^{up} = 3$, $\hat{f}_2^{down} = 0,01$, $n_{elit} = 18$ (μωβ καμπύλη) και από το λογισμικό των εξελικτικών αλγορίθμων (πράσινη καμπύλη).

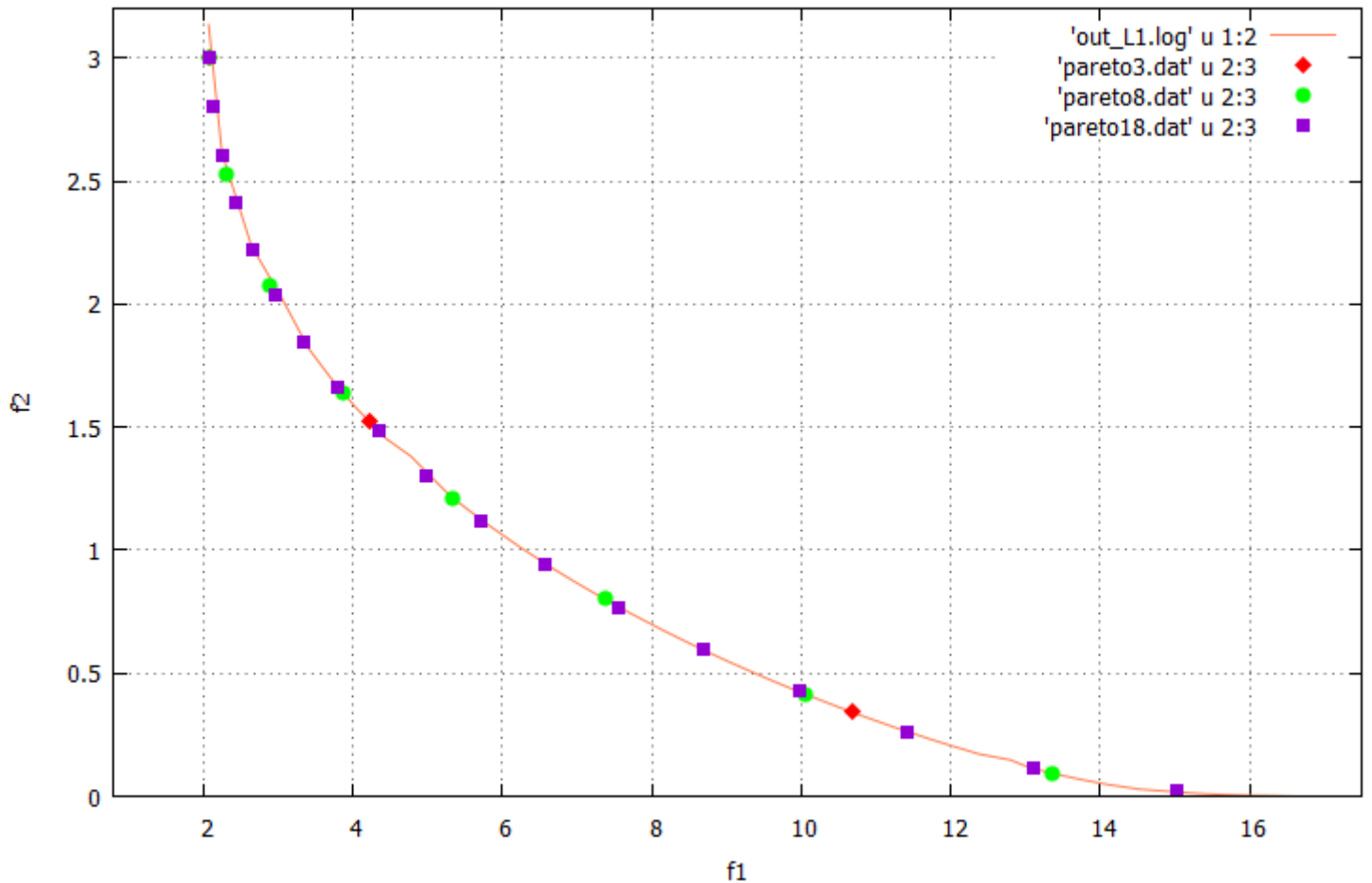


Σχήμα 5.30: Καμπύλες $f_2 - x_1$ και $f_2 - x_2$ που προέκυψαν από τη μέθοδο που βασίζεται στη συνέχιση της καμπύλης των μη-κυριαρχούμενων λύσεων για $\hat{f}_2^{up} = 25$, $\hat{f}_2^{down} = 21$, $n_{elit} = 10$ (γαλάζιο και κόκκινο χρώμα αντίστοιχα) και για $\hat{f}_2^{up} = 3$, $\hat{f}_2^{down} = 0,01$, $n_{elit} = 18$ (σκούρο πράσινο και μωβ χρώμα αντίστοιχα) και από το λογισμικό των εξελικτικών αλγορίθμων (πορτοκαλί και ανοιχτό πράσινο χρώμα αντίστοιχα).

Παρατηρείται ότι όσον αφορά την περίπτωση του πρώτου τμήματος του μετώπου Pareto στα διαγράμματα $x_1 - x_2$ (σχήμα 5.29), $f_2 - x_1$ και $f_2 - x_2$ (σχήμα 5.30) απεικονίζονται καμπύλες γραμμές, σε αντίθεση με την περίπτωση του δεύτερου τμήματος του μετώπου όπου στα διαγράμματα $x_1 - x_2$ και $f_2 - x_1$ πρακτικά απεικονίζονται ευθείες γραμμές (αν εξαιρεθούν τα τμήματα που τείνουν στο $f_2 = 0$, τα οποία όμως δεν συμπεριλαμβάνονται στο διάστημα των τιμών του f_2 που αντιστοιχεί στα αποκτηθέντα από το πρόγραμμα σημεία). Η ανανέωση των τιμών των μεταβλητών σχεδιασμού, όπως δηλώθηκε σε παραπάνω σημείο, γίνεται με βάση τη σχέση 5.1 η οποία αποτελεί μια προσέγγιση πρώτης τάξης στο χώρο $\hat{f}_2 - \bar{z}$. Δεδομένου, λοιπόν, ότι για την περίπτωση του πρώτου τμήματος του μετώπου Pareto οι γραφικές παραστάσεις $f_2 - x_1$ και $f_2 - x_2$ είναι καμπύλες γραμμές, η προσέγγιση που γίνεται ενέχει σφάλμα και για τις δύο μεταβλητές σχεδιασμού. Στην περίπτωση δε του δεύτερου τμήματος του μετώπου το σφάλμα υπάρχει μόνο στον υπολογισμό της δεύτερης μεταβλητής,

καθώς η πρώτη είναι ‘κολλημένη’ στο κάτω όριο της. Το σφάλμα που αφορά στις μεταβλητές σχεδιασμού σαφέστατα ‘μεταφέρεται’ και στον υπολογισμό των συναρτήσεων στόχου. Όσον αφορά το πρώτο τμήμα του μετώπου Pareto, η ολοένα αυξανόμενη τιμή της διαφοράς $f_2 - \hat{f}_2$ που διαπιστώθηκε παραπάνω, λοιπόν, οφείλεται σε ‘μεταφερόμενα’ σφάλματα στον υπολογισμό και των δύο μεταβλητών σχεδιασμού. Αυτό σημαίνει ότι τα προκύπτοντα από το πρόγραμμα σημεία δεν είναι απαραίτητο να βρίσκονται πάνω στο μέτωπο Pareto. Παρ’ όλα αυτά, παρατηρήθηκε μέσω των σχημάτων 5.24 και 5.25 ότι κάτι τέτοιο επιτυγχάνεται (με τα αποκτηθέντα σημεία, βέβαια, να έχουν λίγο μικρότερο f_2 από αυτό που στοχεύθηκε) όταν το δf_2 είναι καταλλήλως μικρό (μέσω της επιλογής n_{elit}). Τα ανωτέρω έρχονται σε αντιπαράθεση με την περίπτωση του δεύτερου τμήματος του μετώπου των μη-κυριαρχούμενων λύσεων όπου η καμπύλη $\hat{f}_2 - x_1$ είναι ουσιαστικά μία οριζόντια ευθεία και άρα η εφαπτομένη σε κάθε σημείο της είναι επίσης η ίδια ευθεία και έτσι το x_1 θα παραμένει ίδιο για όλα τα σημεία του μετώπου Pareto που υπολογίζει το πρόγραμμα. Συνεπώς, η τιμή του f_2 καθορίζεται από την τιμή που λαμβάνει το x_2 . Όμως, για την καμπύλη $\hat{f}_2 - x_2$, η εφαπτομενική προσέγγιση που γίνεται ενέχει σφάλμα. Αυτό σημαίνει ότι οι τιμές των συναρτήσεων στόχου που έχουν τα αποκτηθέντα σημεία από το πρόγραμμα περικλείουν το σφάλμα που έχει γίνει στον υπολογισμό του x_2 . Έτσι, το σφάλμα που αφορά στο x_2 οδηγεί σε διαφορετικές τιμές του f_2 σε σχέση με το \hat{f}_2 που δόθηκε. Όμως, επειδή το x_1 διατηρείται σταθερό, τα προκύπτοντα σημεία θα είναι πάνω στο μέτωπο Pareto. Στο συγκεκριμένο, λοιπόν, πρόβλημα για το δεύτερο τμήμα του μετώπου ουσιαστικά δεν απασχολεί το γεγονός της απόκλισης των τιμών των f_2 και \hat{f}_2 , καθώς τα σημεία που λαμβάνονται θα είναι πάντα πάνω στο μέτωπο Pareto (αν και αυτό είναι κάτι που ο χρήστης ενδέχεται να μη γνωρίζει).

Διερευνάται η συμπεριφορά του προγράμματος στην περίπτωση μικρότερου n_{elit} (άρα μεγαλύτερου δf_2) για το δεύτερο τμήμα του μετώπου Pareto ώστε να επιβεβαιωθεί η ανωτέρω διαπίστωση. Στο σχήμα 5.31 φαίνονται τα σημεία που δίνει το πρόγραμμα για τις περιπτώσεις όπου το n_{elit} παίρνει τις τιμές 18, 8 και 3 (ισχύουν, όπως προηγουμένως, $\hat{f}_2^{up} = 3$ και $\hat{f}_2^{down} = 0,01$).



Σχήμα 5.31: Μη-κυριαρχούμενες λύσεις που προέκυψαν από τη μέθοδο που βασίζεται στη συνέχιση του μετώπου Pareto για $n_{elit} = 18$ (μωβ χρώμα), για $n_{elit} = 8$ (πράσινο χρώμα) και για $n_{elit} = 3$ (κόκκινο χρώμα). Σημειώνεται ότι σε όλες τις περιπτώσεις ισχύουν $\hat{f}_2^{up} = 3$ και $\hat{f}_2^{down} = 0,01$. Η καμπύλη χρώματος πορτοκαλί αποτελεί το αντίστοιχο τμήμα του μετώπου Pareto που αποκτήθηκε μέσω των εξελικτικών αλγορίθμων.

Αν παρατηρηθεί το σχήμα 5.31, καθίσταται φανερό το γεγονός ότι όσο το n_{elit} μειώνεται τόσο μεγαλύτερη είναι η διαφορά $f_2 - \hat{f}_2$ για τα υπολογισθέντα σημεία (χαρακτηριστικά το τελευταίο σημείο βρίσκεται όλο και υψηλότερα στον άξονα των f_2). Ωστόσο, φαίνεται ότι ακόμα και στην περίπτωση όπου $n_{elit} = 3$ (το προκύπτον $\delta\hat{f}_2$ είναι αρκετά μεγάλο), τα υπολογισθέντα σημεία παραμένουν πάνω στο δεύτερο τμήμα του μετώπου Pareto, επιβεβαιώνοντας τα σχολιασθέντα ανωτέρω.

Ακολούθως, κρίνεται απαραίτητο να δηλωθεί ότι το συγκεκριμένο μαθηματικό πρόβλημα, αν ληφθούν οι καλύτεροι συνδυασμοί που επιδεικνύονται στους πίνακες 5.11 και 5.13 οι οποίοι αντιστοιχούν σε 10 σημεία για το πρώτο τμήμα του μετώπου Pareto και 18 για το δεύτερο, δύναται να επιλυθεί με την πραγματοποίηση μόλις $67 + 95 = 162$ αξιολογήσεων. Βέβαια, πρέπει να τονιστεί ότι οι ανάγκες του χρήστη σχετικά με το πόσα σημεία θα αποκτηθούν πάνω στο μέτωπο

Pareto, την περιοχή του μετώπου που αυτά θα καλύπτουν αλλά και την απαιτούμενη ακρίβεια ενδέχεται να μειώσουν ή και να αυξήσουν τον παραπάνω αναφερθέντα αριθμό αξιολογήσεων.

Τέλος, αναφέρεται, ενδεικτικά, σε πόσες περίπου αξιολογήσεις δύναται το λογισμικό των εξελικτικών αλγορίθμων να φτάσει στο μέτωπο Pareto. Για τον εν λόγω υπολογισμό γίνεται ταυτόχρονη χρήση κατανεμημένης ανίχνευσης και ΙΡΕ με τις παραμέτρους του εξελικτικού αλγορίθμου να διατηρούν τις τιμές που παρουσιάστηκαν στην ενότητα 5.4.4.2. Διαπιστώνεται ότι μέσω της εν λόγω επιλογής παραμέτρων το λογισμικό των εξελικτικών αλγορίθμων δύναται να δώσει το μέτωπο Pareto μετά από περίπου 1000 αξιολογήσεις. Σύμφωνα με τα όσα παρουσιάστηκαν στην παρούσα ενότητα, η μέθοδος που βασίζεται στη συνέχιση του μετώπου Pareto φαίνεται να επιτυγχάνει την εύρεση του μετώπου μη-κυριαρχούμενων λύσεων πολύ γρηγορότερα. Αυτό, βέβαια, ισχύει για το δεδομένο πρόβλημα αλλά και για το σετ παραμέτρων που επιλέχθηκε στο λογισμικό των εξελικτικών αλγορίθμων (δεν είναι απαραίτητο αυτό να αντιστοιχεί στο 'ταχύτερο' αποτέλεσμα που μπορεί να δώσει το λογισμικό).

6 Βελτιστοποίηση επιδόσεων υπερηχητικού επιβατικού αεροσκάφους με βάση τη θεωρία της συνέχισης του μετώπου Pareto

6.1 Εισαγωγή

Στο κεφάλαιο που προηγήθηκε, πραγματοποιήθηκε η παρουσίαση της εφαρμογής μεθόδων που βασίζονται στη θεωρία της συνέχισης του μετώπου Pareto σε απλά μαθηματικά προβλήματα βελτιστοποίησης δύο στόχων με το μεγαλύτερο βάρος να δίνεται στην εξέταση της συμπεριφοράς τους όταν απουσιάζει το στάδιο της διόρθωσης. Στο παρόν κεφάλαιο, επιχειρείται η χρήση της θεωρίας της συνέχισης του μετώπου των μη-κυριαρχούμενων λύσεων για τη βελτιστοποίηση των επιδόσεων ενός μικρού υπερηχητικού επιβατικού αεροσκάφους (supersonic business jet).

Αρχικά, παρουσιάζεται συνοπτικά ο αλγόριθμος ανάλυσης των επιδόσεων του εν λόγω αεροσκάφους. Η διατύπωση ενός τέτοιου προβλήματος συνδυάζει πολλούς επιστημονικούς κλάδους (multi-disciplinary problem), καθώς περιλαμβάνει αεροδυναμική ανάλυση, ανάλυση βάρους, ανάλυση συστήματος πρόωσης κλπ. Στον αναφερθέντα αλγόριθμο, τα ανωτέρω υλοποιούνται μέσω εμπειρικών τύπων από τη διεθνή βιβλιογραφία, πρόκειται, δηλαδή, για έναν αλγόριθμο προκαταρκτικού σχεδιασμού. Επισημαίνεται ότι η υλοποίησή του πραγματοποιήθηκε στο πλαίσιο της διδακτορικής διατριβής της Β. Ασούτη, [21], και η λεπτομερής ενασχόληση με τα τμήματα που τον συνθέτουν δεν αποτελεί στόχο της παρούσας εργασίας αλλά θα παρουσιαστούν μόνο ορισμένα ζωτικής σημασίας για την κατανόηση και εφαρμογή του στοιχεία. Η χρήση του σχετικού λογισμικού επιτρέπει τη δοκιμή της θεωρίας της συνέχισης του μετώπου Pareto σε ένα ρεαλιστικό πρόβλημα το οποίο έχει, όμως, αρκετά χαμηλό υπολογιστικό κόστος ανά αξιολόγηση.

Στη συνέχεια, πραγματοποιείται η παρουσίαση του τρόπου υπολογισμού των πρώτων και δεύτερων παραγώγων των συναρτήσεων στόχου του προβλήματος ως προς τις μεταβλητές σχεδιασμού. Για να αποκτηθούν οι εν λόγω πληροφορίες για κάθε τρέχον διάνυσμα των μεταβλητών σχεδιασμού χρησιμοποιείται η μέθοδος των μιγαδικών μεταβλητών (ενώ για να πιστοποιηθεί η ορθότητα των προκυπτουσών τιμών των παραγώγων χρησιμοποιήθηκε και η μέθοδος των πεπερασμένων διαφορών). Για να υπολογιστεί η παράγωγος μιας πραγματικής συνάρτησης (όπως είναι ο αλγόριθμος ανάλυσης επιδόσεων του αεροσκάφους) με τη μέθοδο των μιγαδικών μεταβλητών, είναι απαραίτητο η συνάρτηση αυτή να μετατραπεί σε μιγαδική. Αυτό πρακτικά σημαίνει πως για να υπολογισθεί η παράγωγος του μοντέλου του αεροσκάφους για κάποιο σύνολο μεταβλητών σχεδιασμού, είναι απαραίτητο ο κώδικας του μοντέλου να

συντεθεί σε μιγαδική μορφή, δηλαδή όλες οι ‘διαδικασίες’ πρέπει να μεταβούν από τον άξονα των πραγματικών αριθμών στο επίπεδο των μιγαδικών. Αναφέρεται ότι η ‘μετατροπή’ αυτή του κώδικα πραγματοποιήθηκε από τον Ν. Καζαζάκη στο πλαίσιο της διπλωματικής του εργασίας, [22], και δεν θα αναλυθούν εδώ οι λεπτομέρειες σχετικά με αυτήν. Στο παρόν κεφάλαιο, θα γίνει αναφορά μόνο σε στοιχεία σχετικά με τη μέθοδο που χρησιμοποιήθηκε για την εύρεση των πρώτων και δεύτερων παραγώγων τα οποία αποτελούν αναπόσπαστο κομμάτι της εφαρμογής της θεωρίας της συνέχισης του μετώπου Pareto στο παρόν πρόβλημα. Σημειώνεται ότι το λογισμικό που λήφθηκε μελετήθηκε και τροποποιήθηκε κατάλληλα ώστε να εξυπηρετεί τις ανάγκες της παρούσας εφαρμογής, με βασικότερη τροποποίηση την προσθήκη της διαδικασίας υπολογισμού των παραγώγων των συναρτήσεων στόχου ως προς τις μεταβλητές σχεδιασμού.

Τέλος, επιδεικνύονται και σχολιάζονται τα αποτελέσματα που προέκυψαν από τη βελτιστοποίηση των επιδόσεων του υπερηχητικού επιβατικού αεροσκάφους με τη χρήση της θεωρίας της συνέχισης του μετώπου των μη-κυριαρχούμενων λύσεων. Για την απόκτηση των αποτελεσμάτων αυτών, είναι απαραίτητη η εφαρμογή του παρουσιασθέντος στην ενότητα 5.3.3 αλγορίθμου με κάποιες τροποποιήσεις στις οποίες γίνεται αναφορά σε παρακάτω σημεία.

6.2 Αλγόριθμος ανάλυσης επιδόσεων υπερηχητικού επιβατικού αεροσκάφους – Διατύπωση προβλήματος

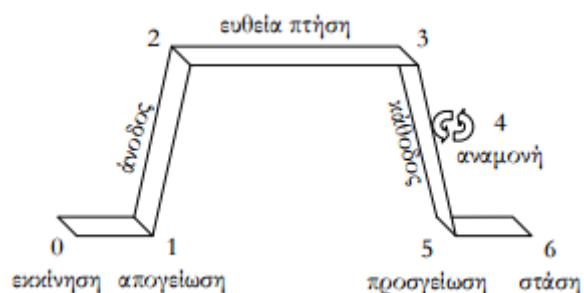
Στην παρούσα ενότητα, παρατίθενται κάποια βασικά στοιχεία του αλγορίθμου ανάλυσης των επιδόσεων ενός μικρού υπερηχητικού επιβατικού αεροσκάφους που αναπτύχθηκε στο [21]. Όπως αναφέρθηκε στην ενότητα που προηγήθηκε, κρίνεται σκόπιμη η παρουσίαση μόνο των απαραίτητων για την κατανόηση του προβλήματος και χρήση του για τους σκοπούς της παρούσας εργασίας στοιχείων του εν λόγω αλγορίθμου και όχι η λεπτομερής περιγραφή των επιμέρους τμημάτων του.

Αρχικά, πρέπει να αναφερθεί ότι στόχο της δημιουργίας του αναφερθέντος αλγορίθμου αποτελεί ο σχεδιασμός ενός μικρού (όσον αφορά τα γεωμετρικά χαρακτηριστικά αλλά και το βάρος του) υπερηχητικού αεροσκάφους με δυνατότητα μεταφοράς έως 12 επιβατών. Το εν λόγω αεροσκάφος, όμως, αν και μικρό, επιθυμείται να έχει μεγάλη εμβέλεια, ώστε να δύναται να χρησιμοποιηθεί για υπερατλαντικές πτήσεις, δηλαδή για μακρινές αποστάσεις. Είναι σαφές ότι οι δύο στόχοι του προβλήματος (μικρό βάρος και μεγάλη εμβέλεια) είναι αντικρουόμενοι, καθώς ο προφανής τρόπος για να επιτευχθεί αύξηση της εμβέλειας πτήσης είναι η αύξηση της

ποσότητας καυσίμου και, άρα, του συνολικού βάρους απογείωσης του αεροσκάφους.

Ένα βασικό χαρακτηριστικό του αεροσκάφους είναι το ότι διαθέτει μόνο κάθετο ουραίο πτερύγιο και όχι οριζόντιο.

Η αποστολή του αεροσκάφους (ή αλλιώς φάκελος πτήσης) είναι τυπική, δηλαδή περιλαμβάνει την απογείωση, την ευθεία οριζόντια πτήση, την κάθοδο, πιθανή αναμονή και την προσγείωση, όπως φαίνεται στο σχήμα 6.1.



Σχήμα 6.1: Αποστολή αεροσκάφους – Φάκελος πτήσης.

Ο δημιουργηθείς κώδικας εκφράζει ουσιαστικά μία διανυσματική συνάρτηση πολλών μεταβλητών. Ως είσοδο δέχεται ένα αρχείο το οποίο στην πρώτη γραμμή του περιέχει τον αριθμό των μεταβλητών σχεδιασμού n και σε καθεμία από τις n επόμενες την τρέχουσα τιμή που λαμβάνει η αντίστοιχη μεταβλητή σχεδιασμού. Η δε έξοδος του εν λόγω αλγορίθμου είναι ένα αρχείο που παρέχει την τιμή της εμβέλειας R (range) σε m και την τιμή του συνολικού βάρους απογείωσης TOW (take-off weight) σε kg για τις δοθείσες τιμές των μεταβλητών εισόδου.

Οι μεταβλητές \vec{x} που χρησιμοποιούνται για το σχεδιασμό και των οποίων τις τιμές απαιτεί ο εν λόγω κώδικας μέσω του αρχείου εισόδου για να δώσει το αρχείο εξόδου είναι 12 και αφορούν στοιχεία της πτήσης, την ποσότητα καυσίμου και γεωμετρικά χαρακτηριστικά της πτέρυγας και του ουραίου πτερυγίου. Πιο συγκεκριμένα,

- για την ευθεία πτήση (δείκτης cr εκ του cruise), στα ζητούμενα του σχεδιασμού είναι ο αριθμός Mach M_{cr} και το υψόμετρο πτήσης z_{cr} ,
- για την πτέρυγα (δείκτης w εκ του wing), επιλέγονται η επιφάνειά της σε κάτοψη S_w , οι γωνίες οπισθόκλισης στην ακμή προσβολής Λ_{LEw} (leading edge sweep angle) και εκφυγής Λ_{TEw} (trailing edge sweep angle), η εκλέπτυνση λ_w (taper ratio) και ο λόγος μέγιστου πάχους προς χορδή $(t/c)_w$ (relative thickness),
- για το κάθετο ουραίο πτερύγιο (δείκτης v εκ του vertical wing), επιλέγονται οι γωνίες οπισθόκλισης στην ακμή προσβολής Λ_{LEv} και εκφυγής Λ_{TEv} , η εκλέπτυνση λ_v και ο λόγος μέγιστου πάχους προς χορδή $(t/c)_v$ και

■ για το καύσιμο ως μεταβλητή σχεδιασμού υιοθετείται το συνολικό του βάρος W_{fuel} .

Στο σημείο αυτό, αναφέρεται ότι για τις ανάγκες της παρούσας εργασίας, λήφθηκε η απόφαση ως ελεύθερες μεταβλητές σχεδιασμού να χρησιμοποιηθούν μόνο 4 εκ των ανωτέρω, οι οποίες είναι το υψόμετρο πτήσης z_{cr} , η εκλέπτυνση της πτέρυγας λ_w , η εκλέπτυνση του κάθετου ουραίου πτερυγίου λ_v και το συνολικό βάρος του καυσίμου W_{fuel} . Οι υπόλοιπες εκ των μεταβλητών που αναφέρθηκαν παραπάνω λαμβάνουν συγκεκριμένες σταθερές τιμές καθ' όλη τη διάρκεια της βελτιστοποίησης. Τα όρια των ελεύθερων μεταβλητών σχεδιασμού που χρησιμοποιούνται κατά τη βελτιστοποίηση αλλά και οι δεδομένες τιμές που λαμβάνουν οι σταθερές μεταβλητές φαίνονται στον πίνακα που ακολουθεί.

Μεταβλητή	Ελάχιστη τιμή	Μέγιστη τιμή	Μονάδες
M_{cr}	2	2	-
z_{cr}	11000	18000	m
S_w	120	120	m^2
Λ_{LEw}	70	70	(°)
Λ_{TEw}	7,5	7,5	(°)
λ_w	0,05	0,5	-
$(t/c)_w$	0,04	0,04	-
Λ_{LEv}	69	69	(°)
Λ_{TEv}	12,4	12,4	(°)
λ_v	0,05	0,5	-
$(t/c)_v$	0,06	0,06	-
W_{fuel}	15000	30000	kg

Πίνακας 6.1: Μεταβλητές σχεδιασμού.

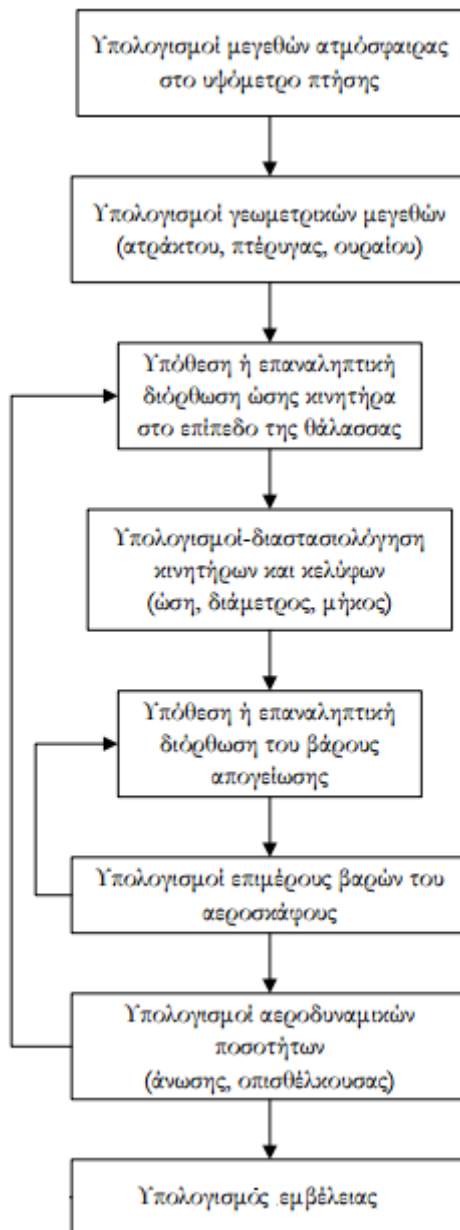
Έτσι, το πρόβλημα βελτιστοποίησης που αποτελεί αντικείμενο του δεδομένου κεφαλαίου διατυπώνεται ως εξής:

$$\begin{bmatrix} \max R(\vec{x}) \\ \min TOW(\vec{x}) \end{bmatrix} = \min \begin{bmatrix} -R(\vec{x}) \\ TOW(\vec{x}) \end{bmatrix}$$

που υπόκειται στους περιορισμούς που υπαγορεύει ο ανωτέρω πίνακας.

Ακολουθως, περιγράφεται εν συντομία ο αλγόριθμος που παρέχει τις τιμές της εμβέλειας και του συνολικού βάρους απογείωσης για ένα δεδομένο διάνυσμα των μεταβλητών σχεδιασμού που αναφέρθηκαν παραπάνω. Αρχικά, υπολογίζονται τα χαρακτηριστικά της τυπικής ατμόσφαιρας στο δεδομένο υψόμετρο πτήσης. Έπειτα, υπολογίζονται τα γεωμετρικά χαρακτηριστικά της ατράκτου, της πτέρυγας και του κάθετου ουραίου πτερυγίου.

Στη συνέχεια, ο αλγόριθμος προβαίνει στην εύρεση της ώσης και του συνολικού βάρους του αεροσκάφους μέσω επαναληπτικής διαδικασίας, έως ότου ικανοποιηθεί η απαίτηση ισότητας της άνωσης με το βάρος και της οπισθέλκουσας με την ώση κατά την ευθεία πτήση. Συγκεκριμένα, σε πρώτη φάση, γίνεται μία υπόθεση για την τιμή της ώσης στο επίπεδο της θάλασσας. Σε δεύτερη φάση, πραγματοποιείται ο υπολογισμός της ώσης του κινητήρα και των γεωμετρικών χαρακτηριστικών του μέσω εμπειρικών σχέσεων οι οποίες συνδέουν τα αναφερθέντα μεγέθη με την ώση στο επίπεδο της θάλασσας, το υψόμετρο και την ταχύτητα πτήσης. Ακολουθεί μια δεύτερη υπόθεση η οποία αφορά στο συνολικό βάρος απογείωσης, συναρτήσει του οποίου υπολογίζονται τα βάρη των επιμέρους τμημάτων του αεροσκάφους. Έπειτα, τα εν λόγω βάρη αθροίζονται και δίνουν μία τιμή για το συνολικό βάρος απογείωσης, η οποία σαφέστατα πρέπει να συμπίπτει με την τιμή που υποτέθηκε, γεγονός που οδηγεί στη διόρθωσή της μέχρι να συγκλίνει η διαδικασία. Έχοντας, πλέον, το συνολικό βάρος του αεροσκάφους αλλά και όλα τα γεωμετρικά του χαρακτηριστικά, ο αλγόριθμος προβαίνει στην εύρεση των αεροδυναμικών μεγεθών. Σε πρώτο στάδιο, υπολογίζεται η άνωση και ο συντελεστής άνωσης κατά την ευθεία πτήση (μέσω της απαίτησης για ισότητα βάρους-άνωσης). Σε δεύτερο στάδιο, πραγματοποιείται ο υπολογισμός του συντελεστή οπισθέλκουσας, ο οποίος οδηγεί σε μία τιμή για την οπισθέλκουσα και άρα για τη συνολική ώση (μέσω της απαίτησης για ισότητα οπισθέλκουσας-ώσης κατά την ευθεία πτήση). Έτσι, υπολογίζεται εκ νέου η ώση για κάθε κινητήρα αρχικά στο υψόμετρο πτήσης και έπειτα στο επίπεδο της θάλασσας. Με τη νέα αυτή τιμή της ώσης στο επίπεδο της θάλασσας, επαναλαμβάνονται τα ανωτέρω περιγραφέντα βήματα και όταν επιτευχθεί η σύγκλιση της διαδικασίας αυτής, διατίθεται, πλέον, η ώση και όλα τα χαρακτηριστικά του αεροσκάφους. Τέλος, ο αλγόριθμος παρέχει την τιμή της εμβέλειας του αεροσκάφους μέσω της εξίσωσης Breguet. Η διαγραμματική απεικόνιση του αλγορίθμου φαίνεται στο σχήμα 6.2.



Σχήμα 6.2: Αλγόριθμος υπολογισμού συνολικού βάρους απογείωσης και εμβέλειας του αεροσκάφους για δεδομένες τιμές των μεταβλητών σχεδιασμού.

Σημειώνεται ότι η εκτενής ανάλυση των σχέσεων που χρησιμοποιούνται για τον υπολογισμό των επιμέρους ποσοτήτων (μεγέθη τυπικής ατμόσφαιρας, γεωμετρικά μεγέθη, βάρη, αεροδυναμικοί υπολογισμοί κλπ) σε κάθε βήμα του αλγορίθμου δεν κρίνεται σκόπιμο να παρουσιαστεί εδώ. Για τις εν λόγω σχέσεις θα πρέπει κανείς να ανατρέξει στη διδακτορική διατριβή [21].

6.3 Εφαρμογή της θεωρίας της συνέχισης του μετώπου Pareto στο πρόβλημα

Σύμφωνα με τα όσα αναπτύχθηκαν στην ενότητα 4.3.2, το παραπάνω πρόβλημα βελτιστοποίησης επαναδιατυπώνεται ως εξής:

$$\min f_1(\vec{x})$$

που υπόκειται σε (subject to – s.t.)

$$f_2(\vec{x}) = \hat{f}_2 \Leftrightarrow f_2(\vec{x}) - \hat{f}_2 = 0$$

όπου $f_1 = -R$, $f_2 = TOW$ και $\hat{f}_2 = \overline{TOW}$.

Από την παραπάνω διατύπωση είναι φανερό ότι επιλέχθηκε $l = 1$. Θα ήταν δυνατό, εναλλακτικά, να επιλεχθεί η ελαχιστοποίηση της δεύτερης συνάρτησης στόχου με την πρώτη να αποτελεί τον περιορισμό ισότητας.

Για λόγους που εξηγήθηκαν στην ενότητα 5.3.2, λήφθηκε η απόφαση οι περιορισμοί που αφορούν στα όρια των μεταβλητών σχεδιασμού (όπως αυτοί παρουσιάστηκαν στον πίνακα της ενότητας που προηγήθηκε) να μην υπεισέρχονται στη συνάρτηση Lagrange άρα ούτε και στις συνθήκες KKT αλλά να εκτελείται από τον αλγόριθμο ένας απλός έλεγχος για το αν ικανοποιούνται κάθε φορά που μεταβάλλονται οι μεταβλητές σχεδιασμού.

Για το δεδομένο, λοιπόν, πρόβλημα, η συνάρτηση Lagrange γράφεται ως εξής:

$$L(\vec{x}, \hat{\lambda}_2) = f_1(\vec{x}) - \hat{\lambda}_2 [f_2(\vec{x}) - \hat{f}_2]$$

Οι δε συνθήκες KKT διατυπώνονται ως εξής:

$$\nabla_x L(\vec{x}, \hat{\lambda}_2) = 0 \Leftrightarrow \nabla_x f_1(\vec{x}) - \hat{\lambda}_2 \nabla_x f_2(\vec{x}) = 0$$

$$f_2(\vec{x}) - \hat{f}_2 = 0$$

Εναλλακτικά, το ανωτέρω σύστημα εξισώσεων γράφεται:

$$\vec{H}(\vec{x}, \hat{\lambda}_2, \hat{f}_2) = 0$$

δηλαδή

$$H_1 = \frac{\partial f_1(\vec{x})}{\partial x_1} - \hat{\lambda}_2 \frac{\partial f_2(\vec{x})}{\partial x_1} = 0$$

$$H_2 = \frac{\partial f_1(\vec{x})}{\partial x_2} - \hat{\lambda}_2 \frac{\partial f_2(\vec{x})}{\partial x_2} = 0$$

$$H_3 = \frac{\partial f_1(\vec{x})}{\partial x_3} - \hat{\lambda}_2 \frac{\partial f_2(\vec{x})}{\partial x_3} = 0$$

$$H_4 = \frac{\partial f_1(\vec{x})}{\partial x_4} - \hat{\lambda}_2 \frac{\partial f_2(\vec{x})}{\partial x_4} = 0$$

$$H_5 = f_2(\vec{x}) - \hat{f}_2 = 0$$

όπου $x_1 = z_{cr}$, $x_2 = \lambda_w$, $x_3 = \lambda_v$ και $x_4 = W_{fuel}$ (τονίζεται ότι σε αυτή τη διαδικασία δεν συμμετέχουν οι μεταβλητές που διατηρούν σταθερή τιμή).

Ορίζεται το διάνυσμα $\vec{z} = (\vec{x}, \hat{\lambda}_2)^T = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ \hat{\lambda}_2 \end{bmatrix}$.

Σύμφωνα με το θεώρημα πεπλεγμένης συνάρτησης έχουμε:

$$\vec{H}(\vec{z}, \hat{f}_2) = 0 \Leftrightarrow \vec{z} = \vec{h}(\hat{f}_2)$$

δηλαδή

$$z_1 = x_1 = h_1(\hat{f}_2)$$

$$z_2 = x_2 = h_2(\hat{f}_2)$$

$$z_3 = x_3 = h_3(\hat{f}_2)$$

$$z_4 = x_4 = h_4(\hat{f}_2)$$

$$z_5 = \hat{\lambda}_2 = h_5(\hat{f}_2)$$

Επιπροσθέτως, λόγω του θεωρήματος αυτού ισχύει:

$$\frac{\partial \vec{h}}{\partial \hat{f}_2}(\hat{f}_2) = - \left\{ \frac{\partial \vec{H}}{\partial \vec{z}} [\vec{h}(\hat{f}_2), \hat{f}_2] \right\}^{-1} \frac{\partial \vec{H}}{\partial \hat{f}_2} [\vec{h}(\hat{f}_2), \hat{f}_2]$$

Από την ανωτέρω σχέση, καθίσταται φανερό το γεγονός ότι για τον υπολογισμό της παραγώγου $\partial \vec{h} / \partial \hat{f}_2$ απαιτείται ο υπολογισμός των παραγώγων $\partial \vec{H} / \partial \vec{z}$ και $\partial \vec{H} / \partial \hat{f}_2$. Έτσι, έχουμε:

$$\frac{\partial \vec{h}}{\partial \hat{f}_2} = \begin{bmatrix} \partial H_1 / \partial \hat{f}_2 \\ \partial H_2 / \partial \hat{f}_2 \\ \partial H_3 / \partial \hat{f}_2 \\ \partial H_4 / \partial \hat{f}_2 \\ \partial H_5 / \partial \hat{f}_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -1 \end{bmatrix}$$

$$\frac{\partial \vec{h}}{\partial \vec{z}} = \begin{bmatrix} \partial H_1 / \partial x_1 & \partial H_1 / \partial x_2 & \partial H_1 / \partial x_3 & \partial H_1 / \partial x_4 & \partial H_1 / \partial \hat{\lambda}_2 \\ \partial H_2 / \partial x_1 & \partial H_2 / \partial x_2 & \partial H_2 / \partial x_3 & \partial H_2 / \partial x_4 & \partial H_2 / \partial \hat{\lambda}_2 \\ \partial H_3 / \partial x_1 & \partial H_3 / \partial x_2 & \partial H_3 / \partial x_3 & \partial H_3 / \partial x_4 & \partial H_3 / \partial \hat{\lambda}_2 \\ \partial H_4 / \partial x_1 & \partial H_4 / \partial x_2 & \partial H_4 / \partial x_3 & \partial H_4 / \partial x_4 & \partial H_4 / \partial \hat{\lambda}_2 \\ \partial H_5 / \partial x_1 & \partial H_5 / \partial x_2 & \partial H_5 / \partial x_3 & \partial H_5 / \partial x_4 & \partial H_5 / \partial \hat{\lambda}_2 \end{bmatrix} =$$

$$\begin{bmatrix} \frac{\partial^2 f_1}{\partial x_1^2} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_1^2} & \frac{\partial^2 f_1}{\partial x_1 \partial x_2} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_1 \partial x_2} & \frac{\partial^2 f_1}{\partial x_1 \partial x_3} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_1 \partial x_3} & \frac{\partial^2 f_1}{\partial x_1 \partial x_4} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_1 \partial x_4} & -\frac{\partial f_2}{\partial x_1} \\ \frac{\partial^2 f_1}{\partial x_1 \partial x_2} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_1 \partial x_2} & \frac{\partial^2 f_1}{\partial x_2^2} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_2^2} & \frac{\partial^2 f_1}{\partial x_2 \partial x_3} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_2 \partial x_3} & \frac{\partial^2 f_1}{\partial x_2 \partial x_4} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_2 \partial x_4} & -\frac{\partial f_2}{\partial x_2} \\ \frac{\partial^2 f_1}{\partial x_1 \partial x_3} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_1 \partial x_3} & \frac{\partial^2 f_1}{\partial x_2 \partial x_3} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_2 \partial x_3} & \frac{\partial^2 f_1}{\partial x_3^2} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_3^2} & \frac{\partial^2 f_1}{\partial x_3 \partial x_4} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_3 \partial x_4} & -\frac{\partial f_2}{\partial x_3} \\ \frac{\partial^2 f_1}{\partial x_1 \partial x_4} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_1 \partial x_4} & \frac{\partial^2 f_1}{\partial x_2 \partial x_4} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_2 \partial x_4} & \frac{\partial^2 f_1}{\partial x_3 \partial x_4} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_3 \partial x_4} & \frac{\partial^2 f_1}{\partial x_4^2} - \hat{\lambda}_2 \frac{\partial^2 f_2}{\partial x_4^2} & -\frac{\partial f_2}{\partial x_4} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_3} & \frac{\partial f_2}{\partial x_4} & 0 \end{bmatrix}$$

Στη συνέχεια, καθίσταται δυνατό, αφού εκτελεστεί ο υπολογισμός της παραγώγου $\partial \vec{h} / \partial \hat{f}_2$, να πραγματοποιηθεί η εύρεση των νέων τιμών των μεταβλητών σχεδιασμού και του $\hat{\lambda}_2$, δηλαδή του επόμενου σημείου πάνω στο μέτωπο Pareto (στην πραγματικότητα, όπως εξηγήθηκε παραπάνω, αποτελεί προσέγγισή του), ως εξής:

$$\vec{z}^{new} = \vec{z}^{old} + \frac{\partial \vec{h}}{\partial \hat{f}_2} (\hat{f}_2^{new} - \hat{f}_2^{old})$$

δηλαδή

$$z_1^{new} = x_1^{new} = x_1^{old} + \frac{\partial h_1}{\partial \hat{f}_2} (\hat{f}_2^{new} - \hat{f}_2^{old})$$

$$z_2^{new} = x_2^{new} = x_2^{old} + \frac{\partial h_2}{\partial \hat{f}_2} (\hat{f}_2^{new} - \hat{f}_2^{old})$$

$$z_3^{new} = x_3^{new} = x_3^{old} + \frac{\partial h_3}{\partial \hat{f}_2} (\hat{f}_2^{new} - \hat{f}_2^{old})$$

$$z_4^{new} = x_4^{new} = x_4^{old} + \frac{\partial h_4}{\partial \hat{f}_2} (\hat{f}_2^{new} - \hat{f}_2^{old})$$

$$z_5^{new} = \hat{\lambda}_2^{new} = \hat{\lambda}_2^{old} + \frac{\partial h_5}{\partial \hat{f}_2} (\hat{f}_2^{new} - \hat{f}_2^{old})$$

6.4 Περιγραφή του αλγορίθμου που αναπτύχθηκε για την επίλυση του προβλήματος

6.4.1 Γενικά

Αρχικά, πρέπει να αναφερθεί ότι για να επιτευχθεί η επίλυση του προβλήματος βελτιστοποίησης που παρουσιάστηκε στην ενότητα 6.2, λαμβάνεται η απόφαση να χρησιμοποιηθεί η μέθοδος που βασίζεται στη συνέχιση του μετώπου Pareto χωρίς τη συμπερίληψη του σταδίου της διόρθωσης. Στην ενότητα αυτή, σχολιάζονται μόνο οι διαφορές που συναντώνται μεταξύ του αναπτυχθέντος για το εν λόγω πρόβλημα αλγορίθμου και των αλγορίθμων που παρουσιάστηκαν στο κεφάλαιο 5.

Όπως δηλώθηκε στην ενότητα 6.2, το πρόβλημα που πραγματεύεται το παρόν κεφάλαιο εμπεριέχει και μεταβλητές σχεδιασμού οι οποίες διατηρούν σταθερή τιμή σε όλη τη διαδικασία της βελτιστοποίησης σε αντίθεση με τα μαθηματικά προβλήματα που παρουσιάστηκαν στο κεφάλαιο που προηγήθηκε στα οποία συναντώνται μόνο μεταβλητές σχεδιασμού που κινούνται μεταξύ των άνω και κάτω ορίων τους. Συνεπώς, απαιτείται να πραγματοποιηθούν κάποιες αλλαγές στον κώδικα ώστε οι μεταβλητές που διατηρούν πάντοτε μία δεδομένη σταθερή τιμή να 'αποκλείονται' από οποιαδήποτε διεργασία του προγράμματος που αφορά σε μεταβολή των μεταβλητών σχεδιασμού.

Το δεύτερο τμήμα που είναι άξιο σχολιασμού είναι αυτό που αφορά στον υπολογισμό των πρώτων και δεύτερων παραγώγων των συναρτήσεων στόχου ως προς τις μεταβλητές σχεδιασμού για το τρέχον \vec{x} . Στην περίπτωση των μαθηματικών προβλημάτων που παρουσιάστηκαν στο κεφάλαιο 5, ο εν λόγω υπολογισμός έγινε αναλυτικά. Κάτι τέτοιο εδώ είναι ανέφικτο, καθώς δεν υπάρχουν καθορισμένες εκφράσεις για τις συναρτήσεις στόχου. Για να αποκτηθούν, λοιπόν, οι επιθυμητές παράγωγοι για κάθε τρέχον διάλυμα των μεταβλητών σχεδιασμού επιλέγεται να χρησιμοποιηθεί η μέθοδος των μιγαδικών μεταβλητών. Όπως αναφέρθηκε στην ενότητα 6.1, ένας τέτοιος υπολογισμός, απαιτεί τη 'μετατροπή' του κώδικα που είναι υπεύθυνος για την εύρεση του συνολικού βάρους απογείωσης και της εμβέλειας σε μιγαδική μορφή, κάτι που πραγματοποιήθηκε από τον Ν. Καζαζάκη στο πλαίσιο της διπλωματικής του εργασίας, [22]. Με βάση την εν λόγω εργασία, αναπτύχθηκε λογισμικό το οποίο επιτρέπει την εύρεση των πρώτων και δεύτερων παραγώγων των συναρτήσεων στόχου ως προς τις

μεταβλητές σχεδιασμού μέσω της μεθόδου των μιγαδικών μεταβλητών. Σημειώνεται ότι για να πιστοποιηθεί η ορθότητα των προκυπτουσών τιμών των παραγώγων που παρέχει το αναφερθέν λογισμικό, δημιουργείται ένα πρόγραμμα το οποίο προβαίνει στην εύρεση των επιθυμητών παραγώγων μέσω της μεθόδου των πεπερασμένων διαφορών. Στις παρακάτω ενότητες ακολουθεί η περιγραφή των μεθόδων που χρησιμοποιήθηκαν και η επίδειξη των αποτελεσμάτων που προέκυψαν.

Τέλος, πρέπει να τονιστεί ότι, παρά το γεγονός ότι για τον υπολογισμό των παραγώγων χρησιμοποιείται η μέθοδος των μιγαδικών μεταβλητών, οι αξιολογήσεις που πραγματοποιούνται από τον αλγόριθμο βελτιστοποίησης μετρώνται όπως και στις περιπτώσεις των μαθηματικών προβλημάτων που παρουσιάστηκαν στο κεφάλαιο 5, δηλαδή σαν να ήταν διαθέσιμες οι παράγωγοι μέσω της συζυγούς μεθόδου. Αυτό εξηγήθηκε ότι γίνεται για το λόγο ότι σε ένα πραγματικό πρόβλημα η πιο πιθανή επιλογή είναι η 'ταχύτερη' μέθοδος και άρα αυτές κρίνονται οι επιθυμητές συνθήκες για τον προσδιορισμό του υπολογιστικού κόστους της μεθόδου που βασίζεται στη συνέχιση του μειώπου Pareto.

6.4.2 Η μέθοδος των μιγαδικών μεταβλητών (complex variable method)

6.4.2.1 Θεωρητικά στοιχεία

Το ανάπτυγμα Taylor της πραγματικής συνάρτησης f μίας μεταβλητής x στο σημείο x έχει ως εξής:

$$f(x + \varepsilon) = f(x) + \varepsilon \left. \frac{\partial f}{\partial x} \right|_x + \frac{\varepsilon^2}{2} \left. \frac{\partial^2 f}{\partial x^2} \right|_x + 0(\varepsilon^3) \quad (6.1)$$

Ανάλογα για μία μιγαδική συνάρτηση f μίας μεταβλητής x της οποίας το φανταστικό μέρος είναι σταθερό και $\rightarrow 0$ ισχύει:

$$f(x + i\varepsilon) = f(x) + i\varepsilon \left. \frac{\partial f}{\partial x} \right|_x - \frac{\varepsilon^2}{2} \left. \frac{\partial^2 f}{\partial x^2} \right|_x + 0[(i\varepsilon)^3] \quad (6.2)$$

Η σχέση (6.2) αποτελεί μία μιγαδική εξίσωση. Αν, λοιπόν, εξισωθούν τα πραγματικά και τα φανταστικά της μέρη λαμβάνονται:

$$Re[f(x + i\varepsilon)] = f(x) - \frac{\varepsilon^2}{2} \left. \frac{\partial^2 f}{\partial x^2} \right|_x$$

$$Im[f(x + i\varepsilon)] = \varepsilon \left. \frac{\partial f}{\partial x} \right|_x$$

Είναι σαφές ότι μέσω της δεύτερης εκ των ανωτέρω σχέσεων δύναται να υπολογιστεί η πρώτη παράγωγος της συνάρτησης f στο σημείο x , ενώ μέσω της πρώτης επιτυγχάνεται ο υπολογισμός της δεύτερης παραγώγου της f στο x ως εξής:

$$\frac{\partial f}{\partial x}(x) = \frac{Im[f(x + i\varepsilon)]}{\varepsilon}$$

$$\frac{\partial^2 f}{\partial x^2}(x) = \frac{2}{\varepsilon^2} \{f(x) - Re[f(x + i\varepsilon)]\}$$

Αν τώρα πρόκειται για μία συνάρτηση n μεταβλητών ισχύουν:

$$\frac{\partial f}{\partial x_k}(\vec{x}) = \frac{Im[f(x_1, x_2, \dots, x_k + i\varepsilon_k, \dots, x_{n-1}, x_n)]}{\varepsilon_k}, k = 1, \dots, n \quad (6.3)$$

$$\frac{\partial^2 f}{\partial x_k^2}(\vec{x}) = \frac{2}{\varepsilon_k^2} \{f(\vec{x}) - Re[f(x_1, x_2, \dots, x_k + i\varepsilon_k, \dots, x_{n-1}, x_n)]\}, \quad (6.4)$$

$$k = 1, \dots, n$$

Στη συνέχεια, παρατίθεται ο τρόπος υπολογισμού των μικτών παραγώγων μίας συνάρτησης n μεταβλητών. Το ανάπτυγμα Taylor της πραγματικής συνάρτησης f δύο μεταβλητών x_1, x_2 στο σημείο \vec{x} έχει ως εξής:

$$f(x_1 + \varepsilon_1, x_2 + \varepsilon_2) = f(\vec{x}) + \varepsilon_1 \left. \frac{\partial f}{\partial x_1} \right|_{\vec{x}} + \varepsilon_2 \left. \frac{\partial f}{\partial x_2} \right|_{\vec{x}} + \frac{1}{2} \left(\varepsilon_1^2 \left. \frac{\partial^2 f}{\partial x_1^2} \right|_{\vec{x}} + 2\varepsilon_1 \varepsilon_2 \left. \frac{\partial^2 f}{\partial x_1 \partial x_2} \right|_{\vec{x}} + \varepsilon_2^2 \left. \frac{\partial^2 f}{\partial x_2^2} \right|_{\vec{x}} \right) + 0(\varepsilon^3) \quad (6.5)$$

Ανάλογα για μία μιγαδική συνάρτηση f δύο μεταβλητών x_1, x_2 των οποίων το φανταστικό μέρος είναι σταθερό και $\rightarrow 0$ ισχύει:

$$f(x_1 + i\varepsilon_1, x_2 + i\varepsilon_2) = f(\vec{x}) + i\varepsilon_1 \left. \frac{\partial f}{\partial x_1} \right|_{\vec{x}} + i\varepsilon_2 \left. \frac{\partial f}{\partial x_2} \right|_{\vec{x}} - \frac{1}{2} \left(\varepsilon_1^2 \left. \frac{\partial^2 f}{\partial x_1^2} \right|_{\vec{x}} + 2\varepsilon_1 \varepsilon_2 \left. \frac{\partial^2 f}{\partial x_1 \partial x_2} \right|_{\vec{x}} + \varepsilon_2^2 \left. \frac{\partial^2 f}{\partial x_2^2} \right|_{\vec{x}} \right) + 0[(i\varepsilon)^3] \quad (6.6)$$

Αν εξισωθούν τα πραγματικά μέρη της μιγαδικής εξίσωσης (6.6) προκύπτει:

$$Re[f(x_1 + i\varepsilon_1, x_2 + i\varepsilon_2)] = f(\vec{x}) - \frac{1}{2} \left(\varepsilon_1^2 \left. \frac{\partial^2 f}{\partial x_1^2} \right|_{\vec{x}} + 2\varepsilon_1 \varepsilon_2 \left. \frac{\partial^2 f}{\partial x_1 \partial x_2} \right|_{\vec{x}} + \varepsilon_2^2 \left. \frac{\partial^2 f}{\partial x_2^2} \right|_{\vec{x}} \right)$$

Έτσι, επιτυγχάνεται ο υπολογισμός της μικτής παραγώγου της f στο \vec{x} ως εξής:

$$\begin{aligned} & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\vec{x}) = \\ & = \frac{1}{2\varepsilon_1 \varepsilon_2} \left\{ 2\{f(\vec{x}) - \operatorname{Re}[f(x_1 + i\varepsilon_1, x_2 + i\varepsilon_2)]\} - \varepsilon_1^2 \left. \frac{\partial^2 f}{\partial x_1^2} \right|_{\vec{x}} - \varepsilon_2^2 \left. \frac{\partial^2 f}{\partial x_2^2} \right|_{\vec{x}} \right\} \end{aligned}$$

Ομοίως αν πρόκειται για μία συνάρτηση n μεταβλητών ισχύουν:

$$\begin{aligned} \frac{\partial^2 f}{\partial x_k \partial x_l}(\vec{x}) &= \frac{1}{\varepsilon_k \varepsilon_l} \{f(\vec{x}) - \operatorname{Re}[f(x_1, \dots, x_k + i\varepsilon_k, \dots, x_l + i\varepsilon_l, \dots, x_n)]\} \\ &\quad - \frac{1}{2\varepsilon_k \varepsilon_l} \left(\varepsilon_k^2 \left. \frac{\partial^2 f}{\partial x_k^2} \right|_{\vec{x}} + \varepsilon_l^2 \left. \frac{\partial^2 f}{\partial x_l^2} \right|_{\vec{x}} \right), \quad k = 1, \dots, n-1, l = k+1, \dots, n \quad (6.7) \\ \frac{\partial^2 f}{\partial x_l \partial x_k}(\vec{x}) &= \frac{\partial^2 f}{\partial x_k \partial x_l}(\vec{x}), \quad k = 1, \dots, n-1, l = k+1, \dots, n \end{aligned}$$

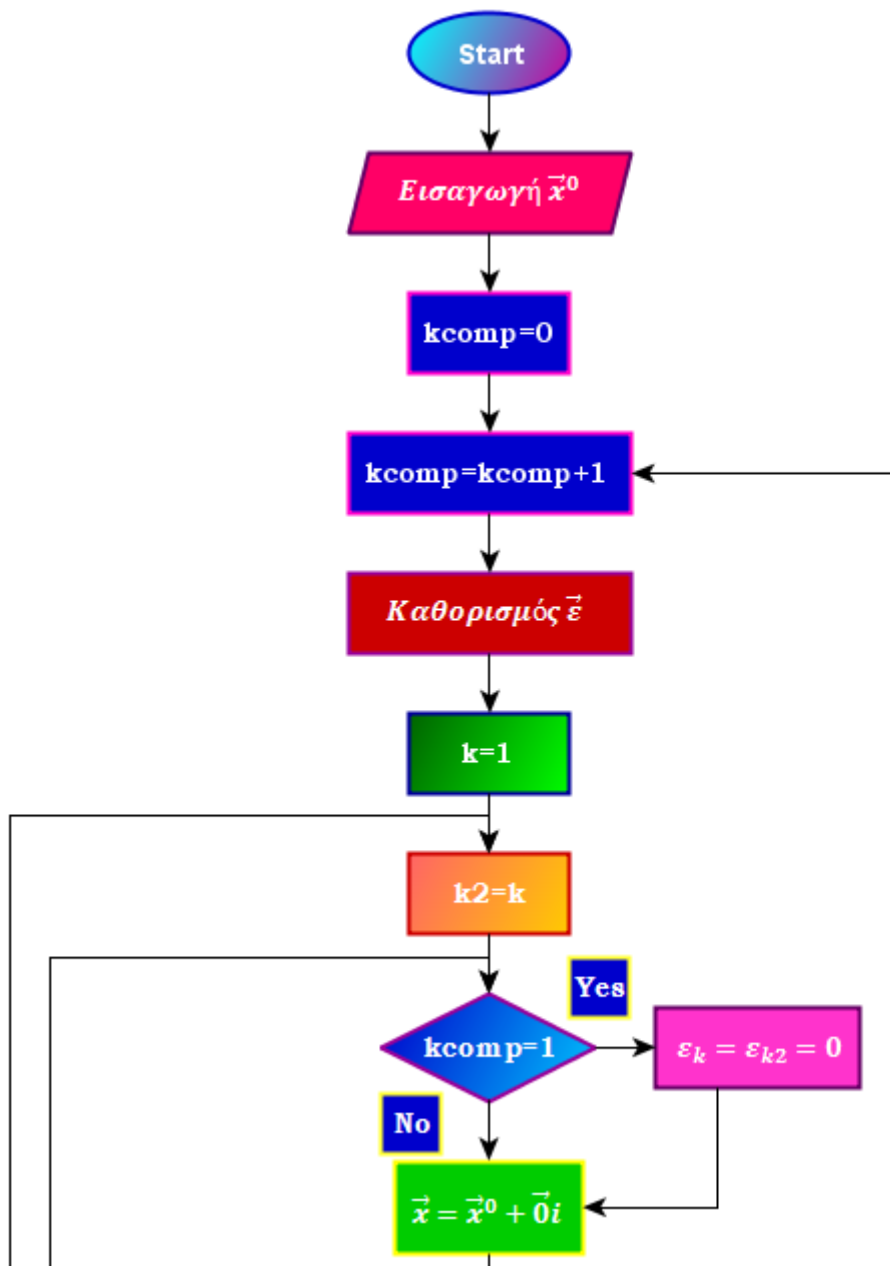
6.4.2.2 Υλοποίηση σε προγραμματιστικό επίπεδο

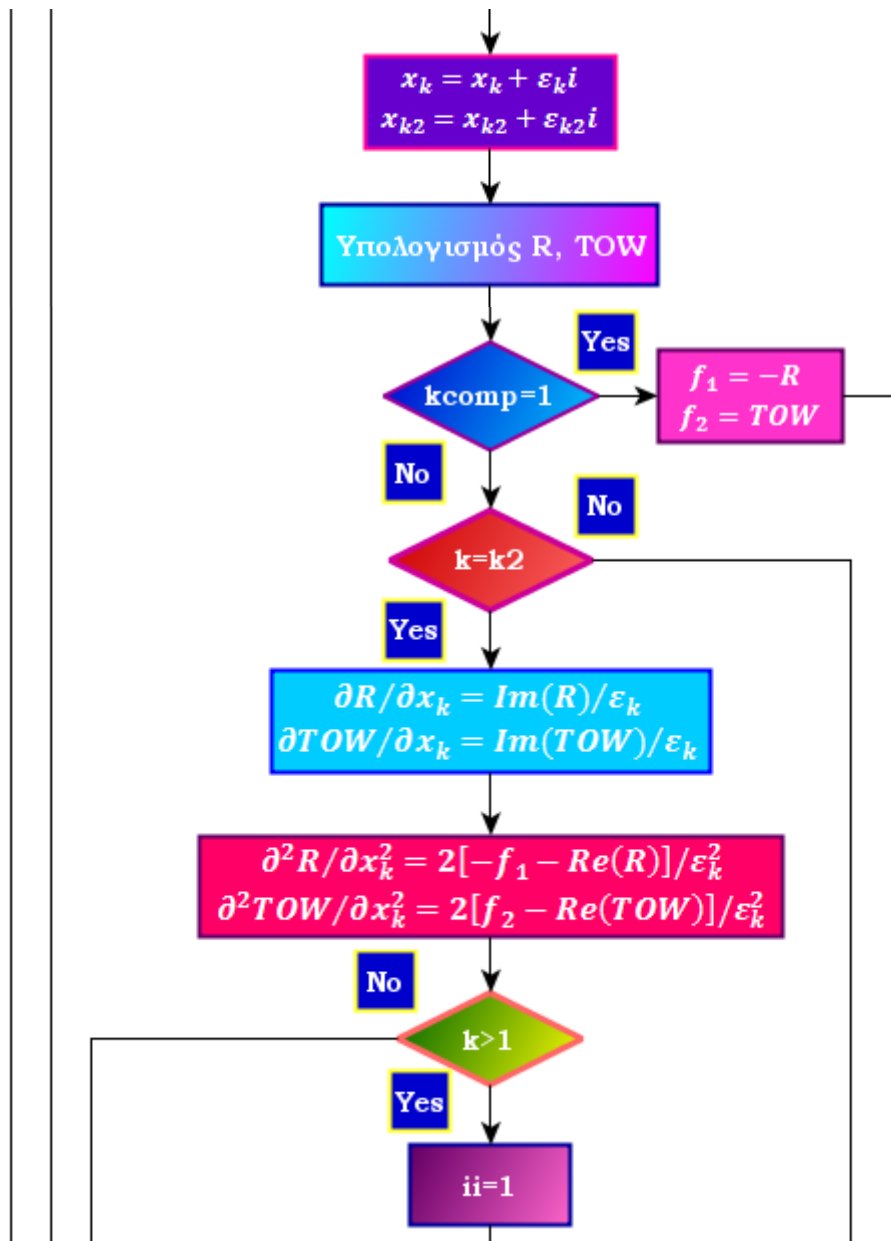
Εξετάζοντας τις σχέσεις (6.3), (6.4) και (6.7) καθίσταται σαφές το γεγονός ότι για να υπολογισθούν οι πρώτες και δεύτερες παράγωγοι της πραγματικής συνάρτησης f μέσω της μεθόδου των μιγαδικών μεταβλητών, πρέπει τόσο οι μεταβλητές σχεδιασμού όσο και όλες οι 'διεργασίες' που εμπεριέχονται στο υπεύθυνο για την εύρεση της εμβέλειας και του συνολικού βάρους του αεροσκάφους κατά την απογείωση πρόγραμμα να γραφούν σε μιγαδική μορφή, δηλαδή να μεταφερθούν από τον άξονα των πραγματικών αριθμών (1 διάσταση), στο επίπεδο των μιγαδικών αριθμών (2 διαστάσεις). Κάτι τέτοιο δεν ενέχει κάποια ιδιαίτερη δυσκολία όταν πρόκειται για απλές σχέσεις αλλά όταν η αναφερθείσα διαδικασία πρέπει να εφαρμοσθεί σε ένα πολύπλοκο μοντέλο όπως αυτό που περιγράφηκε στην ενότητα 6.2, ανακύπτουν δυσκολίες. Όπως αναφέρθηκε παραπάνω, το εν λόγω ζήτημα αντιμετωπίστηκε από τον Ν. Καζαζάκη, [22].

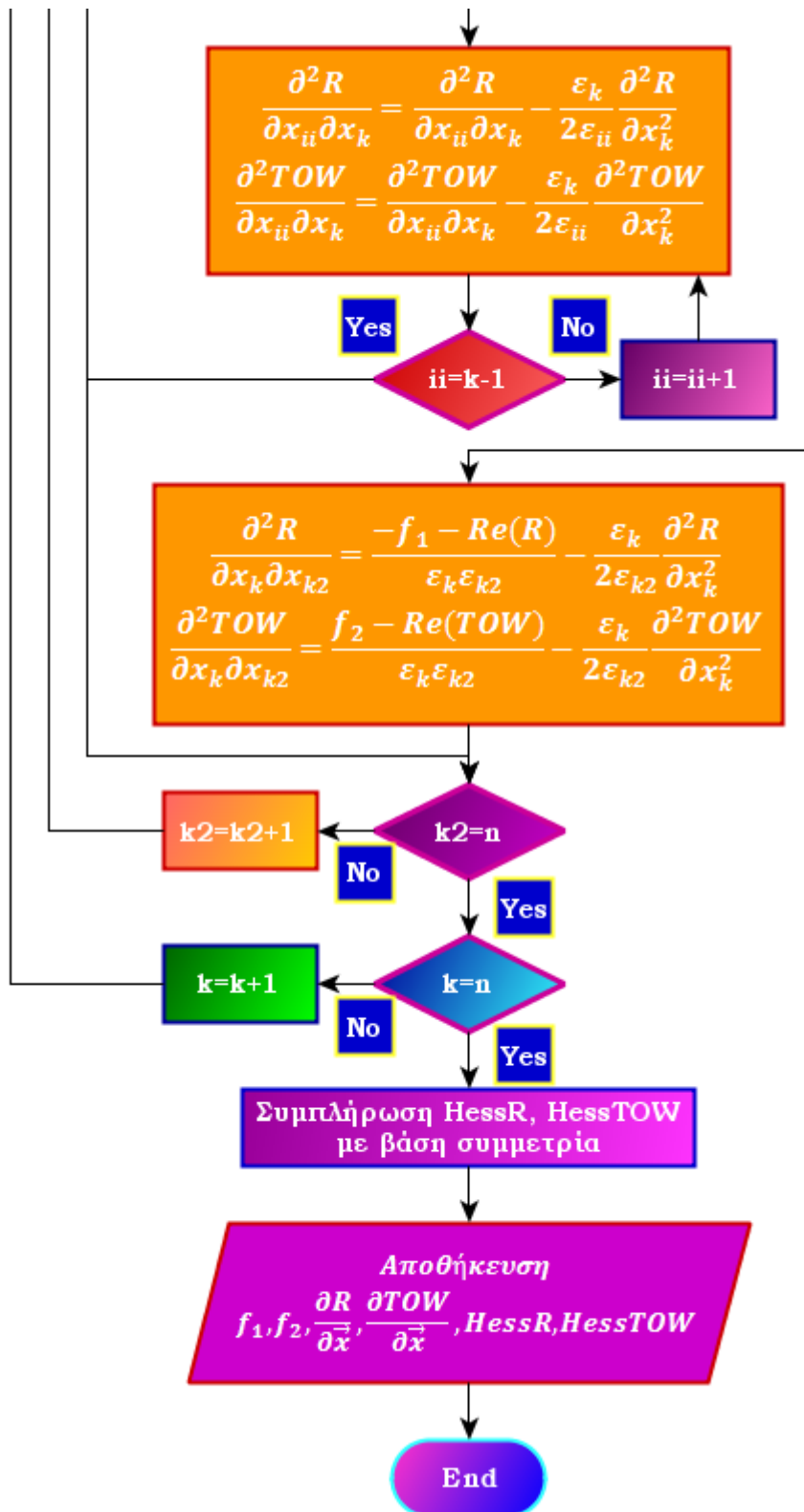
Κρίνεται αναγκαίο, επιπροσθέτως, να αναφερθεί ότι η σχέση (6.3), φανερώνει το γεγονός ότι για την εύρεση των πρώτων παραγώγων μίας συνάρτησης n μεταβλητών για ένα δεδομένο \vec{x} , απαιτείται η 'κλήση' της εν λόγω συνάρτησης n φορές (μία για κάθε μεταβλητή). Μέσω δε της σχέσης (6.4) φαίνεται ότι για την εύρεση των κύριων δευτέρων παραγώγων απαιτούνται τα ίδια ακριβώς δεδομένα με αυτά που είναι απαραίτητα για τον υπολογισμό των πρώτων παραγώγων με την προσθήκη της τιμής της πραγματικής συνάρτησης στο \vec{x} . Τέλος, αν εξετασθεί η σχέση (6.7), διαπιστώνεται ότι για τον υπολογισμό των μικτών δευτέρων παραγώγων, απαιτούνται επιπλέον (της τιμής της πραγματικής συνάρτησης στο \vec{x} και των κύριων δευτέρων παραγώγων)

$n(n - 1)/2$ κλήσεις της συνάρτησης (για το άνω μισό εσσιανό μητρώο ενώ το κάτω μισό συμπληρώνεται με βάση τη συμμετρία). Συνολικά, λοιπόν, είναι απαραίτητες $1 + n + n(n - 1)/2 = 1 + n(n + 1)/2$ κλήσεις της συνάρτησης.

Με βάση τα παραπάνω, αναπτύσσεται ο αλγόριθμος που παρέχει, εκτός των τιμών της εμβέλειας και του συνολικού βάρους κατά την απογείωση, τις πρώτες και δεύτερες παραγώγους των εν λόγω συναρτήσεων ως προς τις μεταβλητές σχεδιασμού για ένα δεδομένο \vec{x} μέσω της μεθόδου των μιγαδικών μεταβλητών, όπως απεικονίζεται στο σχήμα 6.3.







Σχήμα 6.3: Αλγόριθμος που παρέχει τις τιμές $-R$, TOW και τις πρώτες και δεύτερες παραγώγους των R και TOW ως προς τις μεταβλητές σχεδιασμού μέσω της μεθόδου των μιγαδικών μεταβλητών για δεδομένο \vec{x} .

Ο αλγόριθμος του σχήματος 6.3 απαιτεί ως είσοδο ένα αρχείο το οποίο του παρέχει το διάνυσμα \vec{x} για το οποίο ζητούνται οι τιμές $-R$, TOW και οι πρώτες και δεύτερες παράγωγοι των R και TOW .

Ακολουθώντας, εισάγεται ένας μετρητής ($kcomp$) ο οποίος είναι επιφορτισμένος με τον υπολογισμό των τιμών των πραγματικών συναρτήσεων στόχου ($\vec{\varepsilon} = 0$). Στη συνέχεια, καθορίζεται το διάνυσμα $\vec{\varepsilon}$ (περισσότερα σε ακόλουθη ενότητα). Κατόπιν, το πρόγραμμα εισέρχεται σε ένα διπλό επαναληπτικό βρόχο με στόχο την εύρεση των απαιτούμενων πρώτων και δεύτερων παραγώγων. Ο μετρητής k που αντιστοιχεί στον εξωτερικό βρόχο λαμβάνει τιμές από 1 έως n και ουσιαστικά είναι υπεύθυνος για τον υπολογισμό των πρώτων και κύριων δεύτερων παραγώγων, ενώ ο μετρητής $k2$ που αντιστοιχεί στον εσωτερικό βρόχο παίρνει τιμές από k έως n και επιφορτίζεται με την εύρεση των μικτών δεύτερων παραγώγων που αποτελούν το άνω μισό του εσσιανού μητρώου. Το πρώτο βήμα που λαμβάνει χώρα μέσα στον εν λόγω διπλό βρόχο είναι η μετατροπή όλων των μεταβλητών σε μιγαδικές οι οποίες, βέβαια, σε πρώτο στάδιο θα έχουν μηδενικό φανταστικό μέρος. Έπειτα, αποκτούν φανταστικό μέρος (ε_k και ε_{k2}) οι μεταβλητές x_k και x_{k2} (προφανώς αν $k = k2$ μόνο μία μεταβλητή αποκτά φανταστικό μέρος). Ακολουθεί ο υπολογισμός των R και TOW για τις αναδιαμορφωμένες μεταβλητές σχεδιασμού (σύμφωνα με τον αλγόριθμο που περιγράφηκε στην ενότητα 6.2 ο οποίος βέβαια μετατράπηκε σε μιγαδική μορφή). Στη συνέχεια, αν ισχύει $k = k2$, πραγματοποιείται ο υπολογισμός των αντίστοιχων πρώτων και δεύτερων κύριων παραγώγων σύμφωνα με τις σχέσεις (6.3) και (6.4), ενώ σε αντίθετη περίπτωση υπολογίζονται οι αντίστοιχες μικτές δεύτερες παράγωγοι σύμφωνα με τη σχέση (6.7). Πρέπει να τονιστεί ότι όταν το πρόγραμμα φτάσει λ.χ. στην κατάσταση όπου $k = 1$ και $k2 = 2$, διαθέτει όλα τα δεδομένα για την εύρεση των $\partial^2 R / \partial x_1 \partial x_2$ και $\partial^2 TOW / \partial x_1 \partial x_2$ εκτός από τις κύριες δεύτερες παραγώγους $\partial^2 R / \partial x_2^2$ και $\partial^2 TOW / \partial x_2^2$ οι οποίες προφανώς υπολογίζονται όταν $k = 2$ και $k2 = 2$. Για το λόγο αυτό, όταν ισχύει $k = k2$ και $k > 1$ προστίθεται ένας επαναληπτικός βρόχος (μετρητής i που λαμβάνει τιμές από 1 έως $k - 1$) για να συμπληρώνει στην κάθε μικτή δεύτερη παράγωγο το τμήμα που της λείπει μόλις αυτό υπολογιστεί. Τέλος, τα εσσιανά μητρώα συμπληρώνονται με βάση τη συμμετρία και το πρόγραμμα αποθηκεύει όλα τα ζητηθέντα αποτελέσματα σε αρχεία εξόδου.

6.4.3 Η μέθοδος των πεπερασμένων διαφορών (finite difference method)

6.4.3.1 Θεωρητικά στοιχεία

Με κατάλληλη αλγεβρική επεξεργασία των αναπτυγμάτων Taylor (6.1) και (6.5) που παρατέθηκαν στην ενότητα 6.4.2.1 προκύπτουν διαφορετικές εκφράσεις πεπερασμένων διαφορών για τις πρώτες και δεύτερες παραγώγους μίας συνάρτησης f . Για τους σκοπούς της παρούσας εργασίας έγινε χρήση των κεντρικών πεπερασμένων διαφορών ώστε να επιτυγχάνεται ακρίβεια δεύτερης τάξης όπως και

στην περίπτωση της μεθόδου των μιγαδικών μεταβλητών που αναλύθηκε παραπάνω. Οι εκφράσεις που χρησιμοποιήθηκαν είναι οι εξής:

$$\frac{\partial f(\vec{x})}{\partial x_k} = \frac{f(x_1, \dots, x_k + \varepsilon_k, \dots, x_n) - f(x_1, \dots, x_k - \varepsilon_k, \dots, x_n)}{2\varepsilon_k}, \quad (6.8)$$

$$k = 1, \dots, n$$

$$\frac{\partial^2 f(\vec{x})}{\partial x_k^2} = \frac{f(x_1, \dots, x_k + \varepsilon_k, \dots, x_n) - 2f(\vec{x}) + f(x_1, \dots, x_k - \varepsilon_k, \dots, x_n)}{\varepsilon_k^2}, \quad (6.9)$$

$$k = 1, \dots, n$$

$$\frac{\partial^2 f(\vec{x})}{\partial x_k \partial x_l} = \frac{1}{4\varepsilon_k \varepsilon_l} \cdot \{f(x_1, \dots, x_k + \varepsilon_k, \dots, x_l + \varepsilon_l, \dots, x_n) - f(x_1, \dots, x_k - \varepsilon_k, \dots, x_l + \varepsilon_l, \dots, x_n) - f(x_1, \dots, x_k + \varepsilon_k, \dots, x_l - \varepsilon_l, \dots, x_n) + f(x_1, \dots, x_k - \varepsilon_k, \dots, x_l - \varepsilon_l, \dots, x_n)\}$$

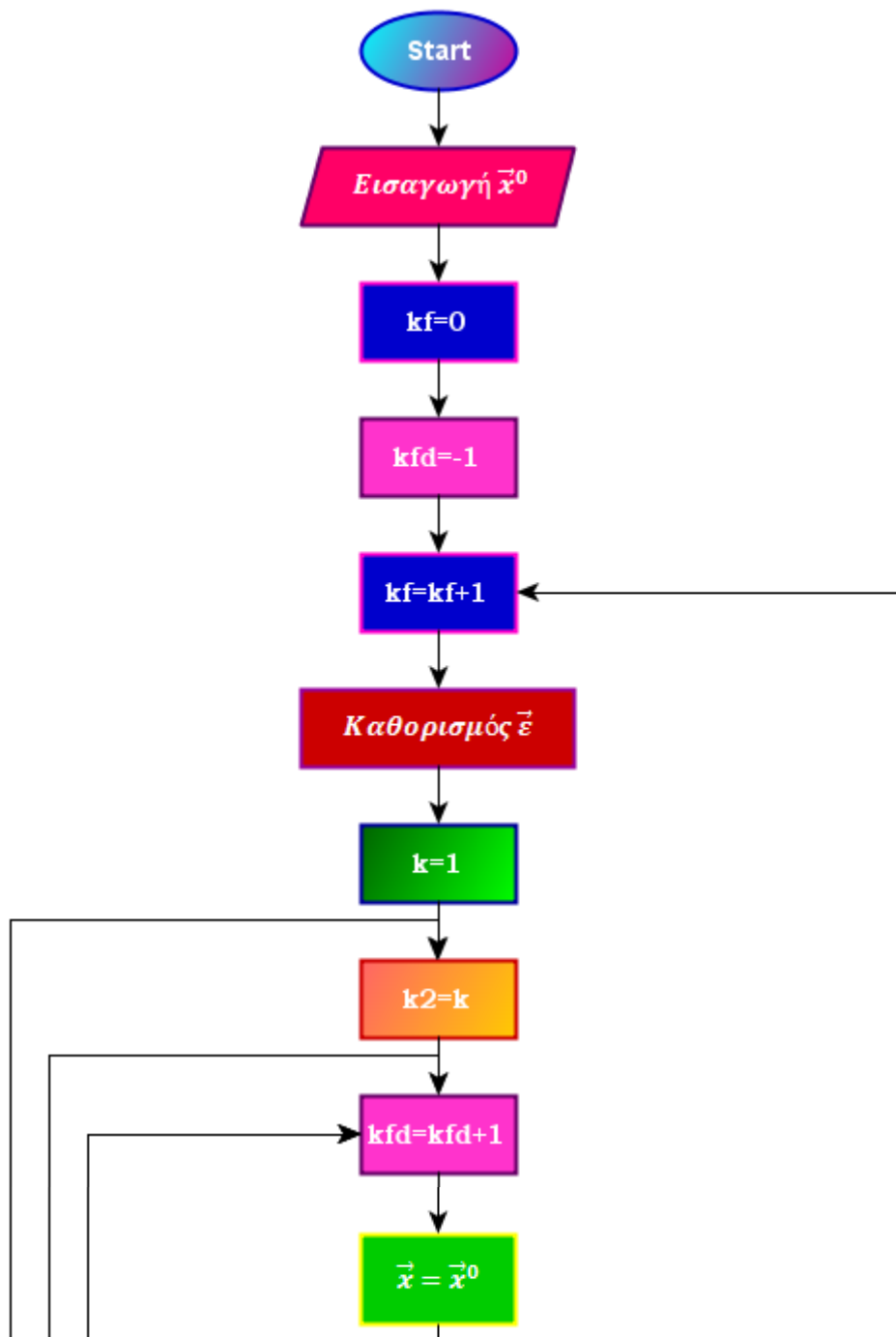
$$k = 1, \dots, n-1, l = k+1, \dots, n \quad (6.10)$$

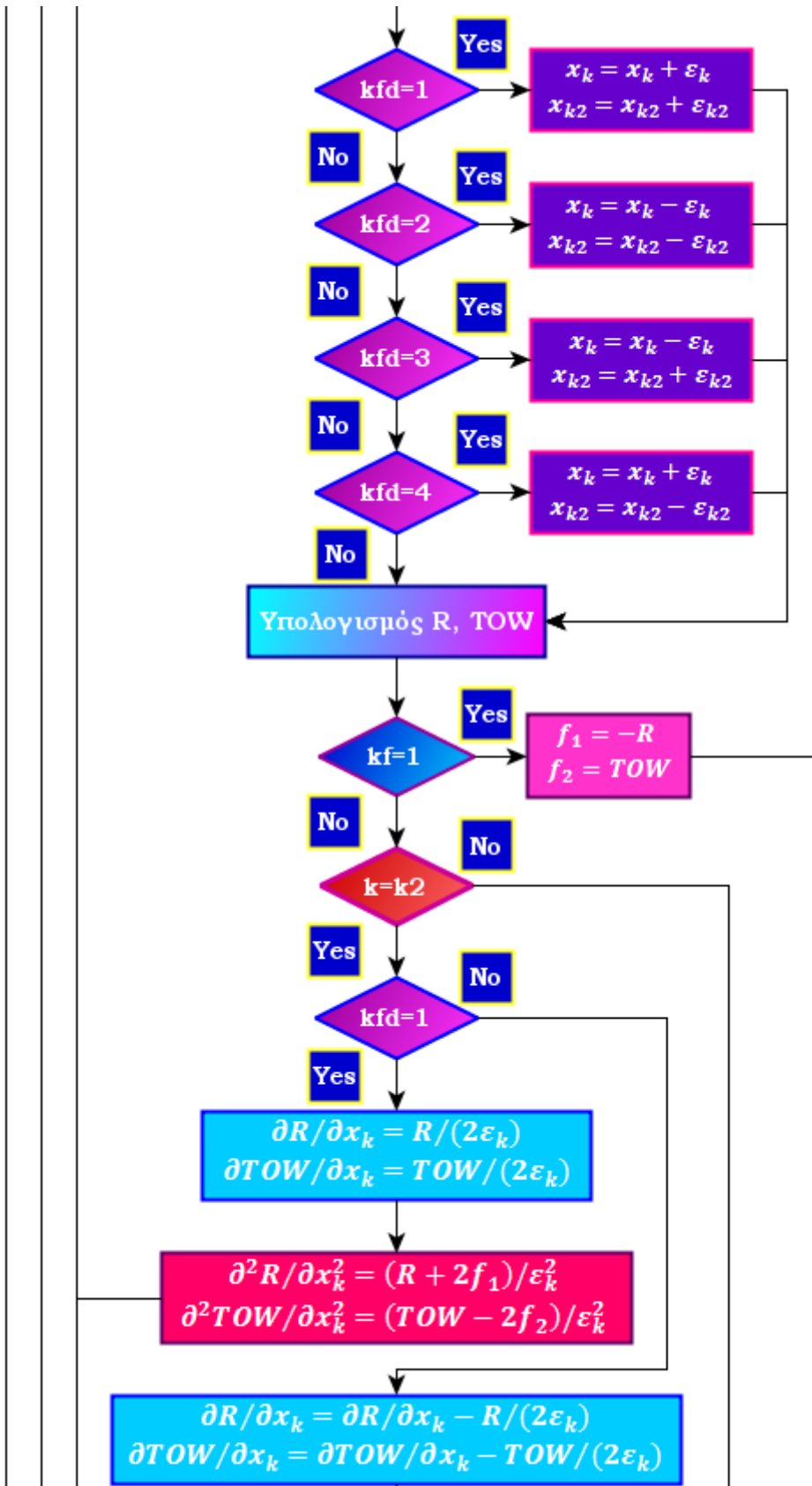
$$\frac{\partial^2 f}{\partial x_l \partial x_k}(\vec{x}) = \frac{\partial^2 f}{\partial x_k \partial x_l}(\vec{x}), \quad k = 1, \dots, n-1, l = k+1, \dots, n$$

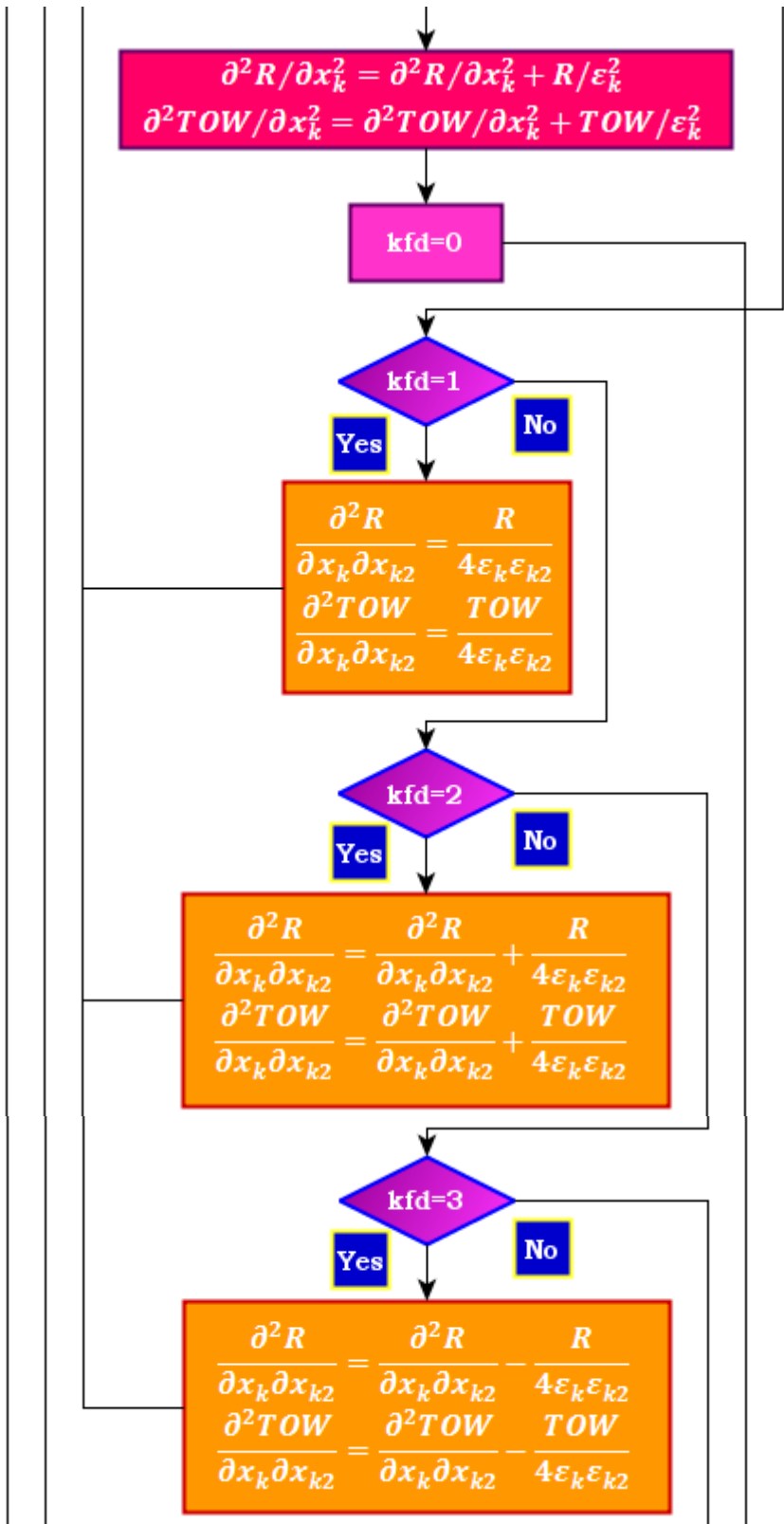
6.4.3.2 Υλοποίηση σε προγραμματιστικό επίπεδο

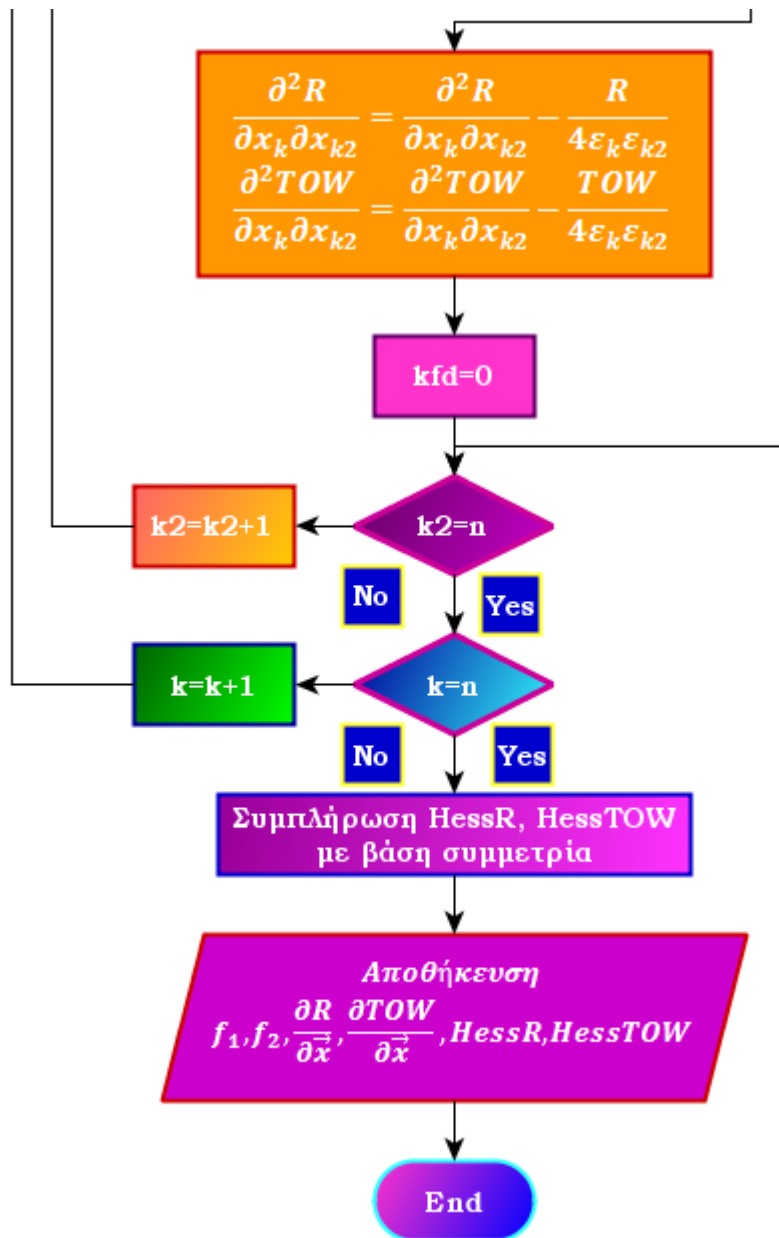
Εξετάζοντας τη σχέση (6.8), καθίσταται σαφές το γεγονός ότι για την εύρεση των πρώτων παραγώγων μίας συνάρτησης n μεταβλητών για ένα δεδομένο \vec{x} μέσω κεντρικών πεπερασμένων διαφορών, απαιτείται η 'κλήση' της εν λόγω συνάρτησης $2n$ φορές (δύο για κάθε μεταβλητή), σε αντίθεση με τη μέθοδο των μιγαδικών μεταβλητών όπου απαιτούνταν n κλήσεις. Μέσω της σχέσης (6.9) φαίνεται ότι για την εύρεση των κύριων δευτέρων παραγώγων απαιτούνται τα ίδια ακριβώς δεδομένα με αυτά που είναι απαραίτητα για τον υπολογισμό των πρώτων παραγώγων με την προσθήκη της τιμής της πραγματικής συνάρτησης στο \vec{x} . Τέλος, αν εξετασθεί η σχέση (6.10), διαπιστώνεται ότι για τον υπολογισμό των μικτών δευτέρων παραγώγων, απαιτούνται $4n(n-1)/2$ κλήσεις της συνάρτησης (για το άνω μισό εσσιανό μητρώο ενώ το κάτω μισό συμπληρώνεται με βάση τη συμμετρία), δηλαδή οι τετραπλάσιες από αυτές που ήταν αναγκαίες μέσω της μεθόδου των μιγαδικών μεταβλητών. Συνολικά, λοιπόν, είναι απαραίτητες $1 + 2n + 4n(n-1)/2 = 1 + 2n^2$ κλήσεις της συνάρτησης.

Με βάση τα παραπάνω, αναπτύσσεται ο αλγόριθμος που παρέχει, εκτός των τιμών της εμβέλειας και του συνολικού βάρους κατά την απογείωση, τις πρώτες και δεύτερες παραγώγους των εν λόγω συναρτήσεων ως προς τις μεταβλητές σχεδιασμού για ένα δεδομένο \vec{x} μέσω κεντρικών πεπερασμένων διαφορών, όπως απεικονίζεται στο σχήμα 6.4.









Σχήμα 6.4: Αλγόριθμος που παρέχει τις τιμές $-R$, TOW και τις πρώτες και δεύτερες παραγώγους των R και TOW ως προς τις μεταβλητές σχεδιασμού μέσω κεντρικών πεπερασμένων διαφορών για δεδομένο \vec{x} .

Η μορφή του αλγορίθμου του σχήματος 6.4 είναι όμοια με αυτή του αλγορίθμου της μεθόδου των μιγαδικών μεταβλητών που περιγράφηκε στην ενότητα 6.4.2.2 (σχήμα 6.3). Η διαφορά έγκειται στις διεργασίες που πραγματοποιούνται στο εσωτερικό του διπλού επαναληπτικού βρόχου, σε κάθε επανάληψη του οποίου υπολογίζεται είτε μία πρώτη παράγωγος συνοδευόμενη από την αντίστοιχη δεύτερη κύρια παράγωγο ($k = k2$) είτε μία δεύτερη μικτή παράγωγος ($k \neq k2$) για κάθε συνάρτηση στόχου. Στη συγκεκριμένη περίπτωση, για να δύναται το πρόγραμμα να 'καλέσει' τη διαδικασία υπολογισμού των R και TOW (όπως αυτή παρουσιάσθηκε στο σχήμα 6.2) όσες φορές απαιτείται σε κάθε επανάληψη του διπλού βρόχου, εισάγεται ένας

επιπλέον μετρητής (kfd). Όταν ισχύει $k = k2$, εφαρμόζονται οι σχέσεις (6.8) και (6.9) για τον υπολογισμό των αντίστοιχων πρώτων και δεύτερων κύριων παραγώγων, οι οποίες υπαγορεύουν δύο 'κλήσεις' της διαδικασίας υπολογισμού των R και TOW , μία για $x_k + \varepsilon_k$ και μία για $x_k - \varepsilon_k$, οπότε ο μετρητής kfd παίρνει τις τιμές 1 και 2 αντίστοιχα. Στην περίπτωση δε όπου $k \neq k2$ υπολογίζονται οι αντίστοιχες μικτές δεύτερες παράγωγοι σύμφωνα με τη σχέση (6.10), η οποία απαιτεί την πραγματοποίηση τεσσάρων κλήσεων των συναρτήσεων στόχου και ο μετρητής kfd υποχρεούται να λάβει 4 τιμές (1 για $x_k + \varepsilon_k, x_{k2} + \varepsilon_{k2}$, 2 για $x_k - \varepsilon_k, x_{k2} - \varepsilon_{k2}$, 3 για $x_k - \varepsilon_k, x_{k2} + \varepsilon_{k2}$ και 4 για $x_k + \varepsilon_k, x_{k2} - \varepsilon_{k2}$).

6.4.4 Σύγκριση αποτελεσμάτων μεθόδου μιγαδικών μεταβλητών και κεντρικών πεπερασμένων διαφορών – Πιστοποίηση ορθότητας αποτελεσμάτων παραγώγων

Οι αλγόριθμοι που περιγράφηκαν στις ενότητες 6.4.2 και 6.4.3 δοκιμάστηκαν για διάφορα διανύσματα μεταβλητών σχεδιασμού \vec{x} και για ποικίλα διανύσματα $\vec{\varepsilon}$ (τα προγράμματα δοκιμάστηκαν επιπλέον σε απλά μαθηματικά προβλήματα αλλά δεν κρίνεται σκόπιμη η παρουσίαση των εν λόγω δοκιμών). Αξίζει να αναφερθεί ότι και για τους δύο αλγόριθμους το $\vec{\varepsilon}$ επιλέχθηκε να καθορίζεται ως εξής:

$$\varepsilon_k = \varepsilon_0 x_k, \quad \varepsilon_0 = \text{σταθ.}, \quad k = 1, \dots, n$$

Η σταθερά ε_0 είναι ένας κατάλληλα μικρός αριθμός και στις πραγματοποιηθείσες δοκιμές αποφασίστηκε να κυμαίνεται μεταξύ των τιμών 10^{-1} και 10^{-7} . Ο λόγος για τον οποίο υιοθετήθηκε η αναφερθείσα προσέγγιση έγκειται στο γεγονός ότι υπάρχει μεγάλη διαφορά στην τάξη μεγέθους των μεταβλητών. Για παράδειγμα, η εκλέπτυνση της πτέρυγας λ_w λαμβάνει τιμές στο διάστημα $[0,05 \ 0,5]$, ενώ το βάρος του καυσίμου W_{fuel} κυμαίνεται μεταξύ των 15000 kg και 30000 kg . Μέσω των δοκιμών που έλαβαν χώρα διαπιστώθηκε ότι πολύ ικανοποιητικά αποτελέσματα δίνουν οι περιπτώσεις όπου $\varepsilon_0 = 10^{-3}$, $\varepsilon_0 = 10^{-4}$ και $\varepsilon_0 = 10^{-5}$.

Παρακάτω, επιδεικνύονται ενδεικτικά τα αποτελέσματα (διπλής ακριβείας) των δύο μεθόδων για $\varepsilon_0 = 10^{-3}$, $\varepsilon_0 = 10^{-4}$ και $\varepsilon_0 = 10^{-5}$, όταν $z_{cr} = 16500 \text{ m}$, $\lambda_w = 0,3$, $\lambda_v = 0,1$ και $W_{fuel} = 26000 \text{ kg}$, ενώ οι λοιπές μεταβλητές σχεδιασμού παίρνουν τις σταθερές τιμές που παρουσιάστηκαν στον πίνακα 6.1 (σημειώνεται ότι τα προγράμματα υπολογίζουν τις παραγώγους των R και TOW και ως προς τις 12 μεταβλητές που φαίνονται στον πίνακα 6.1 αλλά εδώ δεν απασχολούν οι μεταβλητές των οποίων η τιμή διατηρείται σταθερή).

$\left(\frac{\partial R}{\partial \vec{x}}\right)_{Complex}$	$\left(\frac{\partial TOW}{\partial \vec{x}}\right)_{Complex}$
485,91854693333295	0,48558081009548643
-3421736,1491956585	-2415,2074930326421
13144,422896622138	-667,53498275661514
373,62231713283990	1,2784693058465362

Πίνακας 6.2: Πρώτες παράγωγοι των R και TOW ως προς τις μεταβλητές σχεδιασμού z_{cr} , λ_w , λ_v και W_{fuel} μέσω της μεθόδου των μιγαδικών μεταβλητών για $\varepsilon_0 = 10^{-3}$.

$\left(\frac{\partial R}{\partial \vec{x}}\right)_{ED}$	$\left(\frac{\partial TOW}{\partial \vec{x}}\right)_{ED}$
485,91021407645894	0,48558660980779678
-3421749,2480754852	-2415,2017632871866
13144,420066833496	-667,53500282764435
373,62208015180659	1,2784693082899139

Πίνακας 6.3: Πρώτες παράγωγοι των R και TOW ως προς τις μεταβλητές σχεδιασμού z_{cr} , λ_w , λ_v και W_{fuel} μέσω της μεθόδου των κεντρικών πεπερασμένων διαφορών για $\varepsilon_0 = 10^{-3}$.

-0,18728156627264453	-780,83018650009171	28,287915117812879	-2,0592225828405582E-002
-780,83018650009171	25439883,809950620	64612,304170926451	-217,66599363241437
28,287915117812879	64612,304170926451	-2255883,8129043579	11,704380695636701
-2,0592225828405582E-002	-217,66599363241437	11,704380695636701	-9,5532966898919571E-003

Πίνακας 6.4: $HessR$ μέσω της μεθόδου των μιγαδικών μεταβλητών για $\varepsilon_0 = 10^{-3}$.

-0,18728236671449849	-780,83492273092270	28,287949562072754	-2,0592666508491675E-002
-780,83492273092270	25439929,046875000	64612,218750000000	-217,66667747497559
28,287949562072754	64612,218750000000	-2255883,2500000000	11,704372167587280
-2,0592666508491675E-002	-217,66667747497559	11,704372167587280	-9,5532867890142370E-003

Πίνακας 6.5: $HessR$ μέσω της μεθόδου των κεντρικών πεπερασμένων διαφορών για $\varepsilon_0 = 10^{-3}$.

1,5217355283360357E-004	0,41324131848111828	-1,7695490510739161E-002	2,3390478869717749E-005
0,41324131848111828	-5427,9581591900860	38,396683521568775	-5,4255333685805707E-002
-1,7695490510739161E-002	38,396683521568775	3080,6826543994248	-2,8038484742864966E-002
2,3390478869717749E-005	-5,4255333685805707E-002	-2,8038484742864966E-002	8,6749870749948185E-006

Πίνακας 6.6: $HessTOW$ μέσω της μεθόδου των μιγαδικών μεταβλητών για $\varepsilon_0 = 10^{-3}$.

1,5217499878872331E-004	0,41325172921642661	-1,7695766873657703E-002	2,3390911859877406E-005
0,41325172921642661	-5427,9716186523438	38,396850585937500	-5,4255365626886487E-002
-1,7695766873657703E-002	38,396850585937500	3080,6816406250000	-2,8038483578711748E-002
2,3390911859877406E-005	-5,4255365626886487E-002	-2,8038483578711748E-002	8,6749891323734118E-006

Πίνακας 6.7: $HessTOW$ μέσω της μεθόδου των κεντρικών πεπερασμένων διαφορών για $\varepsilon_0 = 10^{-3}$.

$\left(\frac{\partial R}{\partial \vec{x}}\right)_{Complex}$	$\left(\frac{\partial TOW}{\partial \vec{x}}\right)_{Complex}$
485,91442216628752	0,48558368094067356
-3421742,6331562144	-2415,2046568009164
13144,421446640696	-667,53499268443500
373,62219982707609	1,2784693070560442

Πίνακας 6.8: Πρώτες παράγωγοι των R και TOW ως προς τις μεταβλητές σχεδιασμού z_{cr} , λ_w , λ_v και W_{fuel} μέσω της μεθόδου των μιγαδικών μεταβλητών για $\varepsilon_0 = 10^{-4}$.

$\left(\frac{\partial R}{\partial \vec{x}}\right)_{ED}$	$\left(\frac{\partial TOW}{\partial \vec{x}}\right)_{ED}$
485,91433883784339	0,48558373893865792
-3421742,7640991211	-2415,2045994997025
13144,421386718750	-667,53499317169189
373,62219745712355	1,2784693070807407

Πίνακας 6.9: Πρώτες παράγωγοι των R και TOW ως προς τις μεταβλητές σχεδιασμού z_{cr} , λ_w , λ_v και W_{fuel} μέσω της μεθόδου των κεντρικών πεπερασμένων διαφορών για $\varepsilon_0 = 10^{-4}$.

-0,18728196183074161	-780,83386625906428	28,287709662408538	-2,0592473013139823E-002
-780,83386625906428	25439903,967910342	64615,160226821899	-217,66639577272616
28,287709662408538	64615,160226821899	-2255849,5402336116	11,704503916777185
-2,0592473013139823E-002	-217,66639577272616	11,704503916777185	-9,5532931519683294E-003

Πίνακας 6.10: $HessR$ μέσω της μεθόδου των μιγαδικών μεταβλητών για $\varepsilon_0 = 10^{-4}$.

-0,18728196807205677	-780,83393859863281	28,287857055664062	-2,0592477929312736E-002
-780,83393859863281	25439908,000000000	64613,000000000000	-217,66649246215820
28,287857055664062	64613,000000000000	-2255856,0000000000	11,704360961914062
-2,0592477929312736E-002	-217,66649246215820	11,704360961914062	-9,5532918348908424E-003

Πίνακας 6.11: $HessR$ μέσω της μεθόδου των κεντρικών πεπερασμένων διαφορών για $\varepsilon_0 = 10^{-4}$.

1,5217427094189038E-004	0,41324690553463161	-1,7695569884822224E-002	2,3390714617528555E-005
0,41324690553463161	-5427,9452241543268	38,368549818794008	-5,4255509987855555E-002
-1,7695569884822224E-002	38,368549818794008	3080,7859729975457	-2,8039022044350270E-002
2,3390714617528555E-005	-5,4255509987855555E-002	-2,8039022044350270E-002	8,6749907990796322E-006

Πίνακας 6.12: $HessTOW$ μέσω της μεθόδου των μιγαδικών μεταβλητών για $\varepsilon_0 = 10^{-4}$.

1,5217428517644294E-004	0,41324722766876221	-1,7695546150207520E-002	2,3390719434246421E-005
0,41324722766876221	-5427,9687500000000	38,406250000000000	-5,4255306720733643E-002
-1,7695546150207520E-002	38,406250000000000	3080,6875000000000	-2,8038322925567627E-002
2,3390719434246421E-005	-5,4255306720733643E-002	-2,8038322925567627E-002	8,6749887486803345E-006

Πίνακας 6.13: $HessTOW$ μέσω της μεθόδου των κεντρικών πεπερασμένων διαφορών για $\varepsilon_0 = 10^{-4}$.

$\left(\frac{\partial R}{\partial \vec{x}}\right)_{Complex}$	$\left(\frac{\partial TOW}{\partial \vec{x}}\right)_{Complex}$
485,91438091571825	0,48558370964975528
-3421742,6981877889	-2415,2046282795100
13144,420949619509	-667,53499269993631
373,62219865216275	1,2784693070684612

Πίνακας 6.14: Πρώτες παράγωγοι των R και TOW ως προς τις μεταβλητές σχεδιασμού z_{cr} , λ_w , λ_v και W_{fuel} μέσω της μεθόδου των μιγαδικών μεταβλητών για $\varepsilon_0 = 10^{-5}$.

$\left(\frac{\partial R}{\partial \vec{x}}\right)_{ED}$	$\left(\frac{\partial TOW}{\partial \vec{x}}\right)_{ED}$
485,91438009217381	0,48558371022227220
-3421742,6989746094	-2415,2046298980713
13144,419433593750	-667,53499603271484
373,62219863012433	1,2784693070570938

Πίνακας 6.15: Πρώτες παράγωγοι των R και TOW ως προς τις μεταβλητές σχεδιασμού z_{cr} , λ_w , λ_v και W_{fuel} μέσω της μεθόδου των κεντρικών πεπερασμένων διαφορών για $\varepsilon_0 = 10^{-5}$.

-0,18728173468723769	-780,84718699407085	28,244473717429415	-2,0592736128049020E-002
-780,84718699407085	25439593,527052134	62709,053357442201	-217,67635376025470
28,244473717429415	62709,053357442201	-2253800,6305694571	11,684516301521853
-2,0592736128049020E-002	-217,67635376025470	11,684516301521853	-9,5531653017687356E-003

Πίνακας 6.16: $HessR$ μέσω της μεθόδου των μιγαδικών μεταβλητών για $\varepsilon_0 = 10^{-5}$.

-0,18728184700012207	-780,83105468750000	28,279296875000000	-2,0592488348484039E-002
-780,83105468750000	25440128,000000000	65408,000000000000	-217,66552734375000
28,279296875000000	65408,000000000000	-2251776,0000000000	11,708007812500000
-2,0592488348484039E-002	-217,66552734375000	11,708007812500000	-9,5532536506652832E-003

Πίνακας 6.17: $HessR$ μέσω της μεθόδου των κεντρικών πεπερασμένων διαφορών για $\varepsilon_0 = 10^{-5}$.

1,5217481613761244E-004	0,41323029577280534	-1,7726878550919617E-002	2,3390422900813026E-005
0,41323029577280534	-5426,2475007110161	38,805107275643877	-5,4261853130391002E-002
-1,7726878550919617E-002	38,805107275643877	3085,0060284137712	-2,8040421266968485E-002
2,3390422900813026E-005	-5,4261853130391002E-002	-2,8040421266968485E-002	8,6751802322956228E-006

Πίνακας 6.18: $HessTOW$ μέσω της μεθόδου των μιγαδικών μεταβλητών για $\varepsilon_0 = 10^{-5}$.

1.5217415057122707E-004	0.41324234008789062	-1.7700195312500000E-002	2.3390690330415964E-005
0.41324234008789062	-5428.0000000000000	39.000000000000000	-5.4250717163085938E-002
-1.7700195312500000E-002	39.000000000000000	3080.0000000000000	-2.8038024902343750E-002
2.3390690330415964E-005	-5.4250717163085938E-002	-2.8038024902343750E-002	8.6750369518995285E-006

Πίνακας 6.19: $HessTOW$ μέσω της μεθόδου των κεντρικών πεπερασμένων διαφορών για $\varepsilon_0 = 10^{-5}$.

Από την παρατήρηση των πινάκων 6.2 έως 6.19, προκύπτει το συμπέρασμα ότι οι παράγωγοι που υπολογίζονται μέσω της μεθόδου των μιγαδικών μεταβλητών είναι εξαιρετικά όμοιες με αυτές που υπολογίζονται μέσω των κεντρικών πεπερασμένων διαφορών και για τις τρεις παρουσιασθείσες περιπτώσεις διαφορετικού ε_0 , δηλαδή έχει επιτευχθεί η πιστοποίηση των αποτελεσμάτων. Σημειώνεται ότι για τη διαδικασία της βελτιστοποίησης επιλέχθηκε $\varepsilon_0 = 10^{-5}$. Τέλος, μέσω εξέτασης των εν λόγω πινάκων διαπιστώνεται το γνωστό μειονέκτημα της μεθόδου των πεπερασμένων διαφορών που είναι η μεγαλύτερη ευαισθησία στις αλλαγές του $\vec{\varepsilon}$.

6.5 Παρουσίαση αποτελεσμάτων

Έχοντας καλή γνώση του προβλήματος προς επίλυση, λαμβάνεται η απόφαση να τροφοδοτηθεί στο πρόγραμμα ως αρχική εκτίμηση για το \hat{f}_2 η τιμή 48500 kg. Μία αρχική επιλογή των παραμέτρων που σχετίζονται με την ALM δύναται να είναι η εξής: $method = 2$ (μέθοδος Newton), $\vec{x}^0 = [16500 \ 0,3 \ 0,1 \ 26000]^T$, $\hat{\lambda}_2^0 = 0$, $\omega_p^0 = 0,4$, $\gamma = 1,5$, $\omega_p^{max} = 500$. Επιπλέον, έστω ότι οι ανάγκες του χρήστη για το μέτωπο Pareto επιτάσσουν $n_{elit} = 20$ και $\hat{f}_2^{down} = 30000$ kg. Ακολουθεί η παράθεση των αποτελεσμάτων του προγράμματος όσον αφορά το πρώτο σημείο του μετώπου Pareto που λαμβάνεται μέσω της ALM (με δεδομένες τις προαναφερθείσες τιμές των παραμέτρων) για διαφορετικές τιμές των ανοχών με τις οποίες ικανοποιούνται τα κριτήρια σύγκλισης της μεθόδου Newton και της ALM. Τα εν λόγω αποτελέσματα φαίνονται στον πίνακα 6.20, συνοδευόμενα από τις συνολικές αξιολογήσεις που πραγματοποιούνται από το πρόγραμμα (δηλαδή συνυπολογίζονται και οι αξιολογήσεις που γίνονται κατά τη διάρκεια της εύρεσης των υπολοίπων σημείων του Pareto) σε κάθε περίπτωση.

Ανοχή ALM	Ανοχή Newton	$f_1(m)$	$f_2(kg)$	Αξιολογήσεις
1,00E-06	1,00E-06	-10829934,779370	48500,000000	427
1,00E-06	1,00E-04	-10829934,779370	48500,000000	391
1,00E-06	1,00E-03	-10829934,779370	48500,000000	229
1,00E-06	1,00E-01	-10829934,779370	48500,000000	223
1,00E-06	1,00E-00	-10829934,779370	48500,000000	211
1,00E-06	1,00E+01	-10829934,779370	48500,000000	205
1,00E-06	1,00E+03	-10829934,779370	48500,000000	193
1,00E-06	1,00E+04	-10829934,779370	48500,000000	187
1,00E-06	1,00E+05	-10829934,779367	48500,000000	175
1,00E-06	1,00E+06	-10829934,779364	48500,000000	169
1,00E-05	1,00E+06	-10829934,779364	48500,000000	169
1,00E-04	1,00E+05	-10829934,779808	48500,000002	169
1,00E-04	1,00E+06	-10829934,776723	48499,999989	163
1,00E-03	1,00E+06	-10829934,776723	48499,999989	163
1,00E-02	1,00E+06	-10829933,246055	48499,993459	157
1,00E-01	1,00E+06	-10829933,246055	48499,993459	157
1,00E-00	1,00E+06	-10829897,216647	48499,839771	151
1,00E+01	1,00E+06	-10829897,216647	48499,839771	151
1,00E+02	1,00E+06	-10819444,858800	48455,273206	145

Πίνακας 6.20: Πρώτη μη-κυριαρχούμενη λύση και συνολικές αξιολογήσεις που πραγματοποιούνται για την εύρεση του μετώπου Pareto για διαφορετικές τιμές των ανοχών με τις οποίες ικανοποιούνται τα κριτήρια σύγκλισης της μεθόδου Newton και της ALM όταν $\hat{f}_2^{up} = 48500 \text{ kg}$, $\bar{x}^0 = [16500 \ 0,3 \ 0,1 \ 26000]^T$, $\hat{\lambda}_2^0 = 0$, $\omega_p^0 = 0,4$, $\gamma = 1,5$, $\omega_p^{max} = 500$ και $n_{elit} = 20$.

Λαμβάνοντας υπόψη τα δεδομένα του πίνακα 6.20 και τις ανάγκες που επιβάλλει το παρόν πρόβλημα (σε ένα τέτοιο πρόβλημα απασχολεί η εμβέλεια σε km και το βάρος του καυσίμου σε tn , κάτι που σημαίνει ότι διαφορές λίγων μέτρων και της τάξης του ενός κιλού θεωρούνται αμελητέες), επιλέγονται ως ικανοποιητικές ανοχές για τα κριτήρια σύγκλισης της μεθόδου Newton και της ALM οι τιμές 10^6 και 1 αντίστοιχα. Για τις επιλεχθείσες ανοχές αλλά και για τις ανωτέρω αναφερθείσες παραμέτρους, σύμφωνα με τον πίνακα 6.20, οι συνολικές αξιολογήσεις που χρειάζονται για την απόκτηση του μετώπου Pareto είναι 151. Στην εν λόγω περίπτωση, για το πρώτο σημείο του Pareto που λαμβάνεται μετά τη σύγκλιση της ALM ισχύουν (με στρογγυλοποίηση στα 6 δεκαδικά ψηφία) $x_1 = z_{cr} = 17626,902627 \text{ m}$, $x_2 = \lambda_w = 0,104310$, $x_3 = \lambda_v = 0,184203$ και $x_4 = W_{fuel} = 29119,048328 \text{ kg}$.

Ακολουθεί, η παρουσίαση, με βάση αυτή την επιλογή για τις ανοχές, του αριθμού των συνολικών αξιολογήσεων του προγράμματος (για να βρεθούν οι αξιολογήσεις που αντιστοιχούν στην πρώτη μη-κυριαρχούμενη λύση πρέπει να αφαιρεθούν αυτές που αντιστοιχούν στο τμήμα που υπολογίζει τα υπόλοιπα 19 σημεία που είναι σταθερές

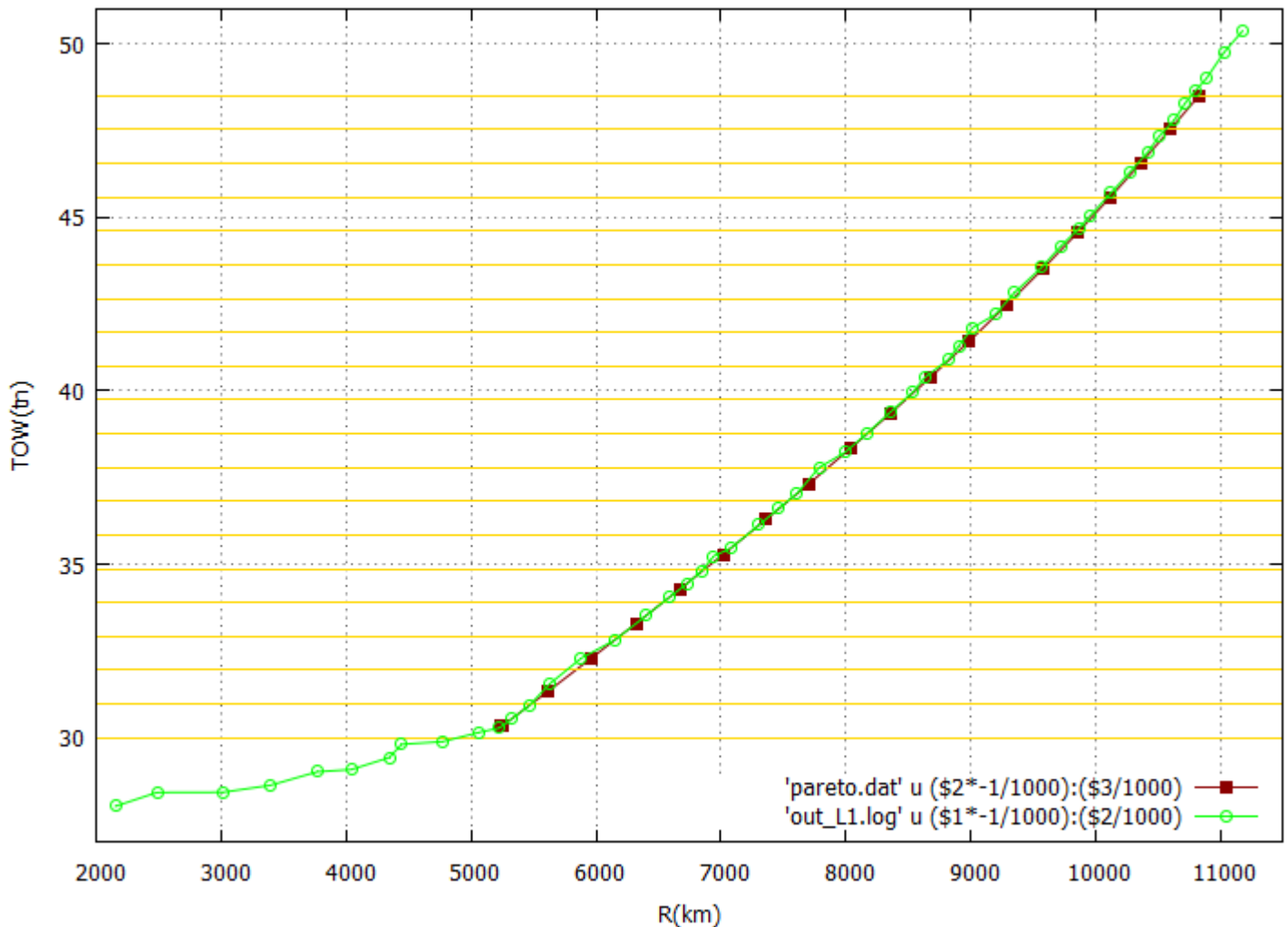
και ίσες με $(n_{elit} - 1)(1 + n + 1) = 114$) για διάφορες τιμές των παραμέτρων της ALM όταν $\bar{x}^0 = [16500 \ 0,3 \ 0,1 \ 26000]^T$ (πίνακας 6.21).

$\hat{\lambda}_2^0$	ω_p^0	γ	Αξιολογήσεις
0	0,4	1,5	151
0	0,3	1,5	δεν συγκλίνει
0	0,5	1,5	145
0	0,6	1,5	151
0	1	1,5	151
0	1,2	1,5	157
0	1,5	1,5	151
0	1,8	1,5	163
0	1,9	1,5	δεν συγκλίνει
0	0,5	1,1	145
0	0,5	1,01	151
0	0,5	2	145
1	0,5	1,5	145
10	0,5	1,5	145

Πίνακας 6.21: Συνολικές αξιολογήσεις που πραγματοποιούνται για την εύρεση του μετώπου Pareto για διαφορετικές τιμές των παραμέτρων της ALM όταν οι ανοχές για τα κριτήρια σύγκλισης της μεθόδου Newton και της ALM λαμβάνουν τις τιμές 10^6 και 1 αντίστοιχα, $\hat{f}_2^{up} = 48500 \text{ kg}$, $\bar{x}^0 = [16500 \ 0,3 \ 0,1 \ 26000]^T$ και $n_{elit} = 20$.

Μέσω του πίνακα 6.21, προκύπτει το συμπέρασμα ότι μία αλλαγή σε οποιαδήποτε από τις παραμέτρους της ALM δεν επιφέρει μεγάλη αυξομείωση στον αριθμό των αξιολογήσεων που πραγματοποιούνται.

Το μέτωπο των μη-κυριαρχούμενων λύσεων που προκύπτει μέσω της μεθόδου που βασίζεται στη συνέχιση του μετώπου Pareto όταν $\hat{f}_2^{up} = 48500 \text{ kg}$, $n_{elit} = 20$ και $\hat{f}_2^{down} = 30000 \text{ kg}$, απεικονίζεται στο σχήμα 6.5, συνοδευόμενο από n_{elit} οριζόντιες γραμμές οι οποίες δηλώνουν την τιμή του \hat{f}_2 για καθένα από τα n_{elit} σημεία των οποίων ζητήθηκε η εύρεση (για να δύναται να διαπιστωθεί η διαφορά μεταξύ της τιμής του \hat{f}_2 και του f_2 για κάθε μη-κυριαρχούμενη λύση που προέκυψε από το πρόγραμμα) και από το μέτωπο που παρέχει το λογισμικό των εξελικτικών αλγορίθμων (για τις τιμές των παραμέτρων που χρησιμοποιήθηκαν στο τρίτο μαθηματικό πρόβλημα όπως αυτό παρουσιάστηκε στην ενότητα 5.5.4).



Σχήμα 6.5: Μέτωπο Pareto που προέκυψε από τη μέθοδο που βασίζεται στη συνέχιση της καμπύλης των μη-κυριαρχούμενων λύσεων για $f_2^{up} = 48500 \text{ kg}$, $f_2^{down} = 30000 \text{ kg}$ και $n_{elit} = 20$ (καφέ καμπύλη) και μέτωπο Pareto που προέκυψε από το λογισμικό των εξελικτικών αλγορίθμων (πράσινη καμπύλη). Οι οριζόντιες κίτρινες γραμμές αντιπροσωπεύουν τις τιμές των \hat{f}_2 για καθένα από τα n_{elit} σημεία που αναζητήθηκαν.

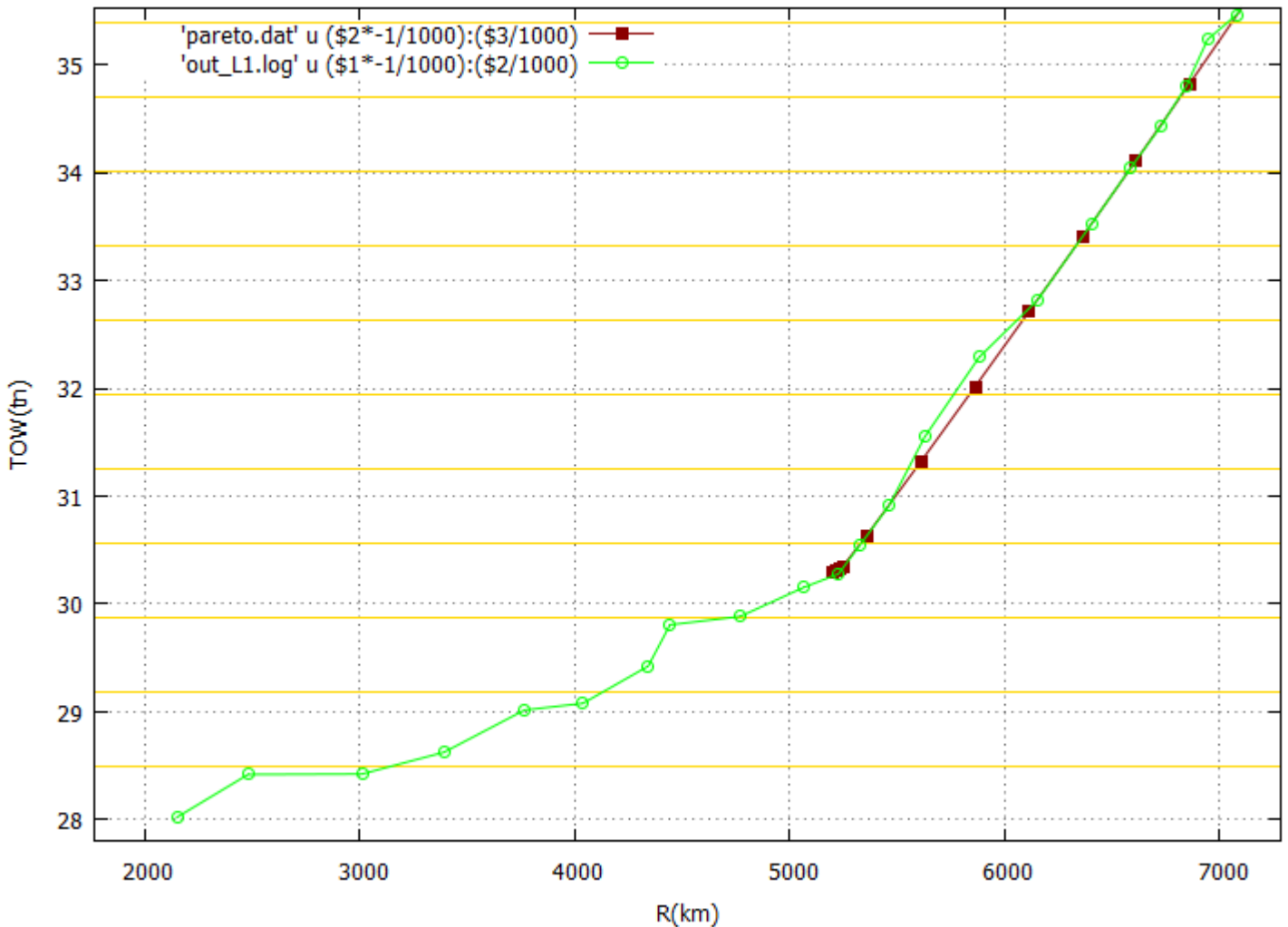
Το σχήμα 6.5, κατά πρώτο λόγο, οδηγεί στο συμπέρασμα ότι υφίστανται διαφορές μεταξύ των τιμών του \hat{f}_2 και του f_2 για τα υπολογισθέντα μέσω της μεθόδου που βασίζεται στη συνέχιση του μετώπου Pareto σημεία, κάτι που παρατηρήθηκε και στα τρία μαθηματικά προβλήματα του κεφαλαίου που προηγήθηκε όπου και εξηγήθηκαν λεπτομερώς οι αιτίες εμφάνισής τους. Παρατηρείται ότι οι εν λόγω διαφορές αρχικά είναι πολύ μικρές, έπειτα αυξάνονται και γίνονται όλο και πιο αισθητές και στα 2 τελευταία σημεία μειώνονται. Μέσω του σχήματος, όμως, καθίσταται φανερό ότι για καθένα από τα αποκτηθέντα σημεία η διαφορά $f_2 - \hat{f}_2$ παραμένει σχετικά – αποδεκτά μικρή. Για την περίπτωση επιλογής του τρίτου συνδυασμού παραμέτρων της ALM που συναντάται στον πίνακα 6.21, στο δεύτερο αποκτηθέν σημείο ισχύει $f_2 - \hat{f}_2 \cong 0,002157 \text{ tn}$, στο τέταρτο $f_2 - \hat{f}_2 \cong$

0,005312 *tn*, στο πέμπτο $f_2 - \hat{f}_2 \cong -0,042240$ *tn*, στο δέκατο $f_2 - \hat{f}_2 \cong -0,383456$ *tn*, στο δέκατο όγδοο $f_2 - \hat{f}_2 \cong -0,622690$ *tn*, στο προτελευταίο $f_2 - \hat{f}_2 \cong -0,611860$ *tn* και στο τελευταίο $f_2 - \hat{f}_2 \cong 0,339763$ *tn*.

Παρά, όμως, τις υφιστάμενες διαφορές $f_2 - \hat{f}_2$, παρατηρείται ότι η καμπύλη που προέκυψε από τη μέθοδο που βασίζεται στη συνέχιση του μετώπου Pareto για $\hat{f}_2^{up} = 48500$ *kg*, $\hat{f}_2^{down} = 30000$ *kg* και $n_{elit} = 20$ ουσιαστικά πέφτει πάνω σε αυτή που προέκυψε από το λογισμικό των εξελικτικών αλγορίθμων. Πρέπει, επίσης, να αναφερθεί ότι η καμπύλη που προέκυψε από το αναπτυχθέν πρόγραμμα είναι ομαλή, σε αντίθεση με αυτή που αποκτήθηκε μέσω του εξελικτικού αλγορίθμου, γεγονός που οδηγεί την πρώτη καμπύλη να βρίσκεται ελαφρώς πιο κάτω – δεξιά από τη δεύτερη σε κάποια μικρά τμήματα (δηλαδή έχει επιτευχθεί η εύρεση ‘καλύτερων’ σημείων στα τμήματα αυτά). Αυτό το πλεονέκτημα της μεθόδου που πραγματεύεται η παρούσα εργασία, επιδείχθηκε και στην περίπτωση των μαθηματικών προβλημάτων του κεφαλαίου 5.

Ένα επιπλέον συμπέρασμα που προκύπτει μετά την παρατήρηση του σχήματος 6.5, είναι ότι η καμπύλη που αποκτήθηκε με βάση τη μέθοδο της συνέχισης του μετώπου Pareto αποτελεί μέρος του μετώπου που αποκτήθηκε μέσω του λογισμικού των εξελικτικών αλγορίθμων, καθώς η πράσινη καμπύλη εκτείνεται και προς τις δύο πλευρές της καφέ (κυρίως προς την πλευρά των μικρότερων *R*). Πρέπει, όμως, να ειπωθεί ότι το αποκτηθέν τμήμα καλύπτει μεγάλο κομμάτι του μετώπου Pareto που δόθηκε μέσω του εξελικτικού αλγορίθμου, οπότε, γενικά, μπορεί να θεωρηθεί ότι καλύπτει τις ανάγκες του χρήστη.

Παρά ταύτα, παρουσιάζει ενδιαφέρον η συμπεριφορά του αναπτυχθέντος αλγορίθμου αν τεθούν $\hat{f}_2^{up} = 48500$ *kg*, $\hat{f}_2^{down} = 28500$ *kg* και $n_{elit} = 30$ (δηλαδή ζητείται μεγαλύτερη περιοχή του μετώπου και περισσότερα σημεία), η οποία απεικονίζεται στο σχήμα 6.6 (σημειώνεται ότι έχουν σχεδιαστεί μόνο τα τελευταία 11 προκύπτοντα σημεία, καθώς ενδιαφέρει η εξέταση των αποτελεσμάτων στην περιοχή των μικρότερων *R* και *TOW*).



Σχήμα 6.6: Τελευταίες 11 μη-κυριαρχούμενες λύσεις που προέκυψαν από τη μέθοδο που βασίζεται στη συνέχιση του μετώπου Pareto για $\hat{f}_2^{up} = 48500 \text{ kg}$, $\hat{f}_2^{down} = 28500 \text{ kg}$ και $n_{elit} = 30$ (καφέ καμπύλη) και τελευταίο τμήμα μετώπου Pareto που προέκυψε από το λογισμικό των εξελικτικών αλγορίθμων (πράσινη καμπύλη). Οι κίτρινες οριζόντιες γραμμές αντιπροσωπεύουν τις τιμές των \hat{f}_2 για καθένα από τα 11 τελευταία σημεία που αναζητήθηκαν.

Μέσω του σχήματος 6.6, καταδεικνύεται ότι ο αναπτυχθείς αλγόριθμος (για τις αναφερθείσες πληροφορίες εισόδου) δεν δύναται να αποδώσει σημεία του μετώπου Pareto τα οποία αντιστοιχούν σε TOW μικρότερα από περίπου $30,3 \text{ tn}$. Συγκεκριμένα, παρατηρείται μία ‘συσσώρευση’ των τελευταίων τεσσάρων αποδιδόμενων σημείων σε μία μικρή περιοχή όπου το TOW κυμαίνεται μεταξύ των $30,297878 \text{ tn}$ και των $30,350961 \text{ tn}$. Ο λόγος για τον οποίο συμβαίνει αυτό φανερώνεται αν παρατηρηθεί η μορφή του μετώπου Pareto που αποδίδεται μέσω του λογισμικού των εξελικτικών αλγορίθμων. Είναι σαφές ότι μετά την μη-κυριαρχούμενη λύση που αντιστοιχεί σε $TOW \cong 30,3 \text{ tn}$, το μέτωπο Pareto φαίνεται να ‘σακίζει’, δηλαδή αλλάζει άρδην η κλίση του. Αυτό οφείλεται στο γεγονός ότι μετά την αναφερθείσα μη-κυριαρχούμενη λύση όλες οι μεταβλητές σχεδιασμού

εκτός του W_{fuel} (το οποίο στην εν λόγω λύση φτάνει την τιμή των 15000 kg και, στη συνέχεια, διατηρείται σταθερό σε όλη την περιοχή που δεν 'καλύπτεται' από τον αναπτυχθέντα αλγόριθμο) υφίστανται 'δραματικές' αλλαγές, δηλαδή οι καμπύλες $\hat{f}_2 - x_i, i = 1,2,3$ παίρνουν πολύ μεγάλες κλίσεις. Τις εν λόγω μεγάλες μεταβολές δεν μπορεί να παρακολουθήσει η μέθοδος που βασίζεται στη συνέχιση του μετώπου Pareto, καθώς η προσέγγιση την οποία πραγματοποιεί για να μεταβεί στο επόμενο σημείο του μετώπου είναι εφαπτομενική (όπως εξηγήθηκε λεπτομερώς στα κεφάλαια 4 και 5). Το γεγονός αυτό σαφέστατα αποτελεί ένα μειονέκτημα της μεθόδου που πραγματεύεται η παρούσα εργασία, καθώς δεν επιτυγχάνεται η εύρεση ολόκληρου του μετώπου Pareto. Ωστόσο, πρέπει να σημειωθεί ότι, γενικά, η τελική επιλογή για το ποιος σχεδιασμός θα κατασκευαστεί δεν αναμένεται να πραγματοποιηθεί από το τμήμα του μετώπου Pareto που αντιστοιχεί σε τόσο μικρές τιμές για το TOW .

Τέλος, αναφέρεται, ενδεικτικά, ότι το λογισμικό των εξελικτικών αλγορίθμων απέδωσε το μέτωπο Pareto σε περίπου 800 αξιολογήσεις (σημειώνεται ότι έγινε ταυτόχρονη χρήση κατανεμημένης ανίχνευσης και IPE με τις παραμέτρους του εξελικτικού αλγορίθμου να διατηρούν τις τιμές που παρουσιάσθηκαν στην ενότητα 5.4.4.2). Σύμφωνα με τον πίνακα 6.21, η μέθοδος που βασίζεται στη συνέχιση του μετώπου Pareto φαίνεται να επιτυγχάνει την εύρεση του μετώπου μη-κυριαρχούμενων λύσεων πολύ γρηγορότερα. Τονίζεται ότι το εν λόγω συμπέρασμα προκύπτει για το δεδομένο πρόβλημα αλλά και για το σετ παραμέτρων που επιλέχθηκε στο λογισμικό των εξελικτικών αλγορίθμων.

7 Επίλογος – Συζήτηση – Συμπεράσματα

Η παρούσα διπλωματική εργασία επικεντρώθηκε στην ανάπτυξη λογισμικού για την επίλυση προβλημάτων βελτιστοποίησης δύο στόχων με βάση τη θεωρία της συνέχισης του μετώπου Pareto που παρουσιάστηκε από τους Schmidt και Schulz, [1]. Το αναπτυχθέν λογισμικό δύναται να αποδώσει το μέτωπο Pareto με και χωρίς το στάδιο της διόρθωσης. Και στις δύο περιπτώσεις, ο αλγόριθμος ξεκινάει από την εύρεση μίας μη-κυριαρχούμενης λύσης κάνοντας χρήση είτε της μεθόδου της απότομης καθόδου είτε της μεθόδου Newton. Έπειτα, στην περίπτωση που αποφασιστεί η μη-συμμετοχή του σταδίου της διόρθωσης, πραγματοποιείται ο υπολογισμός όλων των επόμενων ζητηθέντων μη-κυριαρχούμενων λύσεων με βάση τη θεωρία της συνέχισης του μετώπου Pareto, ενώ, όταν επιλεγεί ο αλγόριθμος πρόβλεψης-διόρθωσης, η κάθε μη-κυριαρχούμενη λύση που προβλέπεται υφίσταται διόρθωση, διδόμενη ως αρχική τιμή στην ίδια μέθοδο βελτιστοποίησης προβλημάτων ενός στόχου που πραγματοποίησε τον υπολογισμό της πρώτης μη-κυριαρχούμενης λύσης. Το μεγαλύτερο βάρος δόθηκε στην εξέταση της 'ποιότητας' των αποτελεσμάτων στην περίπτωση μη-συμπερίληψης του σταδίου της διόρθωσης, με κίνητρο τη μείωση του υπολογιστικού κόστους.

Αρχικά, μελετήθηκαν τρία μαθηματικά προβλήματα βελτιστοποίησης δύο στόχων με μοναδικούς περιορισμούς αυτούς που επιβάλλουν τα όρια των μεταβλητών σχεδιασμού. Για τα δύο πρώτα, το μέτωπο Pareto ήταν συνεχές και αποκτήθηκε τόσο με όσο και χωρίς το στάδιο της διόρθωσης, ενώ ως μέθοδοι προβλημάτων βελτιστοποίησης ενός στόχου χρησιμοποιήθηκαν και η μέθοδος της απότομης καθόδου και η μέθοδος Newton. Όσον αφορά το τρίτο μαθηματικό πρόβλημα, το μέτωπο Pareto ήταν διακοπτόμενο και αποκτήθηκε χωρίς τη συμμετοχή του σταδίου της διόρθωσης, ενώ ως μέθοδος βελτιστοποίησης προβλημάτων ενός στόχου επιλέχθηκε η μέθοδος της απότομης καθόδου. Έπειτα, πραγματοποιήθηκε η βελτιστοποίηση των επιδόσεων υπερηχητικού επιβατικού αεροσκάφους μέσω της μεθόδου που βασίζεται στη συνέχιση του μετώπου Pareto χωρίς το στάδιο της διόρθωσης, με τη μέθοδο Newton να επιλέγεται ως μέθοδος βελτιστοποίησης μονοκριτηριακών προβλημάτων. Για όλες τις εφαρμογές που παρουσιάστηκαν, πραγματοποιήθηκε σύγκριση των αποτελεσμάτων με αυτά ενός αναπτυγμένου από τη ΜΠΥΡΒ/ΕΜΠ εξελικτικού αλγορίθμου.

Κατά πρώτο λόγο, μέσω της προαναφερθείσας σύγκρισης, καταδείχθηκε το γεγονός ότι, για το σύνολο των προβλημάτων που μελετήθηκαν, το λογισμικό που αναπτύχθηκε στο πλαίσιο της παρούσας εργασίας δύναται να αποδώσει το μέτωπο Pareto με σημαντικά μικρότερο αριθμό αξιολογήσεων σε σχέση με το λογισμικό των εξελικτικών αλγορίθμων. Μάλιστα, όπως διαπιστώθηκε μέσω των δύο πρώτων μαθηματικών προβλημάτων που εξετάστηκαν, το ανωτέρω

συμπέρασμα ισχύει και στην περίπτωση εύρεσης του μετώπου Pareto με τη συμμετοχή του σταδίου της διόρθωσης, αν και σαφέστατα η προσθήκη του εν λόγω σταδίου αυξάνει αρκετά τον αριθμό των αξιολογήσεων (από 161 σε 229 για το πρώτο πρόβλημα και από 165 σε 217 για το δεύτερο). Η επίτευξη πολύ μικρότερου υπολογιστικού κόστους αποτελεί το μεγάλο πλεονέκτημα μίας μεθόδου που βασίζεται στη συνέχιση του μετώπου Pareto.

Επιπροσθέτως, για τα προβλήματα που παρουσιάστηκαν, αποδείχθηκε ότι η ποιότητα των αποτελεσμάτων που παρέχονται από το αναπτυχθέν λογισμικό στην περίπτωση αφαίρεσης του σταδίου της διόρθωσης είναι πολύ ικανοποιητική, καθώς οι προκύπτουσες καμπύλες πρακτικά ταυτίζονταν με αυτές του λογισμικού των εξελικτικών αλγορίθμων. Το γεγονός αυτό ήταν δυνατό παρά τις ολοένα αυξανόμενες διαφορές που παρατηρήθηκαν μεταξύ των τιμών του \hat{f}_2 και του f_2 για τα υπολογισθέντα σημεία από το αναπτυχθέν πρόγραμμα και μάλιστα επιτεύχθηκε για σχετικά μεγάλες τιμές του $\delta\hat{f}_2$ (δηλαδή για σχετικά λίγα ζητηθέντα σημεία ανάμεσα στο επιθυμητό τμήμα του μετώπου). Αυτό δείχνει ότι μπορούν να 'εξοικονομηθούν' οι αρκετές επιπλέον αξιολογήσεις που απαιτούνται για το στάδιο της διόρθωσης, καθώς τα σημεία που αποκτώνται θα έχουν μεν 'λίγο' διαφορετικό f_2 σε σχέση με το \hat{f}_2 αλλά θα βρίσκονται πάνω στο μέτωπο Pareto.

Ακόμα, παρατηρήθηκε ότι, όπως ήταν αναμενόμενο λόγω του μαθηματικού υποβάθρου της θεωρίας της συνέχισης του μετώπου Pareto, η προκύπτουσα από το αναπτυχθέν λογισμικό καμπύλη των μη-κυριαρχούμενων λύσεων είναι ομαλή, σε αντίθεση με αυτή του λογισμικού των εξελικτικών αλγορίθμων (λόγω του τρόπου με τον οποίο δημιουργούνται τα νέα άτομα από την εξελικτική διαδικασία). Αυτό είναι ένα πλεονέκτημα της μεθόδου που πραγματεύεται η παρούσα εργασία, καθώς σε κάποιες μικρές περιοχές των καμπυλών επιτεύχθηκε η εύρεση 'καλύτερων' σημείων.

Στον αντίποδα, μία μέθοδος που βασίζεται στη συνέχιση του μετώπου Pareto απαιτεί ο χρήστης να έχει καλή γνώση του προβλήματος που καλείται να επιλύσει, ώστε να δύναται να αποφασίσει σχετικά με τις τιμές των ορίων του μετώπου Pareto \hat{f}_2^{up} και \hat{f}_2^{down} που θα δοθούν στο πρόγραμμα. Στην περίπτωση δε του τρίτου μαθηματικού προβλήματος, η ανάγκη αυτή είναι ακόμα πιο επιτακτική, καθώς ο χρήστης οφείλει να γνωρίζει ότι το μέτωπο Pareto που αναμένεται να προκύψει 'αποτελείται από δύο τμήματα', δηλαδή είναι διακοπτόμενο. Σε αντίθετη περίπτωση, ο χρήστης, ενδεχομένως, θα οδηγηθεί σε εσφαλμένη εντύπωση σχετικά με το μέτωπο Pareto (λ.χ. σχήμα 5.26) ή στην πραγματοποίηση δοκιμών, οι οποίες, βέβαια, συνοδεύονται από αύξηση του υπολογιστικού κόστους. Σημειώνεται, ωστόσο, ότι, γενικά, σε ένα πραγματικό πρόβλημα, ο κληθείς για να το επιλύσει γνωρίζει προσεγγιστικά τα όρια του μετώπου Pareto ή του τμήματος του μετώπου που τον ενδιαφέρει.

Επιπλέον, χαρακτηριστικό του αναπτυσθέντος λογισμικού αποτελεί η αδυναμία του, σε κάποιες περιπτώσεις, να αποδώσει ολόκληρο το μέτωπο Pareto. Πιο χαρακτηριστικές είναι οι περιπτώσεις του πρώτου μαθηματικού προβλήματος και της βελτιστοποίησης των επιδόσεων του υπερηχητικού επιβατικού αεροσκάφους, όπου δεν επιτεύχθηκε η εύρεση 'μικρής' περιοχής σε ένα από τα άκρα του μετώπου (όταν το Pareto όδευε προς οριζοντιοποίηση για την πρώτη περίπτωση και όταν 'τσάκιζε' δηλαδή άλλαζε άρδην η κλίση του για τη δεύτερη). Αυτό, όπως εξηγήθηκε, οφείλεται στο γεγονός ότι η μέθοδος που βασίζεται στη συνέχιση του μετώπου Pareto, δεν μπορεί να 'παρακολουθήσει' τις μεγάλες μεταβολές που παρατηρήθηκαν από ένα σημείο και μετά στις μεταβλητές σχεδιασμού, καθώς η προσέγγιση την οποία πραγματοποιεί για να μεταβεί στο επόμενο σημείο του μετώπου είναι εφαπτομενική. Όμως, πρέπει να ειπωθεί ότι, σε όλα τα προβλήματα που μελετήθηκαν, η καμπύλη που αποκτήθηκε κάλυπτε πολύ μεγάλο μέρος του μετώπου Pareto που απέδωσε το λογισμικό των εξελικτικών αλγορίθμων.

Αυτή η αδυναμία εύρεσης ολόκληρου του μετώπου εξαλείφεται εν μέρει αλλά όχι εξ ολοκλήρου όταν προστίθεται το στάδιο της διόρθωσης, καθώς, όπως παρατηρήθηκε στην περίπτωση του πρώτου μαθηματικού προβλήματος, αφενός στο τμήμα όπου παρατηρούνται οι μεγάλες αλλαγές στις μεταβλητές αυξάνεται ο αριθμός των απαιτούμενων αξιολογήσεων κατά το στάδιο της διόρθωσης (αφού η πρόβλεψη δεν είναι τόσο 'καλή') και αφετέρου τα αποκτηθέντα σημεία σε αυτό το τμήμα θα είναι 'μακριά' το ένα από το άλλο (συγκεκριμένα βρέθηκαν 1-3 σημεία στο εν λόγω τμήμα) ώστε να κρατηθεί το υπολογιστικό κόστος σε χαμηλό επίπεδο.

Το μεγαλύτερο 'εμπόδιο' που θέτει μία μέθοδος που βασίζεται στη συνέχιση του μετώπου Pareto είναι, αδιαμφισβήτητα, η απαίτηση της παροχής των πρώτων και κυρίως των δεύτερων παραγώγων των συναρτήσεων στόχου ως προς τις μεταβλητές σχεδιασμού. Εξηγήθηκε ότι, αν χρησιμοποιηθεί η συζυγής μέθοδος, για κάθε μη-κυριαρχούμενη λύση που υπολογίζεται μέσω της θεωρίας της συνέχισης του μετώπου Pareto απαιτούνται $1 + n + 1$ αξιολογήσεις, ενώ αυτό είναι το κόστος και για κάθε επανάληψη της μεθόδου Newton. Είναι κατανοητό ότι, ιδιαίτερα, οι δεύτερες παράγωγοι αποτελούν ένα 'ακριβό εργαλείο'. Το εν λόγω πρόβλημα ενισχύεται, ενδεχομένως, σε πραγματικά προβλήματα όπου οι δεύτερες παράγωγοι είναι εξαιρετικά δύσκολο να υπολογιστούν. Αναφέρεται, χαρακτηριστικά, ότι, στο πλαίσιο της παρούσας εργασίας, έγινε προσπάθεια επίλυσης ενός προβλήματος αεροδυναμικής βελτιστοποίησης στο οποίο δεν ήταν διαθέσιμες οι δεύτερες παράγωγοι. Επιχειρήθηκε η προσέγγισή τους μέσω της μεθόδου BFGS αλλά, σε αυτό το πρόβλημα, τα αποτελέσματα δεν ήταν ικανοποιητικά.

Ωστόσο, παρά το 'κόστος' που εισάγει ο υπολογισμός των απαιτούμενων παραγώγων στη μέθοδο και παρά την πιθανότητα

απώλειας σε κάποιες περιπτώσεις 'μικρού' τμήματος του μετώπου Pareto, η ταχύτητα με την οποία μπορούν να αποδώσουν ακόμα και μεγάλο αριθμό μη-κυριαρχούμενων λύσεων (λ.χ. στο τρίτο μαθηματικό πρόβλημα αποκτήθηκαν 28 μη-κυριαρχούμενες λύσεις με περίπου το 1/6 των αξιολογήσεων που απαιτήθηκαν από το λογισμικό των εξελικτικών αλγορίθμων) και μάλιστα με εξαιρετικά καλή ακρίβεια, κάνει τις μεθόδους που βασίζονται στη συνέχιση του μετώπου Pareto ένα πολύτιμο εργαλείο στη διάθεση του κάθε επιστήμονα.

Οι μέθοδοι που βασίζονται στη συνέχιση του μετώπου Pareto αποτελούν έναν σχετικά καινούριο και πολλά υποσχόμενο τομέα. Μέσω των ανωτέρω, καθίσταται σαφές ότι μείζον θέμα είναι η αντιμετώπιση της απαίτησης σχετικά με τον υπολογισμό των δευτέρων παραγώγων λόγω των δυσκολιών που αυτός ενέχει αλλά και για να επιτευχθεί περαιτέρω ελάττωση του υπολογιστικού κόστους. Στο πλαίσιο αυτό κινείται η διπλωματική εργασία του Μ. Καφή, [57]. Συγκεκριμένα, στην εν λόγω εργασία λαμβάνεται, επίσης, ως βάση η μελέτη που πραγματοποίησαν οι Schmidt και Schulz σε μία προσπάθεια ανάπτυξης ενός αλγορίθμου ικανού να εντοπίσει το μέτωπο Pareto με χρήση της μεθόδου πρόβλεψης-διόρθωσης. Η διαφοροποίηση σε σχέση με την παρούσα εργασία έγκειται στο ότι ο αναπτυσσόμενος αλγόριθμος συνοδεύεται από τη χρήση της μεθόδου Newton με αποκοπή σε συνδυασμό με τη μέθοδο συζυγών κλίσεων, σε μια προσπάθεια περιορισμού του υψηλού κόστους υπολογισμού του εσσιανού μητρώου.

Τη 'φυσική συνέχεια' των θεμάτων που πραγματεύεται η παρούσα εργασία θα μπορούσε να αποτελέσει, επίσης, η επίλυση προβλημάτων βελτιστοποίησης με περισσότερες μεταβλητές σχεδιασμού και με επιπλέον περιορισμούς πέραν αυτών που επιβάλλουν τα όρια των μεταβλητών σχεδιασμού. Στις προτάσεις μελλοντικής έρευνας μπορεί, ακόμα, να ενταχθεί η επίλυση προβλημάτων περισσότερων των δύο στόχων (σημειώνεται ότι το αναπτυχθέν λογισμικό θα δύναται να αντιμετωπίσει τέτοια προβλήματα με πολύ μικρές αλλαγές), όπου η κίνηση στο μέτωπο Pareto δεν θα γίνεται πλέον στις δύο διαστάσεις. Τέλος, ένα επόμενο βήμα μπορεί να είναι η εφαρμογή της μεθόδου σε πραγματικά προβλήματα της βιομηχανίας.

Βιβλιογραφία

1. S. Schmidt and V. Schulz: *Pareto-curve continuation in multi-objective optimization*, Pacific Journal of Optimization, 2008, Volume 4, pp. 243-257.
2. Α. Ευστρατιάδης: *Μη γραμμικές μέθοδοι σε πολυκριτηριακά προβλήματα βελτιστοποίησης υδατικών πόρων, με έμφαση στη βαθμονόμηση υδρολογικών μοντέλων*, Διδακτορική διατριβή, Σχολή Πολιτικών Μηχανικών, Τομέας Υδατικών Πόρων και Περιβάλλοντος, Ε.Μ.Π., 2008.
3. Α. Πουλιέζος: *Βέλτιστος έλεγχος*, Σημειώσεις μεταπτυχιακού μαθήματος, Τμήμα Μηχανικών Παραγωγής και Διοίκησης, Πολυτεχνείο Κρήτης, Έκδοχή 2.0, 2013, http://pouliezos.dpem.tuc.gr/pdf/notes_OC.PDF.
4. Κ. Χ. Γιαννάκογλου: *Μέθοδοι Βελτιστοποίησης στην Αεροδυναμική*, Εκδόσεις Ε.Μ.Π., 2006.
5. T. Bechtold: *Structure Topology Optimization*, Chapter 4: *Optimality Conditions*.
6. D. Belegundu and T. R. Chandrupatla: *Optimization Concepts and Applications in Engineering*, Prentice-Hall, NJ, ISBN 0-13-031279-7.
7. J. Lötscher: *An augmented Lagrangian method for inequality constrained optimization applied to SPECT reconstruction*, Bachelor Thesis, Universität zu Lübeck, 2011.
8. Κ. Χ. Γιαννάκογλου: *Άλλες (πέρα των ΕΑ) Στοχαστικές Μέθοδοι Βελτιστοποίησης*, <http://velos0.ltt.mech.ntua.gr/research/>.
9. J. C. Principe, N. R. Euliano and W. C. Lefebvre: *Neural and Adaptive Systems: Fundamentals through Simulations*, Chapter 1.8: *Newton's Method*, John Wiley & Sons, Inc. New York, 1999, ISBN: 0471351679.
10. S. Boyd and L. Vandenberghe: *Convex optimization*, Chapter 9: *Unconstrained minimization*, Cambridge University Press, March 2004, ISBN: 9780521833783, http://stanford.edu/~boyd/cvxbook/bv_cvxbook.pdf.
11. Π. Ν. Ανδρεάδης: *Βελτιστοποίηση παραμέτρων έγχυσης σε δίχρονους ναυτικούς κινητήρες Diesel*, Διπλωματική εργασία, Σχολή Ναυτικών Μηχανολόγων Μηχανικών, Τομέας Ναυτικής Μηχανολογίας, Ε.Μ.Π., 2008.
12. Ι. Κ. Δολιανίτης: *Βελτιστοποίηση προγράμματος λειτουργίας θερμικών μονάδων ηλεκτροπαραγωγής με χρήση εξελικτικών αλγορίθμων και μιας νέας παραμετροποίησης*, Διπλωματική εργασία, Σχολή Μηχανολόγων Μηχανικών, Τομέας Ρευστών, Εργαστήριο Θερμικών Στροβιλομηχανών, Ε.Μ.Π., 2011.
13. J. R. J. Rao and P. Y. Papalambros: *A nonlinear Programming Continuation Strategy for One Parameter Design Optimization*

- Problems*, Design Automation Conference American Society of Mechanical Engineers, New York, 1989.
14. O. Forster: *Analysis 2*, Friedr. Vieweg & Sohn, Braunschweig / Wiesbaden, 1982.
 15. *Extremal Behavior of One Parameter Families of Optimal Design Models*, submitted to the ASME Design Automation Conf., 1989.
 16. <http://www.genetic-programming.org/>
 17. http://en.wikipedia.org/wiki/Genetic_programming
 18. Κ. Χ. Γιαννάκογλου, Ι. Αναγνωστόπουλος, Γ. Μπεργελές: *Αριθμητική Ανάλυση για Μηχανικούς*, Εκδόσεις Ε.Μ.Π., 2003.
 19. *The EASY (Evolutionary Algorithms SYstem) software*, <http://velos0.ltt.mech.ntua.gr/EASY>
 20. C. Poloni, G. Mosetti, and S. Contessi: *Multi objective optimization by GAs: Application to system and component design*, in Proc. Comput. Methods in Applied Sciences'96: Invited Lectures and Special Technological Sessions of the 3rd ECCOMAS Comput. Fluid Dynamics Conf. and the 2nd ECCOMAS Conf. Numerical Methods in Engineering, September 1996, pp. 258–264.
 21. Β. Ασούτη: *Μέθοδοι αεροδυναμικής ανάλυσης και σχεδιασμού για ροές υψηλών και χαμηλών ταχυτήτων, σε πολυεπεξεργαστικό περιβάλλον*, Διδακτορική διατριβή, Σχολή Μηχανολόγων Μηχανικών, Τομέας Ρευστών, Εργαστήριο Θερμικών Στροβιλομηχανών, Ε.Μ.Π., 2009.
 22. Ν. Καζαζάκης: *Βελτιστοποίηση υπερηχητικού επιβατικού αεροσκάφους με χρήση της μεθόδου μιγαδικών μεταβλητών*, Διπλωματική εργασία, Σχολή Μηχανολόγων Μηχανικών, Τομέας Ρευστών, Εργαστήριο Θερμικών Στροβιλομηχανών, Ε.Μ.Π., 2008.
 23. J. H. Ferziger and M. Peric: *Computational methods for fluid dynamics*, Springer, NY, Third edition, 2002, Chapter 3: *Finite difference methods*.
 24. H. Lomax, T. H. Pulliam and D. W. Zingg: *Fundamentals of computational fluid dynamics (scientific computation)*, Springer, 2003, Chapter 3: *Finite Difference Approximations*
 25. L. M. Milne-Thomson: *The calculus of finite differences*, AMS Chelsea Pub., Providence, R.I., 2000.
 26. S. K. Lele: *Compact finite difference schemes with spectral-like resolution*, Journal of computational physics, November 1992, Volume 103, Issue 1, pp. 16-42.
 27. P. G. Ciarlet and J. L. Lions: *Handbook of numerical analysis, Volume I: Finite Difference Methods (Part 1)*, North-Holland, January 2000.
 28. J. C. Newman, W. K. Anderson and D. L. Whitefield: *Multidisciplinary sensitivity derivatives using complex*

- variables*, Journal of Aircraft 1999, Volume 36, Issue 1, pp. 87–96.
29. W. Squire and G. Trapp: *Using complex variables to estimate derivatives of real function*, SIAM Review 1998 Volume 40, Issue 1, pp. 110–112.
 30. J. R. R. A. Martins, I. M. Kroo, and J. J. Alonso: *An automated method for sensitivity analysis using complex variables*, Proceedings of the 38th Aerospace Sciences Meeting, AIAA Paper 2000-0689, Reno, NV, January 2000.
 31. W. K. Anderson and J. C. Newman, D. L. Whitfield and E. J. Nielsen: *Sensitivity analysis for the Navier-Stokes equations on unstructured meshes using complex variables*, AIAA Paper 99-3294, Proceedings of the 17th Applied Aerodynamics Conference, 28 June 1999.
 32. S. H. Kima and H. Huh: *Design sensitivity analysis of sheet metal forming processes with a direct differentiation method*, Journal of materials processing technology, 20 December 2002, Volumes 130-131, pp. 504-510.
 33. D. Vieilledent and L. Fourment: *Shape optimization of axisymmetric preform tools in forging using a direct differentiation method*, International journal for numerical methods in engineering, 20 December 2001, Volume 52, Issue 11, pp. 1301–1321.
 34. R. M. Mukherjee, K. D. Bhalerao and K. S. Anderson: *A divide-and-conquer direct differentiation approach for multibody system sensitivity analysis*, Structural and multidisciplinary optimization 2007, Volume 35 pp. 413–429.
 35. Q. Gu and G. Wang: *Direct differentiation method for response sensitivity analysis of a bounding surface plasticity soil model*, Soil dynamics and earthquake engineering, June 2013, Volume 49, pp. 135-145.
 36. C. Zheng, T. Matsumoto, T. Takahashi and H. Chen: *Explicit evaluation of hypersingular boundary integral equations for acoustic sensitivity analysis based on direct differentiation method*, Engineering analysis with boundary elements, November 2011, Volume 35, Issue 11, pp. 1225-1235.
 37. L. B. Rall: *Automatic Differentiation: Techniques and Applications*, Springer Verlag, Berlin, Germany, 1981.
 38. A. Griewank: *On automatic differentiation. Mathematical programming: Recent developments and applications*, Kluwer Academic Publishers, Boston, MA, 1989.
 39. A. Griewank and A. Walther: *Evaluating derivatives: Principles and techniques of algorithmic differentiation*, Second edition, September 2008.
 40. L. L. Sherman, A. C. Taylor III, L. L. Green, P. A. Newman, G. W. Hou and V. M. Korivi: *First- and second-order aerodynamic sensitivity derivatives via automatic differentiation with*

- incremental iterative methods*, Journal of computational physics, 1996, Volume 129, Issue 2, pp. 307–331.
41. V. G. Asouti, A. S. Zymaris, D. I. Papadimitriou and K. C. Giannakoglou: *Continuous and discrete adjoint approaches for aerodynamic shape optimization with low Mach number preconditioning*, International journal for numerical methods in fluids, 2008, Volume 57, Issue 10, pp. 1485–1504.
 42. Δ. Παπαδημητρίου: *Συζυγείς (adjoint) διατυπώσεις για την ανάλυση – σχεδίαση περυγώσεων στροβιλομηχανών και τη βέλτιστη προσαρμογή πλέγματος με a posteriori εκτίμηση σφάλματος*, Διδακτορική διατριβή, Σχολή Μηχανολόγων Μηχανικών, Εργαστήριο Θερμικών Στροβιλομηχανών, Ε.Μ.Π., 2006.
 43. M. S. Campobasso, M. C. Duta and M. B. Giles: *Adjoint calculation of sensitivities of turbomachinery objective functions*, AIAA Journal of propulsion and power, 2003, Volume 19, Issue 4 pp. 693–703.
 44. S. K. Kim, J. J. Alonso and A. Jameson: *Two-dimensional high-lift aerodynamic optimization using the continuous adjoint method*, AIAA Paper 2000-4741, 8th AIAA Symposium on Multidisciplinary analysis and optimization, Long Beach, CA, September 2000.
 45. D. I. Papadimitriou, A. S. Zymaris and K. C. Giannakoglou: *Discrete and continuous adjoint formulations for turbomachinery applications*, EUROGEN 2005, Munich, Germany, 2006.
 46. D. I. Papadimitriou and K. C. Giannakoglou: *Direct, adjoint and mixed approaches for the computation of Hessian in airfoil design problems*, International journal for numerical methods in fluids, 2007, Volume 56, Issue 10, pp. 1929–1943.
 47. J. H. Holland: *Adaptation in natural and artificial systems*, 1975. University of Michigan Press, Second edition, 1992.
 48. I. Rechenberg: *Evolutionsstrategie: Optimierung technischer systeme nach prinzipien der biologischen evolution*, Frommann-Holzboog, Stuttgart, 1973.
 49. T. Bäck, D. Fogel and Z. Michalewicz: *Handbook of Evolutionary Computation*, IOP Publishing Ltd, Bristol, UK, 1997.
 50. V. G. Asouti and K. C. Giannakoglou: *Aerodynamic Optimization Using a Parallel Asynchronous Evolutionary Algorithm Controlled by Strongly Interacting Demes*, Engineering optimization, 2009, Volume 41, Issue 3 pp. 241–257.
 51. I. C. Kampolis, A. S. Zymaris, V. G. Asouti and K. C. Giannakoglou: *Multilevel optimization strategies based on Metamodel – Assisted Evolutionary Algorithms, for computationally expensive problems*, IEEE Congress on

- Evolutionary Computation CEC, 25-28 September 2007, Singapore.
52. K. C. Giannakoglou: *Designing turbomachinery blades using evolutionary methods*, ASME Paper 99-GT-181, 44th ASME Gas Turbine & Aeroengine Congress, Indianapolis, USA, June 1999.
 53. K. C. Giannakoglou, I. C. Kampolis and C. A. Georgopoulou: *Metamodel-assisted evolutionary algorithms (MAEAs)*, von Karman Institute Lectures Series on Introduction to Optimization and Multidisciplinary Design in Aeronautics and Turbomachinery, June 2008.
 54. P. I. K. Liakopoulos, I. C. Kampolis and K. C. Giannakoglou: *Grid enabled, hierarchical distributed meta-model assisted evolutionary algorithms for aerodynamic shape optimization*, Future Generation Computer Systems, 2008, Volume 24, Issue 7, pp. 701–708.
 55. J. R. Koza: *Genetic programming: On the programming of computers by means of natural selection*, The MIT Press, Cambridge, 1992.
 56. J. R. Koza: *Genetic programming II: Automatic discovery of reusable programs*, The MIT Press, Cambridge, 1994.
 57. Μ. Καψής: *Εντοπισμός του μετώπου Pareto στην πολυκριτηριακή αεροδυναμική βελτιστοποίηση με μέθοδο Newton με αποκοπή*, Διπλωματική εργασία, Σχολή Μηχανολόγων Μηχανικών, Τομέας Ρευστών, Εργαστήριο Θερμικών Στροβιλομηχανών, Ε.Μ.Π., 2014.
 58. D. A. Pierre: *Optimization theory with applications*, Dover Publications, New York, 1986
 59. W. H. Marlow: *Mathematics for Operations Research*, Dover Publications, New York, 1993
 60. S. Kirkpatrick, J. C. D. Gelatt, and M. P. Vecchi: *Optimization by simulated annealing*, Science, 1983, Volume 220, Issue 4598, pp. 671–680.
 61. W. Spendley, G. R. Hext, and F. R. Himsworth: *Sequential application of simplex designs in optimization and evolutionary operation*, Technometrics, 1962, Volume 4, pp. 441–461.
 62. D. G. Luenberger: *Linear and nonlinear programming*, Kluwer Academic Publishers, Second edition, 2003.
 63. J. T. Richardson, M. R. Palmer, G. Liepins and M. Hilliard, *Some guidelines for genetic algorithms with penalty functions*, Proceedings of the Third International Conference on Genetic Algorithms, San Mateo, California, 1989.