

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

Τμήμα Μηχανολόγων Μηχανικών

Τομέας Ρευστών

Εργαστήριο Θερμικών Στροβιλομηχανών

Υπολογιστικό θέμα :

« Η βέλτιστη σχεδίαση περύγωσης στροβιλομηχανής και η δημιουργία σχετικού μεταπροτύπου»

Όνοματεπώνυμο	:	Αλέξανδρος Ζυμάρης
Κωδικός	:	mc-02100042
Ακαδημαϊκό έτος	:	2004-2005
Τμήμα	:	Μηχανολόγων Μηχανικών Ε.Μ.Π.
Υπεύθυνος Καθηγητής	:	Δρ. Κυριάκος Γιαννάκογλου

Γενική εισαγωγή.

Η παρούσα εργασία πραγματεύεται την κατασκευή μεταπροτύπων με χρήση τεχνικών της στατιστικής. Στόχος μας είναι να μπορέσουμε να εκτιμήσουμε την απόκριση ενός προβλήματος, στο οποίο εφεξής θα αναφέρουμε ως πρότυπο. Ένα τέτοιο πρότυπο πρόβλημα θα μπορούσε να είναι ένας υπολογιστικός κώδικας που μοντελοποιεί ένα φυσικό ή τεχνολογικό πρόβλημα (π.χ. που να μοντελοποιεί την ροή σε μια πτερύγωση ή που να υπολογίζει τις κατανομές τάσεων σε ένα μηχανολογικό εξάρτημα ή ένα ανθρώπινο οστό). Γενικότερα , πρότυπο πρόβλημα μπορεί να είναι μια οποιοδήποτε φυσική ή τεχνολογική διεργασία, π.χ. η εξάρτηση των μηχανικών ιδιοτήτων ενός υλικού από την σύσταση του μείγματος , αλλά και ένα υπολογιστικό μοντέλο μιας τέτοιας διεργασίας. Πρέπει να διευκρινιστεί ότι η παρούσα εργασία θα πραγματευθεί την κατασκευή μεταπροτύπων σε **υπολογιστικά μοντέλα** . Στο εκάστοτε πρόβλημα ξεχωρίζουμε δύο βασικές συνιστώσες :

- Τις μελετούμενες ιδιότητες, δηλαδή αυτό που πριν ονομάσαμε απόκριση, π.χ. αναφερόμενοι στο προηγούμενο παράδειγμα απόκριση θα μπορούσε να είναι η ελαστικότητα του υλικού, η αντοχή του σε κάμψη ή η αεροδυναμική συμπεριφορά μιας πτερύγωσης. Συχνά μας ενδιαφέρουν περισσότερες από μια ιδιότητες έτσι μελετάμε και περισσότερες από μια αποκρίσεις (εμείς στη παρούσα εργασία θα διαχειριστούμε προβλήματα με μια απόκριση. Στην περίπτωση αυτή κατασκευάζουμε περισσότερα μεταπρότυπα , όσα και οι προς μελέτη ιδιότητες.
- Τις παραμέτρους από τις οποίες εξαρτάται η απόκριση, οι παράμετροι αυτές αποτελούν τις ελεύθερες μεταβλητές του πρότυπου προβλήματος και επομένως και του μεταπροτύπου μας (συχνά μπορούμε να απλοποιήσουμε το πρόβλημα αγνοώντας κάποιες μεταβλητές – αυτό όμως πρέπει να γίνεται μόνο όταν αποδεδειγμένα οι μεταβλητές αυτές έχουν μικρή επίδραση στην μελετούμενη απόκριση). Ο καθορισμός των μεταβλητών είναι καθοριστικός για την σωστή κατασκευή του μεταπροτύπου γιατί μια λάθος επιλογή μπορεί να μας οδηγήσει στην αναζήτηση λύσης σε ένα πρόβλημα διαφορετικό από αυτό το οποίο θέλουμε να επιλύσουμε. Οι μεταβλητές μπορεί να είναι συνεχή μεγέθη ή διακριτά (εμείς θα ασχοληθούμε μόνο με συνεχή) και συχνά οι τιμές που μπορούν να λάβουν περιορίζονται είτε από υπαρκτούς περιορισμούς του προβλήματος είτε επειδή ο μηχανικός ενδιαφέρεται να μελετήσει την απόκριση εντός μιας συγκεκριμένης περιοχής τιμών. Τα όρια των μεταβλητών καθορίζουν το χωρίο μελέτης το οποίο αποτελεί την πειραματική μας περιοχή. Επισημαίνουμε ότι ο μεγάλος αριθμός μεταβλητών οδηγεί συχνά σε προβλήματα με αυξημένη πολυπλοκότητα .

Προκειμένου να κατασκευάσουμε το μεταπρότυπο θα πρέπει να αποκομίσουμε πληροφορία για το πώς εξαρτάται η απόκριση από τις μεταβλητές, αυτό επιτυγχάνεται με την διεξαγωγή πειραμάτων. Τα **πειράματα** διενεργούνται εντός της πειραματικής περιοχής και **πρόκειται για τις αποκρίσεις του πρότυπου προβλήματος σε καθορισμένες τιμές των μεταβλητών**. Οι τιμές αυτές των μεταβλητών ονομάζονται πειραματικά σημεία ή σημεία εκπαίδευσης του μεταπροτύπου και η επιλογή τους είναι σημαντική για την επιτυχία ή όχι του μεταπροτύπου. Η επιλογή των σημείων εκπαίδευσης χαρακτηρίζεται συχνά με τον όρο «σχεδιασμός περαμάτων» (δείτε σχετική παράγραφο). Επισημαίνουμε ότι κατά την χρήση του μεταπροτύπου δεν πρέπει σε καμία περίπτωση να «εγκαταλείπουμε» την πειραματική περιοχή.

Η κατασκευή μεταπροτύπων για υπολογιστικά μοντέλα αποσκοπεί κυρίως στην μερική ή πλήρη αντικατάσταση του πρότυπου υπολογιστικού κώδικα έτσι ώστε να εξοικονομηθεί χρόνος. Επομένως, αφορά πρότυπους κώδικες των οποίων οι απαιτήσεις σε χρόνο CPU είναι ιδιαίτερα υψηλές. Καθίσταται επομένως εμφανές ότι ένα μεταπρότυπο θα πρέπει να έχει ιδιαίτερα περιορισμένες απαιτήσεις σε υπολογιστικούς πόρους ειδάλλως δεν τίθεται θέμα αντικατάστασης του προτύπου. Επίσης, ένα τέτοιο μεταπρότυπο θα μπορούσε να χρησιμεύσει για να μοντελοποιήσει μαθηματικά μια φυσική ή τεχνολογική διεργασία για την οποία δεν υπάρχει γνωστό ή ικανό υπολογιστικό μοντέλο που να βασίζεται σε επαρκή φυσική θεώρηση του προβλήματος.

Μια ιδιαίτερα σημαντική εφαρμογή αυτών των μεταπροτύπων αφορά στην **βελτιστοποίηση**, όπου το μεταπρότυπο χρησιμοποιείται αντικαθιστώντας το πρότυπο πρόβλημα στην αναζήτηση για παράδειγμα, μιας αεροτομής «στόχου», δηλαδή μιας αεροτομής με καθορισμένη π.χ. γεωμετρία. Η χρησιμοποίηση μεταπροτύπου συμφέρει όταν το λογισμικό βελτιστοποίησης πρόκειται να κάνει πολλές κλίσεις της αντικειμενικής συνάρτησης και είναι χρονοβόρο να πάρουμε αυτές τις αποκρίσεις με το πρότυπο πρόβλημα (η αξιολόγηση των υποψήφιων βέλτιστων λύσεων με το μεταπρότυπο στοιχίζει συγκριτικά σε χρόνο πολύ λιγότερο από την αξιολόγηση με το πρότυπο).

Πριν την κατασκευή του μεταπροτύπου πρέπει να ληφθούν αποφάσεις οι οποίες σχετίζονται με :

- Την επιλογή των σημείων του χωρίου μελέτης τα οποία θα χρησιμοποιηθούν για να κατασκευαστεί το μεταπρότυπο (σημεία εκπαίδευσης). Δεδομένου ότι τα φυσικά και τεχνολογικά πειράματα διενεργούνται είτε υπολογιστικά (όταν το πρότυπο είναι υπολογιστικό μοντέλο) είτε με μετρήσεις (όταν μελετάμε απευθείας το πρόβλημα), τα σημεία εκπαίδευσης προκύπτουν από μια «επίπονη» διαδικασία, δαπανηρή χρονικά και χρηματικά. Επομένως, είναι εύλογο να επιθυμούμε μειωμένο αριθμό πειραματικών σημείων. Το «ιδανικό» θα ήταν να μειώναμε τον αριθμό των σημείων εκπαίδευσης χωρίς όμως να αυξήσουμε την ανακρίβεια του μεταπροτύπου. Από τα παραπάνω καθίσταται σαφές ότι πρόκειται για δύο αντικρουόμενους στόχους (μικρός αριθμός σημείων εκπαίδευσης – μικρό σφάλμα), σε κάθε περίπτωση ο μηχανικός είναι αυτός που θα πρέπει να επιτύχει μια ισορροπία μεταξύ των δύο αυτών παραμέτρων. Επίσης, επισημαίνουμε ότι στις εφαρμογές που θα ακολουθήσουν, ότι το πρότυπο πρόβλημα αντιμετωπίζεται σαν να ήταν υπολογιστικά χρονοβόρο παρότι για πρακτικούς λόγους τα μεταπρότυπα που χρησιμοποιήθηκαν δεν είναι απαιτητικά σε χρόνο CPU. Έτσι επιθυμούμε να χρησιμοποιούμε εύλογους αριθμούς σημείων εκπαίδευσης.
- Τη μαθηματική μορφή του μεταπροτύπου, όπου μια πολύπλοκη μορφή θα έχει και αυξημένες απαιτήσεις σε πειράματα, επομένως αν μπορούμε να εφαρμόσουμε λιγότερο πολύπλοκα μεταπρότυπα αυτό είναι σίγουρα επιθυμητό.

Μετά την κατασκευή του μεταπροτύπου ακολουθεί η διαδικασία της επιβεβαίωσης των αποκρίσεων που αυτό δίνει και η σύγκρισή τους με τις αντίστοιχες αποκρίσεις του προτύπου προβλήματος, επιθυμούμε οι δύο αποκρίσεις να είναι «κοντά». Πλέον, έγκειται στον μηχανικό να αποφασίσει κατά πόσο το μεταπρότυπο που κατασκεύασε πληροί τις απαιτήσεις του (να ορίσει το «κοντά»), ένα μεταπρότυπο που κρίνεται π.χ. από εμένα επαρκές για να χρησιμοποιηθεί στην δουλειά Α μπορεί να το κρίνω ανεπαρκές για την δουλειά Β,

ή κάποιος άλλος μηχανικός (με βάση τις απαιτήσεις που αυτός έχει θέσει ή του έχουν τεθεί) να το κρίνει ανεπαρκές και για την δουλειά Α. Η διαδικασία της επιβεβαίωσης είναι πολύ σημαντική και πρέπει να διενεργείται πριν το μεταπρότυπο χρησιμοποιηθεί έτσι ώστε να αποφύγουμε μελλοντικά εσφαλμένα αποτελέσματα. Για παράδειγμα, χρησιμοποιήσουμε μεταπρότυπο για την αντικειμενική συνάρτηση σε ένα πρόβλημα βελτιστοποίησης. Αν το πρόγραμμα βελτιστοποίησης στηρίζεται **αποκλειστικά** στο μεταπρότυπο (δουλειά Α) για την αξιολόγηση των υποψηφίων λύσεων η απόκριση του μεταπρότυπου οφείλει να είναι ιδιαίτερα «κοντά» στην πραγματική απόκριση, πολλές φορές μια απλά κοινή συμπεριφορά των δύο δεν αρκεί. Στην περίπτωση (δουλειά Β) όπου το λογισμικό βελτιστοποίησης εκτός από το μεταπρότυπο κάνει και ορισμένες αξιολογήσεις λύσεων με τον πρότυπο υπολογιστικό κώδικα μπορεί ένα λιγότερο ακριβές μεταπρότυπο να κρίνεται επαρκές.

Τέλος, αξίζει να επισημανθεί, ότι επειδή η παρούσα εργασία αποτελεί το πρώτο εγχείρημά μας στο αντικείμενο αυτό, και επομένως, προς το παρόν, δεν αποσκοπούμε στη βελτιστοποίηση της όλης διαδικασίας.

Εισαγωγή στην μεθοδολογία.

Προτού προβούμε στην παρουσίαση της μορφής του μεταπρότυπου την οποία θα χρησιμοποιήσουμε στην παρούσα εργασία, είναι χρήσιμο να αναφέρουμε ότι αυτό το μεταπρότυπο στηρίζεται στην θεώρηση ότι η προς μοντελοποίηση απόκριση ακολουθεί το μοντέλο των Γκαουσιανών Τυχαίων Συναρτήσεων (Gaussian Random Functions), πράγμα που σημαίνει ότι :

$$Y(x) = \mathbf{f}^T \boldsymbol{\beta} + Z(x) \quad (1)$$

όπου :

\mathbf{x} είναι το διάνυσμα των ελεύθερων μεταβλητών του προβλήματος, η διάσταση (d) του διανύσματος αυτού καθορίζει και τη διάσταση του προβλήματος,

Y είναι η απόκριση στο σημείο \mathbf{x} ,

$\boldsymbol{\beta}$ είναι το διάνυσμα αγνώστων συντελεστών $\boldsymbol{\beta} = [\beta_1, \beta_2, \dots, \beta_p]^T$ (p είναι ο αριθμός των συναρτήσεων αναδρομής – δείτε παρακάτω),

$Z(x)$ μια γκαουσιανή διεργασία με μέση τιμή μηδέν και μεταβλητότητα σ_z^2 ,

και το διάνυσμα \mathbf{f} [$p \times 1$] αποτελεί το διάνυσμα γνωστών συναρτήσεων αναδρομής στο σημείο \mathbf{x} , οι οποίες είναι p στον αριθμό. Στις συναρτήσεις θα επανέλθουμε σύντομα και πιο αναλυτικά.

Άμεση συνέπεια της θεώρησης (1) είναι ότι οι αποκρίσεις, σε ένα πρόβλημα με διάσταση d, με n σημεία εκπαίδευσης του μεταπρότυπου και p συναρτήσεις αναδρομής, θα ακολουθούν πολυμεταβλητή κανονική κατανομή :

$$\begin{pmatrix} Y_0 \\ \mathbf{Y}^n \end{pmatrix} \sim N_{1+n} \left[\begin{pmatrix} \mathbf{f}_0^T \\ \mathbf{F} \end{pmatrix} \boldsymbol{\beta}, \sigma_z^2 \begin{pmatrix} 1 & \mathbf{r}_0^T \\ \mathbf{r}_0 & \mathbf{R} \end{pmatrix} \right] \quad (2)$$

όπου : \mathbf{x}_0 το σημείο στο οποίο ζητάμε την απόκριση Y_0 , \mathbf{f}_0 [$p \times 1$] το διάνυσμα των συναρτήσεων αναδρομής στο σημείο \mathbf{x}_0 , οι οποίες είναι p στον αριθμό, $\mathbf{F}_{i,j}$ [$n \times p$] ο πίνακας που εμπεριέχει ανά γραμμή την τιμή των p συναρτήσεων αναδρομής σε

κάθε ένα από τα n σημεία εκπαίδευσης, ο πίνακας αυτός πρέπει να είναι ομαλός. Το διάνυσμα \mathbf{Y}^n περιέχει όλες τις γνωστές αποκρίσεις στα n σημεία εκπαίδευσης, $\mathbf{Y}^n = [Y(x_1), Y(x_2), \dots, Y(x_n)]^T$. Το διάνυσμα \mathbf{r}_0 είναι το διάνυσμα της συσχέτισης (correlation vector) στην θέση \mathbf{x}_0 , δηλαδή :

$$\mathbf{r}_0 = [R(x_0 - x_1), R(x_0 - x_2), \dots, R(x_0 - x_n)]^T$$

όπου $R(h)$ η συνάρτηση συσχέτισης. Ο πίνακας $\mathbf{R}_{i,j} = [R(x_i - x_j)]_{n,n}$ είναι ο πίνακας των συσχετίσεων (correlation matrix) που κάθε του γραμμή i αποτελεί το εκάστοτε διάνυσμα της συσχέτισης της μεταβλητής i , δηλαδή το $\mathbf{r}_i = [R(x_i - x_1), R(x_i - x_2), \dots, R(x_i - x_n)]_{1,n}$, $\mathbf{R}_{i,j} = [\mathbf{r}_i]_{n,n}$, είναι εμφανές ότι είναι συμμετρικός και πρέπει να είναι ομαλός και θετικά ημιορισμένος. Το διάνυσμα β είναι το διάνυσμα των **αγνώστων** συντελεστών $[\beta]_{1,p}$.

Στην παρούσα εργασία επιλέγουμε συνάρτηση συσχέτισης $R(h)$ που δίνεται από την σχέση :

$$R(h) = \exp \left\{ - \sum_{j=1}^d \left(\frac{h_j}{\theta} \right)^2 \right\} \quad \text{όπου } d \text{ η διάσταση του προβλήματος και}$$

$$h_j = (x_i - x_j), \quad i = 1, \dots, n.$$

Σε μια γενικότερη περίπτωση μπορούμε να χρησιμοποιήσουμε την παρακάτω μορφή για την R :

$$R(h) = \exp \left\{ - \sum_{j=1}^d \left| \frac{h_j}{\theta_j} \right|^{p_j} \right\}$$

Επίσης με μια αναδρομή στην σχετική βιβλιογραφία μπορούμε να βρούμε και άλλες μορφές συναρτήσεων $R(h)$ – στο τέλος της εργασίας μας παραθέτουμε μια την περιγραφή της οικογένειας *Matérn* τέτοιων συναρτήσεων – . Τα μεγέθη $\psi = [\theta_1, \theta_2, \dots, \theta_d, p_1, p_2, \dots, p_d]^T$ στον παρονομαστή και στον εκθέτη, μπορούν να υπολογιστούν με χρήση του κριτηρίου της μέγιστης πιθανοφάνειας (δείτε την σχετική παράγραφο).

Στην σχέση (2) δεν γνωρίζουμε ούτε τους συντελεστές β ούτε την σ_z^2 , έτσι η κατανομή (2) είναι υπό συνθήκη κατανομή με παραμέτρους τα β και σ_z^2 .

Το μεταπρότυπο και οι σχετικές με αυτό αποφάσεις του μηχανικού.

Προκειμένου να προβούμε σε εκτίμηση της απόκρισης $Y(\mathbf{x}_0)$ χρησιμοποιούμε την παρακάτω μαθηματική έκφραση :

$$\hat{Y}_0 = \mathbf{f}_0^T \hat{\beta} + \mathbf{r}_0^T \mathbf{R}^{-1} (\mathbf{Y}^n - \mathbf{F} \hat{\beta}) \quad (3)$$

Όπου η διάσταση d του προβλήματος είναι καθορισμένη, ο αριθμός των σημείων εκπαίδευσης n καθορίζεται από τον μηχανικό με βάση τις απαιτήσεις του προβλήματος, την «δουλεία» όπου θα χρησιμοποιήσει το μεταπρότυπο και πρωτίστως από το κόστος σε CPU που είναι διατεθειμένος να καταβάλει για να πάρει τις αποκρίσεις αυτές. Επίσης οι συναρτήσεις αναδρομής αλλά και ο αριθμός p αυτών καθορίζονται από τον

μηχανικό, συνήθως με βάση την εμπειρία του γύρω από το προς μελέτη πρόβλημα. Επιπροσθέτως, όπως θα δούμε στην παρακάτω παράγραφο, στην πράξη ο μηχανικός αποφασίζει και για τα μητρώα \mathbf{R} και \mathbf{r}_0 , εκλέγοντας την μορφή της συνάρτησης αναδρομής. Δηλαδή, στην έκφραση (3) οι άγνωστοι που θα υπολογιστούν είναι τα $\hat{\beta}$.

Στην περίπτωση που ο πίνακας της συσχέτισης είναι γνωστός αποδεικνύεται ότι η έκφραση (3) αποτελεί, όπως λέμε, την Βέλτιστη Γραμμική Ανεπηρέαστη Εκτίμηση (Best Linear Unbiased Predictor) του $Y(\mathbf{x}_0)$. Δηλαδή είναι το μεταπρότυπο εκείνο το οποίο μας δίνει το ελάχιστο μέσο τετραγωνικό σφάλμα (από το σύνολο των μεταπροτύπων εκείνων όπου ισχύει γραμμική εξάρτηση μεταξύ \hat{Y}_0 και \mathbf{Y}^n και των οποίων η εκτίμηση είναι ανεπηρέαστη). Στην πραγματικότητα όμως δεν γνωρίζω τον πίνακα της συσχέτισης, έτσι όπως φάνηκε και από τα παραπάνω υποθέτουμε ότι η

συνάρτηση συσχέτισης έχει μια συγκεκριμένη μορφή π.χ. $R(h) = \exp\left\{-\sum_{j=1}^d \left|\frac{h_j}{\theta_j}\right|^{p_j}\right\}$,

αυτό σίγουρα αποτελεί προσέγγιση. Στην περίπτωση αυτή το μεταπρότυπό μας εξακολουθεί να δίνεται από την σχέση (3) όμως πλέον κανείς δεν μας εξασφαλίζει την γραμμικότητα μεταξύ \hat{Y}_0 και \mathbf{Y}^n αλλά και το ότι η εκτίμηση είναι ανεπηρέαστη.

Στην πράξη οι εκφράσεις $\hat{\mathbf{R}} = \hat{\mathbf{R}}(\mathbf{Y}^n)$ και $\hat{\mathbf{r}}_0 = \hat{\mathbf{r}}_0(\mathbf{Y}^n)$ είναι μη γραμμικές. Για αυτό λέμε ότι η έκφραση (3) αποτελεί πλέον την Εμπειρική Βέλτιστη Γραμμική Ανεπηρέαστη Εκτίμηση (Empirical Best Linear Unbiased Predictor) του $Y(\mathbf{x}_0)$. Η έκφραση (3) γράφεται ως εξής (για να δηλώσουμε ότι χρησιμοποιούμε εκτίμηση του πίνακα συσχέτισης):

$$\hat{Y}_0 = \mathbf{f}_0^T \hat{\beta} + \hat{\mathbf{r}}_0^T \hat{\mathbf{R}}^{-1} (\mathbf{Y}^n - \mathbf{F} \hat{\beta}) \quad (3a)$$

(στα παρακάτω επειδή πάντα θα αναφερόμαστε σε $\hat{\mathbf{R}}$ και $\hat{\mathbf{r}}_0$ για απλότητα θα παραλείπουμε τα «καπέλα» και θα γράφουμε \mathbf{R} και \mathbf{r}_0)

Οι συναρτήσεις αναδρομής $\mathbf{f}(\mathbf{x})$ εκλέγονται από εμάς και είναι συνήθως πολυωνυμικής μορφής.

π.χ. Έστω ότι το πρόβλημά μου είναι τριδιάστατο ($d=3$, δηλαδή έχω τρεις μεταβλητές x_1, x_2, x_3), θα έχουμε ότι:

$\mathbf{x} = [x_1, x_2, x_3]^T$ μια επιλογή συναρτήσεων αναδρομής θα μπορούσε να ήταν η εξής:

$$\mathbf{f}(\mathbf{x}) = [1, x_1, x_2, x_3, x_1^2, x_2^2, x_3^2, x_1 x_2, x_1 x_3, x_2 x_3]^T$$

Στην παρούσα εργασία χρησιμοποιούνται πολυώνυμα 2^{ου} ή 3^{ου} βαθμού καθώς επίσης και εκθετικές συναρτήσεις.

Προσπαθώντας να δώσουμε μια απλή ερμηνεία στην σχέση (3) θα μπορούσαμε να πούμε ότι ο όρος $\mathbf{r}_0^T \mathbf{R}^{-1} (\mathbf{Y}^n - \mathbf{F} \hat{\beta})$ αποτελεί «διόρθωση» της πρόβλεψης που μας δίνει το μοντέλο $\mathbf{f}_0^T \hat{\beta}$. Επίσης σημασία έχει να μην ξεχνάμε ότι η έκφραση (3) αποτελεί συνάρτηση τόσο του σημείου \mathbf{x}_0 στο οποίο γίνεται η πρόβλεψη όσο και της απόκρισης Y_0 των n σημείων εκπαίδευσης.

Κατασκευή του μεταπροτύπου όταν θεωρούμε σ_z^2 και \mathbf{R} (πίνακας συσχέτισης) γνωστά.

Μέχρι τώρα δεν έγινε λόγος για το πώς υπολογίζονται οι άγνωστες παράμετροι β και σ_z^2 . Αρχικά θα ασχοληθούμε μόνο με τον υπολογισμό των άγνωστων συντελεστών β - ας θεωρήσουμε ότι η μεταβλητότητα σ_z^2 είναι γνωστή - εξάλλου η μεταβλητότητα σ_z^2 δεν υπεισέρχεται στην σχέση (3) θα μας χρησιμεύσει βέβαια στην συνέχεια όπου θα επιχειρήσουμε να υπολογίσουμε την μεταβλητότητα της απόκρισης \hat{Y}_0 , την $\sigma_{0|n}^2$. (Στην περίπτωση αυτή η Y_0 ακολουθεί την κατανομή (2).)

Μπορούμε να θεωρήσουμε ότι οι συντελεστές β ακολουθούν πολυδιάστατη κανονική κατανομή γύρω από μια μέση τιμή \mathbf{b}_0 (που αποτελεί μια εκ των πρότερων δοσμένη ποσότητα) με γνωστή συνδιακύμανση $\tau^2 \mathbf{V}_0$ (\mathbf{V}_0 : *correlation matrix*, τ^2 : μεταβλητότητα κατανομής) :

$$\beta \sim N_p[\mathbf{b}_0, \tau^2 \mathbf{V}_0] \quad (3)$$

Στην γενικότερη περίπτωση όπου το τ^2 είναι πεπερασμένο, λέμε ότι έχουμε *informative* ή *normal prior*, δηλαδή είμαστε στην περίπτωση όπου οι άγνωστοι συντελεστές β ακολουθούν κανονική κατανομή γύρω από την *prior μέση τιμή* \mathbf{b}_0 . Στην περίπτωση αυτή υπολογίζουμε την λεγόμενη *posterior μέση τιμή* :

$$\boldsymbol{\mu}_{\beta|n} = \left(\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F} + \frac{\sigma_z^2 \mathbf{V}_0^{-1}}{\tau^2} \right)^{-1} \left(\mathbf{F}^T \mathbf{R}^{-1} \mathbf{Y}^n + \frac{\sigma_z^2 \mathbf{V}_0^{-1}}{\tau^2} \mathbf{b}_0 \right) \mathbf{R}^{-1} \quad (4a)$$

όποτε μπορούμε να χρησιμοποιήσουμε στην έκφραση (3) την παραπάνω προσέγγιση (4a) για τους άγνωστους συντελεστές β , έχουμε λοιπόν ότι :

$$\hat{Y}_0 = \mathbf{f}_0^T \boldsymbol{\mu}_{\beta|n} + \mathbf{r}_0^T \mathbf{R}^{-1} (\mathbf{Y}^n - \mathbf{F} \cdot \boldsymbol{\mu}_{\beta|n}) \quad (3a)$$

Η μεταβλητότητα της κατανομής (2), για την περίπτωση αυτή, δίνεται από την παρακάτω έκφραση (θεωρούμε ότι η μεταβλητότητα σ_z^2 είναι γνωστή) :

$$\sigma_{0|n}^2 = \sigma_z^2 \left[1 - \mathbf{r}_0^T \mathbf{R}^{-1} \mathbf{r}_0 + (\mathbf{f}_0 - \mathbf{F}^T \mathbf{R}^{-1} \mathbf{r}_0)^T (\mathbf{Q}^T \mathbf{R}^{-1} \mathbf{Q})^{-1} (\mathbf{f}_0 - \mathbf{F}^T \mathbf{R}^{-1} \mathbf{r}_0) \right] \quad (5a)$$

$$\text{όπου: } \mathbf{Q} = \mathbf{F}^T \mathbf{R}^{-1} \mathbf{F} + \frac{\sigma_z^2 \mathbf{V}_0^{-1}}{\tau^2}$$

Θεωρώντας τώρα , ότι $\tau^2 \rightarrow \infty$ (όποτε λέμε ότι έχουμε *non-informative prior*), από την σχέση (4a) προκύπτει ότι :

$$\hat{\beta} = (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{R}^{-1} \mathbf{Y}^n \quad (4b)$$

αντιστοίχως από την σχέση (5a) προκύπτει ότι :

$$\sigma_{0|n}^2 = \sigma_z^2 \left[1 - \mathbf{r}_0^T \mathbf{R}^{-1} \mathbf{r}_0 + (\mathbf{f}_0 - \mathbf{F}^T \mathbf{R}^{-1} \mathbf{r}_0)^T (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} (\mathbf{f}_0 - \mathbf{F}^T \mathbf{R}^{-1} \mathbf{r}_0) \right] \quad (5b)$$

Η παραπάνω σχέση (4b) αποδεικνύεται ότι αποτελεί μέση τιμή μιας υπό συνθήκη πολυμεταβλητής κανονικής κατανομής $[\beta | \mathbf{Y}^n]$, την οποία ακολουθούν οι συντελεστές β . Για την $[\beta | \mathbf{Y}^n]$ ισχύει ότι :

$$[\beta | \mathbf{Y}^n] \sim N_p \left[\hat{\beta}, \sigma_z^2 (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} \right]$$

η έκφραση (4b) για τους συντελεστές $\hat{\beta}$ συχνά αναφέρεται ως μέθοδος των γενικευμένων ελάχιστων τετραγώνων, τα δε $\hat{\beta}$ αποτελούν **MLE** (**M**aximum **L**ikelihood **E**stimator) των β . Έτσι μπορούμε στην σχέση (3) να χρησιμοποιήσουμε την μέση τιμή $\hat{\beta}$ ως την καλύτερη δυνατή προσέγγιση για τα άγνωστα β . Όποτε έχουμε ότι :

$$\hat{Y}_0 = \mathbf{f}_0^T \hat{\beta} + \mathbf{r}_0^T \mathbf{R}^{-1} (\mathbf{Y}^n - \mathbf{F} \hat{\beta}) \quad (3b)$$

Με μεταβλητότητα ίση προς :

$$\sigma_{0|n}^2 = \sigma_z^2 \left[1 - \mathbf{r}_0^T \mathbf{R}^{-1} \mathbf{r}_0 + (\mathbf{f}_0 - \mathbf{F}^T \mathbf{R}^{-1} \mathbf{r}_0)^T (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} (\mathbf{f}_0 - \mathbf{F}^T \mathbf{R}^{-1} \mathbf{r}_0) \right] \quad (5b)$$

Σε όλη την μελέτη που ακολουθεί γίνεται χρήση της *non-informative* υπόθεσης για τα β και επομένως χρησιμοποιούμε την σχέση (1b) για την μέση τιμή της απόκρισης και την έκφραση (5b) για την μεταβλητότητα αυτής. Η έκφραση αυτή αποτελεί ουσιαστικά το μεταμοντέλο μας.

Η περίπτωση όπου είναι άγνωστη και η μεταβλητότητα σ_z^2 (γνωστός ο πίνακας \mathbf{R}).

Η πιο γενική αντιμετώπιση του προβλήματος επιβάλλει η σ_z^2 να είναι και αυτή **άγνωστη** (όπως και τα β), στην περίπτωση αυτή η απόκριση ακολουθεί κατανομή Student με ν_i βαθμούς ελευθερίας :

$$[Y_0 | \mathbf{Y}^n] \sim T(\nu_i, \hat{Y}_0, \sigma_{0|n}^2)$$

Υποθέτοντας ότι η σ_z^2 ακολουθεί κατανομή :

- είτε της μορφής : $[\sigma_z^2] = \frac{c_0}{\chi_{\nu_0}^2}$, c_0 μια σταθερά, $\chi_{\nu_0}^2$ μεταβλητή που ακολουθεί κατανομή χι-τετράγωνο.
- είτε της μορφής : $[\sigma_z^2] = \frac{1}{\sigma_z^2}$

και για την non-informative υπόθεση σε ότι αφορά τα β (με $\tau^2 \equiv \sigma_z^2$), καταλήγουμε στην εκτίμηση της σ_z^2 από την έκφραση:

$$\sigma_z^2 = \frac{1}{n-p} \mathbf{Y}^{nT} \mathbf{R}^{-1} (\mathbf{Y}^n - \mathbf{F} \hat{\beta}) \quad (8a) \quad , \text{ για όταν ισχύει } [\sigma_z^2] = \frac{1}{\sigma_z^2}$$

$$\text{ή } \sigma_z^2 = c_0 + \frac{1}{n-p} \mathbf{Y}^{nT} \mathbf{R}^{-1} (\mathbf{Y}^n - \mathbf{F} \hat{\beta}) \quad (8b) \quad , \text{ για όταν ισχύει } [\sigma_z^2] = \frac{c_0}{\chi_{\nu_0}^2} .$$

(Δεν μελετάμε την περίπτωση της informative υπόθεσης για τα β .)

Το κριτήριο της μέγιστης πιθανοφάνειας.

Μέχρι τώρα θεωρήσαμε «σιωπηλά» ότι η συνάρτηση συσχέτισης είναι γνωστή ή ότι είναι γνωστός ο πίνακας \mathbf{R} . Όμως αυτό συχνά δεν ισχύει, έτσι με την εφαρμογή του κριτηρίου της μέγιστης πιθανοφάνειας μπορούμε να υπολογίσουμε τις παραμέτρους της συνάρτησης συσχέτισεως στην περίπτωση που αυτή δεν είναι πλήρως γνωστή. Έστω ότι ενδιαφερόμαστε να κατασκευάσουμε το παρακάτω μεταπρότυπο :

$$\hat{Y}_0 = \mathbf{f}_0^T \hat{\beta} + \mathbf{r}_0^T \mathbf{R}^{-1} (\mathbf{Y}^n - \mathbf{F} \hat{\beta}) \quad \text{με} \quad \hat{\beta} = (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{R}^{-1} \mathbf{Y}^n$$

η απόκριση του οποίου ακολουθεί την **πολυδιάστατη κανονική κατανομή (2)**, επίσης θεωρούμε ότι δεν γνωρίζουμε σ_z^2 και \mathbf{R} . Για τον προσδιορισμό του μητρώου \mathbf{R} λαμβάνουμε συνάρτηση συσχέτισεως με την παρακάτω μορφή :

$$R(h, \psi) = \exp \left\{ - \sum_{j=1}^d \left| \frac{h_j}{\theta_j} \right|^{p_j} \right\}$$

με άγνωστες παραμέτρους τις : θ_j, p_j , όπου: $j = 1, \dots, d$ ή όπως θα γράφουμε πιο συνοπτικά : $\psi = [\theta_1, \theta_2, \dots, \theta_d, p_1, p_2, \dots, p_d]^T$.

Το εν λόγω κριτήριο αφορά την εύρεση του διανύσματος ψ το οποίο **ελαχιστοποιεί** την παρακάτω έκφραση :

$$\Omega(\psi) = n \log \hat{\sigma}_z^2(\psi) + \log(\det(\mathbf{R}(\psi))) \quad (7)$$

Στην παραπάνω έκφραση η μεταβλητότητα σ_z^2 προσεγγίζεται από την σχέση :

$$\hat{\sigma}_z^2 = \frac{1}{n} (\mathbf{Y}^n - \mathbf{F} \hat{\beta})^T \mathbf{R}^{-1} (\mathbf{Y}^n - \mathbf{F} \hat{\beta}) \quad (6)$$

Η ποσότητα αυτή χρησιμοποιείται αντί της σ_z^2 στην σχέση (7), και υπολογίζεται αφού έχουμε ήδη υπολογίσει από την (4b) τα $\hat{\beta}$. Το διάνυσμα ψ όπως υπολογίζεται από την ελαχιστοποίηση της (7) και την (6) είναι ο **MLE (Maximum Likelihood Estimator)**.

Η διαδικασία επομένως, έχει ως εξής :

- Εκτελούμε την ελαχιστοποίηση της αντικειμενικής συνάρτησης (7) με παράμετρο το διάνυσμα ψ με βάση τα σημεία εκπαίδευσης που διαθέτουμε.
- Αφού υπολογίσουμε το βέλτιστο ψ_{opt} ώστε $\Omega_{min}(\psi_{opt})$ **θεωρούμε γνωστή την συνάρτηση συσχέτισης** και χρησιμοποιούμε την παραπάνω θεωρία για τον υπολογισμό της εκτίμησης \hat{Y}_0 και της μεταβλητότητας της σ_{0n}^2 .

Αντί της έκφρασης (6) μπορούμε να χρησιμοποιήσουμε την :

$$\tilde{\sigma}_z^2 = \frac{1}{n-p} (\mathbf{Y}^n - \mathbf{F}\hat{\beta})^T \mathbf{R}^{-1} (\mathbf{Y}^n - \mathbf{F}\hat{\beta}) \quad (6^a)$$

όταν αντί της (6) χρησιμοποιούμε την (6^a), το διάνυσμα ψ είναι ο **REML** (**R**estricted **M**aximum **L**ikelihood **E**stimator). Οπότε λέμε ότι εφαρμόζουμε το κριτήριο της Περιορισμένης Μέγιστης Πιθανοφάνειας.

Η περίπτωση όπου είναι άγνωστη η μεταβλητότητα σ_z^2 και ο πίνακας \mathbf{R} .

Η περίπτωση αυτή αποτελεί την πλέον γενική και είναι αυτή η οποία έχει πρακτικό ενδιαφέρον και σκόπιμα πριν αναφερθεί αναπτύξαμε τόσο την περίπτωση με γνωστό \mathbf{R} όσο και το κριτήριο της μέγιστης πιθανοφάνειας. Η αντιμετώπιση του προβλήματος θα γίνει με τον πλέον απλό τρόπο, που χρησιμοποιεί το κριτήριο της μέγιστης πιθανοφάνειας για τον υπολογισμό του διανύσματος ψ . Από την στιγμή που καθορίζεται η συνάρτηση συσχέτισης είναι πια υπολογίσιμο το μητρώο \mathbf{R} , και από εδώ και πέρα εφαρμόζουμε τα όσα αναφέρθηκαν για την περίπτωση όπου άγνωστη είναι μόνο η μεταβλητότητα σ_z^2 (και βέβαια τα β), δηλαδή τις σχέσεις (3b) και (5b) καθώς και την (8a) (ή την (8b)). Στο σημείο αυτό οφείλουμε να επισημάνουμε ότι με την μέθοδο αυτή προβαίνουμε στον προσδιορισμό μιας εκτίμησης του διανύσματος $\hat{\psi}$ το οποίο οδηγεί στην εισαγωγή σφάλματος. Γενικά ισχύει ότι :

$$\sigma_{0n}^2(\mathbf{x}_0 | \hat{\psi}) \leq \sigma_{0n}^2(\mathbf{x}_0 | \psi)$$

δηλαδή προβαίνουμε σε υποεκτίμηση του σφάλματος. Η υποεκτίμηση αυτή, είναι συχνά υψηλή όταν η συσχέτιση είναι μικρή, η μεθοδολογία όμως χρησιμοποιείται γιατί σε πολλές περιπτώσεις η υποεκτίμηση είναι αρκετά μικρή. (Υπενθυμίζουμε ότι το **MLE** και το **REML** στηρίζονται στην θεώρηση της πολυδιάστατης κανονικής κατανομής για το $[Y_0 | \mathbf{Y}^n]$ και όχι της κατανομής Student.)

Για λόγους πληρότητας αναφέρουμε ότι ο πιο σωστός τρόπος επίλυσης του προβλήματος είναι η Bayesian αντιμετώπιση (δείτε βιβλιογραφία), με την οποία όμως δεν θα ασχοληθούμε.

Σχεδιασμός Πειραμάτων (Design of Experiments -DoE).

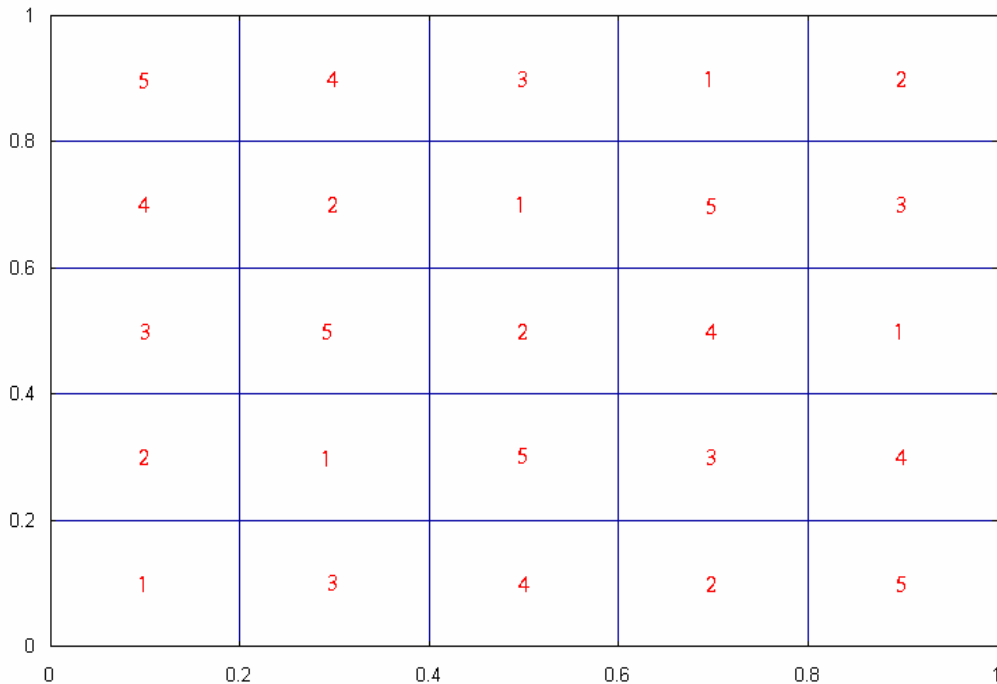
Όπως αναφέραμε και στην αρχή, για την κατασκευή ενός μεταπρότυπου πρέπει εκτός από την μορφή αυτού να αποφασίσουμε και για το πώς θα επιλέξουμε τα σημεία του χωρίου με τα οποία θα εκπαιδύσουμε το μεταπρότυπό μας. Μέχρι τώρα δεν ασχοληθήκαμε καθόλου με αυτό το ζήτημα. Στην συνέχεια όμως, το ζήτημα αυτό θα μας απασχολήσει αρκετά και η επιλογή των σημείων θα γίνεται με πιο εξειδικευμένες μεθόδους.

Πριν προχωρήσουμε στην περιγραφή της μεθόδου των Λατινικών Τετραγώνων (Latin Squares), που ασφαλώς είναι μια από τις πολλές τεχνικές που υπάρχουν, πρέπει να επιστημόνουμε δύο βασικά στοιχεία σχετικά με το σχεδιασμό πειραμάτων. Καταρχάς, κατά την εκτέλεση πειραμάτων στον **υπολογιστή** στην τιμή της απόκρισης του προς μοντελοποίηση προτύπου δεν υπεισέρχονται σφάλματα μέτρησης (συστηματικά ή τυχαία). Επομένως, όσες φορές και να εκτελέσουμε ένα πείραμα, για τα ίδια σημεία εκπαίδευσης, οι αποκρίσεις στα σημεία αυτά δεν θα αλλάζουν. Έτσι δεν χρειάζεται να εισάγουμε σημεία εκπαίδευσης τα οποία να επαναλαμβάνονται. Επιπροσθέτως, οι επιλογή των σημείων εκπαίδευσης πρέπει να μας επιτρέπει να «εξερευνούμε» όλο το χωρίο, και κατά το δυνατόν να μην αφήνουμε περιοχές για τις οποίες υπάρχει ανεπαρκής πληροφόρηση. Στην περίπτωση που δεν έχουμε κάποια εκ των πρότερων γνώση της επιφάνειας – έτσι ώστε π.χ. να επικεντρωθούμε σε κάποια «δύσκολη» περιοχή αυτής – είναι λογικό να επιδιώξουμε τα σημεία εκπαίδευσης να είναι ομοιόμορφα κατανεμημένα στο χωρίο μας (space filling designs). Τέλος, πρέπει ο αριθμός των σημείων εκπαίδευσης να είναι επαρκής (μεγαλύτερος από τον αριθμό των άγνωστων συντελεστών β) έτσι ώστε να περιορίζουμε την μεταβλητότητα του μεταμοντέλου και επίσης, να μπορούμε να κατασκευάζουμε περισσότερα του ενός μεταμοντέλα (για διασταύρωση των αποτελεσμάτων μας).

Η μέθοδος των Λατινικών Τετραγώνων.

Ας επανέλθουμε στην επιλογή σημείων μέσω ενός ορθογωνικού πλέγματος, η μέθοδος αυτή μας εξασφαλίζει ότι εξερευνούμε όλο το χωρίο και ότι τα σημεία εκπαίδευσης είναι «σχετικά» ομοιόμορφα διανεμημένα σε αυτό. Όμως η μέθοδος δεν ενδείκνυται όταν η διάσταση του προβλήματος είναι μεγάλη μιας και ο αριθμός των πειραμάτων που πρέπει να διεξαχθούν γίνεται πλέον υπερβολικά μεγάλος (αν $d=4$ ένα πλέγμα με 9 διαμερίσεις ανά μεταβλητή οδηγεί σε 10^4 σημεία !!!). Η μέθοδος των Λατινικών Τετραγώνων μας επιτρέπει να σχεδιάζουμε πειράματα με «σχετικά» ομοιόμορφα κατανεμημένα σημεία εκπαίδευσης χωρίς όμως να φτάνουμε σε εξωπραγματικούς αριθμούς σημείων.

Παρακάτω θα περιγράψουμε την εφαρμογή της μεθόδου επιλέγοντας n σημεία στο αδιάστατο χωρίο $[0,1] \times [0,1]$. Καταρχάς χωρίζουμε κάθε άξονα σε n τμήματα ίσου μήκους, έτσι χωρίζουμε το χωρίο σε n^2 υποχωρία (κελιά). Τώρα σε κάθε ένα από τα κελιά της πρώτης στήλης αντιστοιχίζουμε έναν ακέραιο από το $1, 2, \dots, n$ χωρίς να υπάρχουν κελιά με κοινή αρίθμηση. Αν επαναλαμβάνουμε την ίδια διαδικασία και για τα κελιά των επόμενων στηλών προσέχοντας όμως ο κάθε αριθμός να εμφανίζεται μόνο μια φορά σε κάθε γραμμή αποκτάμε, όπως λέμε, ένα λατινικό τετράγωνο. Για παράδειγμα, αν επιζητούμε 5 σημεία στο χωρίο $[0,1] \times [0,1]$ δημιουργούμε το παρακάτω λατινικό τετράγωνο της επόμενης σελίδας. Στην συνέχεια, επιλέγουμε στην τύχη ένα ακέραιο από το 1 ως και το n και σε κάθε ένα κελί του λατινικού τετραγώνου το οποίο έχει την αρίθμηση αυτή επιλέγουμε στην τύχη ένα σημείο (μέσα από το κελί). Έτσι αποκτάμε τα n σημεία εκπαίδευσης τα οποία επιζητούσαμε. Εύκολα μπορούμε να φανταστούμε την γενίκευση των παραπάνω στον χώρο των d διαστάσεων.



Σχήμα 1: Λατινικό τετράγωνο για επιλογή 5 σημείων στο χωρίο $[0:1]^2$.

Επιλογή σημείων εκπαίδευσης για προσθήκη στα ήδη υπάρχοντα.

Παρακάτω θα περιγράψουμε μια διαδικασία **σχεδιασμού πειραμάτων** στηριζόμενη σε κάποιο ήδη προϋπάρχον πείραμα. Πρόκειται για «εμπλουτισμό» των ήδη προϋπαρχόντων σημείων εκπαίδευσης με νέα, τα οποία επιλέγονται κάνοντας χρήση της μεταβλητότητας της απόκρισης. Παρότι στην παρούσα εργασία δεν προβαίνουμε στην εφαρμογή αυτής της μεθοδολογίας κρίνουμε ότι παρουσιάζει ενδιαφέρον και για αυτό την καταγράφουμε.

Καταρχάς στόχος μας είναι ξεκινώντας από ένα μικρό αριθμό σημείων εκπαίδευσης να φτάσουμε σε επιθυμητό σφάλμα του μεταπροτύπου προσθέτοντας σταδιακά νέα σημεία εκπαίδευσης. Έτσι εξοικονομούμε σημεία εκπαίδευσης, γεγονός που είναι σημαντικό όταν η εκτέλεση των πειραμάτων είναι απαιτητική σε χρόνο CPU. Η διαδικασία είναι απλή :

1. Επιλογή λίγων σημείων εκπαίδευσης με κάποια σχετική μέθοδο (π.χ. λατινικά τετράγωνα).
2. Κατασκευή μεταπροτύπου και υπολογισμός της μεταβλητότητας της εκτίμησης (ποσότητα που εκφράζει το σφάλμα της απόκρισης του μεταπροτύπου). Στο στάδιο αυτό περιλαμβάνεται ο προσδιορισμός των παραμέτρων θ με χρήση του κριτηρίου της μέγιστης πιθανοφάνειας.
3. Αν το υπάρχον μεταπρότυπο μας ικανοποιεί σταματάμε – ίσως θα έπρεπε να είχαμε επιλέξει λιγότερα σημεία εκπαίδευσης στο πρώτο βήμα ! Η επιλογή αυτή απαιτεί μια σχετική εμπειρία από τον μηχανικό –
4. Αν το υπάρχον μεταπρότυπο δεν μας ικανοποιεί, τότε εντοπίζουμε το σημείο του χωρίου μελέτης με την μέγιστη μεταβλητότητα ή πιο απλά το σημείο

εκείνο με μέγιστο $\frac{\sigma_{0in}^2}{\sigma_z^2}$ (αφού η σ_z είναι μια σταθερή τιμή για το εκάστοτε μεταμοντέλο με δεδομένα σημεία εκπαίδευσης και συγκεκριμένα θ). Η διαδικασία αυτή απαιτεί την χρήση κάποιου αλγόριθμου ανίχνευσης. Επισημαίνουμε ότι για τον υπολογισμό του λόγου $\frac{\sigma_{0in}^2}{\sigma_z^2}$ δεν απαιτείται να γνωρίζουμε την απόκριση του προτύπου στα σημεία εκπαίδευσης.

5. Αφού εντοπιστεί το σημείο του χωρίου όπου έχουμε $\max\left\{\frac{\sigma_{0in}^2}{\sigma_z^2}\right\}$ εκκινούμε μια άλλη διαδικασία – ελαχιστοποίησης αυτή την φορά – κατά την οποία εισάγουμε υπονήφια σημεία εκπαίδευσης με στόχο να επιλέξουμε εκείνα τα οποία οδηγούν σε ελαχιστοποίηση του λόγου $\frac{\sigma_{0in}^2}{\sigma_z^2}$ στο **σημείο εκείνο** του χωρίου όπου κατά το **προηγούμενο βήμα είχαμε βρει το μέγιστο**. Πρόκειται δηλαδή για ένα κριτήριο min-max που έχει ως στόχο την επιλογή των σημείων που θα προστεθούν στα ήδη υπάρχοντα ώστε να ελαχιστοποιούν το σφάλμα της εκτίμησης (το οποίο εδώ εκφράζεται με τον λόγο $\frac{\sigma_{0in}^2}{\sigma_z^2}$) εκεί όπου πριν αυτό ήταν μέγιστο.
6. Μεταβαίνουμε στο 2^ο βήμα κ.ο.κ.

Επισημάνσεις :

- Στο 5^ο βήμα , κάθε φορά που εισάγουμε τα υπονήφια σημεία εκπαίδευσης και για αυτά (μαζί και με τα ήδη υπάρχοντα) υπολογίζουμε τον λόγο $\frac{\sigma_{0in}^2}{\sigma_z^2}$, χρησιμοποιούμε τις παραμέτρους θ , όπως αυτές έχουν υπολογιστεί για τα ήδη υπάρχοντα σημεία εκπαίδευσης. Το γεγονός αυτό δημιουργεί κάποιο σφάλμα και για αυτό είναι προτιμητέο η προσθήκη σημείων εκπαίδευσης να γίνεται σταδιακά με λίγα σημεία κάθε φορά . Το «ιδανικό» θα ήταν για κάθε υπολογισμό του $\frac{\sigma_{0in}^2}{\sigma_z^2}$ να προηγούνταν η εφαρμογή του κριτηρίου της μέγιστης πιθανοφάνειας ώστε να υπολογίσουμε τις παραμέτρους θ , όμως μια τέτοια διαδικασία θα ήταν σημαντικά χρονοβόρα (και ιδίως όσο αυξάνεται η διάσταση του προβλήματος καθώς και τα σημεία εκπαίδευσης). Αυτό το οποίο κάνουμε είναι αφού εντοπίσουμε τα σημεία που θα «εμπλουτίσουν» τα ήδη υπάρχοντα , με την διαδικασία του βήματος 5, μετά να υπολογίσουμε τις νέες παραμέτρους θ . Ανάλογα με την κρίση μας, θα μπορούσαμε να υπολογίζαμε τις νέες παραμέτρους θ π.χ. μετά την εκτέλεση δύο ή τριών κύκλων, αυτό συνήθως εξαρτάται από το πόσο απαιτητικό σε χρόνο CPU είναι το κριτήριο της μέγιστης πιθανοφάνειας αλλά και από πόσα σημεία εισάγουμε με κάθε κύκλο (αν είναι πολλά καλό θα ήταν να ανανεώναμε τις παραμέτρους σε κάθε κύκλο).
- Στην περίπτωση που επιθυμούμε σε κάθε κύκλο να εισάγουμε ένα μόνο σημείο δεν χρειάζεται να εκτελέσουμε την διαδικασία του βήματος 5, αφού το

σημείο αυτό θα είναι σίγουρα αυτό το οποίο εντοπίστηκε στο βήμα 4. Και η νέα τιμή του $\left\{ \frac{\sigma_{0|n}^2}{\sigma_z^2} \right\}$ στο σημείο αυτό θα είναι ίση με μηδέν (0).

- Δεδομένου ότι ασχολούμαστε με «πειράματα» σε υπολογιστή δεν έχουμε δύο ή περισσότερα ίδια σημεία εκπαίδευσης, έτσι απαιτείται κατά την επιλογή των υποψήφιων σημείων να μην επιλέγονται σημεία τα οποία αποτελούν ήδη σημεία εκπαίδευσης. Σίγουρα, στην περίπτωση της εισαγωγής ενός μόνο σημείου σε κάθε κύκλο κάτι τέτοιο δεν είναι απαραίτητο με βάση τα όσα αναφέρθηκαν προηγουμένως (δείτε 3^ο γράφημα).
- Από τη σχέση (5b) έχουμε :

$$\frac{\sigma_{0|n}^2}{\sigma_z^2} = \left[1 - \mathbf{r}_0^T \mathbf{R}^{-1} \mathbf{r}_0 + (\mathbf{f}_0 - \mathbf{F}^T \mathbf{R}^{-1} \mathbf{r}_0)^T (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} (\mathbf{f}_0 - \mathbf{F}^T \mathbf{R}^{-1} \mathbf{r}_0) \right]$$

που αποτελεί την έκφραση υπολογισμού του $\frac{\sigma_{0|n}^2}{\sigma_z^2}$.

Στα παρακάτω θα εφαρμόσουμε την διαδικασία που μόλις αναπτύχθηκε, θα ξεκινήσουμε με 20 σημεία εκπαίδευσης και σε κάθε κύκλο θα προσθέτουμε ένα σημείο, αυτό στο οποίο το μεταμοντέλο του προηγούμενου κύκλου είχε $\max \left\{ \frac{\sigma_{0|n}^2}{\sigma_z^2} \right\}$.

Εφαρμογή με πρότυπο μαθηματική συνάρτηση (d=2).

Στην παρούσα εφαρμογή επιλέγουμε να παρουσιάσουμε μια σχετικά «απλή» περίπτωση, σε ότι αφορά την διάσταση του προβλήματος, έχοντας δύο (2) μόνο ελεύθερες παραμέτρους (x_1, x_2) . Η μικρή διάσταση του προβλήματος μας επιτρέπει να εντοπίσουμε εύκολα την εξάρτηση της μεταβλητότητας στο χωρίο από την επιλογή των σημείων εκπαίδευσης σε αυτό, επίσης μπορούμε να δούμε τη συνεισφορά του κριτηρίου της μέγιστης πιθανοφάνειας στη βελτίωση της ικανότητας πρόβλεψης του μεταπροτύπου μας. Ως πρότυπο χρησιμοποιούμε την παρακάτω μαθηματική συνάρτηση, την απόκριση της οποίας θα προσπαθήσουμε να προσεγγίσουμε δημιουργώντας ένα αντίστοιχο μεταπρότυπο, η επιλογή της δεν είναι τυχαία αφού στο χωρίο μελέτης $[-1,1] \times [-1,1]$ η συνάρτηση διαθέτει πληθώρα τοπικών ακροτάτων (δείτε 3^ο σχήμα):

$$y = 20 + x_1^2 + x_2^2 - 10 \cdot \cos(2 \cdot \pi \cdot x_1) - 10 \cdot \cos(2 \cdot \pi \cdot x_2)$$

όπου $(x_1, x_2) \in [-1,1] \times [-1,1]$.

Στο σημείο αυτό θα επισημάνουμε ότι λόγω της φύσης του προτύπου είναι εξαιρετικά εύκολο να έχουμε όσα στο πλήθος πειράματα (δηλαδή σημεία εκπαίδευσης) θέλουμε, μιας και το υπολογιστικό κόστος είναι ασήμαντο. **Όμως** συχνά κατασκευάζουμε μεταπρότυπα ώστε να αποφύγουμε την χρησιμοποίηση ενός υπολογιστικά χρονοβόρου προτύπου (π.χ. σε μια μέθοδο βελτιστοποίησης). Έτσι σε όλη την ανάλυση μας θα έχουμε αυτό το γεγονός υπόψη μας και βέβαια **δεν** θα καταφύγουμε στην χρησιμοποίηση πολλών σημείων εκπαίδευσης επειδή στην περίπτωσή μας συμβαίνει να είναι κάτι μη απαιτητικό σε χρόνο CPU.

Επιλέγουμε να χρησιμοποιήσουμε συνάρτηση συσχέτισης της μορφής :

$$R(h, \psi) = \exp \left\{ - \sum_{j=1}^d \left| \frac{h_j}{\theta_j} \right|^2 \right\}, \text{ με } d=2$$

Εφαρμόζουμε το κριτήριο της μέγιστης πιθανοφάνειας, για την παραπάνω συνάρτηση συσχέτισης, ώστε να υπολογίσουμε τις άγνωστες παραμέτρους τα θ_1, θ_2 . Έτσι εκτελούμε την ελαχιστοποίηση της παράστασης (7) με άγνωστες παραμέτρους τα θ_1, θ_2 . Τα αποτελέσματα αυτής της διαδικασίας, που γίνεται με χρήση γενετικού αλγορίθμου, παρατίθενται παρακάτω :

$$\theta_1 = 0.636322580645161$$

$$\theta_2 = 0.616774193548387$$

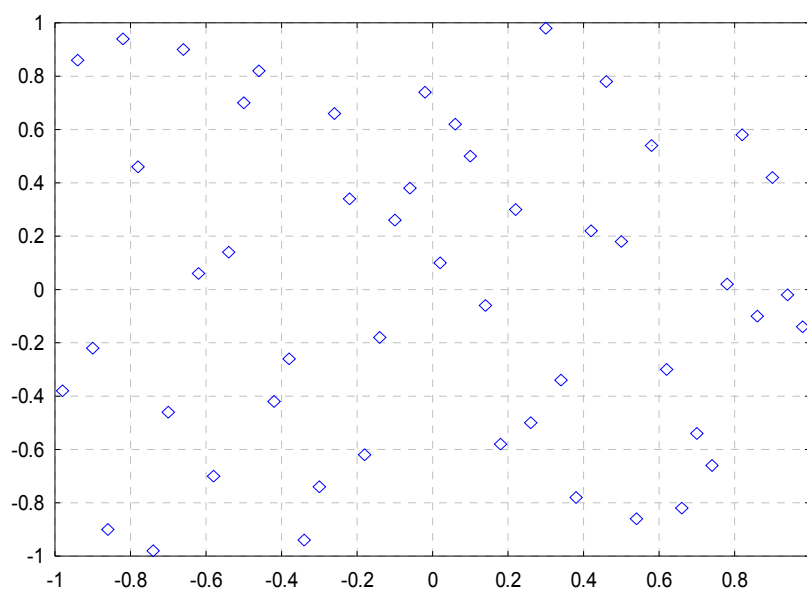
$$\Omega_{\min}(\psi) = 52.0492387997792$$

Με την χρήση γενετικού αλγορίθμου αποφεύγουμε τον εγκλωβισμό σε τοπικά ελάχιστα, τα οποία είναι πολλά, ιδίως όσο η διάσταση του προβλήματος αυξάνει, στη περίπτωση πάντως που χρησιμοποιηθεί κάποια αιτιοκρατική μέθοδος ελαχιστοποίησης απαιτείται η επανεκκίνηση της μεθόδου αρκετές φορές από διαφορετικά σημεία .

Για την εκπαίδευση του μεταπροτύπου εκλέγονται με τη μέθοδο των λατινικών τετραγώνων 50 στο πλήθος σημεία, η επιλογή του πλήθους είναι μεν τυχαία (με την έννοια ότι θα μπορούσαμε να επιλέγαμε για παράδειγμα 40 ή 60) αλλά επιλέχθηκε έτσι ώστε να μας δίνει ένα ικανό μεταπρότυπο από πλευράς δυνατοτήτων πρόβλεψης της πραγματικής απόκρισης. Θα μπορούσαμε να ξεκινήσουμε από λιγότερα στο πλήθος σημεία εκπαίδευσης και μετά σιγά – σιγά να προσθέτουμε καινούργια σημεία ως ότου το μεταπρότυπο μας φτάσει σε ικανά επίπεδα πρόβλεψης – τα επίπεδα αυτά τα καθορίζει ο μηχανικός με βάση την φύση και τις απαιτήσεις του προβλήματος το οποίο καλείται να μοντελοποιήσει. Σε μια τέτοια διαδικασία μπορούμε να αξιοποιήσουμε τη μεταβλητότητα του μεταπροτύπου μας στοχεύοντας στη χρησιμοποίηση του κατά το δυνατόν ελαχίστου πλήθους σημείων εκπαίδευσης (ζήτημα εξαιρετικής σημασίας όταν το πρότυπό μας είναι απαιτητικό σε χρόνο CPU).

Οι συναρτήσεις αναδρομής (\mathbf{f}) τις οποίες χρησιμοποιούμε είναι τα εκθετικά όλων των πολωνύμων, μέχρι και 3^{ου} βαθμού, που μπορούν να μας δώσουν οι συνδυασμοί των τεσσάρων μεταβλητών x_1, x_2 .

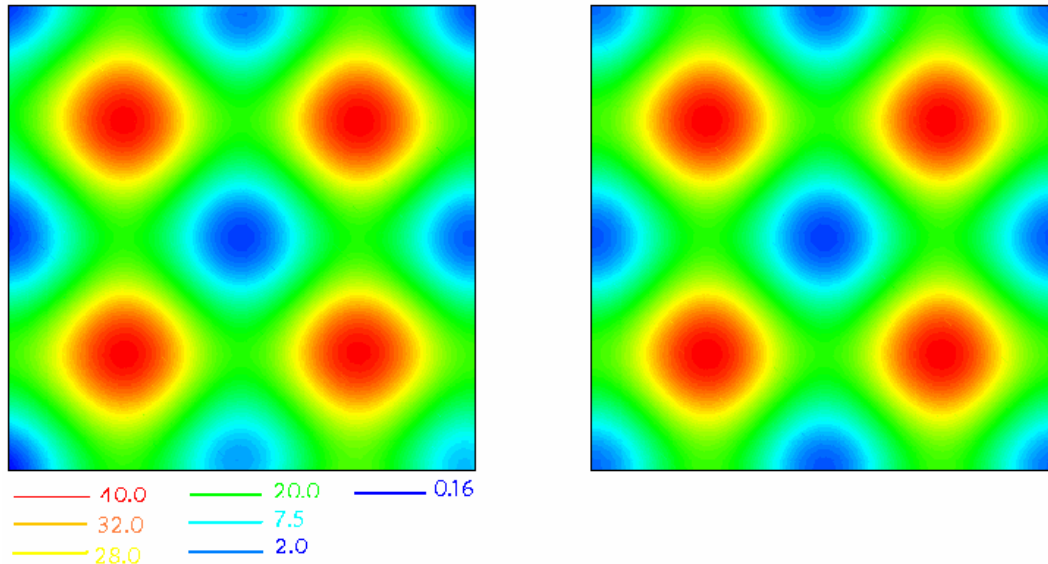
$$\mathbf{f}(\mathbf{x}) = \left[1, \exp(x_1), \exp(x_1^2), \exp(x_1^3), \exp(x_1^2 x_2), \exp(x_1 x_2), \exp(x_1 x_2^2), \exp(x_2), \exp(x_2^2), \exp(x_2^3) \right]^T$$



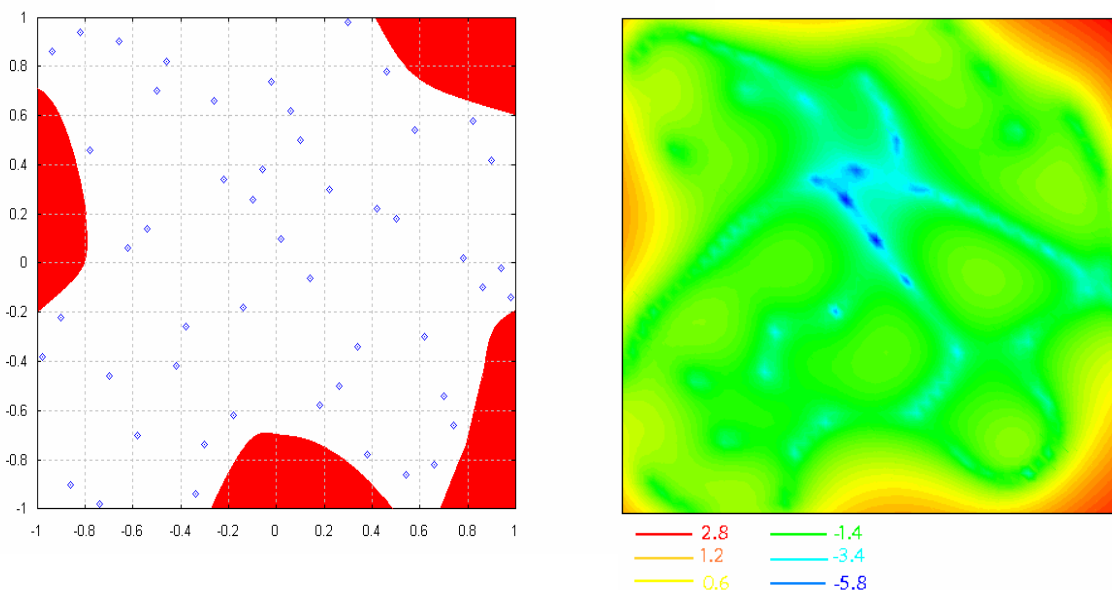
Σχήμα 2: Τα σημεία εκπαίδευσης του μεταπροτύπου. Οριζόντιος άξονας: παράμετρος x1. Κάθετος άξονας : παράμετρος x2.

Στην συνέχεια, παραθέτουμε τα αποτελέσματα που προέκυψαν από την εφαρμογή της μεθοδολογίας.

Αποτελέσματα με εφαρμογή κριτηρίου μέγιστης πιθανοφάνειας, με 50 σημεία εκπαίδευσης (Latin Squares).



Σχήμα 3: Ισοσταθμική x_1 (οριζόντιος)- x_2 (κατακόρυφος). Αριστερά η «πρόβλεψη», δεξιά το «πρότυπο».



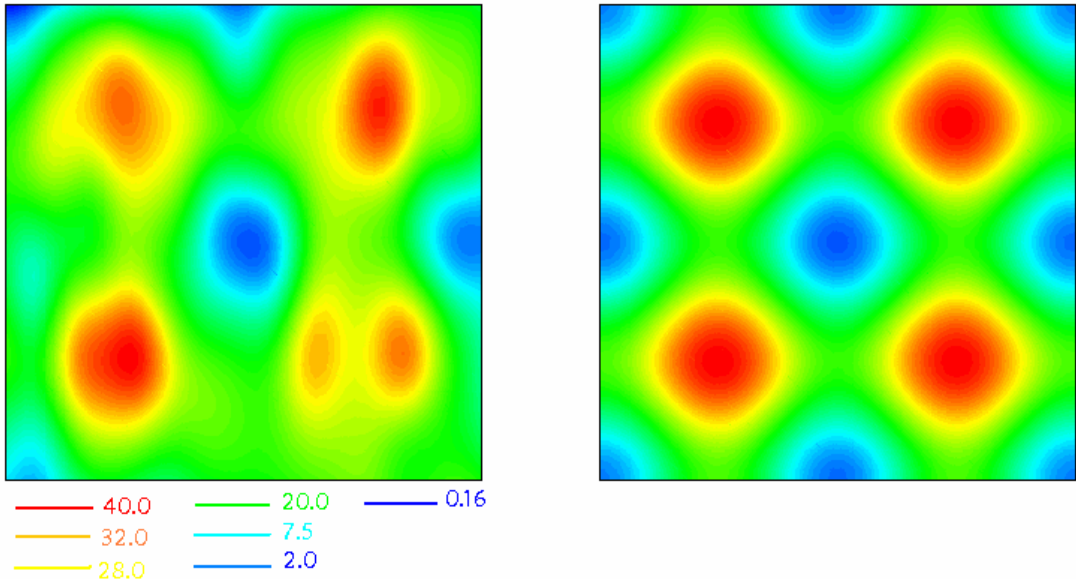
Σχήμα 4: Αριστερά τα σημεία εκπαίδευσης στο χωρίο μελέτης x_1 - x_2 (με κόκκινο σκιαγραφούνται οι περιοχές εκείνες όπου απουσιάζουν σημεία). Δεξιά Ο δεκαδικός λογάριθμος της μεταβλητότητας $\sigma_{0/n}^2$ σε όλο το χωρίο για τις μεταβλητές x_1 - x_2 .

Από τα παραπάνω γραφήματα (2&3) παρατηρούμε αυτό που αναμένουμε, ότι δηλαδή η μεταβλητότητα (η οποία εκφράζει το σφάλμα της εκάστοτε εκτίμησης) είναι μεγαλύτερη στις περιοχές όπου δεν υπάρχει επαρκής πληροφορία (απουσία σημείων εκπαίδευσης) και στις οποίες το μεταπρότυπο δεν επιτυγχάνει τόσο καλή εκτίμηση.

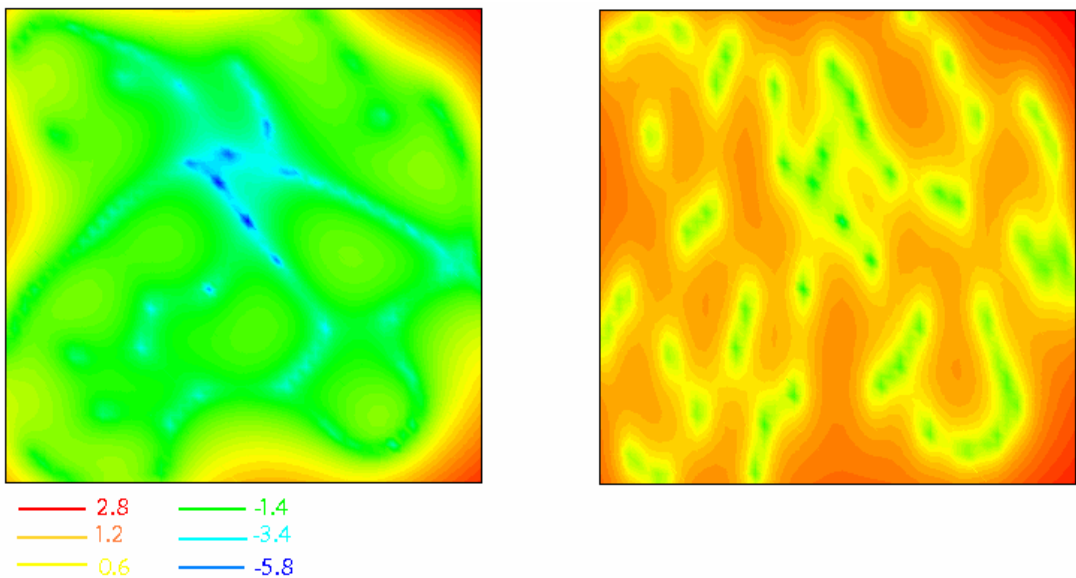
Αποτελέσματα χωρίς την χρήση του κριτηρίου της μέγιστης πιθανοφάνειας, με 50 σημεία εκπαίδευσης (Latin Squares).

Στην συνέχεια προβαίνουμε στην κατασκευή ενός μεταπροτύπου, όπως και πριν, με μόνη διαφορά την εκλογή των παραμέτρων θ_1, θ_2 έτσι ώστε αυτές να μην πληρούν το κριτήριο της μέγιστης πιθανοφάνειας, έχουμε ότι :

$\theta_1 = 0.16$
$\theta_2 = 0.36$
$\Omega(\psi) = 80.66$

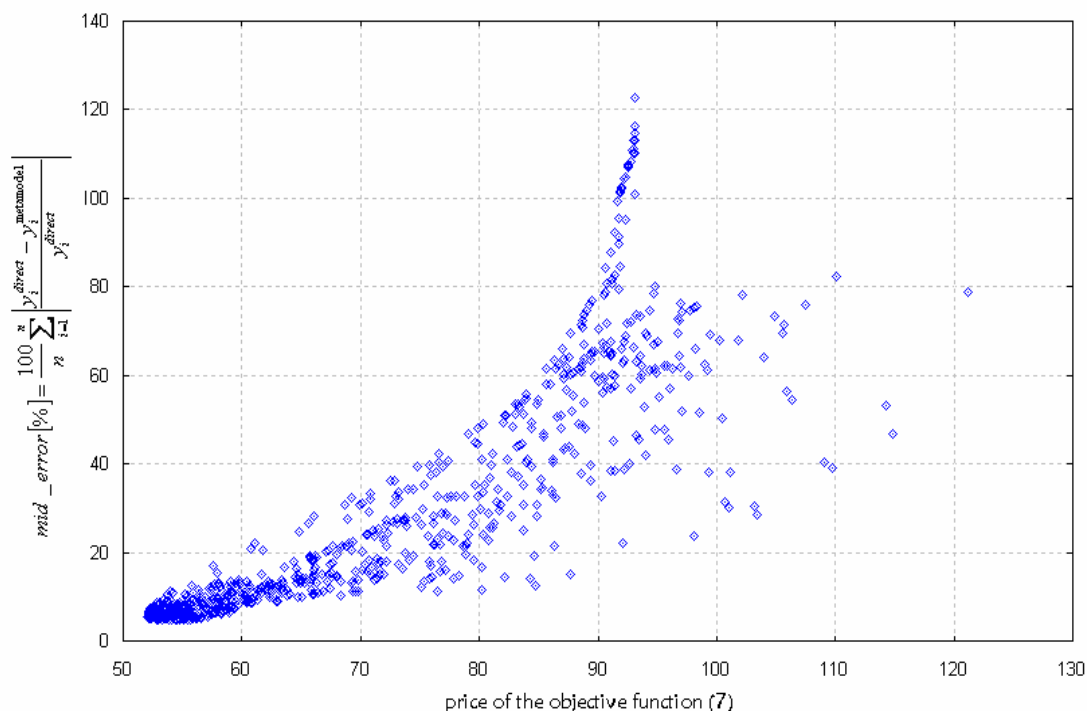


Σχήμα 5: Ισοσταθμική x_1 (οριζόντιος)- x_2 (κατακόρυφος). Αριστερά η «πρόβλεψη», δεξιά το «πρότυπο». Σε αντιπαραβολή με το γράφημα 2 καθίσταται εμφανές το ότι επιτυγχάνουμε καλύτερη εκτίμηση.



Σχήμα 6: Αριστερά ο δεκαδικός λογάριθμος της μεταβλητότητας στο χωρίο για τις μεταβλητές x_1-x_2 , για την περίπτωση που πληρείται το κριτήριο της μέγιστης πιθανοφάνειας (δείτε γράφημα 3). Δεξιά όταν δεν πληρείται . Διακρίνουμε με ευκολία ότι η μεταβλητότητα δεξιά είναι αυξημένη (άρα η εκτίμηση γίνεται με μεγαλύτερο σφάλμα).

Στο παρακάτω γράφημα καθίσταται εμφανές ότι όσο **μειώνεται** η τιμή $\Omega(\psi)$ της παράστασης (7), δηλαδή όσο «περισσότερο» πληρείται το κριτήριο της μέγιστης πιθανοφάνειας (το οποίο εκφράζεται από την σχέση (7)) τόσο μειώνεται το μέσο εκατοστιαίο σφάλμα του μεταπροτύπου μας. Το παρακάτω γράφημα έρχεται σε πλήρη συμφωνία με τα ήδη αναφερθέντα στην προηγούμενη σελίδα .



Σχήμα 7: Οριζόντιος άξονας : η τιμή της αντικειμενικής συνάρτησης $\Omega(\psi)$ για τυχαία επιλεγμένα ζεύγη παραμέτρων $\psi=[\theta_1, \theta_2]^T$.
Κατακόρυφος άξονας : το μέσο εκατοστιαίο σφάλμα όπως αυτό υπολογίστηκε σε $n=50^2$ σημεία στο χωρίο μελέτης, τα σημεία αυτά προέκυψαν από ένα πλέγμα 50×50 .

Το μέσο εκατοστιαίο σφάλμα δίνεται από την παρακάτω έκφραση :

$$mid_error[\%] = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i^{direct} - y_i^{metamodel}}{y_i^{direct}} \right|$$

Επομένως, συμπεραίνουμε ότι η εφαρμογή του κριτηρίου της μέγιστης πιθανοφάνειας, για τον προσδιορισμό των παραμέτρων $\psi = [\theta_1, \theta_2, \dots, \theta_d]^T$, οδηγεί σε δημιουργία μεταπροτύπων με καλύτερες δυνατότητες πρόβλεψης και παρότι συχνά είναι απαιτητικό σε χρόνο CPU (κυρίως όταν αυξάνει η διάσταση (d) του προβλήματος καθώς και όταν αυξάνει ο αριθμός των σημείων εκπαίδευσης) κρίνεται σκόπιμη η εφαρμογή του σε κάθε περίπτωση. Στο σημείο αυτό πρέπει να επισημανθεί ότι με την εφαρμογή του κριτηρίου, ακόμα και όταν δεν έχει επέλθει σύγκλιση της μεθόδου ελαχιστοποίησης (οπότε το διάνυσμα $\psi = [\theta_1, \theta_2, \dots, \theta_d]^T$ δεν αντιστοιχεί σε στο ολικό ελάχιστο), συνήθως εξασφαλίζουμε ότι ο πίνακας R των συσχετίσεων έχει καθορισμένες ιδιότητες. Γιατί μια τυχαία επιλογή των παραμέτρων θ_i δεν οδηγεί πάντα , για παράδειγμα , σε αντιστρέψιμο μητρώο R. (*Επιπρόσθετα , θα πρέπει να αναφέρουμε ότι για να δίνει η έκφραση (3) το ελάχιστο μέσο τετραγωνικό σφάλμα – που είναι ιδιότητα της Βέλτιστης Γραμμικής Ανεπηρέαστης Εκτίμησης (BLUP) και*

υπενθυμίζουμε ότι η σχέση (3) αποτελεί BLUP – θα πρέπει το μητρώο R να είναι θετικά ημιορισμένο. Μια «τυχαία» επιλογή των παραμέτρων θ δεν οδηγεί σε μητρώο θετικά ημιορισμένο.)

Διάστημα εμπιστοσύνης της απόκρισης Y_0 .

Το διάστημα εμπιστοσύνης της απόκρισης Y_0 για επίπεδο εμπιστοσύνης $P = 1 - a$ είναι :

$$\hat{Y}_0 - t_{n-p}^{1-a/2} \cdot \sigma_{0|n} \leq Y_0 \leq \hat{Y}_0 + t_{n-p}^{1-a/2} \cdot \sigma_{0|n}$$

Όπου $t_{n-p}^{1-a/2}$ είναι η τιμή της κατανομής Student για $n-p$ βαθμούς ελευθερίας για την οποία ισχύει ότι : $P(t \leq t^{1-a/2}) = 1 - \frac{a}{2}$, \hat{Y}_0 η απόκριση (εκτίμηση) του μεταπροτύπου και $\sigma_{0|n}$ η τυπική απόκλιση της εκτίμησής μας. (Υπενθυμίζουμε ότι όταν σ_z^2 είναι άγνωστο ισχύει : $[Y_0 | \mathbf{Y}^n] \sim T(v_i, \hat{Y}_0, \sigma_{0|n}^2)$.)

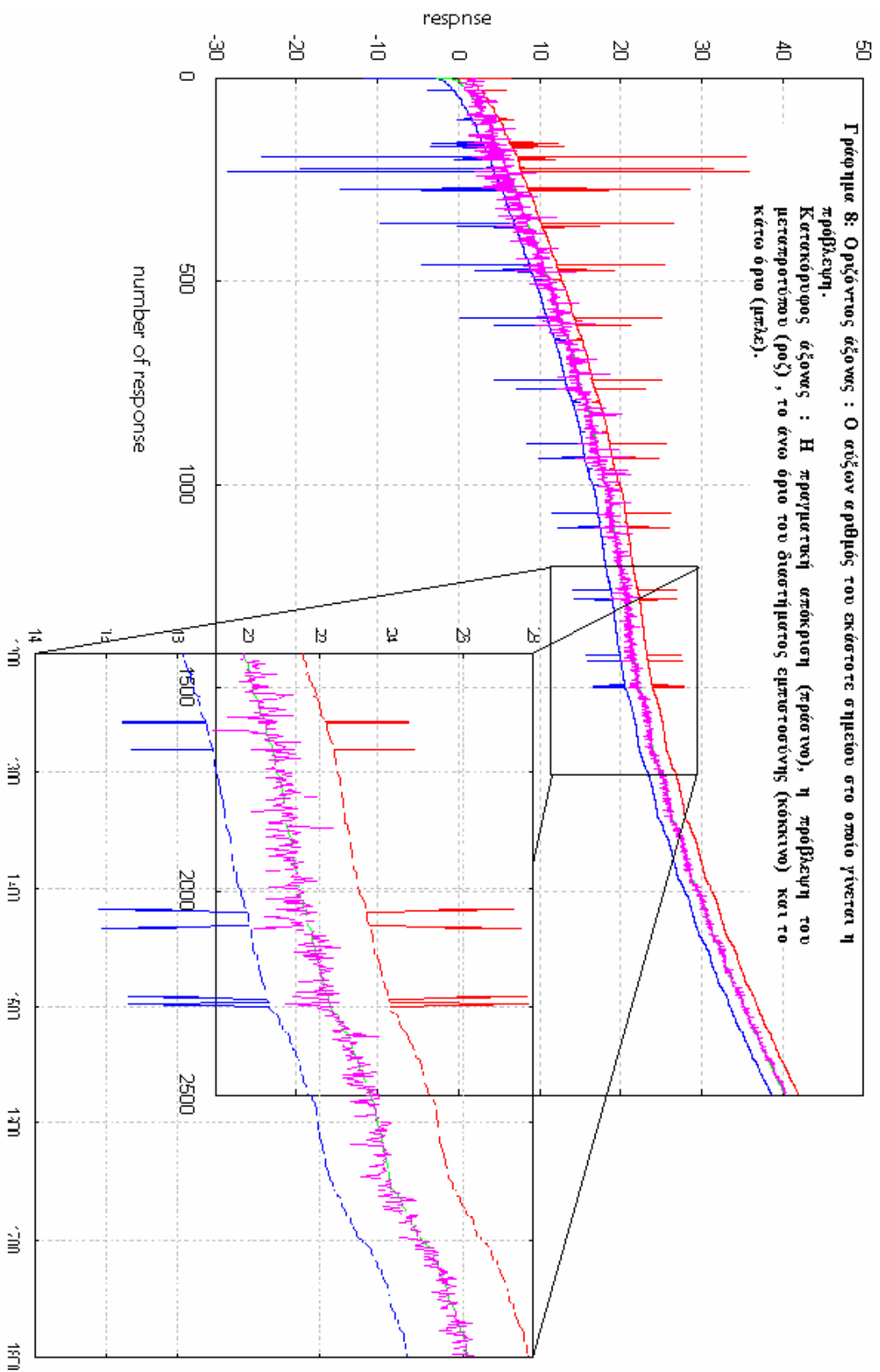
Στη συνέχεια επανερχόμαστε στο πρόβλημά μας και θα παρουσιάσουμε την απόκριση σε 50² σημεία του χωρίου μελέτης μας (που προέκυψαν από ένα 50×50 πλέγμα) καθώς και το διάστημα εμπιστοσύνης αυτής για επίπεδο εμπιστοσύνης 90%. Τα παραπάνω γίνονται αφού έχει εφαρμοστεί το κριτήριο της μέγιστης πιθανοφάνειας και με 50 σημεία εκπαίδευσης, επιλεγμένα με χρήση λατινικών τετραγώνων (για το μεταπρότυπο ισχύουν όσα είπαμε στην αρχή).

Καταρχάς στο πρόβλημά μας έχουμε $v = n - p = 50 - 10 = 40$ βαθμούς ελευθερίας (n : αριθμός σημείων εκπαίδευσης, p : διάσταση διανύσματος \mathbf{f}). Από πινακαποιημένες τιμές της κατανομής Student συναρτήσκει των βαθμών ελευθερίας και του εμβαδού $1 - \frac{a}{2}$, βρίσκουμε ότι για επίπεδο εμπιστοσύνης $P = 1 - a = 0.90$ έχουμε ότι :

$$a = 0.10 \rightarrow 1 - \frac{a}{2} = 0.95 \text{ και άρα η ζητούμενη τιμή είναι : } t_{40}^{0.95} = 1.684 .$$

Στο σχήμα 8 καταγράφουμε την πραγματική απόκριση (δηλαδή αυτή του προτύπου), την εκτίμηση (δηλαδή την απόκριση του μεταπροτύπου) και το διάστημα εμπιστοσύνης της τελευταίας, εντός του οποίου αναμένουμε να βρίσκεται «περίπου» το 90% των πραγματικών αποκρίσεων (για λόγους εποπτικούς έχουμε παρατάξει τις αποκρίσεις κατά αύξουσα σειρά). Παρατηρούμε ότι όντως αυτό συμβαίνει, όμως αδυναμία φαίνεται να αποτελεί το γεγονός ότι σε περιοχές όπου το μεταμοντέλο απέχει από την πραγματική (πρότυπη) απόκριση το διάστημα εμπιστοσύνης δεν είναι απλά αυξημένο αλλά συχνά λαμβάνει αρκετά ιδιαίτερα υψηλές τιμές. Επίσης φαίνεται ότι το μεταπρότυπο αδυνατεί να παρακολουθήσει τις μικρές διακυμάνσεις της πραγματικής απόκρισης (ροζ), θα μπορούσαμε να πούμε ότι αυτές υπερβαίνουν την «διακριτική» δυνατότητα του μεταπροτύπου. Ως θετικό λαμβάνουμε το γεγονός ότι η εκάστοτε εκτίμηση του μεταπροτύπου (πράσινη καμπύλη) «ακολουθεί» την πραγματική απόκριση (ροζ καμπύλη) – δηλαδή δεν έχουμε χονδροειδείς αποκλίσεις. Καθώς επίσης και το ότι εκεί που υπάρχει μεγαλύτερη διαφορά μεταξύ πραγματικής

απόκρισης (ροζ) και εκτίμησης (πράσινο) το διάστημα εμπιστοσύνης αυξάνει (λόγω του ότι αυξάνει η μεταβλητότητα της απόκρισης- δείτε σχήματα 3&4).



Εφαρμογή με πρότυπο μαθηματική συνάρτηση (d=4).

Εφαρμογή με $d=4$, το μαθηματικό μας πρότυπο είναι :

$$y = 40 + x_1^2 + x_2^2 + x_3^2 + x_4^2 - 10 \cdot \cos(2 \cdot \pi \cdot x_1) - 10 \cdot \cos(2 \cdot \pi \cdot x_2) - 10 \cdot \cos(2 \cdot \pi \cdot x_3) - 10 \cdot \cos(2 \cdot \pi \cdot x_4)$$

όπου $(x_1, x_2, x_3, x_4) \in [-1,1] \times [-1,1] \times [-1,1] \times [-1,1]$.

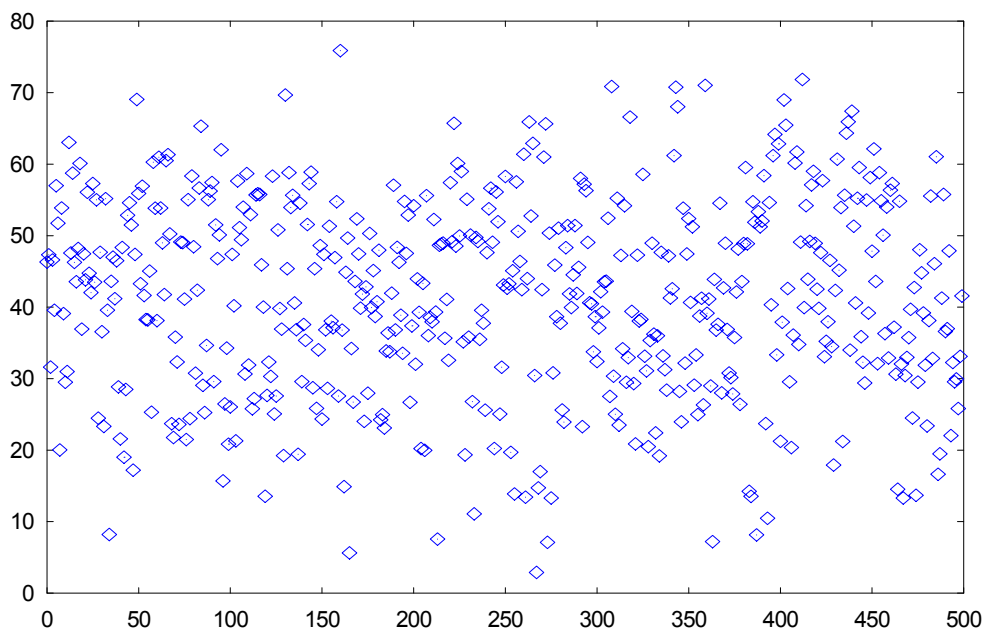
Επιλέγουμε να χρησιμοποιήσουμε συνάρτηση συσχέτισης της μορφής :

$$R(h, \psi) = \exp \left\{ - \sum_{j=1}^d \left| \frac{h_j}{\theta_j} \right|^2 \right\}, \text{ με } d=4$$

Εφαρμόζοντας το κριτήριο της μέγιστης πιθανοφάνειας, για την παραπάνω συνάρτηση συσχέτισης, έτσι εκτελούμε την ελαχιστοποίηση της παράστασης (7) με άγνωστες παραμέτρους τα $\theta_1, \theta_2, \theta_3, \theta_4$. Τα αποτελέσματα αυτής της διαδικασίας παρατίθενται παρακάτω :

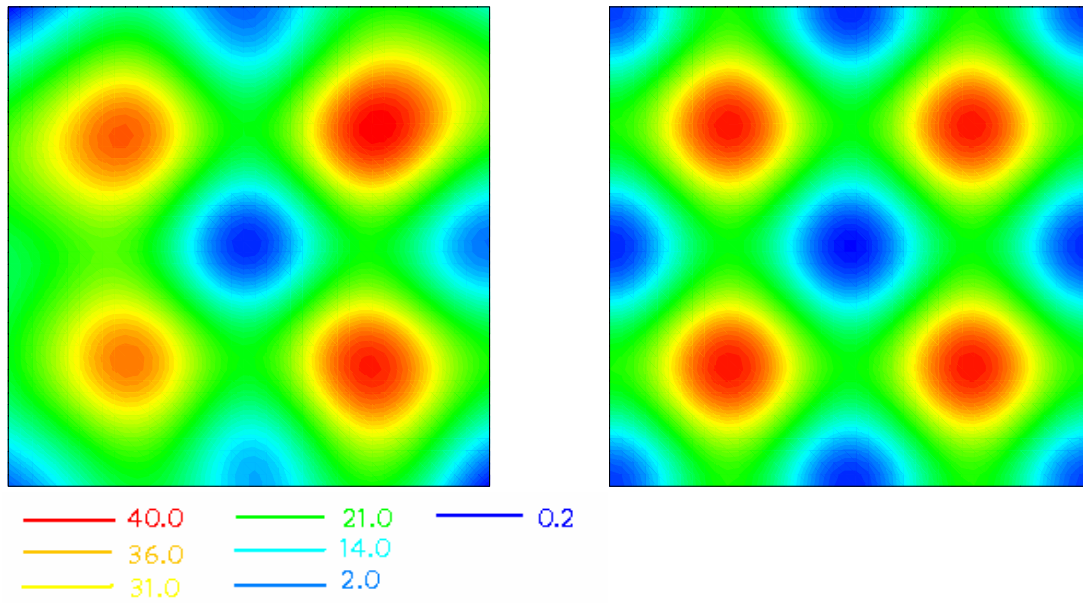
$$\begin{aligned} \theta_1 &= 0.547504105571848 \\ \theta_2 &= 0.537729032258065 \\ \theta_3 &= 0.527953958944282 \\ \theta_4 &= 0.537729032258065 \\ \Omega_{\min}(\psi) &= 819.064615919426 \end{aligned}$$

Οι συναρτήσεις αναδρομής (\mathbf{f}) τις οποίες χρησιμοποιούμε είναι τα εκθετικά όλων των πολωνύμων, μέχρι και 3^{ου} βαθμού, που μπορούν να μας δώσουν οι συνδυασμοί των τεσσάρων μεταβλητών x_1, x_2, x_3, x_4 . Επιπροσθέτως, τα σημεία εκπαίδευσης του μεταμοντέλου μας εκλέγονται με τη μέθοδο των λατινικών τετραγώνων και είναι 500 στο πλήθος.

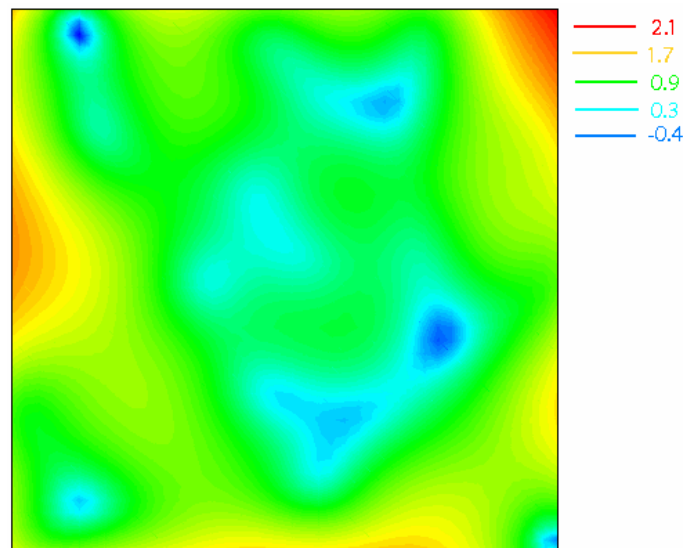


Σχήμα 9: Τα σημεία εκπαίδευσης του μεταπροτύπου. Οριζόντιος άξονας: αύξων αριθμός σημείου εκπαίδευσης. Κάθετος άξονας : η απόκριση του πρότυπου.

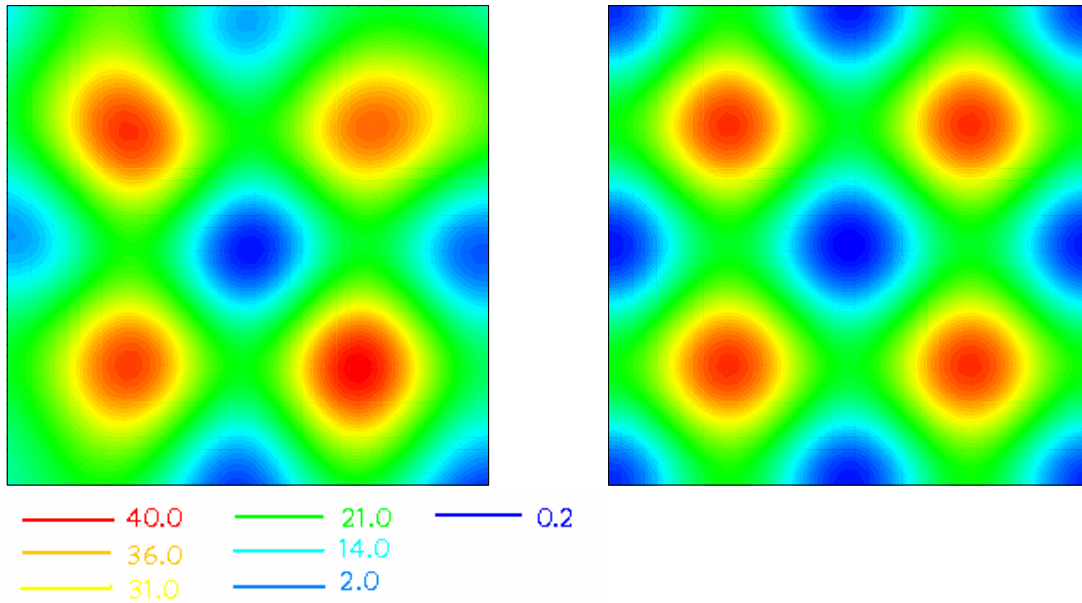
Στη συνέχεια , παραθέτουμε τα αποτελέσματα που προέκυψαν από την εφαρμογή της μεθοδολογίας.



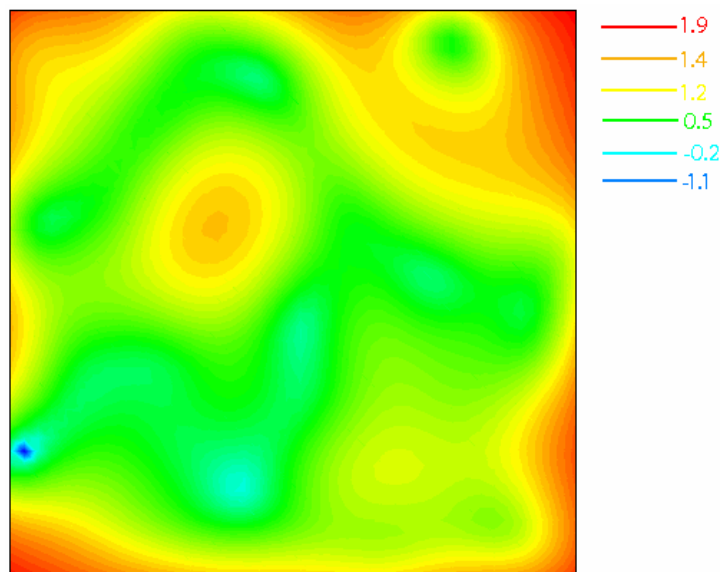
Σχήμα 10: Ισοσταθμική x_1 (οριζόντιος)- x_2 (κατακόρυφος) για $x_3=0$ (αδιάστατη τιμή [-1:1]) και για $x_4=0$ (αδιάστατη τιμή [-1:1]). Αριστερά η «πρόβλεψη», δεξιά το «πρότυπο».



Σχήμα 11: Ο δεκαδικός λογάριθμος της μεταβλητότητας σε όλο το χωρίο για τις μεταβλητές x_1 - x_2 , για $x_3=0$ (αδιάστατη τιμή [-1:1]) και για $x_4=0$ (αδιάστατη τιμή [-1:1]).



Σχήμα 12: Ισοσταθμική x_3 (οριζόντιος)- x_4 (κατακόρυφος) για $x_1=0$ (αδιάστατη τιμή [-1:1]) και για $x_2=0$ (αδιάστατη τιμή [-1:1]). Αριστερά η «πρόβλεψη», δεξιά το «πρότυπο».



Σχήμα 13: Ο δεκαδικός λογάριθμος της μεταβλητότητας σε όλο το χωρίο για τις μεταβλητές x_3 - x_4 , για $x_1=0$ (αδιάστατη τιμή [-1:1]) και για $x_2=0$ (αδιάστατη τιμή [-1:1]).

Τα αποτελέσματα είναι παραπλήσια με αυτά για τις υπόλοιπες ισοσταθμικές έτσι κρίνουμε περιττό το να παρουσιάσουμε και τις υπόλοιπες ισοσταθμικές. Από τα παραπάνω επιβεβαιώνουμε το αναμενόμενο, ότι δηλαδή η υπολογιζόμενη μεταβλητότητα είναι πιο αυξημένη στις περιοχές εκείνες όπου το μεταπρότυπό μας δεν επιτυγχάνει καλή εκτίμηση της πρότυπης απόκρισης. Έτσι η χρήση της μεταβλητότητας μπορεί να εφαρμοστεί, σε δεύτερο στάδιο, για τη βελτίωση του μεταπρότυπου, για παράδειγμα επιλέγοντας να εισάγουμε σημεία εκπαίδευσης στις περιοχές όπου η μεταβλητότητα (και άρα το σφάλμα του μεταπρότυπου) είναι υψηλή.

Εφαρμογή με πρότυπο μοντέλο ανάλυσης συνεκτικών ροών σε πτερυγώσεις στροβιλομηχανών.

Παρακάτω παρουσιάζουμε ένα πρόβλημα εσωτερικής αεροδυναμικής στο οποίο έχουμε έξι (6) μεταβλητές. Το ροϊκό πρόβλημα είναι διδιάστατο και η επίλυση της ροής γίνεται με το λογισμικό Mises. Δηλαδή, το λογισμικό Mises αποτελεί το πρότυπο την απόκριση του οποίου για διάφορες μορφές αεροτομών επιχειρούμε να μοντελοποιήσουμε (οι 6 μεταβλητές του προβλήματος αφορούν την μορφή της πτερύγωσης). Το εν λόγω λογισμικό επιτρέπει την ανάλυση συνεκτικών ροών σε δισδιάστατες πτερυγώσεις με την αλληλεπίδραση ενός επιλύτη ροής ατρίβους – εξωτερικής ροής και μια ολοκληρωματική μέθοδο για τον υπολογισμό του οριακού στρώματος και αποτελείται από ένα σύνολο προγραμμάτων για τη γένεση του πλέγματος, την επίλυση της ροής και για την παρουσίαση των αποτελεσμάτων. Η επίλυση της ροής γίνεται με την χρήση των εξισώσεων Euler σε συνδυασμό με την θεωρία του οριακού στρώματος, η δε παραμετροποίηση των πτερυγώσεων γίνεται με χρήση καμπυλών splines (στις ακμές προσφυγής και εκφυγής χρησιμοποιούμε κύκλους). Η απόκριση, την οποία και καλούμαστε να μοντελοποιήσουμε, είναι ο συντελεστής απωλειών ολικής πίεσης : $\omega = \frac{P_{t2}^{is} - P_{t2}}{P_{t1} - P_1}$ για την πτερύγωση ενός

συμπιεστή με βήμα $\sigma = 0.65$ (αδιάστατη ποσότητα : χορδή (c)/βημα(s)).

Στον παρακάτω πίνακα παρατίθενται συνοπτικά ορισμένα στοιχεία σχετικά με το πρόβλημα και τις συνθήκες ροής του.

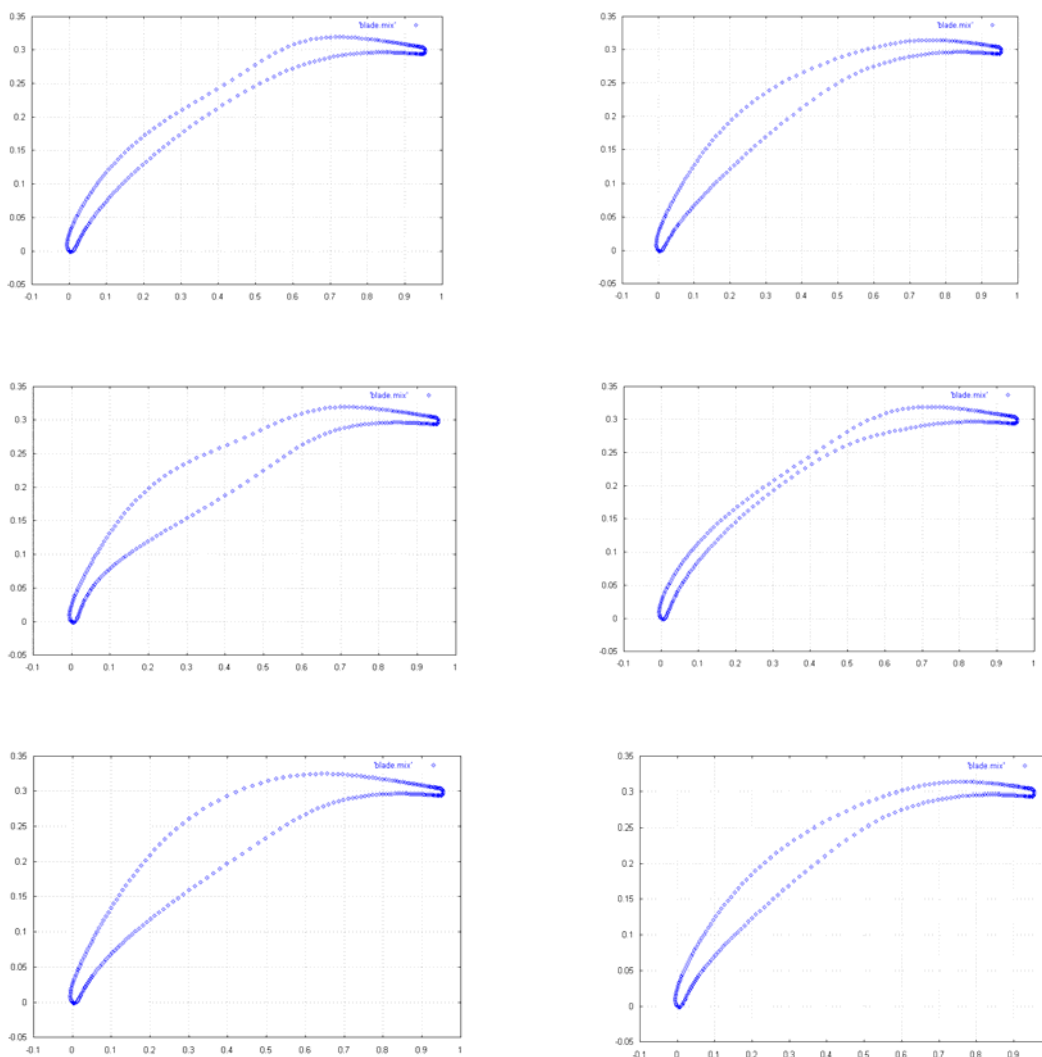
Συνθήκες ροής		
1	αριθμός Mach εισόδου :	M1= 0.7
2	σχετική γωνία εισόδου ροής :	$\beta_1 = 45$ [deg]
3	Κρίσιμος αριθμός Reynolds :	Rec= $4 \cdot 10^6$

Στοιχεία σχετικά με την κατασκευή του μεταπροτύπου.

Η επιλογή των σημείων εκπαίδευσης γίνεται με τη χρήση της μεθόδου των Λατινικών Τετραγώνων (δείτε σχετική παράγραφο), στην συνέχεια εξαιρούμε τα σημεία εκείνα στα οποία ο κώδικας της ροής δεν μας δίνει αποτέλεσμα, το γεγονός αυτό οφείλεται στο ότι βρισκόμαστε εκτός της βαθμονόμησης του κώδικα. Ο αριθμός των σημείων εκπαίδευσης καταλήγει να είναι 470. Η διαχείριση των σημείων όπου ο κώδικας Mises δεν δίνει αποτέλεσμα απαιτεί προσοχής μιας και το μεταμοντέλο μας θα παρέχει πρόλεξη για το συγκεκριμένο σημείο παρότι στην πραγματικότητα αυτή δεν υφίσταται (θεωρώντας πάντα ως πρότυπο τον κώδικα Mises). Η εξαίρεση αυτών των σημείων από την εκπαίδευση του μεταπροτύπου μας «λύνει απλά τα χέρια» στην περίπτωση της εκπαίδευσης του μεταπροτύπου όμως δεν βοηθά στο παραπάνω πρόβλημα (ύπαρξη απόκρισης του μεταμοντέλου την στιγμή που αυτής δεν υφίσταται). Θα μπορούσαμε – ιδίως αν αξιοποιούσαμε το μεταπρότυπο σε εφαρμογή βελτιστοποίησης (ελαχιστοποίησης απωλειών) να χρησιμοποιήσουμε και αυτά τα σημεία θέτοντάς τους μια μεγάλη τιμή (ως ποινή) έτσι ώστε να αποφύγουμε μια πιθανή ύπαρξη ελαχίστου του μεταμοντέλου στην περιοχή αυτή.

Στα επόμενα γραφήματα φαίνονται ορισμένες αεροτομές που αντιστοιχούν σε κάποια από τα σημεία εκπαίδευσης, παρατηρούμε ότι περιλαμβάνουν μεγάλη ποικιλία περιπτώσεων μάλιστα ορισμένες από αυτές μοιάζουν «μη λογικές». Αυτό συμβαίνει γιατί η περιοχή τιμών στην οποία γίνεται η μοντελοποίηση είναι ιδιαίτερα μεγάλη

(δείτε αρχείο datagas.dat στο τέλος). Βέβαια στην πράξη, και ιδίως όταν ενδιαφερόμαστε να εντοπίσουμε μια βέλτιστη -από πλευράς αεροδυναμικής- πτερύγωση, θα ήταν εύκολο και χρήσιμο να περιορίζαμε το χώρο θέτοντας περιορισμούς (π.χ. εξαιρώντας κάποιες πολύ «παχιές» και «περίεργες» αεροτομές). Όμως αυτό δεν είναι πάντα εφικτό σε όλα τα προβλήματα – δηλαδή το να περιορίσουμε το χωρίο μελέτης μας με βάση κάποιο κριτήριο – και εμείς σε αυτό το στάδιο ενδιαφερόμαστε πιο πολύ να επιδείξουμε τη συγκεκριμένη μεθοδολογία παρά να προβούμε σε μια βέλτιστη σχεδίαση με τη χρήση του λογισμικού Mises και του μεταπροτύπου μας.



Σχήμα 2: Αεροτομές που αντιστοιχούν σε έξι τυχαία επιλεγμένα σημεία εκπαίδευσης.

Οι συναρτήσεις αναδρομής οι οποίες χρησιμοποιούμε στην εν λόγω περίπτωση είναι πολωνυμικής μορφής και περιλαμβάνουν όλα τα πολώνυμα μέχρι και 3^{ου} βαθμού που μπορούν να μας δώσουν οι συνδυασμοί των έξι μεταβλητών $x_1, x_2, x_3, x_4, x_5, x_6$. Επιλέγουμε να χρησιμοποιήσουμε συνάρτηση συσχέτισης της μορφής :

$$R(h, \psi) = \exp \left\{ - \sum_{j=1}^d \left| \frac{h_j}{\theta_j} \right|^2 \right\}, \text{ με } d=6$$

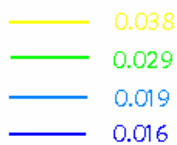
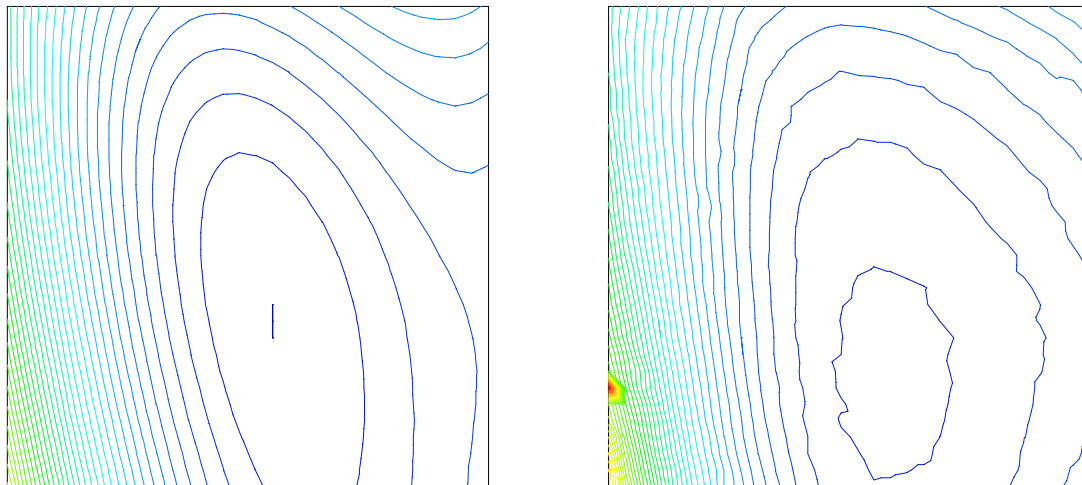
Εφαρμόζοντας το κριτήριο της μέγιστης πιθανοφάνειας, για την παραπάνω συνάρτηση συσχέτισης, έτσι εκτελούμε την ελαχιστοποίηση της παράστασης (7) με άγνωστες παραμέτρους τα $\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6$, η ελαχιστοποίηση αυτή λαμβάνει χώρα με χρήση γενετικού αλγορίθμου.

$$\Omega(\boldsymbol{\psi}) = n \log \hat{\sigma}_z^2(\boldsymbol{\psi}) + \log(\det(\mathbf{R}(\boldsymbol{\psi}))) \quad (7)$$

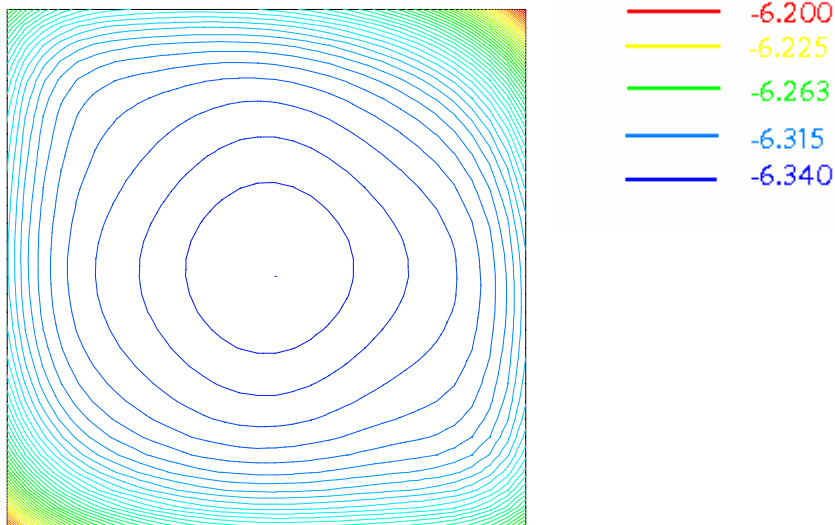
Τα αποτελέσματα αυτής της διαδικασίας παρατίθενται παρακάτω :

$\theta_1 = 0.312903910068426$
$\theta_2 = 1.965024437927664E - 002$
$\theta_3 = 8.07435092864125$
$\theta_4 = 7.66379579667644$
$\theta_5 = 1.E - 004$
$\theta_6 = 1.34906686217009$
ελάχιστη τιμή συνάρτησης (7): $\Omega_{\min} = -3026.62300965936$

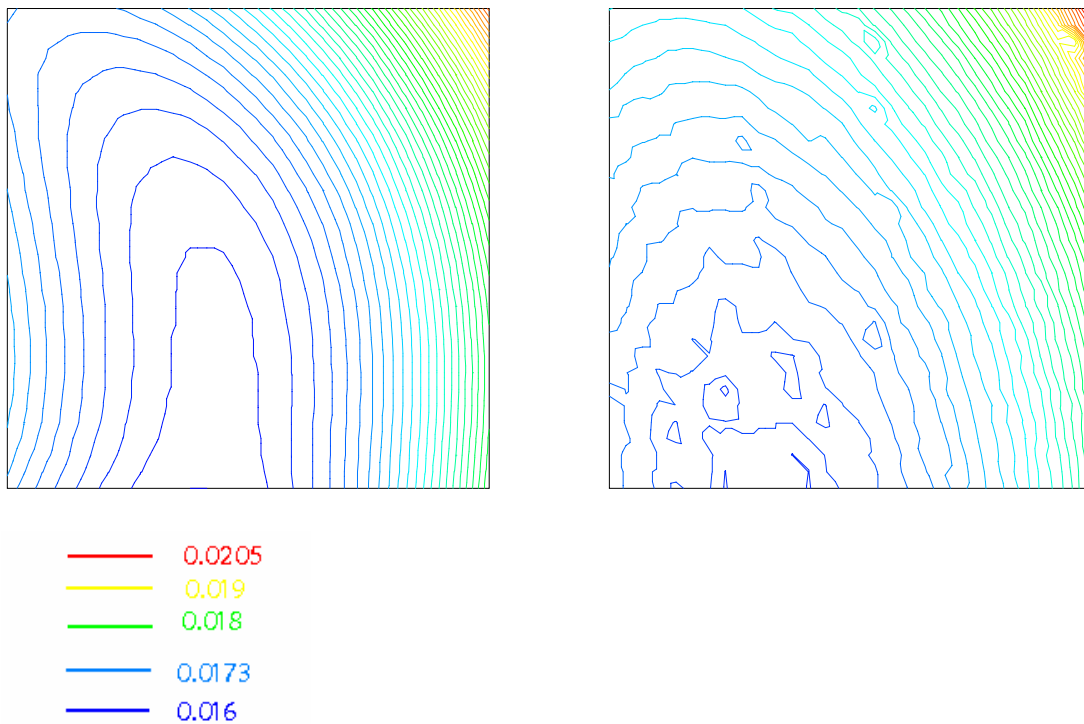
Στην συνέχεια παρουσιάζουμε ορισμένες ισοσταθμικές των μεταβλητών ανά ζεύγη (οι άλλες μεταβλητές διατηρούνται σταθερές στο μέσο του διαστήματος διακύμανσης τους) καθώς και τη μεταβλητότητα στο χωρίο μελέτης μας, έτσι ώστε να προβούμε σε έλεγχο της ικανότητας πρόβλεψης του μεταπροτύπου. Από την μορφή των γραφημάτων της μεταβλητότητας συμπεραίνουμε ότι το μεταπρότυπο επιτυγχάνει καλύτερη πρόλεξη στο κέντρο του χωρίου μελέτης .



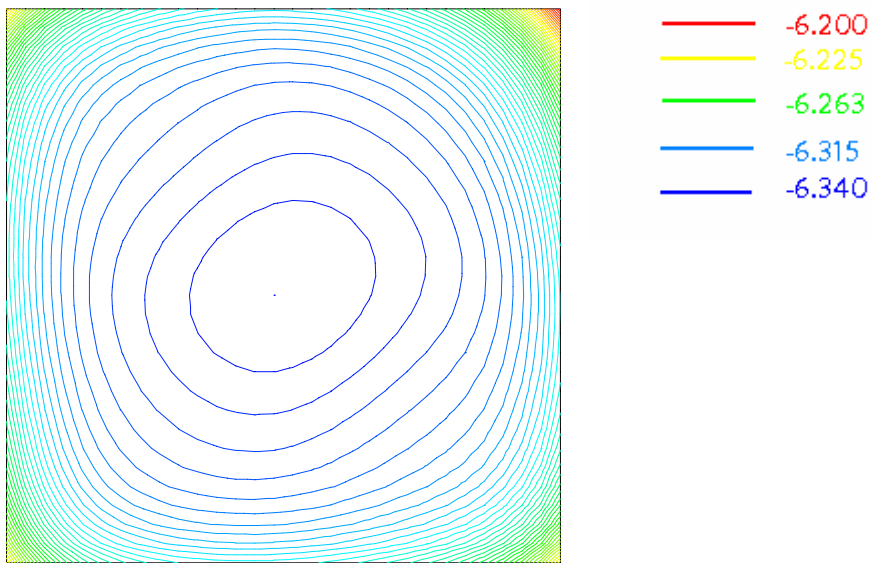
Σχήμα 2: Μεταβλητές x_1 (οριζόντιος) - x_2 (κατακόρυφος). Αριστερά η απόκριση του μεταμοντέλου, δεξιά η πραγματική.



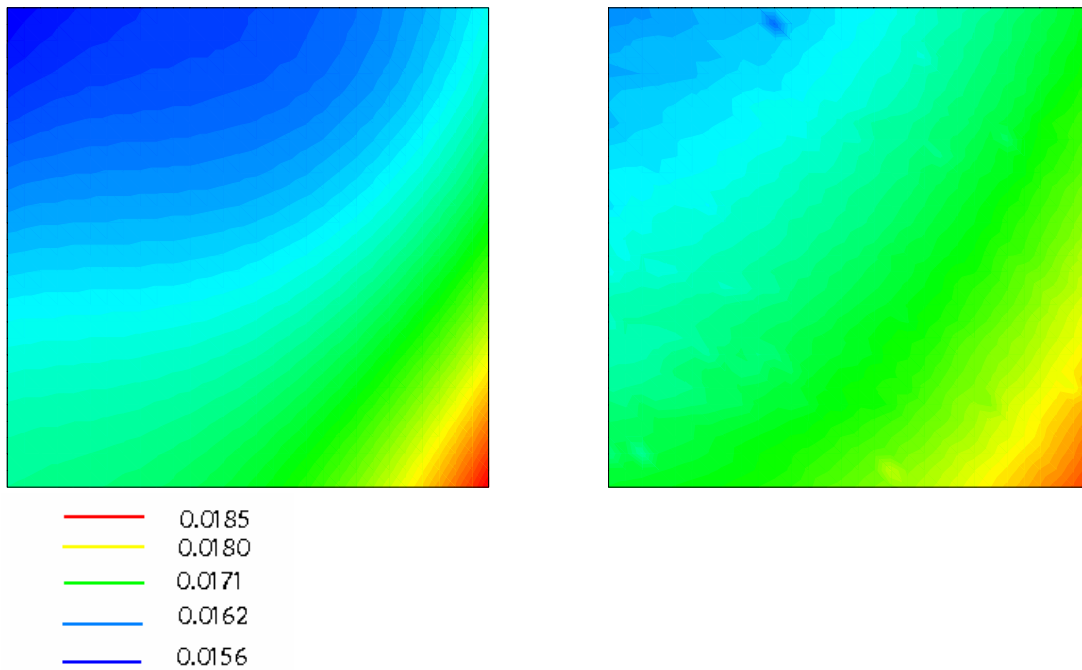
Σχήμα 3: Ο λογάριθμος της μεταβλητότητας σ_z^2 στο χωρίο των μεταβλητών x_1-x_2 . Με βάση αυτό το γράφημα στο κέντρο του χωρίου έχουμε καλύτερη ικανότητα πρόβλεψης.



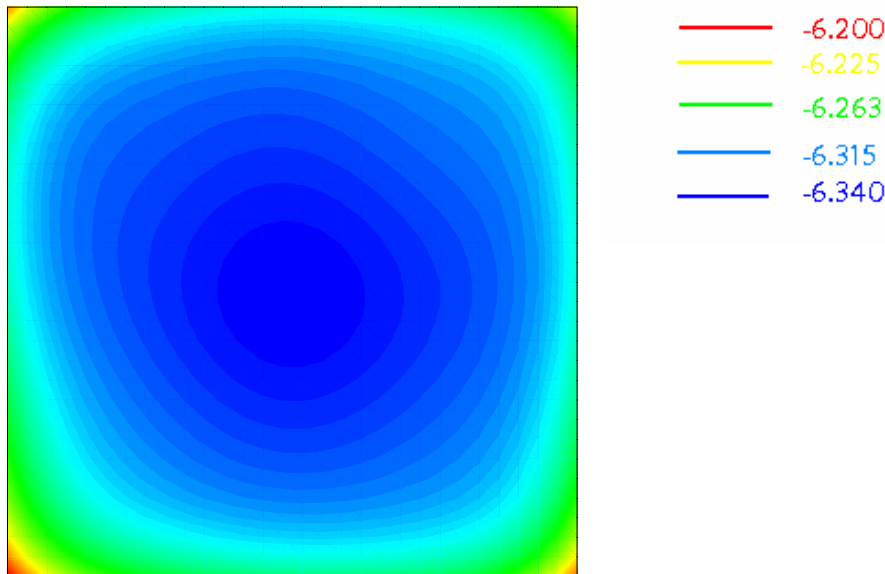
Σχήμα 4: Μεταβλητές x_2 (οριζόντιος) - x_3 (κατακόρυφος). Αριστερά η απόκριση του μεταμοντέλου, δεξιά η πραγματική.



Σχήμα 5: Ο λογάριθμος της μεταβλητότητας σ_z^2 στο χωρίο των μεταβλητών x_2 - x_3 . Με βάση αυτό το γράφημα στο κέντρο του χωρίου έχουμε καλύτερη ικανότητα πρόβλεψης.



Σχήμα 6: Μεταβλητές x_3 (οριζόντιος) - x_5 (κατακόρυφος). Αριστερά η απόκριση του μεταμοντέλου, δεξιά η πραγματική

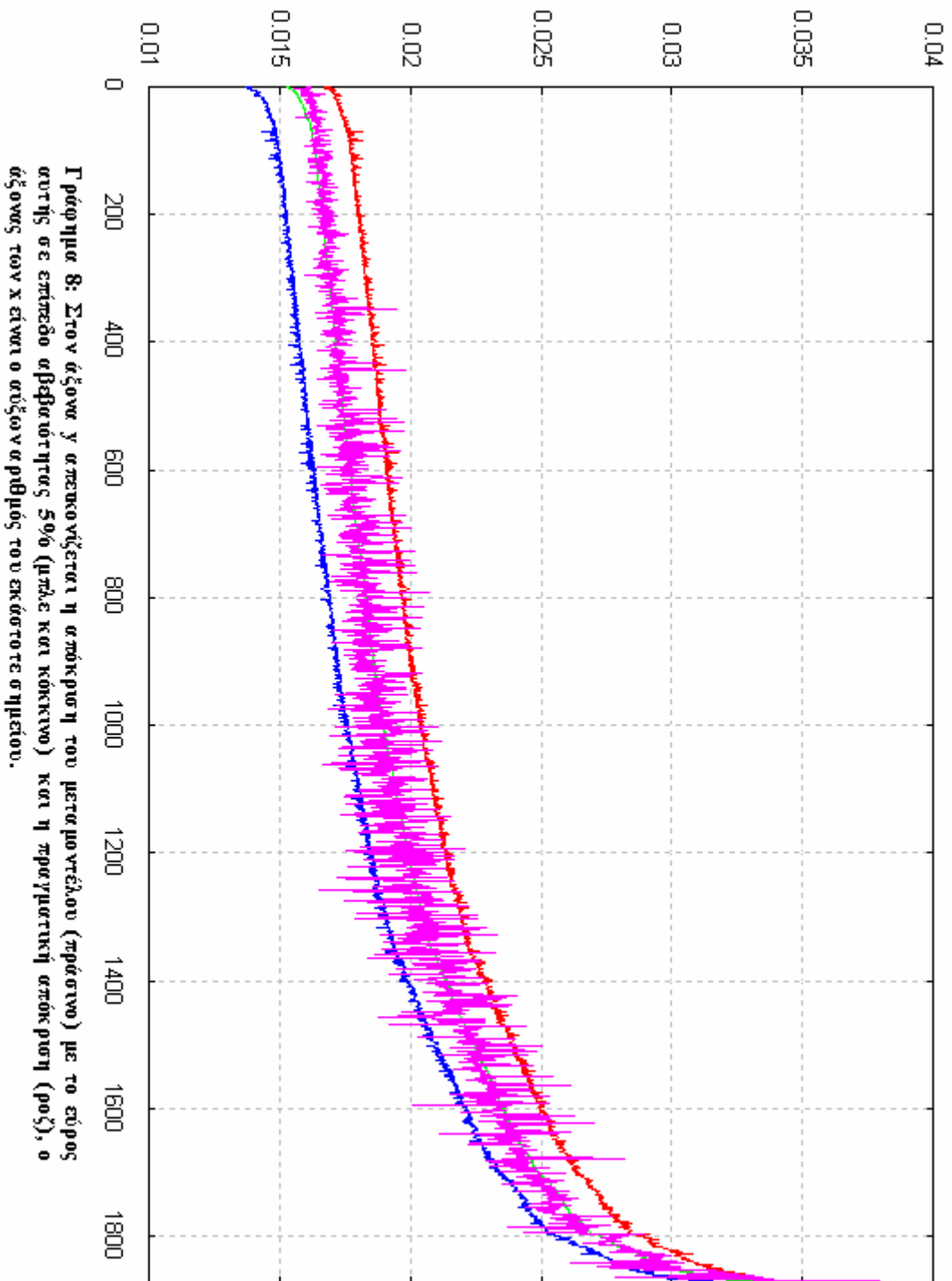


Σχήμα 7: Ο λογάριθμος της μεταβλητότητας σ_z^2 στο χωρίο των μεταβλητών x_2 - x_3 . Με βάση αυτό το γράφημα στο κέντρο του χωρίου έχουμε καλύτερη ικανότητα πρόβλεψης.

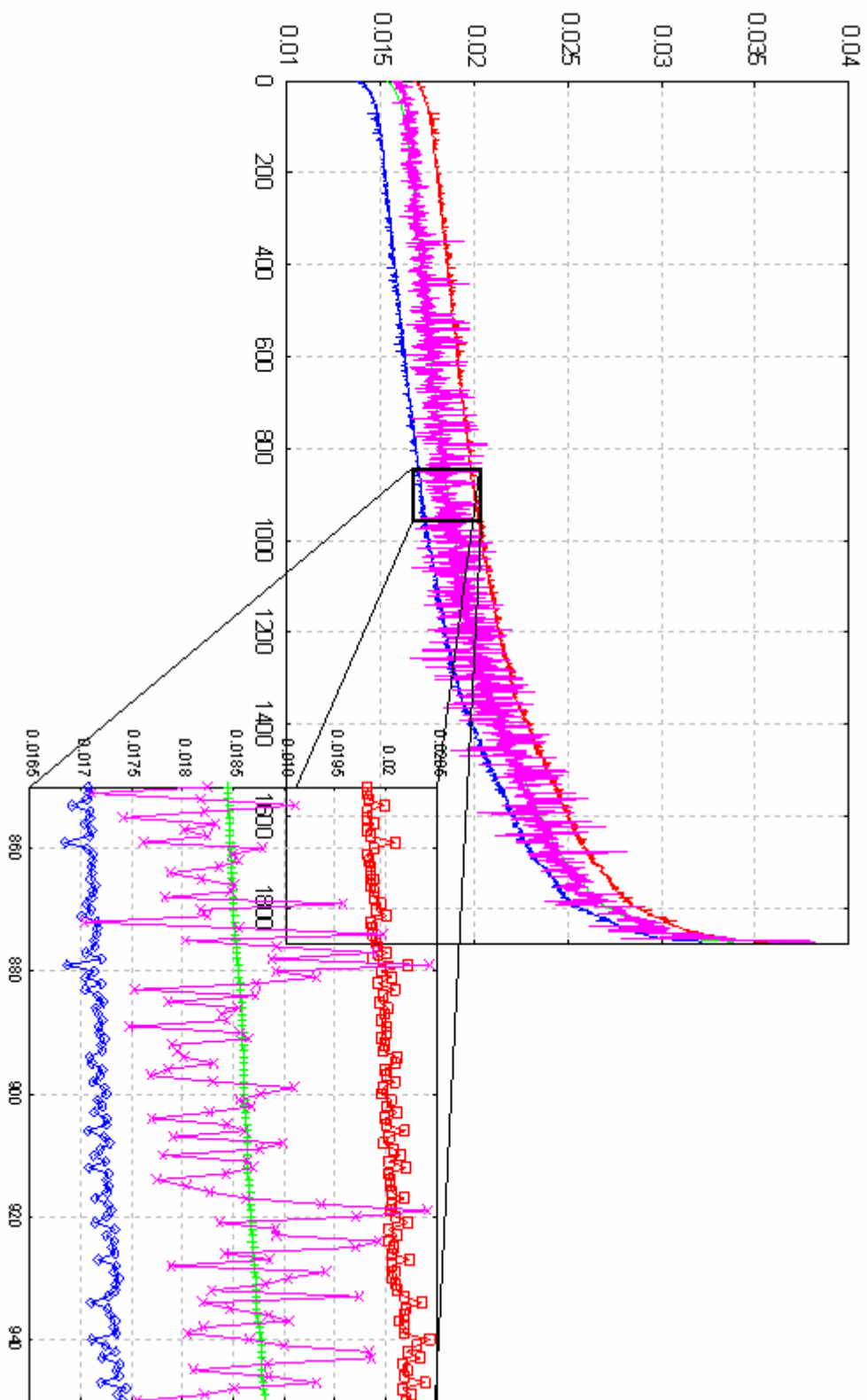
Για την περαιτέρω αξιολόγηση του μεταπροτύπου μας επιλέγουμε ένα σετ από 1875 τυχαία επιλεγμένα και τυχαίως διασκορπισμένα στο χωρίο σημεία (στα οποία το πρότυπο λογισμικό Mises δίνει απόκριση). Σε αυτά υπολογίζουμε την απόκριση του προτύπου κώδικα καθώς και του μεταπροτύπου και στη συνέχεια υπολογίζουμε το μέσο εκατοστιαίο σφάλμα του μεταπροτύπου μας.

$$mid_error[\%] = \left| \frac{y_i^{direct} - y_i^{metamodel}}{n \cdot y_i^{direct}} \right| \times 100 = 2.7\%$$

Παρακάτω παραθέτουμε τα σχήματα 8&9 που στον κατακόρυφο άξονα απεικονίζουν την απόκριση των σημείων αυτών τόσο με τον πρότυπο κώδικα (—) όσο και με το μεταμοντέλο (—) αλλά και το εύρος εντός του οποίου μπορεί να διακυμανθεί η απόκριση του μεταμοντέλου με επίπεδο εμπιστοσύνης 95%. (Υπενθυμίζουμε ότι τα σημεία είναι τυχαία επιλεγμένα και τυχαία διασκορπισμένα στο χωρίο μελέτης ενώ έχουμε επιλέξει για λόγους εποπτικούς να καταγράψουμε τα σημεία αυτά κατά αύξουσα σειρά από αυτό με την μικρότερη απόκριση του μεταμοντέλου προς αυτό με την μεγαλύτερη).



Γράφημα 8: Στον άξονα γ απεικονίζεται η απόκριση του μεταμορφέλου (πρόσγνο) με το εύρος συρής σε επίπεδο αβεβαιότητας 5% (μπλε και κόκκινο) και η πραγματική απόκριση (ροζ), ο άξονας των x είναι ο εύρος αριθμός του εκάστοτε σημείου.



Γράφημα 9: Στον άξονα y απεικονίζεται η απόκριση του μεταμοτέλου (πρόσimo) με το εύρος αυτής σε επίπεδο αβεβαιότητας 50% (μπλε και κόκκινο) και η πραγματική απόκριση (ροζ), ο άξονας των x είναι ο άξονας αριθμός του εκάστοτε σημείου.

Αρχείο **datagas.dat** (καθορίζει το χωρίο εκπαίδευσης δίνοντας τα όρια των ελευθέρων μεταβλητών, ελεύθερες μεταβλητές :10, αμετάβλητες : 0).

46 - Total number of variables			
0	0	0	x_LE (fixed)
0	0	0	y_LE (fixed)
0	0.01	0.01	R-LE
0	44.0196	44.0196	Phi_LE
0	1	1	x_TE (fixed)
0	0	0	y_TE (fixed)
0	0.005	0.005	R-TE
0	-30	-30	Phi_TE
0	5	5	number of control SS
0	4	4	degree of basis functions SS
0	0.15	0.15	interior of the knot vector SS
0	0.6	0.6	interior of the knot vector SS
0	0.83	0.83	interior of the knot vector SS
0	0.0983333	0.0983333	x3 SS
0	0.0891373	0.0891373	y3 SS
0	0.274762	0.274762	x4 SS
10	0.100	0.180	y4 SS
0	0.479048	0.479048	x5 SS
10	0.100	0.180	y5 SS
0	0.685397	0.685397	x6 SS
10	0.10	0.14	y6 SS
0	0.85	0.85	x7 SS
0	0.0692941	0.0692941	y7 SS
0	5	5	number of control PS + 3
0	4	4	degree of basis functions PS
0	0.1	0.1	interior of the knot vector PS
0	0.6	0.6	interior of the knot vector PS
0	0.82	0.82	interior of the knot vector PS
0	0.063254	0.063254	x3 PS
10	0.025	0.06	y3 PS
0	0.245238	0.245238	x4 PS
10	0.04	0.10	y4 PS
0	0.548413	0.548413	x6 PS
10	0.05	0.134	y6 PS
0	0.677619	0.677619	x7 PS
0	0.0828235	0.0828235	y7 PS
0	0.84	0.84	x8 PS
0	0.0541765	0.0541765	y8 PS
0	61	61	see shape.for nnps
0	61	61	see shape.for nnss
0	41	41	see shape.for nnle circle
0	41	41	see shape.for nnte circle
0	101	101	knodPS
0	101	101	knodSS
0	-17.3016	-17.3016	stagger
0	0.65	0.65	pitch

Συμπεράσματα .

Σε αυτήν την εργασία παρουσιάσαμε μια μεθοδολογία κατασκευής μεταπροτύπων. Παρότι πρόκειται για την πρώτη μας προσπάθεια κατασκευής μεταπροτύπων με αυτήν την μεθοδολογία , και στόχος μας δεν ήταν η βελτιστοποίηση της όλης διαδικασίας αλλά περισσότερο η επίδειξη αυτής, τα αποτελέσματα κρίνονται ικανοποιητικά, μιας και σε όλες τις παραπάνω περιπτώσεις τα μεταπρότυπα κατάφεραν να δώσουν καλές εκτιμήσεις των πραγματικών αποκρίσεων. Πλεονέκτημα των μεταπροτύπων αυτών είναι η δυνατότητα εκτίμησης του σφάλματός τους , χωρίς να απαιτείται επιπλέον πληροφορία . Παρουσιάστηκε επίσης (χωρίς να εφαρμοστεί όμως) μια πιθανή εφαρμογή της υπολογιζόμενης μεταβλητότητας για την επιλογή σημείων εκπαίδευσης τα οποία θα προστεθούν στα ήδη υπάρχοντα. Επίσης παρουσιάσαμε την εφαρμογή του κριτηρίου της μέγιστης πιθανοφάνειας για την εκτίμηση των αγνώστων παραμέτρων των συναρτήσεων συσχέτισης, το κριτήριο αυτό μας οδήγησε σε κατασκευή μεταπροτύπων με μικρότερο σφάλμα στην σχεδόν σε ολόκληρο το χωρίο μελέτης μας (χωρίο μελέτης και εκπαίδευσης εδώ ταυτίζονται, γενικά το χωρίο μελέτης μας μπορεί να είναι υποχωρίο του χωρίου εκπαίδευσης). Επιπροσθέτως δείξαμε πως μπορούμε εύκολα να επιλέξουμε σημεία εκπαίδευσης που να είναι κατά το δυνατόν πιο διασκορπισμένα στο χωρίο εκπαίδευσης. Και στο τέλος της εργασίας μας παραθέσαμε μια εφαρμογή από τον χώρο των στροβιλομηχανών. Η κατασκευή τέτοιων μεταπροτύπων μπορεί να βοηθήσει εξοικονομώντας χρόνο σε προβλήματα βελτιστοποίησης στην περίπτωση όπου το πρότυπο υπολογιστικό πρόγραμμα είναι ιδιαίτερα χρονοβόρο και απαιτούνται πολλές κλήσεις της αντικειμενικής συνάρτησης από το πρόγραμμα της βελτιστοποίησης. Δεδομένου ότι το μεταπρότυπο έχει ελάχιστες απαιτήσεις από πλευράς χρόνου CPU.

*****Συναρτήσεις συσχέτισης οικογένειας Matérn*****

Μια συχνά χρησιμοποιούμενη έκφραση για την συνάρτηση συσχέτισης αποτελεί η εξής :

$$R(\mathbf{h}) = \prod_{i=1}^d \frac{1}{\Gamma(\nu) \cdot 2^{\nu-1}} \left(\frac{2\sqrt{\nu}|h_i|}{\theta_i} \right)^\nu \cdot K_\nu \cdot \left(\frac{2\sqrt{\nu}|h_i|}{\theta_i} \right) \quad (1)$$

όπου :

- $\Gamma(\nu)$ είναι η συνάρτηση Γάμα , με $\nu = n + \frac{1}{2}$, $n \in \{0,1,2,\dots\}$
- K_ν είναι η τροποποιημένη συνάρτηση Bessel τάξης ν .

Η παραπάνω οικογένεια συναρτήσεων συσχέτισης αναφέρεται συχνά στην βιβλιογραφία με το όνομα Matérn I. Η τροποποιημένη συνάρτηση Bessel, για την περίπτωση όπου

$\nu = n + \frac{1}{2}$, $n \in \{0,1,2,\dots\}$, δίνεται από την παρακάτω σειρά :

$$K_{n+1/2}(\lambda) = e^{-\lambda} \sqrt{\frac{\pi}{2\lambda}} \cdot \sum_{k=0}^n \frac{(n+k)!}{k!(n-k)!} \frac{1}{(2\lambda)^k} \quad (2)$$

Για την ειδική περίπτωση όπου $\nu = \frac{1}{2}$ (δηλαδή $n = 0$) προκύπτει ότι :

$$K_{1/2}(\lambda) = e^{-\lambda} \sqrt{\frac{\pi}{2\lambda}} \quad \text{και συνεπώς έχουμε ότι : } R(h) = \exp \left\{ - \sum_{i=1}^d \frac{\sqrt{2}|h_i|}{\theta_i} \right\}.$$

Όσο $\nu \rightarrow \infty$ τόσο $R(h) \rightarrow \exp \left\{ - \sum_{i=1}^d \left(\frac{h_i}{\theta_i} \right)^2 \right\}$, δηλαδή στο όριο η κλάση των

συναρτήσεων Matérn I μας δίνει την ως τώρα χρησιμοποιούμενη συνάρτηση συσχέτισης με $p=2$.

Στη περίπτωση όπου η συνάρτηση συσχέτισης Matérn έχει την μορφή :

$$R(\mathbf{h}) = \prod_{i=1}^d \frac{1}{\Gamma(\nu_i) \cdot 2^{\nu_i-1}} \left(\frac{2\sqrt{\nu_i}|h_i|}{\theta_i} \right)^{\nu_i} \cdot K_{\nu_i} \cdot \left(\frac{2\sqrt{\nu_i}|h_i|}{\theta_i} \right) \quad (1')$$

λέμε ότι ανήκει στην οικογένεια Matérn II.

Βιβλιογραφία.

- 1) *The Design and Analysis of Computer Experiments.*
Συγγραφείς : *Thomas J. Santer, Brian J. Williams, William I. Notz.*
Εκδόσεις : *Springer*
- 2) *Response Surface methodology.*
Process and Product Optimization Using Designed Experiments.
Συγγραφείς : *Raymond H. Myers, Douglas C. Montgomery.*
Εκδόσεις : *John Wiley & Sons, INC*