



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ
ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΣΗΜΑΤΩΝ ΕΛΕΓΧΟΥ & ΡΟΜΠΟΤΙΚΗΣ

Μοντελοποίηση κι εκτίμηση των παραμέτρων του συστήματος
ινσουλίνης γλυκόζης του ανθρώπινου οργανισμού

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Αλέξανδρος Π. Ηλιάδης

Επιβλέπων: Γεώργιος Π. Παπαβασιλόπουλος
Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2006



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΗΜΑΤΩΝ ΕΛΕΓΧΟΥ & ΡΟΜΠΟ-
ΤΙΚΗΣ

Μοντελοποίηση και εκτίμηση των παραμέτρων του συστήματος
ινσουλίνης γλυκόζης του ανθρώπινου οργανισμού

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Αλέξανδρος Π. Ηλιάδης

Επιβλέπων: Γεώργιος Παπαβασιλόπουλος
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την -εισάγετε ημερομηνία-.

.....
Γ. Παπαβασιλόπουλος
Καθηγητής Ε.Μ.Π.

.....
Τ. Κουσιουρής
Καθηγητής Ε.Μ.Π.

.....
Σ. Τζαφέστας
Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2006.

.....
Αλέξανδρος Π. Ηλιάδης
Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών

Περίληψη

Ο διαβήτης είναι μια πολύ σοβαρή μακροχρόνια ασθένεια από την οποία πάσχουν εκατομμύρια άνθρωποι σε ολόκληρο τον κόσμο. Κατά συνέπεια, η έρευνα σχετικά με τη διάγνωση και τη θεραπεία της αποτελεί πρωταρχική μέριμνα για τους επιστήμονες. Ως προς αυτό, η μοντελοποίηση του συστήματος ελέγχου της συγκέντρωσης της γλυκόζης στο αίμα, αποτελεί βασική προτεραιότητα της έρευνας για δύο λόγους: Πρώτον, η εκτίμηση των παραμέτρων ενός μοντέλου που προσομοιώνει τη συμπεριφορά του συστήματος μπορεί να αξιοποιηθεί για τη διάγνωση της πάθησης και των ειδικών χαρακτηριστικών της και, κατά δεύτερον, γνώση του μοντέλου θα βοηθήσει καταλυτικά στη θεραπεία της ασθένειας με την κατασκευή ενός αυτοματοποιημένου συστήματος ελέγχου του σακχάρου (που συχνά αναφέρεται ως τεχνητό πάγκρεας).

Οι περισσότερες δημοφιλείς προσπάθειες μοντελοποίησης του συστήματος προτείνουν την προσέγγιση του συστήματος από γραμμικά, στάσιμα μοντέλα μίας εισόδου (συγκέντρωση ινσουλίνης) και μίας εξόδου (συγκέντρωση γλυκόζης). Στην παρούσα εργασία, εισάγεται και μελετάται ένα μη γραμμικό, περίπου στάσιμο παραμετρικό μοντέλο μίας εισόδου και μίας εξόδου. Συγκεκριμένα, το μοντέλο αποτελείται από 2 συμμετρικά τμήματα στα οποία η είσοδος αρχικά φιλτράρεται από ένα γραμμικό σύστημα, η έξοδος του οποίου απεικονίζεται μη γραμμικά στην έξοδο του τμήματος, ενώ η έξοδος του συστήματος προκύπτει από την άθροιση των εξόδων των 2 τμημάτων. Παράλληλα μοντελοποιείται η επίδραση εξωγενών παραγόντων στην ποσότητα ινσουλίνης στον οργανισμό (disturbance-διαταραχής) με ένα Auto-Regressive AR μοντέλο. Η συνολική ποσότητα γλυκόζης στον οργανισμό είναι το άθροισμα της εξόδου του μοντέλου ινσουλίνης-γλυκόζης και της διαταραχής.

Σκοπός της παρούσας εργασίας είναι η εκτίμηση των παραμέτρων του μοντέλου ινσουλίνης-γλυκόζης διαδικασία θεμελιώδης τόσο για τη διάγνωση της αρρώστιας όσο και για τον έλεγχο της γλυκόζης σε επιθυμητά επίπεδα. Τα δεδομένα μας προέρχονται από πραγματικές μετρήσεις του επιπέδου γλυκόζης στο αίμα σε ασθενείς οι οποίες πραγματοποιούνταν κάθε 5 λεπτά, καθώς κι οι χρονικές στιγμές και η ποσότητα των ενέσεων ινσουλίνης.

Στο πλαίσιο αυτό, εξετάζουμε και εφαρμόζουμε τεχνικές εκτίμησης για μη γραμμικά συστήματα με βάση το κριτήριο των ελαχίστων τετραγώνων καθώς και τη χρησιμότητα τροποποίησης του μοντέλου. Υλοποιούμε κι συγκρίνουμε δημοφιλείς αλγόριθμους που υπάρχουν στη βιβλιογραφία (Gauss-Newton, Levenberg-Marquardt, Extended Kalman Filtering) και παραθέτουμε προβλήματα τα οποία συναντήσαμε κατά την εφαρμογή τους στα πραγματικά δεδομένα και πως τροποποιήσαμε τους αλγόριθμους ώστε να τα ξεπεράσουμε. Επιπλέον αναφέρουμε την τελική μορφή των αλγορίθμων που χρησιμοποιήσαμε και περιγράφουμε αναλυτικά και συγκρίνουμε τα αποτελέσματα τα οποία πήραμε με τις διάφορες μεθόδους.

Λέξεις Κλειδιά

Διαβήτης
Ινσουλίνη
Γλυκόζη
Εκτίμηση παραμέτρων
Μη γραμμικό σύστημα
Μη γραμμικό μοντέλο
Extended Kalman Φίλτρο
Levenberg-Marquardt Μέθοδος

Abstract

Diabetes is a very serious long-lasting disease counting million of patients around the world. As a consequence, research considering the diagnoses and cure of diabetes is constantly gaining more and more attention among scientists. On that basis, modelling of the glucose control system in blood, is of major priority in the field for two important reasons: Firstly, parameter estimation of a simulating model for the system's behaviour could be exploited for the diagnoses of diabetes and secondly, knowledge of the model could contribute to the cure through an automated glucose control system known as "artificial pancreas".

Until now the most popular attempts for the modelling of the system propose a linear stationary single input (insulin infusions)-single output (glucose concentration) abbreviated SISO, approximation. In this thesis, we introduce a non-linear SISO parametric model. In particular, the model consists of two symmetric compartments in each of which the input is initially filtered by a linear system, whose output is transformed nonlinearly. The insulin-glucose's output is the sum of the aforementioned nonlinearities of the two compartments. On the same time we model the effect of all exogenous factors to the system namely the disturbance, of the system as an Auto-Regressive (AR) model. Blood Glucose concentration is the sum of the output of the system plus the disturbance.

Our goal in this thesis is to estimate the parameters of the insulin-glucose model a fundamental task not only for the diagnosis of the disease but for the control of the glucose to acceptable level as well. The data that we use originate from actual measurements on the blood glucose level and the insulin infusions on human subjects, taken every five minutes.

In this context, we analyze and apply optimal estimation techniques for non-linear systems for the minimization of the least square criterion and additionally the usefulness of the implementation for a real time application. We adjust to the current problem, implement and compare popular algorithms such as Gauss-Newton, Levenberg-Marquardt and Extended Kalman Filtering and we give the obstacles for each implementation (if any) when it comes to the real data and how we modified the algorithms to overpass them. Last but not the least we give the final form for the algorithms used and we comment extensively the results of the various methods.

KEYWORDS

Diabetes
Insulin
Glucose
Parameter Estimation
Nonlinear system
Nonlinear model
Extended Kalman Filter
Levenberg-Marquardt Method

Ευχαριστίες

Η συγγραφή της παρούσας διπλωματικής εργασίας έγινε υπό την επίβλεψη του κ. Γεώργιου Παπαβασιλόπουλου, καθηγητή Ε.Μ.Π., τον οποίο θα ήθελα να ευχαριστήσω θερμά για την υποστήριξη και τις καθοριστικές παρατηρήσεις του όλο αυτό το διάστημα.

Επιπλέον, θα ήθελα να ευχαριστήσω τον κ. Βασίλη Μαρμαρέλη, καθηγητή στο πανεπιστήμιο USC, ο οποίος παρείχε όχι μόνο τα βασικά στοιχεία της μοντελοποίησης αλλά και σημαντικές πληροφορίες σχετικά με την πρακτική του προβλήματος καθοδηγώντας με στην πρωτόγνωρη αυτή περιοχή έρευνας.

Τέλος θα ήθελα να ευχαριστήσω τον κ.Μιχάλη Μαρκάκη, Διπλωματούχο Ηλεκτρολόγο Μηχανικό κι Μηχανικό Υπολογιστών, τόσο για τις εύστοχες υποδείξεις του όσο και για τις σημαντικές παρατηρήσεις του.

Ηλιάδης Αλέξανδρος

Περιεχόμενα

1	Εισαγωγικά για τον Διαβήτη- Έρευνα στην περιοχή	5
1.1	Εισαγωγικά Για τον Διαβήτη-Μηχανισμοί του Διαβήτη	5
1.1.1	Γενικά για την Ινσουλίνη	5
1.1.2	Σακχαρώδης διαβήτης και η παθολογία του	6
1.1.3	Είδη του διαβήτη	7
1.2	Ανασκόπηση της έρευνας στην περιοχή	7
2	Τοποθέτηση του Προβλήματος	10
2.1	Γενικά για το μοντέλο	10
2.2	Τυπική Μορφή του μοντέλου	13
2.3	Εισαγωγικά Για Το Πρόβλημα-Αρχική Θεώρηση Του Προβλήματος .	16
2.4	Πλήρης Θεώρηση Του Προβλήματος-Προσαρμογή στα Πραγματικά Δεδομένα	22
3	Αλγόριθμοι Βελτιστοποίησης	24
3.1	Γραμμικά ελάχιστα τετράγωνα	24
3.2	Εκτίμηση για μη γραμμικά προβλήματα	26
3.2.1	Αναγκαίες κι ικανές συνθήκες για τοπικά ελάχιστα χωρίς περιορισμούς	27
3.2.2	Γενική Μορφή και Στοιχεία Αλγορίθμων Βελτιστοποίησης . .	27
3.2.3	Κριτήρια Τερματισμού	29
3.2.4	Έρευνα γραμμής ή ελαχιστοποίηση συνάκρτησης μιας μεταβλητής	29
3.2.5	Γενικά για την κατεύθυνση έρευνας	31
3.3	Μέθοδος της ταχύτερης κλίσης-Steepst Descent Method	31
3.3.1	Γενική μορφή αλγόριθμου Steepst Descent	31
3.3.2	Μορφή αλγορίθμου στην συγκεκριμένη εφαρμογή	32
3.4	Newton's Method	36
3.4.1	Γενικά για την μέθοδο του Newton	36
3.4.2	Η μέθοδος Newton στο συγκεκριμένο πρόβλημα	37
3.5	Levenberg-Marquardt	37
3.5.1	Γενικά για την μέθοδο Levenberg-Marquardt	37
3.5.2	Μέθοδος Levenberg-Marquardt στο συγκεκριμένο πρόβλημα	38
4	Φίλτρο Kalman	39
4.1	Διακριτό Φίλτρο Kalman	39
4.1.1	Υπολογιστική Προέλευση του Διακριτού Φίλτρου	39
4.1.2	Αλγόριθμος Διακριτού Φίλτρου Kalman	40

4.2	EXTENDED KALMAN FILTER	41
4.2.1	Εισαγωγή	41
4.2.2	Υπολογιστική Προέλευση του Φίλτρου	41
4.3	Προσαρμογή Στο Συγκεκριμένο Πρόβλημα	44
4.4	Διαφοροποιήσεις Στον Αλγόριθμο	48
4.4.1	Time Update Equations	48
4.4.2	Measurement Update Equations	49
5	Ζητήματα Που Παρουσιάστηκαν	50
5.1	Αστάθεια	50
5.2	Μέγεθος Δεδομένων	56
6	Αποτελέσματα	58
6.1	Δείγματα 350-520	58
6.1.1	Τάξη Μοντέλου 25	59
6.1.2	Τάξη Μοντέλου 20	70
6.1.3	Σύγκριση της Τάξης των 2 μοντέλων	77
6.2	Δείγματα 1-150	77
6.2.1	Τάξη Μοντέλου 25	77
6.2.2	Τάξη Μοντέλου 20	88
6.2.3	Σύγκριση της Τάξης των 2 μοντέλων	94

Κεφάλαιο 1

Εισαγωγικά για τον Διαβήτη-Έρευνα στην περιοχή

Στο κεφάλαιο αυτό συνοψίζουμε τα κύρια χαρακτηριστικά του διαβήτη και δίνουμε αρκετά στοιχεία για την πάθηση αυτή. Προκειμένου να γίνει αυτή η εισαγωγή όσο το δυνατό πιο κατανοητή παραθέτουμε αρκετά στοιχεία για τη φυσιολογία της ασθένειας. Αρχικά μετά από μία σύντομη παρουσίαση των χαρακτηριστικών της νόσου στην πρώτη παράγραφο περιγράφουμε την βασική λειτουργία των 2 ουσιών που θα μας απασχολήσουν την γλυκόζη και την ινσουλίνη και την μεταξύ τους σχέση. Στην επόμενη παράγραφο συνοψίζονται τα συμπτώματα του διαβήτη και τρόποι θεραπείας του. Στο τέλος της ενότητας ακολουθεί μια γρήγορη ανασκόπηση της έρευνας στην περιοχή αυτή και τα ήδη υπάρχοντα αποτελέσματα.

1.1 Εισαγωγικά Για τον Διαβήτη-Μηχανισμοί του Διαβήτη

Ο σακχαρώδης διαβήτης είναι μια χρόνια πάθηση η οποία οφείλεται στην μειωμένη ικανότητα των κυτάρων να απορροφήσουν γλυκόζη. Το φαινόμενο αυτό οφείλεται στο ότι ο οργανισμός των ασθενών παράγει ελαττωμένη ποσότητα ινσουλίνης. Αυτό που θα επιχειρήσουμε σε αυτή την παράγραφο είναι να παρουσιάσουμε κάποια γενικά χαρακτηριστικά της λειτουργίας του οργανισμού ώστε να γίνει καλύτερα κατανοητή η φυσιολογία της ασθένειας.

1.1.1 Γενικά για την Ινσουλίνη

Το πάγκρεας ανατομικά αποτελείται από δύο διαφορετικούς τύπους ιστών. Ο ένας εξ'αυτών (που ονομάζονται (κύτταρα)) νησίδια του Langerhans εκρίνουν και μία ορμόνη την ινσουλίνη κατευθείαν προς το αίμα. Η ινσουλίνη είναι ένα μεγάλο πολυπεπτίδιο και για να μπορέσει να ασκήσει την επίδρασή της σε κάποιο κύτταρο πρέπει να συνδεθεί με μια πρωτεΐνη υποδοχέα στην κυτταρική μεμβράνη. Για την έναρξη των επιδράσεων της ινσουλίνης στα κύτταρα στόχους η ινσουλίνη συνδέεται πρώτα με μια δεκτική πρωτεΐνη που βρίσκεται στην κυτταρική μεμβράνη. Αυτό που πρέπει να τονίσουμε είναι ότι οι επακόλουθες επιδράσεις προκαλούνται από τον υποδοχέα και όχι από την ινσουλίνη. Ο υποδοχέας της ινσουλίνης αποτελείται από τέσσερις

ξεχωριστές υπομονάδες που συνδέονται μεταξύ τους ,δύο άλφα υπομονάδες που βρίσκονται έξω από την κυτταρική μεμβράνη και δύο βήτα υπομονάδες που διαπερνούν τη μεμβράνη και το ένα άκρο τους προεξέχει μέσα στο κυτόπλασμα του κυττάρου.Η ινσουλίνη συνδέεται με τις άλφα υπομονάδες που βρίσκονται στο εξωτερικό του κυττάρου .Εξαιτίας όμως των δεσμών με τις βήτα υπομονάδες προκαλεί μια αλυσωτή αντίδραση στο εσωτερικό του κυττάρου η οποία (αν και στο μεγαλύτερο μέρος της μας είναι παντελώς άγνωστη)έχει σαν αποτέλεσμα οι μεμβράνες των μυικών κυττάρων ,των λιποκυττάρων και πολλών άλλων τύπων κυττάρων σε ολόκληρο το σώμα να γίνονται να γίνονται σε πολύ μεγάλο βαθμό διαβατές από γλυκόζη.Με άλλα λόγια λοιπόν η ινσουλίνη είναι αυτή που βοηθάει την γλυκόζη να μπει στο εσωτερικό των κυττάρων όπου θα μπορέσει να χρησιμοποιηθεί.

1.1.2 Σακχαρώδης διαβήτης και η παθολογία του

Στις περισσότερες περιπτώσεις ο σακχαρώδης διαβήτης οφείλεται σε ελαττωμένη έκκριση ινσουλίνης από τα κύτταρα των νησιδίων του langerhans. Η κληρονομικότητα συνήθως διαδραματίζει τον κύριο ρόλο για το ποιος θα εμφανίσει ζακχαρώδη διαβήτη και ποιος δεν θα εμφανίσει την νόσο.Σε μερικές περιπτώσει αυτό γίνεται με την αύξηση της ευαισθησίας των βήτα κυττάρων σε ιούς,είτε με την ανάπτυξη αυτοάνοσων αντισωμάτωνέναντι των βήτα κυττάρων ,με αποτέλεσμα την καταστροφή τους(ο οργανισμός δηλαδή αναγνωρίζει τα βήτα κύτταρα ως ξένα και παράγει αντισώματα για να τα σκοτώσει).Σε άλλες περιπτώσεις μπορεί να υπάρχει μια απλή τάση για εκφύλιση των βήτα κυττάρων.Η παχυσαρκία διαδραματίζει επίσης κάποιο ρόλο στην ανάπτυξη του σακχαρώδους διαβήτη με την κλινική του μορφή. Ένας απο τους λόγους είναι ότι κατά την παχυσαρκία επέρχεται μείωση του αριθμού των υποδοχέων ινσουλίνης στα κύτταρα-στόχους σε ολόκληρο το σώμα , και με αυτό τον τρόπο τα ποσά της ινσουλίνης που διατίθενται καθίσταντια λιγότερο αποτελεσματικά όσον αφορά τις συνήθειες της επιδράσεις στον μεταβολισμό.

Τα περισσότερο παθολογικά χαρακτηριστικά του σακχαρώδους διαβήτη μπορούν να αποδοθούν σε μια από τις ακόλουθες μείζονες επιδράσεις της έλλειψης ινσουλίνης:

- ελαττωμένη χρησιμοποίηση γλυκόζης από τα κύτταρα του σώματος,με συνακόλουθη αύξηση της συγκεντρώσεως της γλυκόζης του αίματος μέχρι τα 300 έως 1.200 mg/dL αίματος
- μεγάλη αύξηση της κινητοποίησης λιπών από όλες τις λιπαροθήκες του σώματος(αφού η γλυκόζη δεν μπορεί να μπει στα κύτταρα) με αποτέλεσμα τον παθολογικό μεταβολισμό λίπους και την εναπόθεση χοληστερόλης στα τοιχώματα των αρτηριών.
- εξάντληση των πρωτεϊνών από τους ιστούς του σώματος

Οι στόχοι της θεραπείας του διαβήτη είναι οι ίδιοι σχεδόν σε όλους τους τυπους διαβήτη και δεν είναι άλλοι από την διατήρηση των συγκεντρώσεων της γλυκόζης του αίματος σε σχεδόν φυσιολογικά όρια. Οι συγκεντρώσεις της γλυκόζης στο αίμα πριν από τα γεύματα πρέπει να είναι 80-140 mg/dl(και μπορεί να είναι 70-120 mg/dl όταν ο ασθενής προσπαθεί με επιμονή να πετύχει επίπεδα σχεδόν ευγλυκαιμίας).Επιπλέον στους ασθενείς που παίρνουν ινσουλίνη , ο στόχος για την ινσουλίνη πριν από το ελαφρό βραδυνο γεύμα θα πρέπει να είναι 100-140 mg/dl ώστε να αποφεύγεται η νυχτερινή υπογλυκαιμία. Τα όρια αυτά καθώς και οι μετρήσεις (για την ποσότητα της γλυκόζης στο αίμα

που πραγματοποιούν μόνοι τους οι ασθενείς) μπορούν να διαφέρουν ανάλογα με την ηλικία των ασθενών και με το στάδιο που βρίσκεται η νόσος.

1.1.3 Είδη του διαβήτη

- *Διαβήτης τύπου 1* Οι ασθενείς με διαβήτη τύπου 1 έχουν γενετική επιρρέπεια προς ανάπτυξη της νόσου, είναι ευαίσθητοι σε διάφορα αίτια (ιογενή, περιβαλλοντικά ή τοξικά) που διεγείρουν την ανοσολογική καταστροφή του παγκρεατικού κυττάρου Β. Όταν καταστραφεί το 80-90% των κυττάρων Β αναπτύσσεται υπεργλυκαιμία και τίθεται η διάγνωση του διαβήτη. Σε πρόσφατης εγκατάστασης διαβήτη τύπου 1 έχει επιχειρηθεί η πρόληψη της εξέλιξης της ανοσολογικής καταστροφής των κυττάρων Β με την εφαρμογή ανοσοκατασταλτικής θεραπείας, η οποία βρέθηκε ότι είναι αποτελεσματική αλλά ο διαβήτης επανέρχεται αμέσως μόλις διακοπεί το ανοσοκατασταλτικό φάρμακο και, επιπλέον, θεωρείται ότι οι παρενέργειες και οι κίνδυνοι της μακροχρόνιας ανοσοκαταστολής υπερβαίνουν τους κινδύνους από τον διαβήτη. Εξ' ορισμού, οι ασθενείς με διαβήτη τύπου 1 χρειάζονται ισόβια θεραπεία με ινσουλίνη. Τα υπάρχοντα σκευάσματα ινσουλίνης διακρίνονται σε βραχείας μακράς και ενδιάμεσης δράσης ανάλογα με το πότε αρχίζει η επίδραση της χορηγούμενης ινσουλίνης στον ασθενή.
- *Διαβήτης τύπου 2* Ο διαβήτης τύπου 2 έχει ισχυρή γενετική προδιάθεση αν και η γενετική του βάση είναι άγνωστη. Πολλοί ασθενείς με διαβήτη τύπου 2 είναι ασυμπτωματικοί και η διάγνωση του διαβήτη τους τίθεται με την ευκαιρία είτε αδρού ελέγχου αποκλεισμού της νόσου είτε εξέτασης για άλλο, άσχετο παθολογικό πρόβλημα. Οι ασθενείς που αναπτύσσουν διαβήτη τύπου 2 παρουσιάζουν περιφερική αντίσταση στην ινσουλίνη, σε συνδυασμό με ανεπαρκή έκκριση ινσουλίνης από τα κύτταρα Β του παγκρέατος. Οι περισσότεροι (90% των) από τους ασθενείς που αναπτύσσουν διαβήτη τύπου 2 είναι παχύσαρκοι και η ίδια η παχυσαρκία συνοδεύεται από αντίσταση στην ινσουλίνη επιδεινώνοντας περισσότερο τη διαβητική κατάσταση. Για πολλούς ασθενείς με διαβήτη τύπου 2, και ιδιαίτερα για τους παχύσαρκους, οι πρώτοι τρόποι θεραπείας που πρέπει να δοκιμαστούν είναι η δίαιτα και η άσκηση. Εάν δεν καρποφορήσει αυτός ο τρόπος τότε οδηγούνται σε θεραπεία της οποίας ο στόχος είναι αυτός που περιέγραψα και προηγουμένως.
- *Ο Διαβήτης των νέων* Αρχίζει σε ώριμη ηλικία και είναι μορφή διαβήτη που προσβάλλει πολλές γενεές της ίδιας οικογένειας με πρώιμη εγκατάσταση της νόσου. Το σύνδρομο αυτό είναι δυνατό να προκαλέσουν διάφορα γενετικά ελλείματα.
- *Ο Διαβήτης της εγχυμοσύνης* Είναι μια μορφή ελαττωμένης ανοχής της γλυκόζης, που αναπτύσσεται στη διάρκεια της κύησης όταν η κληρονομική ελάττωση της εφεδρείας των κυττάρων Β συνεπάγεται αδυναμία του παγκρέατος να εκκρίνει επαρκείς ποσότητες ινσουλίνης ώστε να υπερνικηθεί η δημιουργούμενη από τις ορμόνες του πλακούντα αντίσταση στην ινσουλίνη

1.2 Ανασκόπηση της έρευνας στην περιοχή

Η ανάγκη λοιπόν συνεχούς παρακολούθησης του επιπέδου γλυκόζης στον οργανισμό έχει υπάρξει το αντικείμενο εκτεταμένης έρευνας. Οι ασθενείς με διαβήτη αντιμετω-

πίζουν ένα διττό κίνδυνο όταν πρέπει να αποφασισθεί το επιθυμητό επίπεδο γλυκόζης στον οργανισμό. Από την μια μεριά αυστηρός έλεγχος των επιπέδων γλυκόζης έχει αποδειχθεί καθοριστικός στην αποφυγή μακροπρόθεσμων επιπλοκών, από την άλλη μεριά όμως πολύ σφιχτός έλεγχος οδηγεί σε αυξημένο κίνδυνο ιατρογενούς υπογλυκαιμίας.

Η ανάπτυξη συστημάτων συνεχούς παρακολούθησης της γλυκόζης έχει δώσει την δυνατότητα όχι μόνο να ανιχνεύουμε υπογλυκαιμικά επεισόδια αλλά και να κάνουμε προβλέψεις χρησιμοποιώντας θεωρία βελτίστων εκτιμητών (optimal estimation theory) που θα επιτρέπουν στον ασθενή να δρα εκ των προτέρων ώστε να αποφεύγει εντελώς το επεισόδιο.

Έχοντας λοιπόν υπόψιν μας τα παραπάνω φαίνεται φυσικό επακόλουθο ότι απώτερος στόχος της συνεχούς παρακολούθησης γλυκόζης στο αίμα δεν είναι άλλος από το να παρέχει ένα σήμα από τους αισθητήρες μέτρησης γλυκόζης το οποίο θα μπορεί να χρησιμοποιηθεί σε ένα κλειστού βρόχου "τεχνητό πάγκρεας".

Απώτερος και τελικός σκόπος δηλαδή θα ήταν να καταστεί δυνατό να χρησιμοποιηθούν οι μετρήσεις των αισθητήρων παρακολούθησης ώστε να μοντελοποιηθεί το σύστημα ινσουλίνης-γλυκόζης του οργανισμού και να μπορέσει να αναπτυχθεί ένα σύστημα αυτομάτου ελέγχου για το σύστημα γλυκόζης ινσουλίνης του οργανισμού (αυτό που είχαμε αναφέρει προηγουμένως ως "τεχνητό πάγκρεας").

Προγενέστερες κλινικές μελέτες έχουν κινηθεί σε διαφορετικές κατευθύνσεις προσπαθώντας να προβλέψουν τις πιθανές τιμές της γλυκόζης στα μετά των γευμάτων διαστήματα χρησιμοποιώντας σαν δεδομένα την εκάστοτε κατάσταση του οργανισμού των ασθενών και την ευαισθησία του σε ινσουλίνη καθώς και τις αποκρίσεις του οργανισμού των ασθενών σε παρόμοια γεύματα τα οποία είχαν προηγηθεί.

Η πρώτη γενιά συσκευών τεχνητού παγκρέατος περιελάμβαναν σαν παραμέτρους δύο χειριζόμενες μεταβλητές (ινσουλίνη -γλυκόζη) και μη γραμμικές συναρτήσεις λάθους (διαφορά μεταξύ της επιθυμητής και της μετρούμενης τιμής γλυκόζης) για να ελαχιστοποιήσουν την υπεργλυκαιμία προσπαθώντας ταυτόχρονα να αποφύγουν την υπογλυκαιμία. Οι χρονικές καθυστερήσεις μεταξύ των ενέσεων γλυκόζης και των μετρήσεων ινσουλίνης για τα ενδοφλέβια αυτά συστήματα ήταν σχετικά μικρές.

Η πρόοδος όμως τόσο σε μηχανισμούς που προσφέρουν συνεχή απεικόνιση τιμών γλυκόζης, όσο και σε ανάλογα ινσουλίνης που δρουν πολύ γρήγορα καθώς και μια αναπτυσσόμενη αγορά μικρο-αντλιών προωθούν την δημιουργία ενός κλειστού-βρόχου "τεχνητού παγκρέατος".

Ένα βασικό όμως πρόβλημα το οποίο παραμένει ακόμη αφορά τον χειρισμό της "άγνωστης" διαταραχής γλυκόζης η οποία υπεισέρχεται στις μετρήσεις παντός είδους και έχει ως κύρια αιτία της τα γεύματα. Μέχρι στιγμής έχουν προταθεί αρκετές προσεγγίσεις ώστε να βασίζεται η εκάστοτε έγχυση ινσουλίνης στο προβλεπόμενο αποτέλεσμα της επίδρασης ενός γεύματος στην συνακόλουθη αύξηση της γλυκόζης. Το γεγονός αυτό αποτελεί και βασικό μειονέκτημα πολλών μεθόδων οι οποίες βασίζονται τα αποτελέσματά τους και τις προτεινόμενες χορηγήσεις ινσουλίνης στην συμπεριφορά του οργανισμού δεδομένου του ότι το επίπεδο γλυκόζης του χωρίς τις ενέσεις θα παρέμενε σταθερό.

Το πως θα προβλέπεται όμως το αποτέλεσμα αυτό, δηλαδή της επίδρασης ενός γεύματος, το οποίο είναι φυσικά διαφορετικό για τον εκάστοτε ασθενή, ακόμη είναι άγνωστο και αναζητούνται καλύτερα μοντέλα από αυτά που ήδη έχουν προταθεί. Επιπλέον, όπως ανέφερα και προηγουμένως, έντονες διακυμάνσεις σε τιμές της γλυκόζης θα μπορούσαν να δημιουργηθούν ακόμη και από πλήθος εξωτερικών παραγόντων (άγχος, σωματική άσκηση κ.α.). Η μοντελοποίηση λοιπόν αυτή του "dis-

turbance”, θα αποτελέσει σημαντικό ζήτημα στην υπάρχουσα εργασία. Σε αντίθεση με τις λοιπές υπάρχουσες προσεγγίσεις το μοντέλο το οποίο θα χρησιμοποιήσουμε βασίζεται αποκλειστικά στα δεδομένα για τον εκάστοτε ασθενή, που δεν είναι άλλα από τις εγχύσεις ινσουλίνης και τις μετρήσεις γλυκόζης του αίματος και προσπαθεί να μοντελοποιήσει όλους τους εξωγενείς παράγοντες με ένα AR μοντέλο. Θα ασχοληθούμε αναλυτικότερα με το μοντέλο αυτό σε επόμενο κεφάλαιο.

Συνοψίζοντας τα παραπάνω, το πλαίσιο στο οποίο κινείται η δημιουργία του τεχνητού παγκρέατος περιλαμβάνει τους εξής τομείς :

- Εκτίμηση των παραμέτρων του συστήματος αλληλεπίδρασης ινσουλίνης γλυκόζης και δυναμική μοντελοποίηση των εξωγενών διαταραχών γλυκόζης (Disturbance)
- Υλοποίηση αυτοματοποιημένου ελέγχου με βάση το μοντέλο αυτό με χρήση μικροαντλίας ινσουλίνης (insulin micropump) με ταυτόχρονη εκτίμηση των παραμέτρων του μοντέλου (καθώς το σύστημα είναι γενικά μη στάσιμο) εφαρμόζοντας κάποιο είδος προσαρμοστικού ελέγχου (adaptive control).

Κεφάλαιο 2

Τοποθέτηση του Προβλήματος

Στην ενότητα αυτή θα παρουσιάσουμε με ακρίβεια το πρόβλημα με το οποίο θα ασχοληθούμε. Αρχικά στην πρώτη ενότητα θα προσπαθήσουμε να δείξουμε όσο το δυνατό πιο περιεκτικά και περιληπτικά τους λόγους για τους οποίους, το μοντέλο πάνω στο οποίο θα τοποθετήσουμε το πρόβλημα αναμένουμε να έχει καλύτερη συμπεριφορά από τα είδη χρησιμοποιούμενα μοντέλα. Στην συνέχεια θα δώσουμε την γενική μορφή του μοντέλου και θα οπτικοποιήσουμε την συμπεριφορά του και στην τρίτη ενότητα θα τοποθετήσουμε το πρόβλημα και θα παρουσιάσουμε αναλυτικότερα ορισμένα σημεία του.

Το μοντέλο στο οποίο θα βασιστούμε κα με το οποίο θα ασχοληθούμε οφείλεται στην ερευνητική ομάδα του καθηγητή του University of Southern California κ.Μαρμαρέλη.

2.1 Γενικά για το μοντέλο

Κάθε φορά που κάνουμε μια ένεση ινσουλίνης σε κάποιο διαβητικό ή αντίστοιχα υπάρχει κάποια ποσότητα ινσουλίνης στο αίμα, τίθεται σε κίνηση μια ολόκληρη μεταβολική διεργασία η οποία έχει ως τελικό αποτέλεσμα την αλλαγή στην συγκέντρωση της γλυκόζης στο αίμα.

Οι βασικοί φυσιολογικοί μηχανισμοί έκκρισης ινσουλίνης και μεταβολισμού γλυκόζης έχουν μελετηθεί εκτεταμένα μέσω διαμερισματικών μοντέλων αλλά μέχρι στιγμής δεν έχει αναπτυχθεί κάποιο μοντέλο που να ερμηνεύει σε ικανοποιητικό βαθμό τα υπάρχοντα δεδομένα.

Ένας από τους λόγους για αυτό είναι ότι πολλαπλοί ρυθμιστικοί μηχανισμοί υπεισέρχονται στις αλληλεπιδράσεις ινσουλίνης-γλυκόζης (ινσουλινοεξαρτώμενη χρησιμοποίηση γλυκόζης στους μυς και στον λιπώδη ιστό,έγχυση ινσουλίνης από το πάγκρεας εξαρτώμενη από την ποσότητα γλυκόζης ,παραγωγή γλυκόζης από το συκώτι) .Επιπλέον, υπάρχει χρησιμοποίηση γλυκόζης στο κεντρικό νευρικό σύστημα και στα ερυθρά κύτταρα του αίματος η οποία εξαρτάται από την συγκέντρωση της ινσουλίνης .Είναι επίσης γνωστό ότι η συγκέντρωση ελεύθερων λιπαρών οξέων κι επινεφρίνης επηρεάζουν την μη-γραμμική δυναμική σχέση μεταξύ ινσουλίνης και γλυκόζης.

Έχοντας να αντιμετωπίσουν αυτό το μάλλον αρκετά περίπλοκο μοντέλο πολλοί ερευνητές κατέφυγαν σε απλουστεύσεις προκειμένου να επιτύχουν κάποιο μοντέλο απλό μεν αλλά που θα έδινε χρησιμοποιήσιμα αποτελέσματα.Το πιο επιτυχημένο από αυτά ήταν το λεγόμενο minimal μοντέλο το οποίο προσέφερε χρήσιμα αποτελέσματα

στον τομέα εξειδικευμένων κλινικών τεστ "ανοχής στην γλυκόζη", αλλά παραμένει ανεπαρκές για να περιγράψει την πλήρη πολυπλοκότητα του συστήματος υπο φυσιολογικές συνθήκες λειτουργίας του οργανισμού.

Η προσέγγιση που χρησιμοποιήθηκε από την ερευνητική ομάδα του κ.Μαρμαρέλη δεν λαμβάνει υπόψιν τους περιορισμούς που λαμβάνονται από τα διαμερισματικά μοντέλα και υιοθετεί μια προσέγγιση για την περιγραφή του συστήματος η οποία βασίζεται αποκλειστικά στα data. Οι παραπάνω πολλοί και αρκετά πολύπλοκοι μηχανισμοί μπορούν να χωριστούν σε δύο κατηγορίες μηχανισμών:

1. *Γλυκόληψη* Αφορά την βοηθούμενη από την ινσουλίνη κατανάλωση γλυκόζης από διάφορους ιστούς κι όργανα.
2. *Γλυκογένεση* Αφορά την δημιουργία νέας γλυκόζης από διάφορους ιστούς κι όργανα σε απάντηση στην μείωση της γλυκόζης.

Οι μηχανισμοί αυτοί μπορούν να περιγραφούν δυναμικά από ένα σύνολο (ελάχιστο σύνολο) γραμμικών φίλτρων τα οποία μπορούν να δώσουν την δυναμική του συστήματος κι ονομάζονται PDMs (Principal Dynamic Modes). Η ακριβής μορφή των PDMs είναι διαφορετική για κάθε άνθρωπο. Το σημαντικό το οποίο πρέπει να τονίσουμε στην προσέγγιση αυτή είναι ότι τα PDMs παρουσιάζουν ακριβή αντιστοιχία στην ερμηνεία τους σε σχέση με τους μηχανισμούς γλυκογένεσης κι γλυκόληψης.

Συγκεκριμένα, μετά από επεξεργασία σε πολλαπλά άτομα της σχέσης ινσουλίνης-γλυκόζης (ενέσεις ινσουλίνης κι έλεγχο της ποσότητας γλυκόζης στο αίμα κάθε 5 λεπτά), το αποτέλεσμα ήταν ακριβώς 2 PDMs από τα οποία το πρώτο αντιπροσωπεύει την βοηθούμενη από την ινσουλίνη λήψη γλυκόζης από τον οργανισμό (γλυκόληψη) και το δεύτερο αντιπροσωπεύει την δημιουργούμενη από τον οργανισμό γλυκόζη εξαιτίας της ινσουλίνης (γλυκογένεση).

Το πλήρες μοντέλο λοιπόν το οποίο προαναφέραμε και το οποίο θα χρησιμοποιήσουμε περιγράφεται από δύο τμήματα καθένα εκ των οποίων είναι ένα απλό L-N φίλτρο δηλαδή αποτελείται από ένα γραμμικό φίλτρο (PDM γλυκόληψης, PDM γλυκογένεσης) του οποίου η έξοδος αποτελεί την είσοδο σε μια μη γραμμική απεικόνιση.

Η έξοδος του συστήματος προκύπτει από την άθροιση των εξόδων των 2 τμημάτων. Σαν γενικό διάγραμμα για αυτή την μοντελοποίηση που κάναμε θα μπορούσαμε να είχαμε το εξής :



Σχήμα 2.1: General Model

2.2 Τυπική Μορφή του μοντέλου

Κάθε φίλτρο έχει μια κρουστική απόκριση η οποία είναι διαφορετική για το μεταβολικό σύστημα κάθε ανθρώπου. Ισοδύναμα με την κρουστική απόκριση τα 2 PDMS μπορούν να περιγραφούν κι από ένα ARMA μοντέλο το οποίο έχει την εξής μορφή :

$$u_1(n) = a_{11}u_1(n-1) + a_{12}u_1(n-2) + b_{10}\chi(n-1) + b_{11}\chi(n-1) \quad (2.1)$$

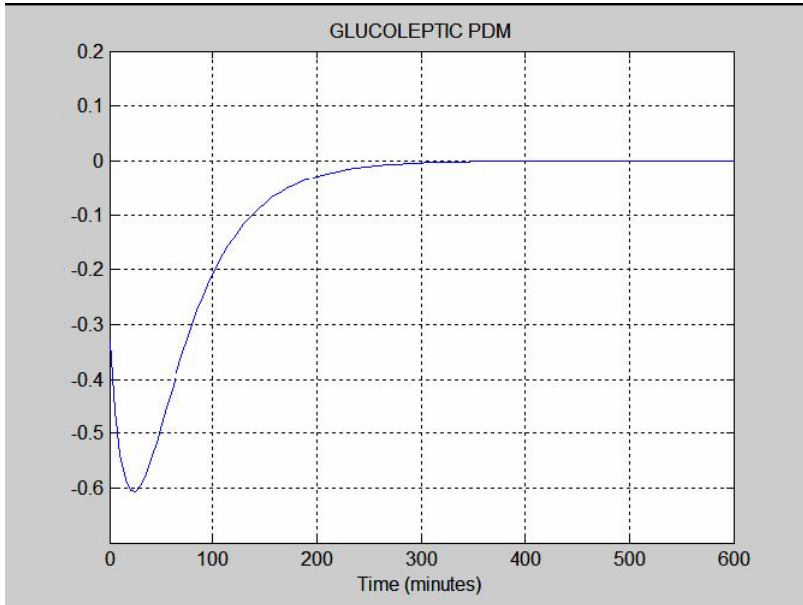
$$u_2(n) = a_{12}u_2(n-1) + a_{22}u_2(n-2) + b_{20}\chi(n-1) + b_{21}\chi(n-1) \quad (2.2)$$

$$y_1(n) = c_{11}u_1(n) + c_{12}u_1(n)^2 + c_{13}u_1(n)^3 \quad (2.3)$$

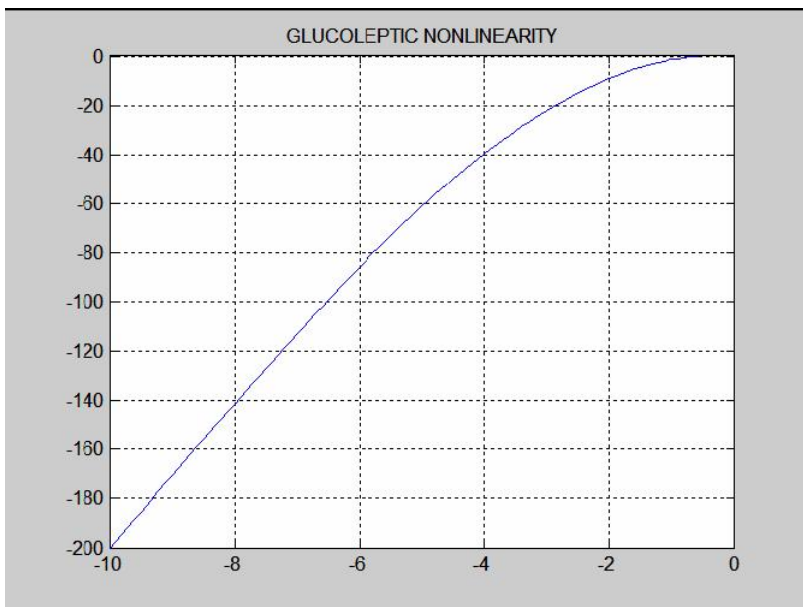
$$y_2(n) = c_{12}u_2(n) + c_{22}u_2(n)^2 + c_{23}u_2(n)^3 \quad (2.4)$$

$$y(n) = y_1(n) + y_2(n) \quad (2.5)$$

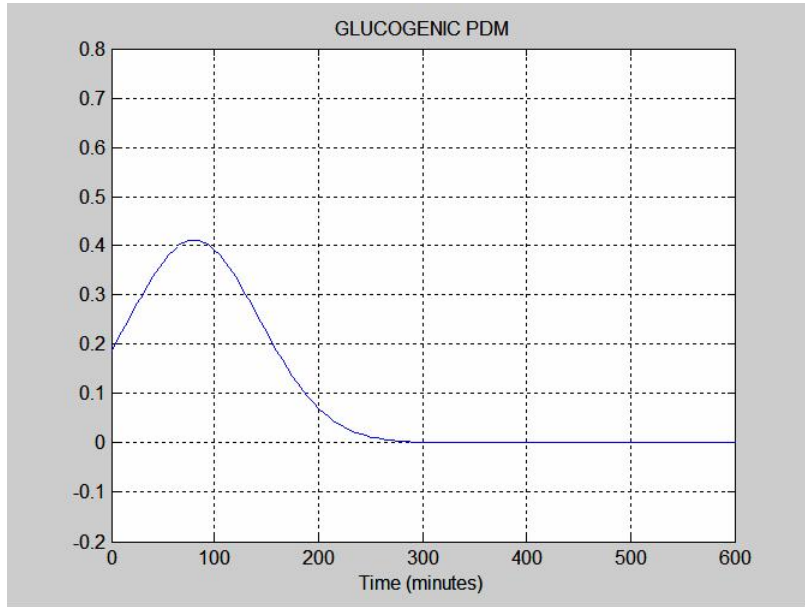
Μία τυπική απόκριση αυτού του μοντέλου είναι η εξής :



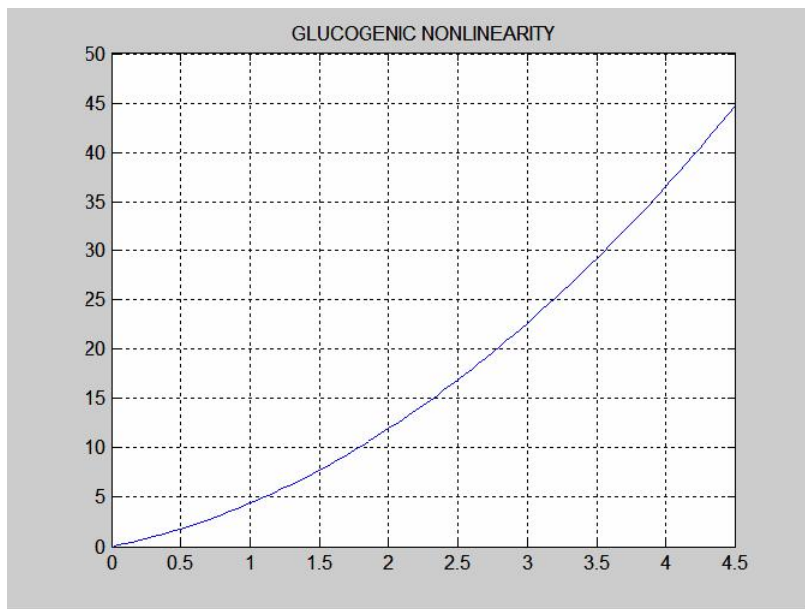
Σχήμα 2.2: Glucoleptic PDM



Σχήμα 2.3: Glucoleptic Nonlinearity



Σχήμα 2.4: Glucogenic PDM



Σχήμα 2.5: Glucogenic Nonlinearity

2.3 Εισαγωγικά Για Το Πρόβλημα-Αρχική Θεώρηση Του Προβλήματος

Σύμφωνα με τα προηγούμενα σε αυτήν την παράγραφο θα κάνουμε μια γενική εισαγωγή για το πρόβλημα στην πιο απλουστευμένη του μορφή ώστε στην επόμενη ενότητα να δικαιολογήσουμε την πλήρη μορφή του μοντέλου η οποία αφορά τα πραγματικά δεδομένα.

Ας υποθέσουμε λοιπόν ότι έχουμε ένα ασθενή στον οργανισμό του οποίου η γλυκόζη βρίσκεται συνεχώς σε ένα σταθερό επίπεδο (τουλάχιστον για την χρονική διάρκεια των μετρήσεων), και παρατηρούμε την αυξομειώσή της ποσότητας ινσουλίνης ενώ του εισάγουμε ινσουλίνη. Το πρόβλημα αυτό είναι καθαρά ιδεατό αλλά θα βοηθήσει στην μετέπειτα κατανόηση του πλήρους προβλήματος.

Σύμφωνα με όσα περιγράψαμε προηγουμένως το μοντέλο το οποίο θα θέλαμε να χρησιμοποιήσουμε θα περιγράφεται μόνο από τα δεδομένα του συστήματος τα οποία δεν είναι άλλα από τις χρονικές στιγμές και την ποσότητα των ενέσεων ινσουλίνης και από την διακύμανση της γλυκόζης εξ' αιτίας της εγγερόμενης ινσουλίνης.

Με βάση μόνο αυτά τα δεδομένα θα πρέπει να βρούμε τις τιμές των παραμέτρων του μοντέλου (Σχέσεις 2.1, 2.2, 2.3, 2.4, 2.5). Δηλαδή θα πρέπει με βάση τα δεδομένα να βρούμε τις τιμές των παραμέτρων του μοντέλου οι οποίες περιγράφουν καλύτερα το μοντέλο.

Βέβαια η έννοια καλύτερα είναι σχετική. Καλύτερα ως προς κάποιο κριτήριο. Το κριτήριο που θα έχουμε εμείς θα είναι το κριτήριο ελαχίστων τετραγώνων. Δηλαδή αν έχουμε μετρήσεις για N χρονικές στιγμές $y = [y_1 y_2 \dots y_N]$ θα θέλαμε να εκτιμήσουμε τις παραμέτρους του μοντέλου, τις οποίες τις τοποθετούμε σε ένα διάνυσμα παραμέτρων

$$\hat{x} = [a_{11} a_{12} a_{21} a_{22} b_{10} b_{11} b_{20} b_{21} c_{11} c_{12} c_{13} c_{21} c_{22} c_{23}]$$

, κατά τέτοιο τρόπο ώστε αυτό το διάνυσμα παραμέτρων να δίνει ως έξοδο ένα διάνυσμα $z' = [z'_1 z'_2 \dots z'_N]$ το οποίο να ελαχιστοποιεί την ποσότητα

$$l_{LS} = \frac{1}{2} \sum_{i=1}^N (y_i - z'_i)^2 \quad (2.6)$$

Ουσιαστικά λοιπόν ψάχνουμε για το ολικό ελάχιστο της προηγούμενης σχέσης. Εάν μπορούσαμε να λύσουμε τις αναγκαίες συνθήκες ελαχίστου αναλυτικά ώστε να βρούμε τα στάσιμα σημεία και από εκεί και έπειτα να βρούμε τα τοπικά ελάχιστα και τέλος το ολικό ελάχιστο τότε το πρόβλημα θα είχε λυθεί.

Το πρόβλημα με αυτή την προσέγγιση είναι ότι δεν μπορούμε να πάρουμε αναλυτική λύση και πρέπει να χρησιμοποιήσουμε άλλη προσέγγιση. Κι εδώ πάλι όμως όπως θα αναφέρω και στην συνέχεια υπάρχει ένα πρόβλημα που αφορά την χρήση της εφαρμογής: Εάν το αρχείο δεδομένων έχει αυστηρά καθορισμένο μήκος τότε θα μπορούσαμε να εφαρμόσουμε κάποιον αλγόριθμο βελτιστοποίησης ώστε να λύσουμε το πρόβλημα. Εάν όμως πρέπει να λύσουμε το πρόβλημα σε μια εφαρμογή πραγματικού χρόνου στην οποία έχουμε καινούρια μέτρηση κάθε 5 λεπτά δεν μπορούμε να εκτελούμε εκ νέου ένα τέτοιο αλγόριθμο κάθε φορά ο οποίος πιθανότατα και να διαρκέσει πάνω από 5 λεπτά. Σε αυτή την περίπτωση θα πρέπει να βρούμε κάποια άλλη λύση.

Ένα επίσης σημαντικό θέμα στο οποίο θα επανέλθουμε και αργότερα είναι η συμπεριφορά του συστήματος. Ας προχωρήσουμε λίγο παραπάνω τις σχέσεις που περιγράφουν τα 2 PDMs και πάρουμε την συνάρτηση μεταφοράς του συστήματος: εφαρμόζοντας μετασχηματισμό Z και στα δύο μέλη και λαμβάνοντας υπόψιν ότι $u_1(-1) =$

$$0, u_1(-2) = 0$$

$$Z[u_1(n)] = Z[a_{11}u_1(n-1) + a_{12}u_1(n-2) + b_{10}\chi(n) + b_{11}\chi(n-1)]$$

$$U_1(z) = a_{11}z^{-1}U_1(z) + a_{12}z^{-2}U_1(z) + b_{10}X(z) + b_{11}z^{-1}X(z)$$

από όπου παίρνουμε ότι η συνάρτηση μεταφοράς έχει την μορφή

$$\frac{U_1(z)}{X(z)} = \frac{b_{10} + b_{11}z^{-1}}{1 - a_{11}z^{-1} - a_{12}z^{-2}}$$

$$\frac{U_1(z)}{X(z)} = \frac{z^2b_{10} + b_{11}z}{z^2 - a_{11}z - a_{12}}$$

$$\frac{U_1(z)}{X(z)} = \frac{z(zb_{10} + b_{11})}{z^2 - a_{11}z - a_{12}} \quad (2.7)$$

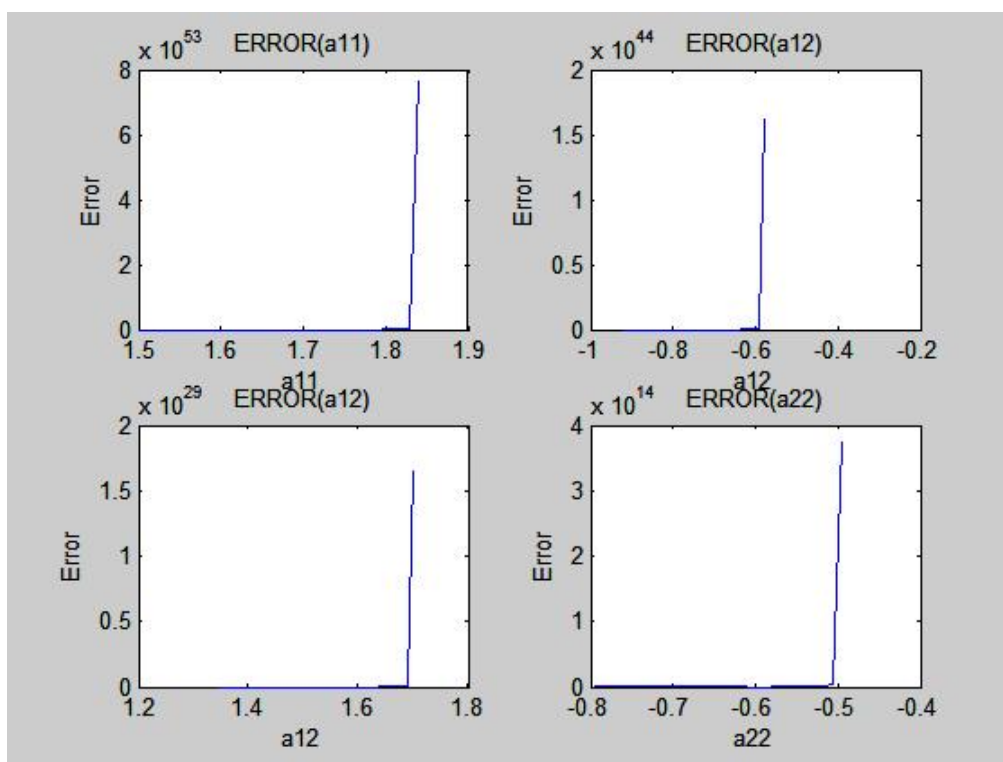
βλέπουμε δηλαδή ότι η συνάρτηση μεταφοράς έχει ένα μηδενικό κι 2 πόλους. Την συνάρτηση μεταφοράς την υπολογίζουμε σε αυτό το σημείο και θα την χρησιμοποιήσουμε στην συνέχεια για να αναλύσουμε κάποια προβλήματα τα οποία σχετίζονται με την ευστάθεια του συστήματος.

Επιπλέον σε αυτήν την ενότητα θα προσπαθήσουμε να οπτικοποιήσουμε ορισμένες δυσκολίες της εκτίμησης των παραμέτρων του συστήματος. Θεωρούμε πραγματικές τιμές παραμέτρων όπως φαίνονται στον ακόλουθο πίνακα:

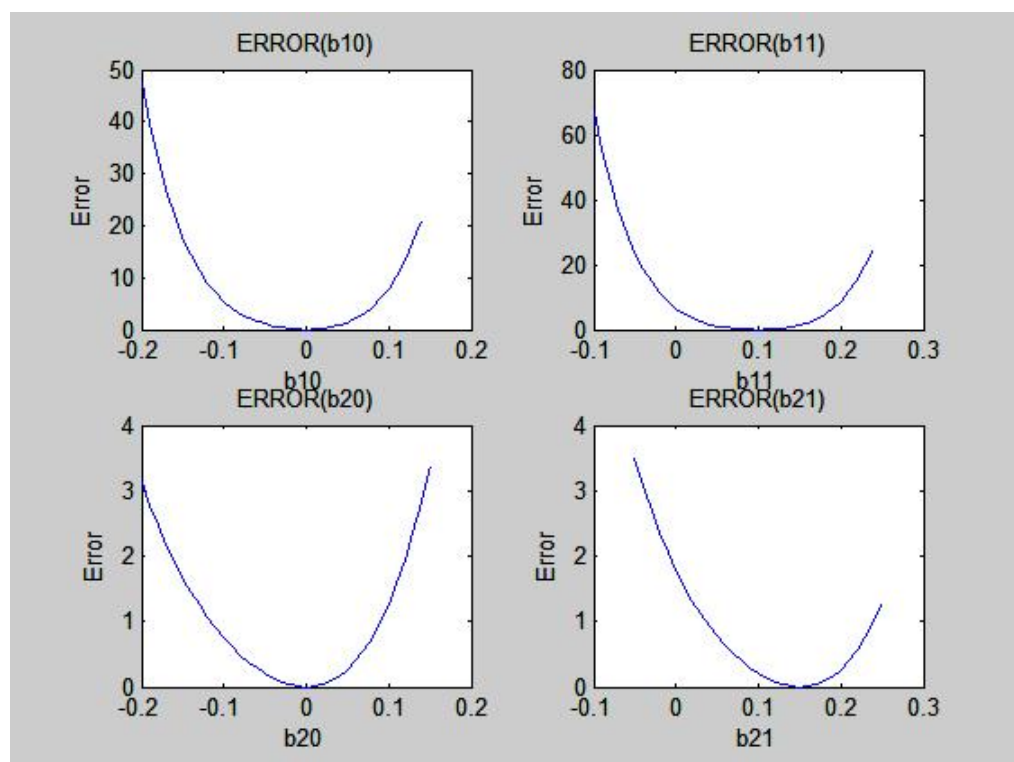
Όνομα Παραμέτρου	a_{11}	a_{12}	a_{21}	a_{22}	b_{10}	b_{11}	b_{20}	b_{21}	c_{11}	c_{12}
Πραγματική τιμή	1,7	-0.72	1,55	-0,595	0	0,1	0	0,15	-2	-3,5

Όνομα Παραμέτρου	c_{13}	c_{21}	c_{22}	c_{23}
Πραγματική τιμή	-0,2	1	2	0

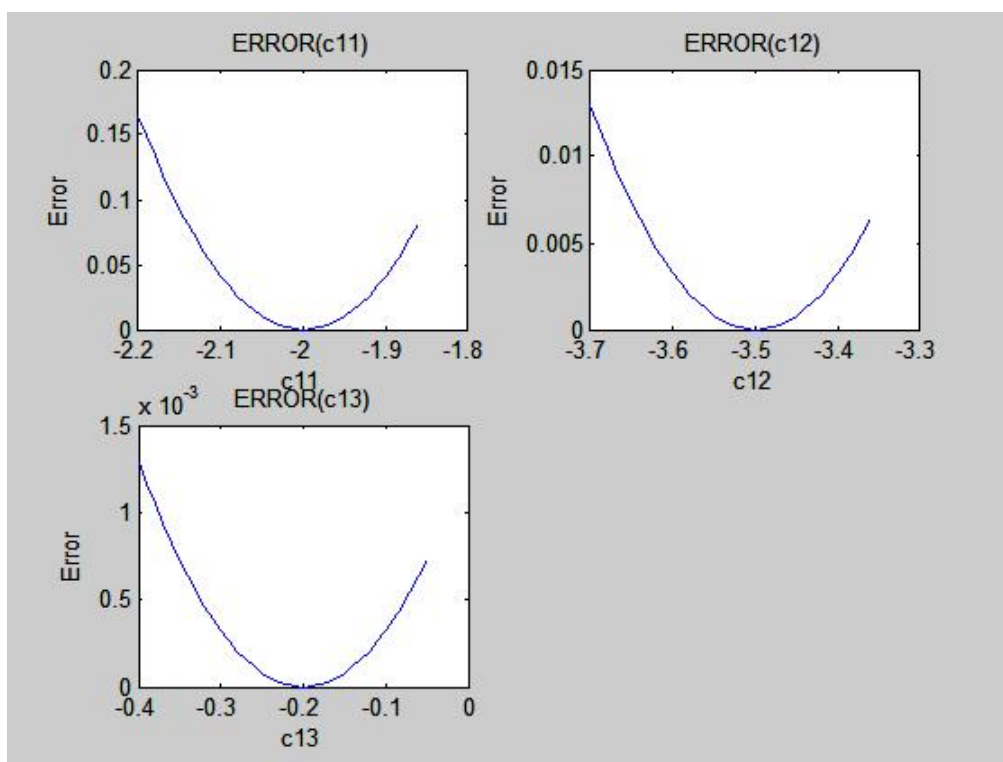
Διατηρώντας κάθε φορά το σύνολο των παραμέτρων σταθερό πλην μίας παραμέτρου εξετάζουμε πως μεταβάλλεται το συνολικό σφάλμα. Στο εκάστοτε διάγραμμα αναφέρουμε την παράμετρο που μεταβάλλεται σε ένα εύρος τιμών γύρω από την πραγματική.



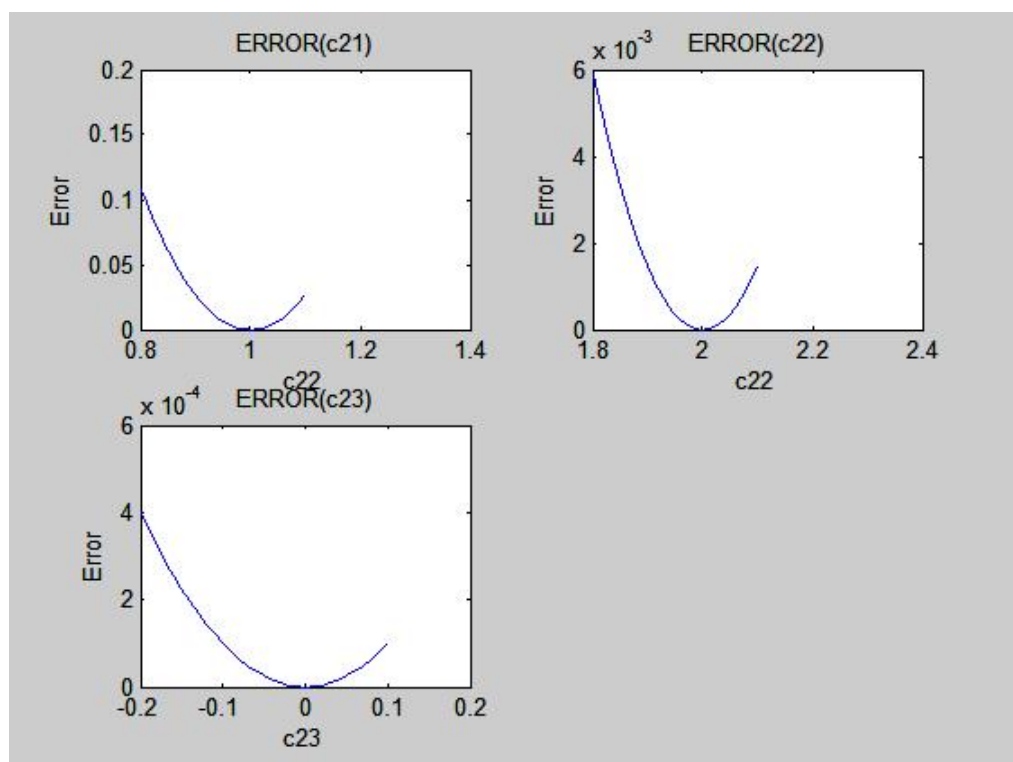
Σχήμα 2.6: Παράμετροι a



Σχήμα 2.7: Παράμετροι b



Σχήμα 2.8: Παράμετροι c_{1i}



Σχήμα 2.9: Παράμετροι c_{2i}

Αυτό που μπορούμε με μια πρώτη ματιά να παρατηρήσουμε είναι ότι όσον αφορά τις παραμέτρους a για πολύ μικρές αποκλίσεις από τις πραγματικές τους τιμές κι ενώ όλες οι λοιπές παράμετροι του συστήματος έχουν τις πραγματικές τους τιμές το τετραγωνικό σφάλμα είναι τεράστιο. Προκαταβολικά θα αναφέρουμε ότι το γεγονός αυτό οφείλεται στο ότι για την εκάστοτε τιμή της παραμέτρου a το τεράστιο τετραγωνικό σφάλμα είναι απόρεια της αστάθειας του συστήματος. Στο γεγονός όμως αυτό θα αναφερθούμε αναλυτικά σε επόμενη ενότητα.

2.4 Πλήρης Θεώρηση Του Προβλήματος-Προσαρμογή στα Πραγματικά Δεδομένα

Στην προηγούμενη παράγραφο είχαμε χρησιμοποιήσει την υπόθεση ότι στον οργανισμό του ατόμου από τον οποίο παίρνουμε τα δείγματα το επίπεδο γλυκόζης μεταβάλλεται μόνο από τις ενέσεις ινσουλίνης, κι έτσι ήταν δυνατό να εκτιμήσουμε τις παραμέτρους του μοντέλου.

Στην πραγματικότητα όμως κάτι τέτοιο δεν ισχύει. Η ποσότητα γλυκόζης στον οργανισμό μεταβάλλεται εξαιτίας όλων των παραγόντων που ανέφερα στις προηγούμενες ενότητες κι από ένα ακόμη πολύ σημαντικό παράγοντα που δεν είναι άλλος από τα γεύματα που θα έχει ο ασθενής ανάμεσα στις μετρήσεις. Όλη αυτή η διαταραχή είναι αυτό που ονομάζουμε Disturbance. Το πρόβλημα πλέον είναι πως θα καταφέρουμε να αναγνωρίσουμε το Disturbance του κάθε οργανισμού και στην συνέχεια να εκτιμήσουμε την τιμή των συντελεστών του συστήματος του κάθε ατόμου.

Ένα πραγματικό αρχείο ινσουλίνης είναι αυτό της εικόνας 2.10 Στο αρχείο αυτό για παράδειγμα έχουμε ενέσεις ινσουλίνης την χρονική στιγμή 2,23 στις οποίες όμως χρονικές στιγμές βλέπουμε ότι η συγκέντρωση γλυκόζης αυξάνεται. Αντιθέτως σε άλλες χρονικές στιγμές ενέσεων ινσουλίνης βλέπουμε ότι η ποσότητα γλυκόζης ελαττώνεται.

Είναι πολύ σημαντικό λοιπόν να μπορέσουμε να εκτιμήσουμε σωστά την διαταραχή. Το πρόβλημα λοιπόν έγκειται και σε σωστή μοντελοποίηση του Disturbance.

Για την μοντελοποίηση του Disturbance χρησιμοποιούμε ένα AR μοντέλο έστω p τάξης, το οποίο σημαίνει ότι η μορφή της εξόδου του μοντέλου την χρονική στιγμή n θα είναι η εξής :

$$y(n) = y_1(n) + y_2(n) + d(n) \quad (2.8)$$

όπου τα $y_1(n), y_2(n)$ είναι οι αποκρίσεις των μη γραμμικοτήτων γλυκόληψης κι γλυκογέννησης κι $d(n)$ η διαταραχή την εκάστοτε χρονική στιγμή n η οποία δίνεται από

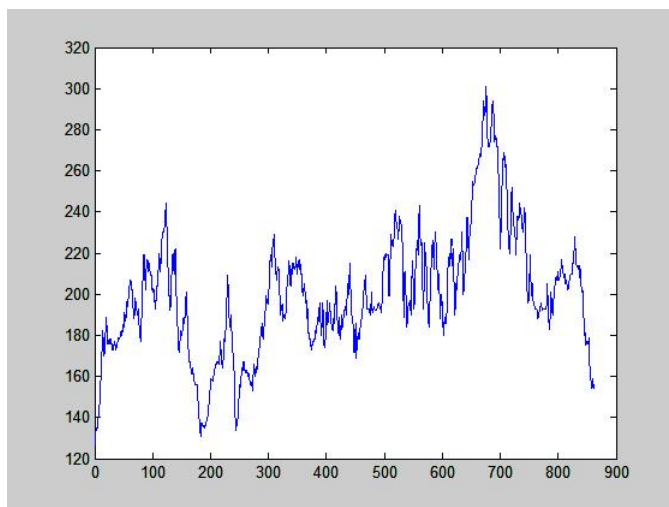
$$d(n) = \sum_{i=1}^p k_i d(n-i) \quad (2.9)$$

Οι συντελεστές του μοντέλου για την διαταραχή τους οποίους τοποθετούμε σε ένα διάνυσμα

$$\mathbf{k} = [k_1 k_2 k_3 \dots k_n]$$

θα πρέπει να εκτιμηθούν κι αυτοί μαζί με τις παραμέτρους των PDMs κι των μη γραμμικοτήτων της γλυκογέννησης και της γλυκόληψης.

Το πρόβλημα μας λοιπόν έγκειται στο να εκτιμήσουμε τους συντελεστές που προανέφερα δηλαδή το διάνυσμα παραμέτρων $\mathbf{x} = [\mathbf{x} \ \mathbf{k}]$ κατά τέτοιο τρόπο ώστε αυτό



Σχήμα 2.10: Τιμές Γλυκόζης

το διάνυσμα παραμέτρων να δίνει ως έξοδο ένα διάνυσμα $z = [z_1 z_2 \dots z_N]$ το οποίο να ελαχιστοποιεί την ποσότητα

$$l_{LS} = \frac{1}{2} \sum_{i=1}^N (y_i - z_i)^2 \quad (2.10)$$

Οι δυσκολίες σε αυτήν την πλήρη θεώρηση του προβλήματος είναι ίδιες με πριν και συγκεκριμένα τα προβλήματα είναι πολύ πιο δύσκολα καθώς το διάνυσμα παραμέτρων είναι πολύ μεγάλο. Ενδεικτική τιμή για το p για να δώσουμε το μέγεθος του διανύσματος είναι $p=16$ οπότε έχουμε 30 παραμέτρους προς εκτίμηση.

Αναλυτική λύση στο πρόβλημα μέσω των ικανών συνθηκών δεύτερης τάξης δεν μπορεί να δοθεί (σε απλή χρησιμοποιήσιμη μορφή) και αναζητούμε και πάλι κάποιον αλγόριθμο βελτιστοποίησης.

Όσον αφορά τώρα κάποιον αλγόριθμο βελτιστοποίησης εάν το αρχείο δεδομένων έχει αυστηρά καθορισμένο μήκος τότε θα μπορούσε πρακτικά να εφαρμοστεί ώστε να λυθεί το πρόβλημα. Εάν όμως πρέπει να λύσουμε το πρόβλημα σε μια εφαρμογή πραγματικού χρόνου στην οποία έχουμε καινούρια μέτρηση κάθε 5 λεπτά δεν μπορούμε να εκτελούμε εκ νέου ένα τέτοιο αλγόριθμο κάθε φορά ο οποίος πιθανότατα και να διαρκέσει πάνω από 5 λεπτά καθώς το μέγεθος του αρχείου μεγαλώνει συνεχώς.

Κεφάλαιο 3

Αλγόριθμοι Βελτιστοποίησης

Όπως ορίσαμε στην προηγούμενη παράγραφο το αντικείμενό μας πρόβλημα είναι να εκτιμήσουμε τις παραμέτρους του μοντέλου έγχυσης ινσουλίνης και της συνακόλουθης συμπεριφοράς του στην γλυκόζη του οργανισμού. Στην εκάστοτε χρονική στιγμή n έχουμε ως δεδομένα όλες τις προηγούμενες χρονικές στιγμές που είχαμε έγχυση ινσουλίνης και την ποσότητά τους καθώς και την τιμή της γλυκόζης στον οργανισμό μέχρι εκείνη την χρονική στιγμή. Με βάση αυτά τα στοιχεία την χρονική στιγμή n μπορούμε να εκτιμήσουμε τις παραμέτρους του μοντέλου κατά τρόπο τέτοιο ώστε να ελαχιστοποιείται το τετραγωνικό σφάλμα μεταξύ της προβλεπόμενης από το μοντέλο ποσότητας γλυκόζης στον οργανισμό και της πραγματικής μετρούμενης ποσότητας γλυκόζης. Δηλαδή να επιλέξουμε τις τιμές των παραμέτρων τέτοιες ώστε να ελαχιστοποιείται το κριτήριο

$$\min \sum_{i=1}^n y_i - z_i(x) \quad (3.1)$$

όπου y_i είναι η μετρούμενη τιμή της γλυκόζης την χρονική στιγμή i και $z_i(x)$ είναι η προσδοκώμενη τιμή της γλυκόζης την χρονική στιγμή i με βάση το μοντέλο.

Την επόμενη χρονική στιγμή όμως $n+1$ στην οποία πέραν των τιμών των οποίων ήταν διαθέσιμες την χρονική στιγμή n θα έχουμε και την μέτρηση για την χρονική στιγμή $n+1$ καθώς και την ινσουλίνη που εισάγαμε την χρονική στιγμή n θα πρέπει να εφαρμόσουμε εκ νέου τον αλγόριθμο για όλα τα δεδομένα.

Στην επόμενη παράγραφο θα εξετάσουμε το φίλτρο Kalman το οποίο εκτελεί ακολουθιακή εκτίμηση παραμέτρων δηλαδή σε κάθε χρονική στιγμή χρησιμοποιεί την προηγούμενη εκτίμηση και με βάση την εκάστοτε μέτρηση διορθώνει την εκτίμηση ώστε να έχουμε όσο το δυνατό καλύτερη προσαρμογή με την νέα μέτρηση.

Αρχίζουμε την παράγραφο αυτή παρουσιάζοντας τα γραμμικά ελάχιστα τετράγωνα στο δεύτερο μέρος παραθέτουμε τα βασικά χαρακτηριστικά και μεθόδους επίλυσης για μη γραμμικά προβλήματα, ώστε να δείξουμε την γενική μορφή των αλγορίθμων επίλυσης κι από εκεί κι έπειτα παρουσιάζουμε την πλήρη μορφή συγκεκριμένων αλγορίθμων και πως προσαρμόζονται στο πρόβλημά μας.

3.1 Γραμμικά ελάχιστα τετράγωνα

Η γενική διατύπωση του προβλήματος είναι η εξής:

Υποθέτουμε ότι έχουμε n παραμέτρους, που τις δηλώνουμε σαν το n -διάστατο διάνυσμα \bar{x} , οι οποίες πρόκειται να εκτιμηθούν από M μετρήσεις, που αντιπροσωπεύονται από το M -διάστατο διάνυσμα \bar{z}

. Οι παράμετροι \bar{x} και οι μετρήσεις \bar{z} συνδέονται σύμφωνα με την σχέση :

$$\bar{z} = H\bar{x} + \bar{v} \quad (3.2)$$

Ο $M \times n$ πίνακας H θεωρείται ότι είναι γνωστός και επίσης ο αριθμός των μετρήσεων θεωρείται ότι είναι τουλάχιστον τόσος όσο και ο αριθμός των αγνώστων παραμέτρων, δηλαδή το μέγεθος του διανύσματος \bar{x} . Επιπλέον ο πίνακας H θεωρείται ότι έχει βαθμό n (είναι πλήρους βαθμού). Το διάνυσμα \bar{v} αντιπροσωπεύει τυχαία σφάλματα τα οποία γίνονται κατά την διάρκεια των μετρήσεων.

Από τα παραπάνω δεν είναι δυνατόν να καθορίσουμε μονοσήμαντα το \bar{x} εξαιτίας των σφαλμάτων \bar{v} . Όμως δεδομένου ότι υπάρχουν περισσότερες μετρήσεις από αγνώστους ($M > n$) μπορούμε να επιχειρήσουμε να διαλέξουμε μία εκτίμηση για το \bar{x} που ελαχιστοποιεί με κάποιο αυθαίρετα επιλεγμένο τρόπο (σύμφωνα με κάποιο κριτήριο δηλαδή) την επίδραση αυτών των σφαλμάτων.

Για την εκτίμηση ελαχίστων τετραγώνων ο εκτιμητής επιλέγεται να ελαχιστοποιεί το άθροισμα των τετραγώνων των σφαλμάτων. Ακριβέστερα, ο \hat{x}_{LS} ορίζεται ως ο εκτιμητής ελαχίστων τετραγώνων του \bar{x} δοσμένων των μετρήσεων \bar{z} , εάν ελαχιστοποιεί το l_{LS} :

$$\begin{aligned} l_{LS} &\equiv \frac{1}{2} \sum_{i=1}^M v_i^2 = \frac{1}{2} \bar{v}^T \bar{v} \\ &= \frac{1}{2} (\bar{z} - H\bar{x})^T (\bar{z} - H\bar{x}) \end{aligned} \quad (3.3)$$

Ο εκτιμητής ελαχίστων τετραγώνων \hat{x}_{LS} μπορεί να προσδιοριστεί βρίσκοντας το στάσιμο σημείο της σχέσης (3.3). Έτσι σχηματίζουμε την πρώτη τάξης μερική παράγωγο $\frac{\partial l_{LS}}{\partial \bar{x}}$ (δηλαδή το gradient) το εξισώνουμε με το μηδέν και λύνουμε το ακόλουθο σύστημα γραμμικών εξισώσεων. Από την (3.3), έπεται ότι

$$\frac{\partial l_{LS}}{\partial \bar{x}} = -(\bar{z} - H\bar{x})^T H \quad (3.4)$$

και ο εκτιμητής ελαχίστων τετραγώνων ικανοποιεί

$$-(\bar{z} - H\hat{x}_{LS})^T H = 0 \quad (3.5)$$

Το προηγούμενο μπορεί να ξαναγραφεί ως:

$$H^T (\bar{z} - H\hat{x}_{LS}) = 0 \quad (3.6)$$

Επιπλέον η μήτρα δευτέρων παραγώγων είναι :

$$\frac{\partial^2 l_{LS}}{\partial \bar{x}^2} = H^T H \quad (3.7)$$

Αυτός ο πίνακας είναι θετικά ορισμένος αρκεί η μήτρα H να είναι πλήρους βαθμού (που εξ' αρχής έχουμε υποθέσει ότι ισχύει). Έτσι η λύση της (3.6) είναι μοναδική και ελαχιστοποιεί το l_{LS} . Η εξίσωση (3.6) μπορεί να ξαναγραφεί ως :

$$(H^T H) \hat{x}_{LS} = H^T \bar{z} \quad (3.8)$$

Αυτό το σύστημα αναφέρεται ως κανονική εξίσωση(normal equation).Επειδή η H έχει υποθεθεί ότι έχει πλήρη βαθμό,ο αντίστροφος του πίνακα ($H^T H$) υπάρχει και ο εκτιμητής ελαχίστων τετραγώνων βρίσκειται ότι είναι :

$$\hat{x}_{LS} = (H^T H)^{-1} H^T \bar{z} \quad (3.9)$$

Από την προηγούμενη σχέση βλέπουμε ότι το \hat{x}_{LS} είναι γραμμική συνάρτηση των μετρήσεων \bar{z} .

Επιπλέον υπάρχουν κάποιες ιδιότητες τις οποίες αξίζει να αναφέρουμε:

1. Το λάθος στον εκτιμητή \hat{x}_{LS} είναι γραμμική συνάρτηση του διανύσματος των σφαλμάτων μέτρησης \bar{v} .Αυτό έπεται από:

$$x_{\tilde{L}S} \equiv \bar{x} - \hat{x}_{LS} = \bar{x} - (H^T H)^{-1} H^T [H\bar{x} + \bar{v}] \quad (3.10)$$

$$x_{\tilde{L}S} = -(H^T H)^{-1} H^T \bar{v} \quad (3.11)$$

2. Χρησιμοποιώντας την (3.11),προκύπτει αμέσως ότι το residual (σφάλμα) μπορεί να γραφεί :

$$\begin{aligned} \bar{r} &\equiv \bar{z} - H\hat{x}_{LS} \\ &= H\tilde{x}_{LS} + \bar{v} \\ &= [I - H(H^T H)^{-1} H^T] \bar{v} \end{aligned} \quad (3.12)$$

3. Μέχρι στιγμής δεν έχουμε κάνει καμία υπόθεση σχετικά με το \bar{v} .Μια σχετικά απλή υπόθεση που μπορούμε να κάνουμε (ίσως κι η απλούστερη) είναι ότι το διάνυσμα \bar{v} έχει μηδενική μέση τιμή κι γνωστή θετικά ορισμένη μεταβλητότητα(covariance) R .Δηλαδή

$$E[\bar{v}] = 0, cov[\bar{v}] = 0$$

Με αυτές τις επιπλέον υποθέσεις για τον θόρυβο η μέση τιμή του λάθους για τον εκτιμητή ελαχίστων τετραγώνων βλέπουμε ότι είναι

$$\begin{aligned} E[\bar{x} - \hat{x}_{LS}] &= -E[(H^T H)^{-1} H^T \bar{v}] \\ &= -(H^T H)^{-1} H^T E[\bar{v}] \end{aligned} \quad (3.13)$$

όπου θεωρούμε ότι το H είναι γνωστό.Έχοντας την προηγούμενη σχέση προκύπτει αμέσως από τον ορισμό για το \bar{v} ότι

$$E[\tilde{x}_{LS}] = 0 \quad (3.14)$$

Βλέπουμε δηλαδή ότι ο \hat{x}_{LS} είναι ένας unbiased εκτιμητής του διανύσματος \bar{x} .

3.2 Εκτίμηση για μη γραμμικά προβλήματα

Στην αρχή αυτής της παραγράφου θα επισημάνω για ακόμη μια φορά ώστε να υπάρχει άμεση συσχέτιση μεταξύ των μεθόδων που θα αναφέρω και του πραγματικού προβλήματος ότι η προς ελαχιστοποίηση συνάρτηση είναι η

$$f(x) = \frac{1}{2}(y - z(x))^T (y - z(x)) \quad (3.15)$$

όπου x είναι το διάνυσμα παραμέτρων

$$x = [a_{11} a_{12} a_{21} a_{22} b_{10} b_{11} b_{20} b_{21} c_{11} c_{12} c_{13} c_{21} c_{22} c_{23} k_1 k_2 \dots k_p] \quad (3.16)$$

και $z(k)$ είναι η εκτιμώμενη ακολουθία εξόδων που παίρνουμε με βάση το διάνυσμα x . Το ζητούμενο λοιπόν είναι η ελαχιστοποίηση της συνάρτησης (3.15).

3.2.1 Αναγκαίες κι ικανές συνθήκες για τοπικά ελάχιστα χωρίς περιορισμούς

Θεώρημα 3.2.1 (Αναγκαίες Συνθήκες Πρώτης Τάξης). Αν η $f : R^n \rightarrow R$ είναι συνεχώς διαφορίσιμη και $\hat{x} \in R^n$ είναι ένα τοπικό ελάχιστο της f , τότε

$$\nabla f(\hat{x}) = 0 \quad (3.17)$$

Θεώρημα 3.2.2. Αναγκαίες συνθήκες Δεύτερης Τάξης Αν η $f : R^n \rightarrow R$ είναι συνεχώς διαφορίσιμη και $\hat{x} \in R^n$ είναι ένα τοπικό ελάχιστο της f , τότε

$$\nabla f(\hat{x}) = 0 \quad (3.18)$$

και

$$y^T \frac{\partial^2 f(\hat{x})}{\partial x^2} y \geq 0 \quad (3.19)$$

δηλαδή η μήτρα δευτέρων παραγώγων της f είναι θετικά ημιορισμένη

Θεώρημα 3.2.3. Ικανές Συνθήκες Δεύτερης Τάξης Αν η $f : R^n \rightarrow R$ είναι δύο φορές συνεχώς διαφορίσιμη και για κάποιο $\hat{x} \in R^n$ έχουμε

- $\nabla f(\hat{x}) = 0$
- $\frac{\partial^2 f(\hat{x})}{\partial x^2}$

τότε το \hat{x} είναι αυστηρό τοπικό ελάχιστο της f .

3.2.2 Γενική Μορφή και Στοιχεία Αλγορίθμων Βελτιστοποίησης

Στην γενική περίπτωση οι συνθήκες που αναφέραμε προηγουμένως είναι συστήματα μη-γραμμικών εξισώσεων των οποίων η πρακτική επίλυση μπορεί να γίνει μόνο αριθμητικά. Τις αριθμητικές μεθόδους για επίλυση του προβλήματος βελτιστοποίησης ονομάζουμε (αλγόριθμους βελτιστοποίησης).

Υπάρχει μεγάλη ποικιλία αλγορίθμων βελτιστοποίησης, όλοι όμως έχουν ορισμένα βασικά κοινά χαρακτηριστικά : είναι

- επαναληπτικές μέθοδοι
- μέθοδοι καθόδου

Αναλυτικότερα

- οι αλγόριθμοι είναι επαναληπτικές μέθοδοι σημαίνει ότι ξεκινούν από ένα δεδομένο αρχικό σημείο x_0 και κατασκευάζουν μια ακολουθία σημείων $x_{i=0}^{\inf}$ σύμφωνα με τον εξής κανόνα : κάθε νέο σημείο x_{i+1} της ακολουθίας προκύπτει από το προηγούμενο μέσω μιας απεικόνισης $A : R^n \rightarrow R^n$: $x_{i+1} = A(x_i)$. Η απεικόνιση A χαρακτηρίζει τον αλγόριθμο και διαφέρει από αλγόριθμο σε αλγόριθμο
- οι αλγόριθμοι είναι μέθοδοι καθόδου (descent methods) σημαίνει ότι κάθε νέο σημείο x_{i+1} της ακολουθίας που κατασκευάζει ένας αλγόριθμος μειώνει την τιμή $f(x_{i+1})$ μιας συνεχούς συνάρτησης f που χαρακτηρίζει το προς επίλυση πρόβλημα βελτιστοποίησης σε σχέση με το προηγούμενο σημείο. Δηλαδή $f(x_{i+1}) < f(x_i)$

Η ιδιότητα της καθόδου παίζει σημαντικό ρόλο για την σύγκλιση ενός αλγορίθμου. Κάθε αλγόριθμος βελτιστοποίησης εξετάζεται αναφορικά με δύο βασικές ιδιότητες

- γενική σύγκλιση
- ταχύτητα σύγκλισης

Ας θεωρήσουμε ένα πρόβλημα βελτιστοποίησης ,τις λύσεις του ,διάφορα αρχικά σημεία και έναν αλγόριθμο A . Αν για κάθε αρχικό σημείο η αντίστοιχη ακολουθία $x_{i=0}^{\inf}$ που κατασκευάζεται από τον αλγόριθμο συγκλίνει σε κάποιο σημείο \hat{x} που είναι λύση του προβλήματος βελτιστοποίησης τότε ο αλγόριθμος έχει την ιδιότητα της (γενικής σύγκλισης). Η ταχύτητα σύγκλισης είναι ένα μέτρο του πόσο γρήγορα η ακολουθία $x_{i=0}^{\inf}$ συγκλίνει στο \hat{x} .

Σχεδόν πάντοτε ένας αλγόριθμος απαρτίζεται από δύο στοιχεία :την κατεύθυνση έρευνας $h(x_i)$ και το μήκος βήματος a_i με τρόπο ώστε $x_{i+1} = x_i + a_i h(x_i)$.

Εδώ $h(x_i)$ είναι ένα διάνυσμα n διαστάσεων (στο συγκεκριμένο πρόβλημα είναι ένα διάνυσμα $(14+p) \times 1$, όπου p είναι η τάξη του AR μοντέλου) και a_i ένας θετικός πραγματικός αριθμός. Οι διάφοροι αλγόριθμοι διαφέρουν μεταξύ τους κυρίως ως προς τον τρόπο που ορίζουν την κατεύθυνση έρευνας. Το μήκος βήματος συνήθως βρίσκεται σαν το ελάχιστο της συνάρτησης κατά μήκος της κατεύθυνσης h .

Σύμφωνα με τα παραπάνω η γενική πρακτική μορφή ενός αλγορίθμου για το πρόβλημα $\min[f(x) : x \in F]$ είναι η εξής:

1. Αρχικό σημείο $x_0 \in F$,μηδενισμός δείκτη επαναλήψεων
2. Υπολογισμός κατεύθυνσης έρευνας $h_i = h(x_i)$
3. Υπολογισμός μήκους βήματος λ_i έτσι ώστε $f(x_i + \lambda_i h_i) = \min_{\lambda} (f(x_i + \lambda h_i)) : \lambda \geq 0, x_i + \lambda h_i \in F$
4. $x_{i+1} = x_i + \lambda_i h_i$
5. Αν ικανοποιείται το κριτήριο τερματισμού τότε τερμάτισε αλλιώς θέσε $i=i+1$ και πήγαινε στο βήμα 1.

Το παραπάνω σχήμα είναι αρκετά γενικό. Η βασική διαφορά των αλγορίθμων έγκειται στον τρόπο επιλογής της κατεύθυνσης έρευνας h . Πριν φτάσουμε όμως σε αυτό θα επεκταθούμε λίγο σε σχέση με τα βήματα 3 και 5 δηλαδή την επιλογή βήματος για ελαχιστοποίηση κατά μήκος της κατεύθυνσης h και την επιλογή κριτηρίων τερματισμού.

3.2.3 Κριτήρια Τερματισμού

Για τον πρακτικό τερματισμό του αλγορίθμου χρησιμοποιείται ένα από τα ακόλουθα κριτήρια

1. $-\varepsilon < f(x_{i+1}) - f(x_i) < 0$ δηλαδή τερματίζουμε τις επαναλήψεις όταν η βελτίωση της συνάρτησης καθόδου γίνει μικρότερη από μια σταθερά ε .
2. $\|x_{i+1} - x_i\| < \varepsilon$ δηλαδή τερματίζουμε τις επαναλήψεις όταν η μετατόπιση του τρέχοντος σημείου x_i από επανάληψη σε επανάληψη γίνει μικρότερη από ε
3. $\|\nabla f(x_i)\| < \varepsilon$ δηλαδή τερματίζουμε τις επαναλήψεις όταν το τρέχον σημείο x_i είναι κατά προσέγγιση επιθυμητό.

3.2.4 Έρευνα γραμμής ή ελαχιστοποίηση συνάκρησης μιας μεταβλητής

Τις μεθόδους ελαχιστοποίησης τις χωρίζουμε σε δύο κατηγορίες:αυτές που χρησιμοποιούν μόνο τιμές της συνάρτησης $f(\lambda)$ και αυτές που χρησιμοποιούν τιμές της $f(\lambda)$ και των παραγώγων της $f'(\lambda)$ και $f''(\lambda)$.

Έρευνα γραμμής χρησιμοποιώντας μόνο τιμές της $f(\lambda)$

1. Τετραγωνική Προσαρμογή

Δεδομένα οι τιμές της συνάρτησης f σε τρία σημεία $\lambda_1, \lambda_2, \lambda_3$, δηλαδή $f_1 = f(\lambda_1), f_2 = f(\lambda_2), f_3 = f(\lambda_3)$. Ζητείται να προσαρμοσθεί μια τετραγωνική συνάρτηση $q(\lambda) = a + b\lambda + c\lambda^2$ στα σημεία αυτά ,δηλαδή να ισχύει $q(\lambda_i) = f_i, i=1,2,3$. Η λύση αυτού του προβλήματος είναι :

$$q(\lambda) = f_1 \frac{(\lambda - \lambda_2)(\lambda - \lambda_3)}{(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3)} + f_2 \frac{(\lambda - \lambda_1)(\lambda - \lambda_3)}{(\lambda_2 - \lambda_1)(\lambda_2 - \lambda_3)} + f_3 \frac{(\lambda - \lambda_1)(\lambda - \lambda_2)}{(\lambda_3 - \lambda_1)(\lambda_3 - \lambda_2)} \quad (3.20)$$

Το νέο σημείο λ_4 ορίζεται σαν το ελάχιστο της συνάρτησης $q(\lambda)$ δηλαδή σαν το σημείο όπου $\frac{dq(\lambda_4)}{d\lambda} = 0$. Το σημείο αυτό δίνεται από τον τύπο :

$$\lambda_4 = \frac{1}{2} \frac{b_{23}f_1 + b_{31}f_2 + b_{12}f_3}{a_{23}f_1 + a_{31}f_2 + a_{12}f_3} \quad (3.21)$$

όπου $a_{ij} = \lambda_i - \lambda_j$, $b_{ij} = \lambda_i^2 - \lambda_j^2$ με $i,j=1,2,3$ Από τα τέσσερα σημεία $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ τα τρία εκλέγονται για την επόμενη τετραγωνική προσαρμογή με τρόπο ώστε να εξασφαλίζεται η σύγκλιση της μεθόδου.

Απαραίτητη προϋπόθεση για την σύγκλιση είναι : Τα σημεία $\lambda_1, \lambda_2, \lambda_3$ που γίνεται η προσαρμογή να ικανοποιούν $\lambda_1 < \lambda_2 < \lambda_3$.

2. Fibonacci Search

Η μέθοδος αυτή στηρίζεται στη σειρά Fibonacci. Η μόνη ιδιότητα που υποθέτουμε για την f είναι ότι θα έχει ένα μόνο ελάχιστο στο προς έρευνα διάστημα (unimodal). Το αρχικό διάστημα αβεβαιότητας θεωρείται δεδομένο .Επίσης δεδομένος θεωρείται ο αριθμός N των σημείων $\lambda_1, \lambda_2, \dots, \lambda_N$ στα οποία θα υπολογισθεί η συνάρτηση f κατά την διάρκεια της έρευνας με $c_1 \leq \lambda_1 < \lambda_2 \dots < \lambda_{N-1} < \lambda_N \leq c_2$. Το τελικό διάστημα αβεβαιότητας μέσα στο οποίο θα βρισκείται το ελάχιστο της συνάρτησης θα είναι $[\lambda_{k-1}, \lambda_{k+1}]$ όπου το λ_k είναι το σημείο

στο οποίο η συνάρτηση έχει την ελάχιστη τιμή δηλαδή $f(\lambda_k) = \min[f(\lambda_i)] : i = 0, 1, \dots, N + 1$ όπου ορίζουμε $c_1 = \lambda_0$ και $c_2 = \lambda_N + 1$. Το ζήτημα πλέον είναι να τοποθετηθούν τα σημεία αυτά εντός του διαστήματος ώστε το τελικό διάστημα αβεβαιότητας να είναι το ελάχιστο. Η λύση του προβλήματος αυτού είναι η εξής: Αν είναι

$d_1 = c_2 - c_1$ το πλάτος του αρχικού διαστήματος αβεβαιότητας και

d_k = το πλάτος του διαστήματος αβεβαιότητας μετά από k υπολογισμούς της συνάρτησης f . Τότε:

$$d_k = \left(\frac{F_{N-k+1}}{F_N}\right)d_1 \quad (3.22)$$

όπου F_k είναι τα αντίστοιχα στοιχεία της σειράς Fibonacci που παράγεται από την αναδρομική ακολουθία

$$F_N = F_{N-1} + F_{N-2}, F_0 = F_1 = 1$$

Η διαδικασία επιλογής των σημείων για να ελαττώσουμε το πλάτος του διαστήματος είναι η εξής: Οι πρώτες δύο μετρήσεις γίνονται συμμετρικά προς τα άκρα του διαστήματος σε απόσταση $(\frac{F_{N-1}}{F_N})d_1$. Σύμφωνα με το που παρουσιάζεται η μικρότερη τιμή έχουμε ένα νέο διάστημα αβεβαιότητας μήκους $d_2 = (\frac{F_{N-1}}{F_N})d_1$. Ο τρίτος υπολογισμός εκτελείται συμμετρικά μέσα στο νέο διάστημα αβεβαιότητας και σε απόσταση $d_3 = \frac{F_{N-2}}{F_N}d_1$ από το αντίστοιχο άκρο του. Έτσι λοιπόν παίρνουμε ένα νέο διάστημα αβεβαιότητας μήκους d_3 . Ομοίως ακολουθεί συνεχής ελάττωση του διαστήματος αβεβαιότητας ανάλογα με το μέγεθος βημάτων που έχουμε επιλέξει να κάνουμε.

3. Έρευνα Χρυσής Τομής

Εάν ο αριθμός των επαναλήψεων στην έρευνα Fibonacci τείνει στο άπειρο παίρνουμε την έρευνα χρυσής τομής. Τότε $\lim_{N \rightarrow \infty} \frac{F_{N-1}}{F_N} = \frac{1}{\tau} \approx 0,618$ όπου $\tau = \frac{1+\sqrt{5}}{2}$ είναι ο λόγος της χρυσής τομής. Η έρευνα της χρυσής τομής εκτελείται με παρόμοιο τρόπο προς την έρευνα Fibonacci αλλά το πλάτος του διαστήματος αβεβαιότητας μετά από i υπολογισμούς της f είναι : $d_k = (\frac{1}{\tau})^{k-1}d_1$.

Επειδή από την προηγούμενη σχέση βλέπουμε ότι $\frac{d_{k+1}}{d_k} = \frac{1}{\tau} = 0,618$ βλέπουμε ότι η έρευνα χρυσής τομής συγκλίνει γραμμικά στο ελάχιστο με βαθμό σύγκλισης 0,618

Έρευνα γραμμής χρησιμοποιώντας τιμές της f και παραγώγων

1. Κυβική Προσαρμογή

Δεδομένα : $f(\lambda_{k-1}), f'(\lambda_{k-1}), f(\lambda_k), f'(\lambda_k)$

Στα δεδομένα αυτά προσαρμόζεται ένα τριτοβάθμιο πολυώνυμο του λ . Στη συνέχεια το πολυώνυμο αυτό ελαχιστοποιείται και το ελάχιστο δίνει το νέο σημείο λ_{k+1} :

$$\lambda_{k+1} = \lambda_k - (\lambda_k - \lambda_{k-1}) \left[\frac{f'(\lambda_k) + u_2 - u_1}{f'(\lambda_k) - f'(\lambda_{k-1}) + 2u_2} \right] \quad (3.23)$$

όπου

$$u_1 = f'(\lambda_{k-1}) + f'(\lambda_k) - 3 \frac{f(\lambda_{k-1}) - f(\lambda_k)}{\lambda_{k-1} - \lambda_k} \quad (3.24)$$

$$u_2 = [u_1^2 - f'(\lambda_{k-1})f'(\lambda_k)]^{\frac{1}{2}} \quad (3.25)$$

Η τάξη σύγκλισης της μέθοδου είναι τετραγωνική. Για να πετύχουμε γενική σύγκλιση η μέθοδος χρειάζεται αναπροσαρμογή ώστε $\lambda_{k+1} \in [\lambda_{k-1}, \lambda_k]$

3.2.5 Γενικά για την κατεύθυνση έρευνας

Σε αυτή την ενότητα θα αναφερθούμε κυρίως στην περιγραφή βασικών μεθόδων που έχουν προταθεί για την ορισμό της κατεύθυνσης έρευνας $h(x_i)$. Αυτούς τους αλγόριθμους μπορούμε να τους κατηγοριοποιήσουμε ως προς τις υπολογιστικές απαιτήσεις που έχουν για την συνάρτηση f , σε δύο κατηγορίες :

- Αλγόριθμοι που προϋποθέτουν ότι εάν δοθεί ένα $x \in R^n$ είναι διαθέσιμο ένα υποπρόγραμμα που παίρνει το δεδομένο x σαν είσοδο και επιστρέφει την τιμή $f(x)$. Αυτού του είδους οι αλγόριθμοι λέγονται αλγόριθμοι άμεσης έρευνας (direct search methods)
- Αλγόριθμοι που χρησιμοποιούν όχι μόνο την τιμή της συνάρτησης στο σημείο αλλά x αλλά και της κλίσης $\frac{\partial f}{\partial x}$ και σε μερικές περιπτώσεις και της Hessian $\frac{\partial^2 f(x)}{\partial x^2}$

3.3 Μέθοδος της ταχύτερης κλίσης-Steepest Descent Method

Έστω ότι η f είναι η προς ελαχιστοποίηση συνάρτηση και έστω επίσης ότι η f έχει συνεχείς μερικές παραγώγους πρώτης τάξης. Ο αλγόριθμος της κλίσης ορίζει σαν κατεύθυνση έρευνας $h(x_i)$ την κατεύθυνση :

$$h(x_i) = -\nabla f(x_i) \quad (3.26)$$

3.3.1 Γενική μορφή αλγόριθμου Steepest Descent

Σύμφωνα με την γενική μορφή των επαναληπτικών αλγορίθμων (3.2.2) ο αλγόριθμος της ταχύτερης κλίσης είναι ο εξής :

1. Επιλογή αρχικού σημείου $x_0 \in R^n$
2. $h_i = -\nabla f(x_i)$
3. Έρευνα γραμμής Αναζητούμε λ_i : $f(x_i + \lambda_i h_i) = \min_{\lambda} (f(x_i + \lambda h_i)) : \lambda \geq 0$
Επιλέγουμε κάποιο αλγόριθμο από αυτούς που ανέφερα στην ενότητα (3.2.4)
4. $x_{i+1} = x_i + \lambda_i h_i$
5. Έλεγχος αν ικανοποιείται κάποιο κριτήριο τερματισμού. Αν όχι επιστροφή στο βήμα 2

3.3.2 Μορφή αλγορίθμου στην συγκεκριμένη εφαρμογή

Στην συγκεκριμένη περίπτωση η προς ελαχιστοποίηση συνάρτηση είναι η 3.15 και το προς εκτίμηση διάνυσμα παραμέτρων είναι το (3.16).

Σύμφωνα λοιπόν με τα παραπάνω εάν την χρονική στιγμή k ο αλγόριθμός μου βρίσκεται στο σημείο \bar{x}_k η κατεύθυνση έρευνας για αυτή την χρονική στιγμή θα είναι $h(\bar{x}_k) = -\nabla f(\bar{x}_k)$ η οποία δηλαδή είναι:

$$\nabla f(\bar{x}_k) = \left[\frac{\partial f}{\partial a_{11}} \frac{\partial f}{\partial a_{12}} \frac{\partial f}{\partial a_{21}} \frac{\partial f}{\partial a_{22}} \frac{\partial f}{\partial b_{10}} \frac{\partial f}{\partial b_{11}} \frac{\partial f}{\partial b_{20}} \frac{\partial f}{\partial b_{21}} \frac{\partial f}{\partial c_{11}} \frac{\partial f}{\partial c_{12}} \frac{\partial f}{\partial c_{13}} \frac{\partial f}{\partial c_{21}} \frac{\partial f}{\partial c_{22}} \frac{\partial f}{\partial c_{23}} \frac{\partial f}{\partial k_1} \dots \frac{\partial f}{\partial k_p} \right]_{(\bar{x}_k)} \quad (3.27)$$

Αυτό που μένει πλέον να δούμε για να καθορίσουμε το διάνυσμα κατεύθυνσης είναι να υπολογίσουμε αναλυτικά την τιμή της εκάστοτε μερικής παραγώγου στο σημείο \bar{x}_k . Να σημειώσουμε εδώ ότι τις εισόδους του συστήματος δηλαδή τις εγχύσεις ινσουλίνης τις συμβολίζουμε με το ελληνικό χ π.χ. $\chi(i)$ είναι η είσοδος ινσουλίνης την χρονική στιγμή i . Για να γίνουν πιο ευπαρουσίαστοι οι υπολογισμοί ξαναγράφουμε το τετραγωνικό κόστος ως εξής:

$$f(x) = I_{LS} = \frac{1}{2} \sum_{i=1}^M (y_i - z_i(x))^2$$

.Αναλυτικά οι υπολογισμοί δίνονται στον πίνακα (3.3.2) στο τέλος της υποενότητας. Η γενική μορφή του αλγορίθμου που χρησιμοποιήσαμε στην συγκεκριμένη εφαρμογή είναι η εξής:

1. Επιλογή αρχικού σημείου $x_0 \in R^n$
2. $h_i = -\nabla f(x_i)$
3. Έρευνα γραμμής Αναζητούμε λ_i : $f(x_i + \lambda_i h_i) = \min_{\lambda} (f(x_i + \lambda h_i))$: $\lambda \geq 0$
Επιλέγουμε τον αλγόριθμο κυβικής προσαρμογής (1) για τον υπολογισμό του βήματος λ_i
4. $x_{i+1} = x_i + \lambda_i h_i$
5. Χρησιμοποιούμε σαν κριτήριο τερματισμού το εξής: $\|x_{i+1} - x_i\| < \varepsilon$ (2) σε συνδυασμό με ταυτόχρονο έλεγχο εάν έχουμε υπερβεί ένα μέγιστο αριθμό επαναλήψεων. Αν δεν ικανοποιείται το κριτήριο τερματισμού επιστρέφουμε στο βήμα 2

$$\begin{aligned}\frac{\partial f(x)}{\partial a_{11}} &= \frac{1}{2} \frac{\sum_{i=1}^M (y_i - z_i(x))^2}{\partial a_{11}} = - \sum_{i=1}^M (y_i - z_i(x)) \frac{\partial z_i(x)}{\partial a_{11}} \\ &= - \sum_{i=1}^M (y_i - z_i(x)) [c_{11} + 2c_{12}u_1(i) + 3c_{13}u_1(i)^2] [u_1(i-1)]\end{aligned}$$

$$\begin{aligned}\frac{\partial f(x)}{\partial a_{12}} &= \frac{1}{2} \frac{\sum_{i=1}^M (y_i - z_i(x))^2}{\partial a_{12}} = - \sum_{i=1}^M (y_i - z_i(x)) \frac{\partial z_i(x)}{\partial a_{12}} \\ &= - \sum_{i=1}^M (y_i - z_i(x)) [c_{11} + 2c_{12}u_1(i) + 3c_{13}u_1(i)^2] [u_1(i-2)]\end{aligned}$$

$$\begin{aligned}\frac{\partial f(x)}{\partial a_{21}} &= \frac{1}{2} \frac{\sum_{i=1}^M (y_i - z_i(x))^2}{\partial a_{21}} = - \sum_{i=1}^M (y_i - z_i(x)) \frac{\partial z_i(x)}{\partial a_{21}} \\ &= - \sum_{i=1}^M (y_i - z_i(x)) [c_{21} + 2c_{22}u_2(i) + 3c_{23}u_2(i)^2] [u_2(i-1)]\end{aligned}$$

$$\begin{aligned}\frac{\partial f(x)}{\partial a_{22}} &= \frac{1}{2} \frac{\sum_{i=1}^M (y_i - z_i(x))^2}{\partial a_{22}} = - \sum_{i=1}^M (y_i - z_i(x)) \frac{\partial z_i(x)}{\partial a_{22}} \\ &= - \sum_{i=1}^M (y_i - z_i(x)) [c_{21} + 2c_{22}u_2(i) + 3c_{23}u_2(i)^2] [u_2(i-2)]\end{aligned}$$

$$\begin{aligned}\frac{\partial f(x)}{\partial b_{10}} &= \frac{1}{2} \frac{\sum_{i=1}^M (y_i - z_i(x))^2}{\partial b_{10}} = - \sum_{i=1}^M (y_i - z_i(x)) \frac{\partial z_i(x)}{\partial b_{10}} \\ &= - \sum_{i=1}^M (y_i - z_i(x)) [c_{11} + 2c_{12}u_1(i) + 3c_{13}u_1(i)^2] [\chi(i)]\end{aligned}$$

$$\begin{aligned}\frac{\partial f(x)}{\partial b_{11}} &= \frac{1}{2} \frac{\sum_{i=1}^M (y_i - z_i(x))^2}{\partial b_{11}} = - \sum_{i=1}^M (y_i - z_i(x)) \frac{\partial z_i(x)}{\partial b_{11}} \\ &= - \sum_{i=1}^M (y_i - z_i(x)) [c_{11} + 2c_{12}u_1(i) + 3c_{13}u_1(i)^2] [\chi(i-1)]\end{aligned}$$

$$\begin{aligned}\frac{\partial f(x)}{\partial b_{20}} &= \frac{1}{2} \frac{\sum_{i=1}^M (y_i - z_i(x))^2}{\partial b_{20}} = - \sum_{i=1}^M (y_i - z_i(x)) \frac{\partial z_i(x)}{\partial b_{20}} \\ &= - \sum_{i=1}^M (y_i - z_i(x)) [c_{21} + 2c_{22}u_2(i) + 3c_{23}u_2(i)^2] [\chi(i)]\end{aligned}$$

$$\begin{aligned}\frac{\partial f(x)}{\partial b_{21}} &= \frac{1}{2} \frac{\sum_{i=1}^M (y_i - z_i(x))^2}{\partial b_{21}} = - \sum_{i=1}^M (y_i - z_i(x)) \frac{\partial z_i(x)}{\partial b_{21}} \\ &= - \sum_{i=1}^M (y_i - z_i(x)) [c_{21} + 2c_{22}u_2(i) + 3c_{23}u_2(i)^2] [\chi(i-1)]\end{aligned}$$

$$\begin{aligned}\frac{\partial f(x)}{\partial c_{11}} &= \frac{1}{2} \frac{\sum_{i=1}^M (y_i - z_i(x))^2}{\partial c_{11}} = - \sum_{i=1}^M (y_i - z_i(x)) \frac{\partial z_i(x)}{\partial c_{11}} \\ &= - \sum_{i=1}^M (y_i - z_i(x)) u_1 i\end{aligned}$$

$$\begin{aligned}\frac{\partial f(x)}{\partial c_{12}} &= \frac{1}{2} \frac{\sum_{i=1}^M (y_i - z_i(x))^2}{\partial c_{12}} = - \sum_{i=1}^M (y_i - z_i(x)) \frac{\partial z_i(x)}{\partial c_{12}} \\ &= - \sum_{i=1}^M (y_i - z_i(x)) u_1 i^2\end{aligned}$$

$$\begin{aligned}\frac{\partial f(x)}{\partial c_{13}} &= \frac{1}{2} \frac{\sum_{i=1}^M (y_i - z_i(x))^2}{\partial c_{13}} = - \sum_{i=1}^M (y_i - z_i(x)) \frac{\partial z_i(x)}{\partial c_{13}} \\ &= - \sum_{i=1}^M (y_i - z_i(x)) u_1 i^3\end{aligned}$$

$$\begin{aligned}\frac{\partial f(x)}{\partial c_{21}} &= \frac{1}{2} \frac{\sum_{i=1}^M (y_i - z_i(x))^2}{\partial c_{21}} = - \sum_{i=1}^M (y_i - z_i(x)) \frac{\partial z_i(x)}{\partial c_{21}} \\ &= - \sum_{i=1}^M (y_i - z_i(x)) u_2 i\end{aligned}$$

$$\begin{aligned}\frac{\partial f(x)}{\partial c_{22}} &= \frac{1}{2} \frac{\sum_{i=1}^M (y_i - z_i(x))^2}{\partial c_{22}} = - \sum_{i=1}^M (y_i - z_i(x)) \frac{\partial z_i(x)}{\partial c_{22}} \\ &= - \sum_{i=1}^M (y_i - z_i(x)) u_2 i^2\end{aligned}$$

$$\begin{aligned}\frac{\partial f(x)}{\partial c_{23}} &= \frac{1}{2} \frac{\sum_{i=1}^M (y_i - z_i(x))^2}{\partial c_{23}} = - \sum_{i=1}^M (y_i - z_i(x)) \frac{\partial z_i(x)}{\partial c_{23}} \\ &= - \sum_{i=1}^M (y_i - z_i(x)) u_2 i^3\end{aligned}$$

$$\begin{aligned}\frac{\partial f(x)}{\partial k_1} &= \frac{1}{2} \frac{\sum_{i=1}^M (y_i - z_i(x))^2}{\partial k_1} = - \sum_{i=1}^M (y_i - z_i(x)) \frac{\partial z_i(x)}{\partial k_1} \\ &= - \sum_{i=1}^M (y_i - z_i(x)) d(i-1)\end{aligned}$$

$$\begin{aligned}\frac{\partial f(x)}{\partial k_j} &= \frac{1}{2} \frac{\sum_{i=1}^M (y_i - z_i(x))^2}{\partial k_j} = - \sum_{i=1}^M (y_i - z_i(x)) \frac{\partial z_i(x)}{\partial k_j} \\ &= - \sum_{i=1}^M (y_i - z_i(x)) d(i-j)\end{aligned}$$

$$\begin{aligned}\frac{\partial f(x)}{\partial k_p} &= \frac{1}{2} \frac{\sum_{i=1}^M (y_i - z_i(x))^2}{\partial k_p} = - \sum_{i=1}^M (y_i - z_i(x)) \frac{\partial z_i(x)}{\partial k_p} \\ &= - \sum_{i=1}^M (y_i - z_i(x)) d(i-p)^1\end{aligned}$$

3.4 Newton's Method

3.4.1 Γενικά για την μέθοδο του Newton

Η μέθοδος Newton υποθέτει ότι η συνάρτηση f είναι δύο φορές παραγωγίσιμη και απαιτεί η μήτρα δευτέρων παραγώγων $\frac{\partial^2 f(x)}{\partial x^2}$ να είναι διαθέσιμη υπολογιστικά. Η ιδέα πίσω από την μέθοδο του Newton είναι να προσεγγίσουμε την προς ελαχιστοποίηση συνάρτηση με μία τετραγωνική συνάρτηση και ύστερα να ελαχιστοποιήσουμε ακριβώς την τετραγωνική αυτή συνάρτηση.

Κοντά στο σημείο x_k μπορούμε να προσεγγίσουμε την συνάρτηση f με μια σειρά Taylor ως εξής:

$$f(x) \approx f(x_k) + \nabla f(x_k)(x - x_k) + \frac{1}{2}(x - x_k)^T \frac{\partial^2 f(x_k)}{\partial x^2} (x - x_k) \quad (3.28)$$

Το δεξιό μέλος της προηγούμενης σχέσης ελαχιστοποιείται στο σημείο x_{k+1}

$$x_{k+1} = x_k - \left(\frac{\partial^2 f(x_k)}{\partial x^2}\right)^{-1} \nabla f(x_k) \quad (3.29)$$

και αυτή η εξίσωση είναι η μορφή της μεθόδου του Newton. Η κατεύθυνση λοιπόν προς την οποία η μέθοδος αναζητά το ελάχιστο είναι η

$$\left(\frac{\partial^2 f(x_k)}{\partial x^2}\right)^{-1} \nabla f(x_k) \quad (3.30)$$

Η μέθοδος αυτή έχει τετραγωνική ταχύτητα σύγκλισης. Αυτό που πρέπει να αναφέρουμε όμως εδώ σαν προβλήματα της μεθόδου του Newton είναι ότι η ακολουθία σημείων που κατασκευάζεται με την μέθοδο αυτή συγκλίνει στο ελάχιστο μόνο εάν το αρχικό σημείο βείσκεται κοντά στο ελάχιστο. Για αρχικά σημεία μακριά από το ελάχιστο δεν υπάρχει καμία εγγύηση ότι η μέθοδος θα συγκλίνει. Για τον λόγο αυτό τροποποιούμε την μέθοδο Newton με τους ακόλουθους δύο τρόπους.

1. *Εισαγωγή μήκους βήματος στην κατεύθυνση h_k .* Για μη τετραγωνικές συναρτήσεις f είναι δυνατόν η τιμή $f(x_{k+1})$ της f στο νέο σημείο x_{k+1} που δίνεται από την σχέση (3.29) να είναι μεγαλύτερη από $f(x_k)$ λόγω των μη τετραγωνικών όρων της f . Για το λόγο αυτό κάνουμε μια έρευνα γραμμής κατά την κατεύθυνση h_k .
2. Η ιδιότητα της καθόδου που είναι απαραίτητη για την σύγκλιση (δηλ. σε κάθε βήμα ελαττώνεται η τιμή του κριτηρίου) εξασφαλίζεται μόνο εάν η μήτρα $\frac{\partial^2 f(x_k)}{\partial x^2}$ είναι θετικά ορισμένη. Όμως σε σημεία που βρίσκονται μακριά από ένα τοπικό ελάχιστο της f η μήτρα αυτή είναι δυνατόν να μην είναι θετικά ορισμένη. Για το λόγο αυτό εισάγουμε μια επιπλέον τροποποίηση στην κατεύθυνση έρευνας η οποία οδηγεί στην μέθοδο Levenberg-Marquardt (που περιγράφεται αναλυτικά στην επόμενη ενότητα).

Η ταχύτητα της μεθόδου του Newton στην καθαρή της μορφή είναι τετραγωνική. Στην τροποποιημένη μορφή που χρησιμοποιούμε όμως εισάγοντας και μια έρευνα γραμμής στην επιθυμητή κατεύθυνση αποδυναμώνεται ότι δεν έχει τετραγωνική ταχύτητα σύγκλισης παρά μόνο εάν από κάποια επανάληψη και μετά έχουμε ο συντελεστής λ_i που προκύπτει από την έρευνα γραμμής είναι μονάδα, δηλαδή $\lambda_i = 1$.

3.4.2 Η μέθοδος Newton στο συγκεκριμένο πρόβλημα

Τα στοιχεία όλα της μεθόδου έχουν παρουσιαστεί προηγουμένως. Η μήτρα δευτέρων παραγώγων υπολογίζεται με παρόμοιες πράξεις με αυτές για την μήτρα πρώτων παραγώγων παραγωγίζοντας κάθε όρο σύμφωνα με τον κανόνα τις αλυσίδας. Επιπλέον χρησιμοποιούμε κι την τροποποίηση να πραγματοποιήσουμε μια έρευνα γραμμής στο βήμα 3. Η γενική μορφή του αλγορίθμου είναι η εξής:

1. Επιλογή αρχικού σημείου $x_0 \in R^n$
2. $h_i = -\frac{\partial^2 f(x_k)}{\partial x^2}^{-1} \nabla f(x_k)$
3. Έρευνα γραμμής Αναζητούμε λ_i : $f(x_i + \lambda_i h_i) = \min_{\lambda} (f(x_i + \lambda h_i))$: $\lambda \geq 0$
Επιλέγουμε τον αλγόριθμο κυβικής προσαρμογής (1) για τον υπολογισμό του βήματος λ_i
4. $x_{i+1} = x_i + \lambda_i h_i$
5. Χρησιμοποιούμε σαν κριτήριο τερματισμού το εξής: $\|x_{i+1} - x_i\| < \epsilon$ (2) σε συνδυασμό με ταυτόχρονο έλεγχο εάν έχουμε υπερβεί ένα μέγιστο αριθμό επαναλήψεων. Αν δεν ικανοποιείται το κριτήριο τερματισμού επιστρέφουμε στο βήμα 2

3.5 Levenberg-Marquardt

3.5.1 Γενικά για την μέθοδο Levenberg-Marquardt

Η μέθοδος αυτή αποτελεί μια βελτίωση της μεθόδου του Newton. Όπως επισημάνεται και στα προβλήματα της μεθόδου του Newton απαραίτητη προϋπόθεση για να έχει η μέθοδος την ιδιότητα της καθόδου είναι η μήτρα $\frac{\partial^2 f(x_k)}{\partial x^2}$ να είναι θετικά ορισμένη, ιδιότητα όμως που μπορεί να μην ικανοποιείται εάν το σημείο x_k είναι μακριά από το τοπικό ελάχιστο. Η μέθοδος Levenberg-Marquardt ορίζει μια τροποποιημένη κατεύθυνση έρευνας, όπως περιγράφεται ακολούθως:

Έστω $\delta > 0$ ένας μικρός θετικός αριθμός. Εκτελούμε ανάλυση ιδιοτιμών της μήτρας $\frac{\partial^2 f(x_k)}{\partial x^2}$ και ως είναι $\lambda_{min} = \lambda_1 \leq \lambda_2 \leq \lambda_3 \dots \leq \lambda_n$ οι ιδιοτιμές της μήτρας. Αν $\lambda_{min} \geq \delta$ τότε χρησιμοποιούμε κανονικά την κατεύθυνση της μεθόδου του Newton (3.30). Αν $\lambda_{min} < \delta$ τότε αντί της $\frac{\partial^2 f(x_k)}{\partial x^2}$ στην κατεύθυνση (3.30) χρησιμοποιούμε μια τροποποιημένη μήτρα που έχει ελάχιστη ιδιοτιμή ίση με δ . Η τροποποιημένη μέθοδος Levenberg-Marquardt ορίζει σαν κατεύθυνση έρευνας

$$h_k = -\left[\epsilon_k I + \frac{\partial^2 f(x_k)}{\partial x^2}\right]^{-1} \nabla f(x_k) \quad (3.31)$$

όπου

$$\begin{aligned} \epsilon_k &= 0 \text{ αν } \lambda_{min} \geq \delta \\ \epsilon_k &= \delta - \lambda_{min} \text{ αν } \lambda_{min} < \delta \end{aligned} \quad (3.32)$$

3.5.2 Μέθοδος Levenberg-Marquardt στο συγκεκριμένο πρόβλημα

Αναλυτικότερα η μέθοδος Levenberg-Marquardt στο συγκεκριμένο πρόβλημα η οποία είναι και η μέθοδος που χρησιμοποιήθηκε περισσότερο έχει την εξής μορφή:

1. Επιλογή αρχικού σημείου $x_0 \in R^n$
2. Επιλογή κατεύθυνσης έρευνας σύμφωνα με την σχέση (3.31)
3. Έρευνα γραμμής Αναζητούμε λ_i : $f(x_i + \lambda_i h_i) = \min_{\lambda} (f(x_i + \lambda h_i)) : \lambda \geq 0$
Επιλέγουμε τον αλγόριθμο κυβικής προσαρμογής (1) για τον υπολογισμό του βήματος λ_i
4. $x_{i+1} = x_i + \lambda_i h_i$
5. Χρησιμοποιούμε σαν κριτήριο τερματισμού το εξής: $\|x_{i+1} - x_i\| < \varepsilon$ (2) σε συνδυασμό με ταυτόχρονο έλεγχο εάν έχουμε υπερβεί ένα μέγιστο αριθμό επαναλήψεων. Αν δεν ικανοποιείται το κριτήριο τερματισμού επιστρέφουμε στο βήμα 2

Κεφάλαιο 4

Φίλτρο Kalman

4.1 Διακριτό Φίλτρο Kalman

Το απλό διακριτό φίλτρο Kalman αναφέρεται στο γενικό πρόβλημα εκτίμησης παραμέτρων μιας διακριτού χρόνου διαδικασίας στην οποία τόσο η εξίσωση "κατάστασης" όσο κι η εξίσωση εξόδου είναι γραμμικές. Η γενική μορφή των εξισώσεων είναι η εξής :

$$x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1} \quad (4.1)$$

$$z_k = Hx_k + v_k \quad (4.2)$$

Στο εξής θα αναφέρομαι στην (4.1) η οποία υπολογίζει την τιμή του προς εκτίμηση διανύσματος x σαν process equation (εξίσωση ΔΙΑΔΙΚΑΣΙΑΣ) στην (4.2) η οποία δίνει το μετρούμενο διάνυσμα z_k σαν (εξίσωση μέτρησης) measurement equation. Οι τυχαίες μεταβλητές w_k και v_k αναπαριστούν τον θόρυβο της διαδικασίας και των μετρήσεων αντίστοιχα. Υποθέτουμε ότι είναι ανεξάρτητες και ακολουθούν κατανομή

$$p(w) \sim N(0, Q) \quad (4.3)$$

και

$$p(v) \sim N(0, R) \quad (4.4)$$

Ο $n \times n$ πίνακας A στην εξίσωση διαφορών (4.1) συνδέει την κατάσταση την προηγούμενη χρονική στιγμή $k-1$ με την κατάσταση την χρονική στιγμή k , απουσία του θορύβου ή της εισόδου u . Ο $n \times l$ πίνακας B συνδέει την είσοδο ελέγχου u με $u \in R^l$ με την κατάσταση x . Ο $m \times n$ πίνακας H στην εξίσωση μέτρησης συνδέει (4.2) συνδέει την κατάσταση με το μετρούμενο διάνυσμα z_k .

4.1.1 Υπολογιστική Προέλευση του Διακριτού Φίλτρου

Ορίζουμε ως $\hat{x}_k^- \in R^n$ να είναι η a priori εκτίμηση της κατάστασης στο βήμα k δοσμένης οποιασδήποτε πληροφορίας (στην προκειμένη περίπτωση των μετρήσεων) πριν την χρονική στιγμή k , και $\hat{x}_k \in R^n$ να είναι η a posteriori εκτίμηση της κατάστασης την χρονική στιγμή k δοσμένης της μέτρησης z_k . Μπορούμε λοιπόν να ορίσουμε τα a priori και a posteriori σφάλματα εκτίμησης ως

$e_k^- \equiv x_k - \hat{x}_k^-$ και
 $e_k \equiv x_k - \hat{x}_k$.

Η συμμεταβλητότητα του a priori λάθους εκτίμησης είναι

$$P_k^- = E[e_k^- e_k^{-T}] \quad (4.5)$$

και του a posteriori λάθους εκτίμησης είναι

$$P_k = E[e_k e_k^T] \quad (4.6)$$

. Για να πάρουμε τις εξισώσεις για το Kalman Filter ,αρχίζουμε προσπαθώντας να βρούμε μια εξίσωση που να υπολογίσει μια a posteriori εκτίμηση του διανύσματος κατάστασης \hat{x}_k σαν γραμμικό συνδυασμό μιας a priori εκτίμησης \hat{x}_k^- και μιας σταθμισμένης διαφοράς μεταξύ της πραγματικής μέτρησης z_k και μιας πρόβλεψης της μέτρησης H όπως φαίνεται στην εξίσωση που ακολουθεί.

$$\hat{x}_k = \hat{x}_k^- + K(z_k - H\hat{x}_k^-) \quad (4.7)$$

Η διαφορά ($z_k - H\hat{x}_k^-$) ονομάζεται καινοτομία (innovation) ή απλά ως σφάλμα μέτρησης (residual). Η διαφορά αυτή αντικατοπτρίζει το πόσο απέχει η εκτιμώμενη μέτρηση από την πραγματική. Σφάλμα μηδέν σημαίνει ότι η μέτρηση είναι ίδια με την προβλεπόμενη τιμή και ως εκ τούτου δεν γίνεται καμία αλλαγή στην εκτίμηση του διανύσματος κατάστασης ως προς την a priori τιμή του. Ο πίνακας K ο οποίος λέγεται κέρδος Kalman επιλέγεται έτσι ώστε να ελαχιστοποιεί την a posteriori συμμεταβλητότητα του σφάλματος (4.6). Η μορφή του K δίνεται από την

$$K_k = P_k^- H^T (H P_k^- H^T + R)^{-1} \quad (4.8)$$

Παρατηρώντας την προηγούμενη σχέση βλέπουμε ότι καθώς ο πίνακας R τείνει στο μηδέν ,το κέρδος K δίνει περισσότερο βάρος στο residual. Συγκεκριμένα ,

$$\lim_{(R_k) \rightarrow 0} K_k = H^{-1} \quad (4.9)$$

Από την άλλη μεριά όμως, καθώς ο πίνακας συμμεταβλητότητας του λάθους πλησιάζει το μηδέν ,το κέρδος K δίνει λιγότερο βάρος στο residual. Συγκεκριμένα

$$\lim_{P_k^- \rightarrow 0} K_k = H^{-1} \quad (4.10)$$

4.1.2 Αλγόριθμος Διακριτού Φίλτρου Kalman

Το φίλτρο Kalman εκτιμάει μια διαδικασία χρησιμοποιώντας μια μορφή ελέγχου ανάδρασης: το φίλτρο εκτιμάει την κατάσταση κάποια χρονική στιγμή και με βάση τις μετρήσεις ανανεώνει την εκτίμηση κατάστασης για την εκάστοτε χρονική στιγμή με βάση την οποία γίνεται και ο υπολογισμός της επόμενης εκτιμώμενης εξόδου. Ως εκ τούτου λοιπόν οι εξισώσεις για το φίλτρο Kalman χωρίζονται σε δύο ομάδες :time update equations και measurement update equations. Οι χρονικές εξισώσεις ανανέωσης είναι υπεύθυνες για να προβάλλουν απευθείας χρονικά την τρέχουσα κατάσταση και τις εκτιμήσεις της συμμεταβλητότητας του λάθους για να δώσουν τις a priori εκτιμήσεις για το επόμενο χρονικό βήμα. Οι εξισώσεις ανανέωσης με βάση την μέτρηση (measurement update equations) είναι υπεύθυνες για την ανάδραση, δηλαδή για να ενσωματώσουν την καινούρια μέτρηση στις a priori εκτιμήσεις και να δώσουν τις βελτιωμένες a posteriori εκτιμήσεις.

Οι χρονικές εξισώσεις ανανέωσης (time update equations) μπορούν να θεωρηθούν διαισθητικά σαν εξισώσεις πρόβλεψης(predictor equations) ενώ οι εξισώσεις ανανέωσης με βάση την μέτρηση μπορούν να θεωρηθούν σαν εξισώσεις διόρθωσης(corrector equations). Πραγματι ο τελικός αλγόριθμος εκτίμησης μοιάζει με ένα σχήμα predictor-corrector .

Discrete Kalman Filter Time updates Equations

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1} \quad (4.11)$$

$$P_k^- = AP_{k-1}A^T + Q \quad (4.12)$$

Discrete Kalman Filter measurement update equations

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1} \quad (4.13)$$

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-) \quad (4.14)$$

$$P_k = (1 - K_k H)P_k^- \quad (4.15)$$

Πίνακας 4.1: Discrete Kalman Filter Algorithm

4.2 EXTENDED KALMAN FILTER

4.2.1 Εισαγωγή

Στην παρούσα περίπτωση όμως η εξίσωση μέτρησης είναι μη γραμμική συνάρτηση των παραμέτρων του συστήματος. Σε αυτή την περίπτωση μπορούμε να γραμμικοποιήσουμε την εκτίμηση γύρω από την παρούσα τιμή της χρησιμοποιώντας τις μερικές παραγώγους της "διαδικασίας" και των μετρήσεων, μια μέθοδος που είναι γνωστή ως Extended Kalman Filter. Στο συγκεκριμένο πρόβλημα η process equation είναι γραμμική αλλά η μέθοδος θα παρουσιασθεί γενικά και έπειτα θα προσαρμοσθεί στο συγκεκριμένο πρόβλημα.

4.2.2 Υπολογιστική Προέλευση του Φίλτρου

Ας υποθέσουμε αρχικά ότι η διαδικασία μας έχει ένα διάνυσμα κατάστασης $x \in \mathbf{R}^n$ αλλά η διαδικασία μας αυτή τη φορά υπόκειται στην μη γραμμική στοχαστική εξίσωση διαφορών

$$x_k = f(x_{k-1}, u_{k-1}, w_{k-1}) \quad (4.16)$$

με διάνυσμα μετρήσεων $z \in (R^m)$

το οποίο υπολογίζεται από την μη γραμμική σχέση

$$z_k = H(x_k, v_k), \quad (4.17)$$

όπου οι τυχαίες μεταβλητές w_k και v_k αναπαριστούν το θόρυβο στην διαδικασία και στις μετρήσεις. Στην πράξη όμως δεν γνωρίζουμε τις ακριβείς τιμές του θορύβου (αντίστοιχα της διαδικασίας ή των μετρήσεων). Μπορούμε όμως να προσεγγίσουμε το διάνυσμα κατάστασης x_k και το διάνυσμα των μετρήσεων z_k χωρίς θόρυβο ως εξής:

$$\tilde{x}_k = f(x_{k-1}, u_{k-1}, 0) \quad (4.18)$$

$$\tilde{z}_k = h(\tilde{x}_k, 0) \quad (4.19)$$

Για να εκτιμήσουμε μια διαδικασία η οποία είναι μη γραμμική τόσο ως προς την εξίσωση διαφορών $f(x_{k-1}, u_{k-1}, w_{k-1})$ αλλά και ως προς την εξίσωση των μετρήσεων $H(x_k, v_k)$, αρχίζουμε γραμμικοποιώντας τις προηγούμενες εξισώσεις γύρω από τις τιμές \tilde{x}_k και \tilde{z}_k που υπολογίσαμε.

$$x_k = \tilde{x}_k + A(x_{k-1} - \hat{x}_{k-1}) + W w_{k-1} \quad (4.20)$$

$$z_k = \tilde{z}_k + H(x_k - \tilde{x}_k) + V v_k \quad (4.21)$$

όπου

- x_k και z_k είναι το πραγματικό διάνυσμα κατάστασης και το πραγματικό μετρούμενο διάνυσμα
- \tilde{x}_k και \tilde{z}_k είναι αντίστοιχα τα εκτιμώμενα διανύσματα κατάστασης και μέτρησης αντίστοιχα
- \hat{x}_k είναι μια a posteriori εκτίμηση της κατάστασης την χρονική στιγμή k
- οι τυχαίες μεταβλητές w_k και v_k αντιπροσωπεύουν το θόρυβο της διαδικασίας και των μετρήσεων
- A είναι η Ιακωβιανή μήτρα μερικών παραγώγων ως προς το x ,

$$A_{[i,j]} = \frac{\partial f^{[i]}}{\partial x^{[j]}}(x_{k-1}, u_{k-1}, 0) \quad (4.22)$$

- W είναι η Ιακωβιανή μήτρα μερικών παραγώγων ως προς το w

$$W_{[i,j]} = \frac{\partial f^{[i]}}{\partial w^{[j]}}(x_{k-1}, u_{k-1}, 0) \quad (4.23)$$

- H είναι η Ιακωβιανή μήτρα μερικών παραγώγων του h ως προς το x

$$H_{[i,j]} = \frac{\partial h^{[i]}}{\partial x^{[j]}}(\tilde{x}_k, 0) \quad (4.24)$$

- V είναι η Ιακωβιανή μήτρα μερικών παραγώγων της h ως προς το v

$$V_{[i,j]} = \frac{\partial h^{[i]}}{\partial v^{[j]}}(\tilde{x}_k, 0) \quad (4.25)$$

Τώρα ορίζουμε το λάθος πρόβλεψης ως εξής:

$$e_{x_k}^- \equiv x_k - \tilde{x}_k \quad (4.26)$$

και το residual της μέτρησης ως εξής

$$e_{z_k}^- \equiv z_k - \tilde{z}_k \quad (4.27)$$

Χρησιμοποιώντας τις 2 προηγούμενες εξισώσεις μπορούμε να γράψουμε τις γραμμικοποιημένες εξισώσεις ως εξής :

$$e_{x_k}^- \approx A(x_{k-1} - \hat{x}_{k-1}) + \epsilon_k \quad (4.28)$$

$$e_{z_k}^- \approx H e_{x_k}^- + \eta_k \quad (4.29)$$

όπου τα ϵ_k , η_k αντιπροσωπεύουν ανεξάρτητες τυχαίες μεταβλητές που έχουν μηδενική μέση τιμή και covariances $WQ_k W^T$ και $VR_k V^T$. Οι προηγούμενες εξισώσεις είναι γραμμικές και μοιάζουν αρκετά με τις αρχικές εξισώσεις του διακριτού γραμμικού Φίλτρου Kalman. Έτσι οδηγούμαστε στο να χρησιμοποιήσουμε το residual της μέτρησης $e_{z_k}^-$ και το γραμμικό φίλτρο Kalman πλέον ώστε βάση των προηγούμενων σχέσεων να εκτιμήσουμε το λάθος πρόβλεψης $e_{x_k}^-$ που δίνεται από τη σχέση (4.28). Αυτή την εκτίμηση την οποία θα την ονομάσουμε \hat{e}_k μπορούμε να την χρησιμοποιήσουμε μαζί με την (4.26) για να πάρουμε την a posteriori εκτίμηση της κατάστασης για την αρχική μη γραμμική εξίσωση κατάστασης της διαδικασίας.

$$\hat{x}_k = \tilde{x}_k + \hat{e}_k \quad (4.30)$$

Οι τυχαίες μεταβλητές των προηγούμενων σχέσεων έχουν τις εξής κατανομές πυκνότητας πιθανότητας:

- $p(e_{x_k}^-) \sim N(0, E[e_{x_k}^- e_{x_k}^{-T}])$
- $p(\epsilon_k) \sim N(0, WQ_k W^T)$
- $p(\eta_k) \sim N(0, VR_k V^T)$

Έχοντας λοιπόν τις προηγούμενες προσεγγίσεις για τις κατανομές και αφήνοντας την προβλεπόμενη τιμή για το \hat{e}_k να τείνει στο 0 η εξίσωση του Φίλτρου Kalman που χρησιμοποιείται για να εκτιμήσουμε το \hat{e}_k είναι

$$\hat{e}_k = K_k e_{z_k}^- \quad (4.31)$$

Αντικαθιστώντας την προηγούμενη σχέση στην (4.30) και χρησιμοποιώντας την (4.27) παίρνουμε

$$\hat{x}_k = \tilde{x}_k + K_k e_{z_k}^- = \tilde{x}_k + K_k (z_k - \tilde{z}_k) \quad (4.32)$$

Η προηγούμενη εξίσωση τώρα μπορεί να χρησιμοποιηθεί για την ανανέωση στο Extended Kalman Filter με τα \tilde{x}_k και \tilde{z}_k να τα παίρνουμε από τις εξισώσεις (4.18) ,(4.19) αντίστοιχα και το κέρδος K_k να δίνεται κατά αντιστοιχία με το γραμμικό φίλτρο από (4.62). Το πλήρες σύνολο εξισώσεων του EKF δίνεται στους ακόλουθους δύο πίνακες. Στα set αυτά εξισώσεων έχουμε προσάψει τον δείκτη k στις Ιακωβιανές A, W, H και V για να τονίσω ότι είναι διαφορετικές σε κάθε βήμα και πρέπει ως εκ τούτου να υπολογίζονται σε κάθε βήμα. Στον πίνακα αυτό χρησιμοποιούμε στην θέση του \tilde{x}_k το σύμβολο \hat{x}_k^- κατ'αντιστοιχία με όσα αναφέραμε για το διακριτό φίλτρο Kalman για

Time updates Equations

$$\hat{x}_k^- = f(x_{k-1}^-, u_{k-1}, 0) \quad (4.33)$$

$$\hat{P}_k^- = A_k P_{k-1} A_k^T + W_k O_{k-1} W_k^T \quad (4.34)$$

EKF measurement update equations

$$K_k = P_k^- H_k^T (H_k P_k^- H_k^T + V_k R_k V_k^T)^{-1} \quad (4.35)$$

$$\hat{x}_k = \hat{x}_k^- + K_k (z_k - h(\hat{x}_k^-, 0)) \quad (4.36)$$

$$P_k = (I - K_k H_k) P_k^- \quad (4.37)$$

Πίνακας 4.2: Extended Kalman Filter Algorithm

να δηλώσουμε την a priori εκτίμηση για το διάνυσμα κατάστασης.

Όπως και με το γραμμικό διακριτό Kalman Filter ,οι Time update equations υπολογίζουν την εκτίμηση για την κατάσταση και για τον P βασιζόμενοι μόνο στην προηγούμενη κατάσταση (k-1)για την παρούσα χρονική στιγμή k,ενώ οι measurement update equations διορθώνουν τις εκτιμήσεις για την κατάσταση και τον P με την παρούσα μέτρηση z_k .

4.3 Προσαρμογή Στο Συγκεκριμένο Πρόβλημα

Συγκεκριμενοποίηση του αλγορίθμου στο παρών πρόβλημα. Στο συγκεκριμένο πρόβλημα ονομάζουμε x το προς εκτίμηση διάνυσμα το οποίο είναι

$$x = [a_{11} a_{12} a_{21} a_{22} b_{10} b_{11} b_{20} b_{21} c_{11} c_{12} c_{13} c_{21} c_{22} c_{23} k_1 \dots k_p] \quad (4.38)$$

όπου p είναι η τάξη του μοντέλου πρόβλεψης της διαταραχής. Οι τιμές των παραμέτρων είναι μη χρονικά μεταβαλλόμενες οπότε οι τιμές τους είναι ίδιες για οποιαδήποτε χρονική στιγμή και επομένως έχουμε ότι

$$x_k = x_{k-1} \quad (4.39)$$

Συγκρίνοντας με την εξίσωση (4.16) βλέπουμε ότι πέραν του μη μεταβαλλόμενου έχουμε γραμμική σχέση και επιπλέον δεν έχουμε καθόλου θόρυβο.

Όσον αφορά τώρα τις μετρήσεις έχουμε ότι η μέτρηση του συστήματος την εκάστοτε χρονική στιγμή k είναι συνάρτηση του διανύσματος x καθώς κι των εισόδων του συστήματος τις οποίες ονομάζουμε με το ελληνικό χ για να μην υπάρχει σύγχυση.

Έχουμε λοιπόν ότι η έξοδος του συστήματος την εκάστοτε χρονική στιγμή είναι :

$$z_k = h_k(x, \chi^k) + w_k \quad (4.40)$$

Όπου w_k είναι ο θόρυβος στην εκάστοτε μέτρηση. Έχουμε υποθέσει αρχικά ότι ο θόρυβος είναι λευκός Gaussian με μέση τιμή 0 και variance σ^2 . Αυτό που μένει να πούμε είναι ποια είναι η συνάρτηση h_k και το διάνυσμα χ^k . Αρχικά ο λόγος που βάλουμε τον δείκτη k στην συνάρτηση h είναι ότι δεν θα την εκφράσω αναλυτικά συναρτήσει των χ^k και του x αλλά με την βοήθεια των $u_1(k)$, $u_2(k)$, ώστε να προκύψει μια πιο εύκολα χρησιμοποιήσιμη μορφή. Το διάνυσμα χ^k περιλαμβάνει όλες τις εισόδους μέχρι την χρονική στιγμή k δηλαδή είναι το διάνυσμα $\chi^k = [\chi(0)\chi(1)\dots\chi(k)]$. Έτσι λοιπόν την χρονική στιγμή k η έξοδος του συστήματος υπολογίζεται, υπολογίζοντας αναδρομικά τα $u_1(k)$, $u_2(k)$ από τις σχέσεις

$$u_1(k) = a_{11}u_1(k-1) + a_{12}u_2(k-2) + b_{10}\chi(k) + b_{11}\chi(k-1) \quad (4.41)$$

$$u_2(k) = a_{12}u_2(k-1) + a_{22}u_2(k-2) + b_{20}\chi(k) + b_{21}\chi(k-1) \quad (4.42)$$

και έπειτα υπολογίζοντας

$$h_k(x, \chi^k) = c_{11}u_1(k) + c_{12}u_1(k)^2 + c_{13}u_1(k)^3 + c_{21}u_2(k) + c_{22}u_2(k)^2 + c_{32}u_2(k)^3 + \sum_{i=1}^p k_k d(k-i) \quad (4.43)$$

Σύμφωνα λοιπόν με τα παραπάνω και ανατρέχοντας στον πίνακα (4.2.2) ο οποίος περιέχει τις εξισώσεις για ανανέωση της πρόβλεψης για το διάνυσμα x βλέπουμε ότι αυτές παίρνουν την παραπάνω μορφή. Ο λόγος για την μορφή αυτή των εξισώσεων

Time update equations

$$\hat{x}_k^- = x_{k-1} \quad (4.44)$$

$$\hat{P}_k^- = P_{k-1} \quad (4.45)$$

είναι ότι στην συγκεκριμένη εφαρμογή ο πίνακας A οριζόμενος ως 4.22 είναι ο μοναδιαίος πίνακας και ο πίνακας W (οριζόμενος όπως στην 4.23 είναι ο μηδενικός. Προχωράμε τώρα στο να βρούμε την μορφή που θα έχουν οι measurement update equations. Θα προσδιορίσουμε αρχικά τον πίνακα H ο οποίος υπολογίζεται από την σχέση 4.24. Στην συγκεκριμένη περίπτωση έχουμε ότι

$$H_k = \left[\frac{\partial h_k}{\partial a_{11}} \frac{\partial h_k}{\partial a_{12}} \frac{\partial h_k}{\partial a_{21}} \frac{\partial h_k}{\partial a_{22}} \frac{\partial h_k}{\partial b_{10}} \frac{\partial h_k}{\partial b_{11}} \frac{\partial h_k}{\partial b_{20}} \frac{\partial h_k}{\partial b_{21}} \frac{\partial h_k}{\partial c_{11}} \frac{\partial h_k}{\partial c_{12}} \frac{\partial h_k}{\partial c_{13}} \frac{\partial h_k}{\partial c_{21}} \frac{\partial h_k}{\partial c_{22}} \frac{\partial h_k}{\partial c_{23}} \frac{\partial h_k}{\partial k_1} \dots \frac{\partial h_k}{\partial k_p} \right] (\tilde{x}_k, \chi^k) \quad (4.46)$$

Οι επιμέρους μερικές παράγωγοι υπολογίζονται όπως δείχνει ο πίνακας (4.4)

Απομένει πλέον να καθορισθεί η τιμή του πίνακα V_k και του πίνακα R_k . Ο πίνακας V όπως ορίστηκε από την σχέση 4.25 είναι πίνακας στοιχείο και έχει την τιμή 1, (η εξίσωση μέτρησης για το z_k (4.40) είναι γραμμική ως προς το w_k) και επιπλέον ο πίνακας R είναι επίσης 1x1 δηλαδή στοιχείο και η τιμή του είναι σ^2 . Με βάση τα προηγούμενα ο αλγόριθμος για το Extended Kalman Filter για την συγκεκριμένη εφαρμογή παίρνει την μορφή του πίνακα (4.3):

Time update(Predict)

$$\hat{x}_k^- = \hat{x}_{k-1} \quad (4.47)$$

$$\hat{P}_k^- = P_{k-1} \quad (4.48)$$

Mesaurement update

$$K_k = \tilde{P}_k H_k^T (H_k \tilde{P}_k H_k^T + R_k)^{-1} \quad (4.49)$$

$$P_k = (I - K_k H_k) \tilde{P}_k \quad (4.50)$$

$$\hat{x}_k = \tilde{x}_k + K_k (z_k - h_k(\tilde{x}_k, \chi^k)) \quad (4.51)$$

Πίνακας 4.3: Extended Kalman Filter Algorithm Για Το Συγκεκριμένο Πρόβλημα

$$\frac{\partial h_k}{\partial a_{11}} = \frac{\partial h_k}{\partial u_1(k)} \frac{\partial u_1(k)}{\partial a_{11}} = [c_{11} + 2c_{12}u_1(k) + 3c_{13}u_1(k)^2][u_1(k-1)]$$

$$\frac{\partial h_k}{\partial a_{12}} = \frac{\partial h_k}{\partial u_1(k)} \frac{\partial u_1(k)}{\partial a_{12}} = [c_{11} + 2c_{12}u_1(k) + 3c_{13}u_1(k)^2][u_1(k-2)]$$

$$\frac{\partial h_k}{\partial a_{21}} = \frac{\partial h_k}{\partial u_2(k)} \frac{\partial u_2(k)}{\partial a_{21}} = [c_{21} + 2c_{22}u_2(k) + 3c_{23}u_2(k)^2][u_2(k-1)]$$

$$\frac{\partial h_k}{\partial a_{22}} = \frac{\partial h_k}{\partial u_2(k)} \frac{\partial u_2(k)}{\partial a_{22}} = [c_{21} + 2c_{22}u_2(k) + 3c_{23}u_2(k)^2][u_2(k-2)]$$

$$\frac{\partial h_k}{\partial b_{10}} = \frac{\partial h_k}{\partial u_1(k)} \frac{\partial u_1(k)}{\partial b_{10}} = [c_{11} + 2c_{12}u_1(k) + 3c_{13}u_1(k)^2][\chi(k)]$$

$$\frac{\partial h_k}{\partial b_{11}} = \frac{\partial h_k}{\partial u_1(k)} \frac{\partial u_1(k)}{\partial b_{11}} = [c_{11} + 2c_{12}u_1(k) + 3c_{13}u_1(k)^2][\chi(k-1)]$$

$$\frac{\partial h_k}{\partial b_{20}} = \frac{\partial h_k}{\partial u_2(k)} \frac{\partial u_2(k)}{\partial b_{20}} = [c_{21} + 2c_{22}u_2(k) + 3c_{23}u_2(k)^2][\chi(k)]$$

$$\frac{\partial h_k}{\partial b_{21}} = \frac{\partial h_k}{\partial u_2(k)} \frac{\partial u_2(k)}{\partial b_{21}} = [c_{21} + 2c_{22}u_2(k) + 3c_{23}u_2(k)^2][\chi(k-1)]$$

$$\frac{\partial h_k}{\partial c_{11}} = u_1(k)$$

$$\frac{\partial h_k}{\partial c_{12}} = u_1(k)^2$$

$$\frac{\partial h_k}{\partial c_{13}} = u_1(k)^3$$

$$\frac{\partial h_k}{\partial c_{21}} = u_2(k)$$

$$\frac{\partial h_k}{\partial c_{22}} = u_2(k)^2$$

$$\frac{\partial h_k}{\partial c_{23}} = u_2(k)^3$$

και για τις παραμέτρους που περιγράφουν την διαταραχή έχουμε

$$\frac{\partial h_k}{\partial k_j} = d(k-j)$$

4.4 Διαφοροποιήσεις Στον Αλγόριθμο

Δοκιμάζουμε να γραμμικοποιήσουμε τις σχέσεις (4.20) και (4.21) σε διαφορετικό σημείο από το \tilde{x}_k . Ονομάζουμε αυτό το σημείο ως προς το οποίο γραμμικοποιούμε \acute{x}_k και το ορίζουμε ως

$$\acute{x}_k = \sum_{i=1}^5 \frac{x_{k-i}}{5} \quad (4.52)$$

δηλαδή ορίζουμε την γραμμικοποίηση γύρω από την μέση τιμή των 5 τελευταίων μετρήσεων. Οι δύο λοιπόν αυτές εξισώσεις παίρνουν την εξής μορφή :

$$x_k = f(\acute{x}_k) + A(x_{k-1} - \acute{x}_k) + W w_{k-1} \quad (4.53)$$

$$z_k = \acute{z}_k + H(x_k - \acute{x}_k) + V v_k \quad (4.54)$$

όπου

•

$$\acute{z}_k = h(\acute{x}_k, 0) \quad (4.55)$$

•

$$A_{[i,j]} = \frac{\partial f_{[i]}}{\partial x_{[j]}}(\acute{x}_k, u_{k-1}, 0) \quad (4.56)$$

•

$$W_{[i,j]} = \frac{\partial f_{[i]}}{\partial w_{[j]}}(\acute{x}_k, u_{k-1}, 0) \quad (4.57)$$

•

$$H_{[i,j]} = \frac{\partial h_{[i]}}{\partial x_{[j]}}(\acute{x}_k, 0) \quad (4.58)$$

•

$$V_{[i,j]} = \frac{\partial h_{[i]}}{\partial v_{[j]}}(\acute{x}_k, 0) \quad (4.59)$$

αντίστοιχα με τις σχέσεις (4.22),(4.23),(4.24),(4.25).Κάνοντας παρόμοια εργασία με πριν θεωρώντας δηλαδή τον πίνακα A τον μοναδιαίο πίνακα και τον πίνακα W εξ'ορισμού μηδενικό καθώς επίσης και ότι ο πίνακας V είναι μονάδα τότε οι εξισώσεις αυτές παίρνουν την μορφή:

$$x_k = x_{k-1} \quad (4.60)$$

και

$$z_k = \acute{z}_k + H(x_k - \acute{x}_k) + V v_k \quad (4.61)$$

με τους πίνακες H και V όπως ορίστηκαν προηγουμένως. Με βάση την προηγούμενη αλλαγή οι εξισώσεις στο συγκεκριμένο πρόβλημα τροποποιούνται ως εξής :

4.4.1 Time Update Equations

Εδώ επειδή ουσιαστικά η f στην σχέση (4.16) είναι η ταυτοτική συνάρτηση δηλαδή $x_k = x_{k-1}$ οι εξισώσεις αυτές έχουν την ίδια μορφή με πριν(4.3). (Τόσο ο πίνακας A (Ο οποίος είναι μοναδιαίος διάστασης 14×14) αλλά και ο W (που είναι μηδέν αφού δεν υπάρχει θόρυβος) είναι ανεξάρτητοι ως προς το σημείο το οποίο γραμμικοποιούμε και είναι ίδιοι με πριν.) Αυτό δηλαδή σημαίνει ότι η πρόβλεψη για την τιμή του διανύσματος την χρονική στιγμή k πριν επεξεργαστούμε την τελευταία διαθέσιμη μέτρηση z_k , είναι ίση με τη βέλτιστη πρόβλεψη που είχαμε για την στιγμή $k-1$ δηλαδή την \hat{x}_{k-1}

4.4.2 Measurement Update Equations

Οι εξισώσεις αυτές διαφοροποιούνται δεδομένου ότι επιλέγουμε διαφορετικό σημείο ως προς το οποίο θα γραμμικοποιήσουμε τις εξισώσεις. Έτσι λοιπόν οι εξισώσεις αυτές γίνονται :

$$K_k = P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1} \quad (4.62)$$

$$\hat{x}_k = \hat{x}_k^- + K_k (z_k - h_k(\hat{x}_k, 0)) \quad (4.63)$$

$$P_k = (I - K_k H_k) P_k^- \quad (4.64)$$

Αν και μορφολογικά οι παραπάνω εξισώσεις έχουν προφανείς ομοιότητες με τις αντίστοιχες (4.3) θα επισημάνω για ακόμη μια φορά ότι οι πίνακες H_k και K_k καθώς και η τιμή της συνάρτησης h_k είναι διαφορετικές με πριν δεδομένου ότι υπολογίζονται σε διαφορετικό σημείο.

Κεφάλαιο 5

Ζητήματα Που Παρουσιάστηκαν

Στο κεφάλαιο αυτό θα παρουσιάσουμε ορισμένα από τα σημαντικότερα τα ζητήματα τα οποία αντιμετωπίσαμε στην πορεία του αλγορίθμου και τα οποία θα πρέπει να αναφερθούν πριν προχωρήσουμε στα αποτελέσματα. Επιπλέον θα αναφέρουμε ορισμένες τροποποιήσεις που έγιναν στους αλγορίθμους κι τις ανάγκες που τις επέβαλλαν.

5.1 Αστάθεια

Ένα από τα βασικότερα ζητήματα τα οποία υπήρξαν κατά την υλοποίηση του φίλτρου Kalman και την προσαρμογή του στα πραγματικά δεδομένα ήταν το πρόβλημα της αστάθειας.

Πριν προχωρήσουμε όμως σε πρακτικά ζητήματα θα εξετάσουμε αναλυτικότερα το σύστημά μας ώστε να δούμε πως μπορεί να οδηγηθεί σε αστάθεια και σε ποιές παραμέτρους οφείλεται ένα τέτοιο φαινόμενο.

Όπως είχαμε αναφέρει και στο δεύτερο κεφάλαιο αν εξετάσουμε την συνάρτηση μεταφοράς που προκύπτει από την έξοδο του πρώτου PDM πριν την είσοδό της στην μη γραμμικότητα βλέπουμε ότι έχει την εξής μορφή:

$$\frac{U_1(z)}{X(z)} = \frac{z(zb_{10} + b_{11})}{z^2 - a_{11}z - a_{12}z}$$

Για να είναι ευσταθής θα πρέπει οι πόλοι της να βρίσκονται εντός του μοναδιαίου κύκλου. Στο συμπέρασμα αυτό θα μπορούσαμε να καταλήξουμε επίσης γράφοντας την σχέση εισόδου-εξόδου του πρώτου PDM σε μορφή εξισώσεων κατάστασης ως εξής:

$$\begin{bmatrix} u_1(n) \\ w(n) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ 1 & 0 \end{bmatrix} \begin{bmatrix} u_1(n-1) \\ w(n-1) \end{bmatrix} + \begin{bmatrix} b_{10} & b_{11} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \chi(n) \\ \chi(n-1) \end{bmatrix} \quad (5.1)$$

όπου $w(n) = u_1(n-1)$. Απαιτώντας για ευστάθεια οι ιδιοτιμές τις μήτρας $A = \begin{bmatrix} a_{11} & a_{12} \\ 1 & 0 \end{bmatrix}$ να βρίσκονται εντός του μοναδιαίου κύκλου καταλήγουμε στο ίδιο αποτέλεσμα με πριν.

Στο δεύτερο κεφάλαιο είχαμε παρουσιάσει με σχετικά διαγράμματα ότι ακόμη και αν έχουμε πολύ μικρή είσοδο (συγκεκριμένα χρησιμοποιήσαμε λευκό Gaussian θόρυβο) και για πολύ μικρές αποκλίσεις από τις πραγματικές τιμές τιμές το σφάλμα μπορεί να γίνει τεράστιο.

Ο λόγος για τα τεράστια αυτά σφάλματα είναι ότι για τις τιμές παραμέτρων που παρατηρούνται τεράστια σφάλματα το σύστημα είναι ασταθές.

Kalman Filter

Όσον αφορά την συγκεκριμένη εφαρμογή χρησιμοποιώντας την κοινή λογική καταλαβαίνουμε ότι το σύστημα ινσουλίνης-γλυκόζης στον ανθρώπινο οργανισμό δεν θα μπορούσε ποτέ να είναι ασταθές.

Κατά την διαδικασία εκτίμησης των παραμέτρων όμως με το φίλτρο Kalman θα μπορούσε ο αλγόριθμός μας ενώ ξεκινάει από κάποιο ευσταθές σημείο να εγκλωβιστεί σε κάποιο σημείο που οδηγεί σε ασταθές σύστημα. Στα επόμενα θα προσπαθήσουμε να δώσουμε μια εικόνα του λόγου που συμβαίνει αυτό και να παρουσιάσουμε την λύση που χρησιμοποιήσαμε ώστε να αποφύγουμε αυτά τα σημεία.

Για να συγκρίνουμε το φίλτρο Kalman με τους αλγορίθμους βελτιστοποίησης που παρουσιάσαμε στο κεφάλαιο 2 και να δείξουμε γιατί το πρόβλημα αυτό είναι τόσο έντονο στο φίλτρο Kalman θα εξετάσουμε την διαδικασία εκτίμησης από την χρονική στιγμή N στην επόμενη χρονική στιγμή N+1.

Ας υποθέσουμε αρχικά ότι βρισκόμαστε στην χρονική στιγμή N και έχουμε το διάνυσμα μετρήσεων $\hat{z}_N = [z_1 \dots z_N]$ και την βέλτιστη εκτίμηση για αυτή την χρονική στιγμή θ_{opt}^N . Την επόμενη χρονική στιγμή N+1 παίρνουμε την νέα μέτρηση z_{N+1} και έχουμε το νέο διάνυσμα παραμέτρων $\hat{z}_{N+1} = [\hat{z}_N z_{N+1}]$.

Ας πάρουμε σαν παράδειγμα αλγορίθμου βελτιστοποίησης την μέθοδο Levenberg-Marquardt την οποία χρησιμοποιήσαμε κατά κόρον. Για να εκτιμήσουμε με την μέθοδο αυτή το σημείο που ελαχιστοποιεί το τετραγωνικό κριτήριο ο αλγόριθμος κατασκευάζει μια ακολουθία σημείων

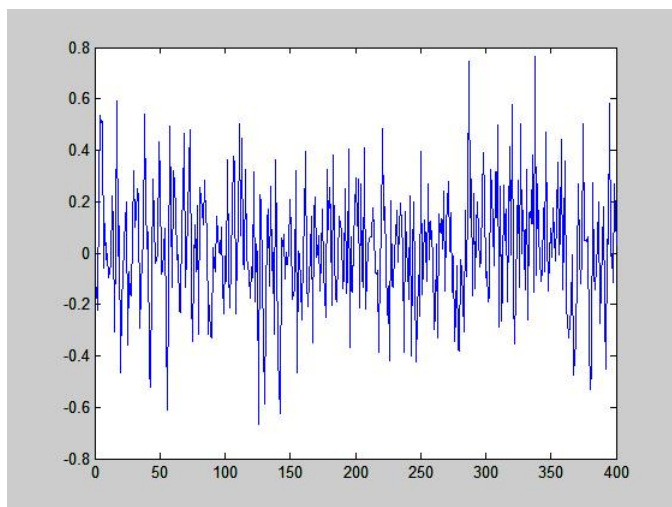
$$\theta_{opt}^N \equiv \theta_0^{N+1} \rightarrow \theta_1^{N+1} \rightarrow \theta_2^{N+1} \rightarrow \dots \rightarrow \theta_{opt}^{N+1}$$

η οποία συγκλίνει στην βέλτιστη λύση.

Όσον αφορά το Kalman φίλτρο τώρα με βάση την νέα μέτρηση η τιμή του διανύσματος x ανανεώνεται προς μια και μοναδική κατεύθυνση σε ένα και μόνο βήμα. Αν θα μπορούσαμε να φανταστούμε μια υποθετική ακολουθία σημείων που θα οδηγούσε στην βέλτιστη λύση αυτό που κάνει το φίλτρο Kalman ουσιαστικά είναι να πάρει ως βέλτιστη λύση για την χρονική στιγμή N+1 το πρώτο σημείο αυτής της ακολουθίας και όταν έρθει η νέα μέτρηση την χρονική στιγμή N+2 χρησιμοποιώντας ως βέλτιστη λύση αυτή την οποία πήρε από την προηγούμενη χρονική στιγμή κάνει μια και μοναδική ανανέωση και δίνει αυτή ως βέλτιστη λύση. Αυτή η πορεία δείχνεται στο παρακάτω σχήμα.

$$\begin{array}{ccc} z_{N+1} & \theta_{opt}^N \equiv \theta_0^{N+1} \rightarrow \theta_1^{N+1} & \\ & \Downarrow & \\ z_{N+2} & \theta_0^{N+2} \rightarrow \theta_1^{N+2} & \\ & \Downarrow & \\ z_{N+3} & \theta_0^{N+3} \rightarrow \theta_1^{N+3} & \end{array}$$

Με αυτή τη διαδικασία λοιπόν είναι πιθανό μετά από κάποια επανάληψη η ανανέωση να οδηγήσει σε ασταθές σύστημα. Τα αποτελέσματα για ένα τέτοιο σύστημα είναι μετά από κάποιες επαναλήψεις να "χαθεί" ο αλγόριθμος και να επιστρέφονται



Σχήμα 5.1: Είσοδος

τιμές "τεράστιας" τάξης οι οποίες δεν είναι μέσα στο όριο τιμών που μπορεί να επεξεργαστεί ο υπολογιστής.

Θα παρουσιάσουμε ένα παράδειγμα εξέλιξης ενός datarecord το οποίο υποπέττει σε αυτή την κατάσταση. Θεωρούμε την ακόλουθη είσοδο και την ακόλουθη έξοδο του συστήματος και πραγματικές τιμές των παραμέτρων αυτές του πίνακα του κεφαλαίου 2.

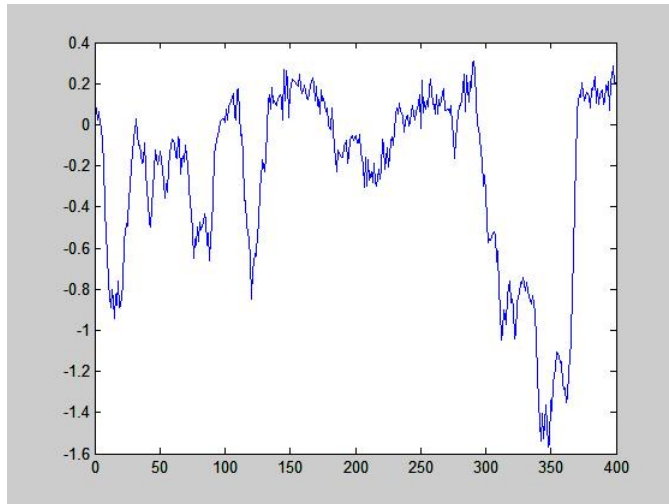
Για το συγκεκριμένο παράδειγμα χρησιμοποιήσαμε μόνο την διαταραχή με πειραματικά δεδομένα κι όχι την πλήρη θεώρηση του πρόβληματος με το μοντέλο για το Disturbance. Ακόμη στο μοντέλο ινσουλίνης-γλυκόζης χρησιμοποιήθηκαν πολύ μικρές τιμές για την είσοδο. Ο λόγος ήταν ότι θέλουμε να δείξουμε το πρόβλημα σε όλη του την έκταση και που μπορεί να οδηγήσει.

Σαν είσοδο όπως προείπα χρησιμοποιήσαμε λευκό Gaussian θόρυβο ο οποίος έδωσε την έξοδο που φαίνεται στο σχήμα που ακολουθεί.

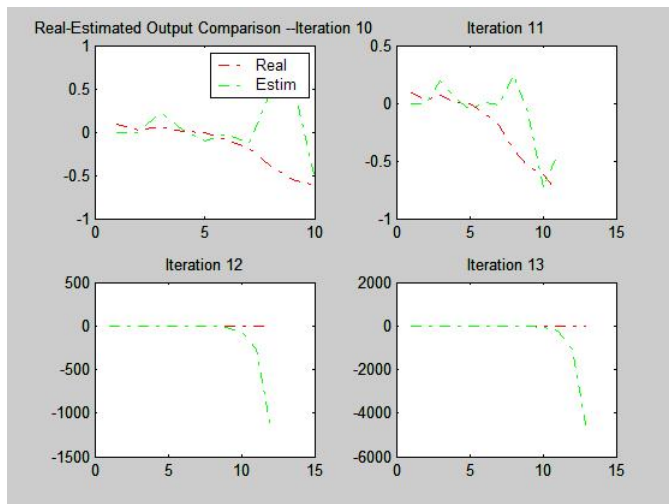
Οι αρχικές τιμές των παραμέτρων φαίνονται στον ακόλουθο πίνακα στον οποίο παρατίθενται κι οι πραγματικές τιμές. Βλέπουμε ότι όλες οι παράμετροι βρίσκονται πολύ κοντά στις πραγματικές τους τιμές.

Όνομα Παραμέτρου	a_{11}	a_{12}	a_{21}	a_{22}	b_{10}	b_{11}	b_{20}	b_{21}
Πραγματική τιμή	1,7	-0.72	1,55	-0,595	0	0,1	0	0,15
Αρχική Τιμή	1,72	-0.74	1,57	-0,615	0,04	0,05	0,01	0,075

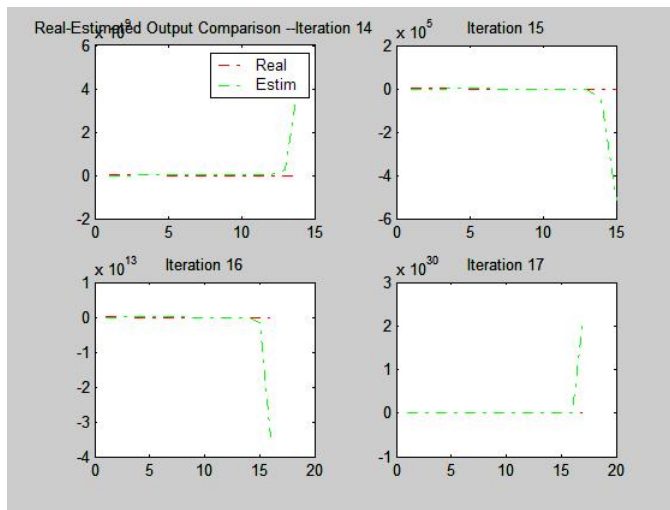
Όνομα Παραμέτρου	c_{11}	c_{12}	c_{13}	c_{21}	c_{22}	c_{23}
Πραγματική τιμή	-2	-3,5	-0,2	1	2	0
Αρχική Τιμή	-2,14	-3,54	-0,23	1,1	2,04	0,1



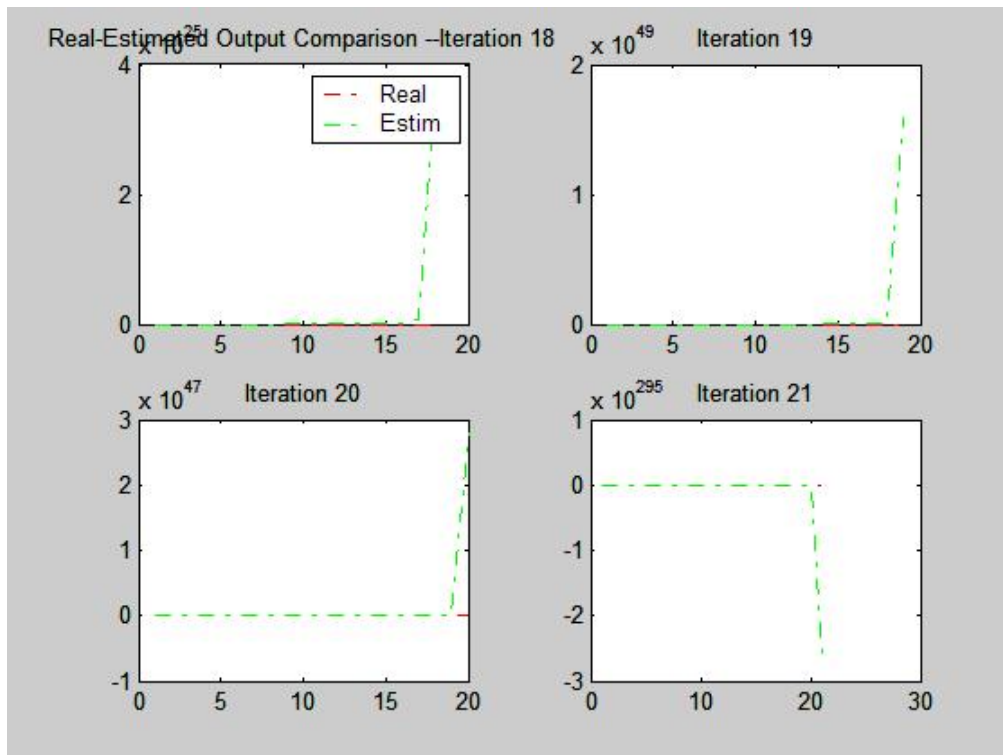
Σχήμα 5.2: Έξοδος Μοντέλου



Σχήμα 5.3: Επαναλήψεις 10-13



Σχήμα 5.4: Επαναλήψεις 14-17



Σχήμα 5.5: Επαναλήψεις 18-21

Αυτό που θα πρέπει να τονίσουμε εδώ είναι ότι από την έβδομη επανάληψη και μετά (από το έβδομο δείγμα δηλαδή) όλες οι επαναλήψεις δίνουν ασταθές σύστημα.

Το αποτέλεσμα είναι μετά από την εικοστή δεύτερη επανάληψη να έχουμε πολύ μεγάλες εξόδους οι οποίες δίνουν τιμές παραμέτρων της τάξης του 10^5 και τόσο μεγάλους αριθμούς οι οποίοι ξεφεύγουν από το εύρος του συστήματος. Βέβαια το μέχρι ποιους αριθμούς μπορεί να επεξεργαστεί ένα σύστημα διαφέρει μεταξύ των υπολογιστών η ουσία όμως είναι ότι δεν θα θέλαμε να έχουμε τέτοιες καταστάσεις οι οποίες αφενός μπορεί να δημιουργήσουν υπολογιστικά προβλήματα αφετέρου οδηγούν τις παραμέτρους του συστήματος σε μεγάλες τιμές που δίνουν ασταθές σύστημα.

Προκειμένου να αντιμετωπίσουμε αυτό το φαινόμενο το οποίο εάν σκεφτούμε το πλήρες μοντέλο θα ήταν πολύ πιο έντονο αφενός επειδή οι τιμές της εγγεόμενης ινσουλίνης δεν θα ήταν τόσο μικρές κι αφετέρου επειδή θα είχαμε κι το Disturbance, προσθέσαμε μια έρευνα γραμμής στην κατεύθυνση που ορίζει το φίλτρο. Ουσιαστικά λοιπόν η μορφή που παίρνουν οι εξισώσεις του τροποποιημένου φίλτρου Kalman θα είναι αυτές που φαίνονται στον ακόλουθο πίνακα όπου λ_k προέρχεται από έρευνα γραμμής στην κατεύθυνση που ορίζει το φίλτρο.

Time update(Predict)

$$\hat{x}_k^- = \hat{x}_{k-1}$$

$$\hat{P}_k^- = P_{k-1}$$

Mesaurement update

$$K_k = \tilde{P}_k H_k^T (H_k \tilde{P}_k H_k^T + R_k)^{-1}$$

$$P_k = (I - K_k H_k) \tilde{P}_k$$

$$\hat{x}_k = \tilde{x}_k + \lambda_k * K_k (z_k - h_k(\tilde{x}_k, \chi^k))$$

5.2 Μέγεθος Δεδομένων

Όπως είχα αναφέρει και στο δεύτερο κεφάλαιο τα δεδομένα είναι οι ενέσεις ινσουλίνης και η χρονική στιγμή στην οποία γίνονται καθώς και η ποσότητα της γλυκόζης στο αίμα η οποία μετράται ανα πεντάλεπτο. Εδώ προκύπτει ένας αντικειμενικός περιορισμός ότι υπο φυσιολογικές συνθήκες θα μπορούσε ένας ασθενής να είναι καλωδιωμένος περίπου 8-10 ώρες οπότε θα είχαμε περίπου 100-120 μετρήσεις.

Θα υπενθυμίσουμε για ακόμη μια φορά ότι αντικειμενικός σκοπός είναι η εκτίμηση των συντελεστών του μοντέλου ινσουλίνης-γλυκόζης. Σε περίπτωση τώρα όπως συμβαίνει και με το υπάρχον δείγμα (σχήμα 5.6) το οποίο είναι πολύ μεγάλου μεγέθους (863) θα θέλαμε το μέγεθος των δεδομένων που θα πάρουμε δηλαδή τα 100-150 δείγματα να είναι τέτοια ώστε να περιέχουν όσο το δυνατόν περισσότερες εγχύσεις ινσουλίνης ώστε να μπορέσουμε να εκτιμήσουμε το μοντέλο όσο το δυνατό καλύτερα.

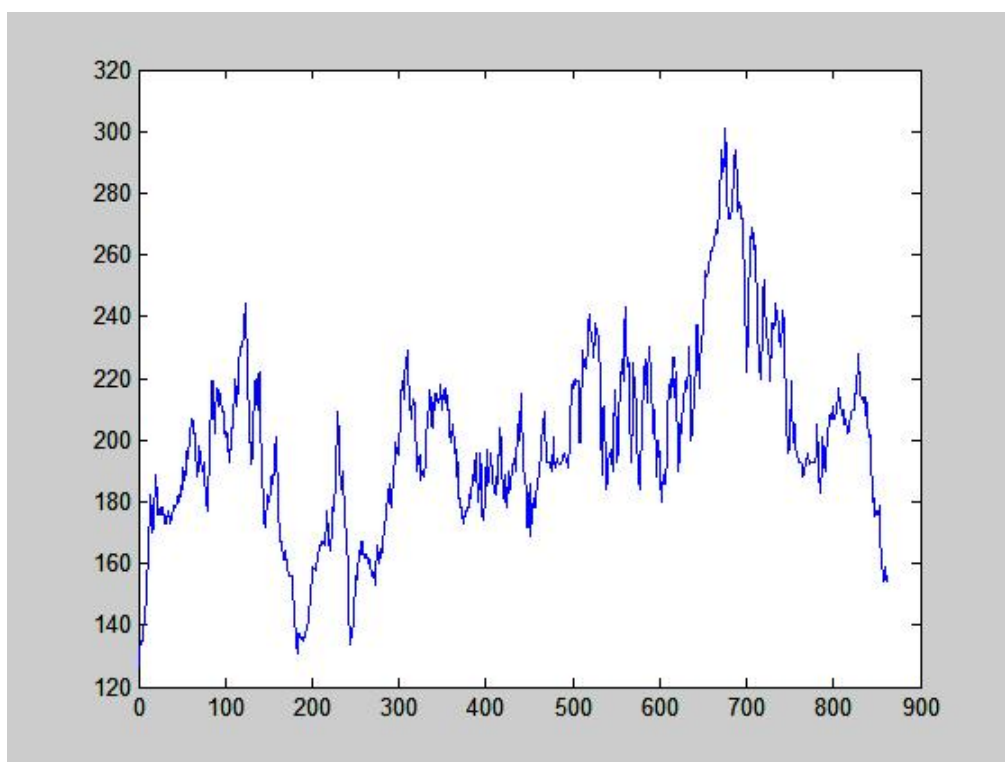
Στο παρόν δείγμα έχουμε 19 εγχύσεις ινσουλίνης τις χρονικές στιγμές που φαίνονται στον ακόλουθο πίνακα :

Χρονική Στιγμή	2	23	36	48	64	167	297	307	351	430	451
Ποσότητα Ινσ.	7	4	2	4	12	8	6	12	9,7	3,5	5

Χρονική Στιγμή	474	482	551	621	625	724	737	760
Ποσότητα Ινσ.	6	5,5	8	5	8	6	9,1	11,5

Δεδομένου λοιπόν ότι θα χρησιμοποιήσουμε ένα αρχείο μεγέθους περίπου 120-150 δειγμάτων κι ότι θέλουμε αυτό να περιέχει όσο το δυνατόν περισσότερα δείγματα μπορούμε να επιλέξουμε είτε την περιοχή 1-150 είτε την περιοχή 350-500.

Επιπλέον ένα ακόμη σημείο το οποίο πρέπει να αναφέρουμε είναι ότι προκειμένου να "εκπαιδεύσουμε" το σύστημά μας όσο το δυνατό καλύτερα θα κάνουμε διαδοχικές προσπελάσεις στο datarecord . Αυτό σημαίνει ότι όταν φτάσουμε στο τέλος του αρχείου κρατώντας τις τιμές που έχουμε καταλήξει και χρησιμοποιώντας τις ως αρχικές εφαρμόζουμε εκ νέου τον αλγόριθμο σε όλο το αρχείο. Το πόσες φορές θα περάσουμε το datarecord εξαρτάται από το ποια κριτήρια σύγκλισης του Kalman φίλτρου έχουμε επιλέξει ή από το εάν έχουμε ορίσει ένα μέγιστο αριθμό επαναλήψεων



Σχήμα 5.6: Τιμές Γλυκόζης

Κεφάλαιο 6

Αποτελέσματα

6.1 Δείγματα 350-520

Στην προηγούμενη ενότητα παρουσιάσαμε αναλυτικά το πρόβλημα της αστάθειας και που μπορεί να οδηγήσει καθώς επίσης και τις περιοχές μετρήσεων στις οποίες θα εφαρμόσουμε τον αλγόριθμο. Στην δεύτερη ενότητα είχαμε κάνει μια προσπάθεια να οπτικοποιήσουμε πόσο μεγάλες αποκλίσεις στα τετραγωνικά σφάλματα μπορούν να προκύψουν από μικρές αλλαγές των παραμέτρων. Πολύ σημαντικά επίσης όσον αφορά την σύγκλιση του αλγορίθμου είναι και τα αρχικά σημεία.

Όπως είναι ανεμενόμενο σε χώρο διάστασης ίσο με το άθροισμα της τάξης του μοντέλου που χρησιμοποιούμε συν το μέγεθος των παραμέτρων του μοντέλου γλυκόζης -ινσουλίνης(14) περιμένουμε να έχουμε πολλά τοπικά ελάχιστα τα οποία θα ικανοποιούν τις αναγκαίες συνθήκες κι τα οποία θα μπορούσαν να αποτελούν σημεία σύγκλισης των αλγορίθμων. Όπως γίνεται κατανοητό ένας αλγόριθμος ανάλογα με την επιλογή του αρχικού σημείου μπορεί να μην συγκλίνει στο ολικό ελάχιστο της συνάρτησης αλλά σε κάποιο τοπικό ελάχιστο.

Το γεγονός αυτό εξαρτάται και από την τάξη του μοντέλου για την διαταραχή. Αναμένουμε ότι χρησιμοποιώντας μοντέλο μεγαλύτερης τάξης θα πάρουμε τουλάχιστον εξίσου καλά αποτελέσματα με οποιοδήποτε μοντέλο μικρότερης τάξης. Αυτό εξηγείται αν απλά σκεφτούμε ότι μπορούμε να θέσουμε όλους τους συντελεστές των όρων μεγαλύτερης τάξης ίσους με μηδέν και να πάρουμε το βέλτιστο αποτέλεσμα που θα έδινε το μικρότερης τάξης μοντέλο.

Επιπλέον χρειαζόμαστε ένα τρόπο ώστε να αξιολογήσουμε κατά πόσο το λάθος το οποίο παίρνουμε είναι ικανοποιητικό και κατά πόσο το τελικό σημείο στο οποίο καταλήξαμε είναι ικανοποιητικό ή όχι. Σαν κριτήριο θα έχουμε την τιμή του κανονικοποιημένου μέσου τετραγωνικού σφάλματος η οποία δεν είναι τίποτα άλλο παρά το άθροισμα των τετραγώνων των residuals διά το άθροισμα των τετραγώνων των μετρήσεων από τις οποίες έχουμε αφαιρέσει την μέση τιμή τους, δηλαδή ορίζεται ως :

$$l_{NLS} = \frac{\sum_{i=1}^N (y_i - z_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (6.1)$$

όπου κατά τα γνωστά y_i είναι η μέτρηση την χρονική στιγμή i $\bar{y} = \frac{1}{M} \sum_{i=1}^M y_i$, η μέση τιμή των μετρήσεων και z_i είναι η εκτιμώμενη τιμή για το σύστημα που θα μας έδινε το μοντέλο με τις υπάρχουσες τιμές συντελεστών. Για αξιόπιστες πρακτικές

εφαρμογές θα ήταν αποδεκτή μια τιμή του σφάλματος κοντά στο 10%.

Αυτό που αντικειμενικά θα επιθυμούσαμε θα ήταν να επιτύχουμε τον σκοπό μας με όσο το δυνατό μικρότερης τάξης μοντέλο. Επιπλέον δεν θα θέλαμε να είχαμε αριθμό παραμέτρων μεγαλύτερο από το $1/3$ των εξισώσεων (δηλαδή των δειγμάτων). Έτσι δεδομένου ότι θα χρησιμοποιήσουμε τουλάχιστον 120 δείγματα και μέχρι 150 για τον λόγο που προανέφερα είναι επιτρεπτό ένα μοντέλο το πολύ τάξης 30.

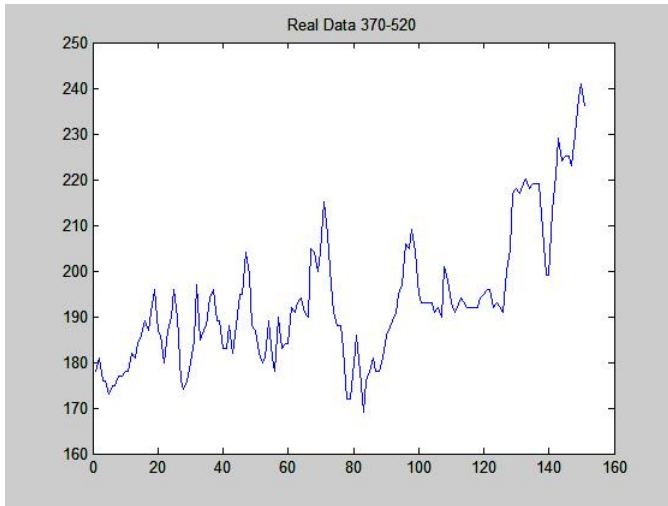
Όσον αφορά την επιλογή των αρχικών σημείων για τον αλγόριθμο μας αυτή έγινε τυχαία από ένα σύνολο αρχικών σημείων με βάση ποιο σημείο έδινε το μικρότερο τετραγωνικό σφάλμα. Αυτό το σημείο χρησιμοποιούνταν ως αρχικό σημείο και έπειτα εφαρμόζαμε το φίλτρο Kalman καθώς και την μέθοδο Levenberg-Marquardt.

6.1.1 Τάξη Μοντέλου 25

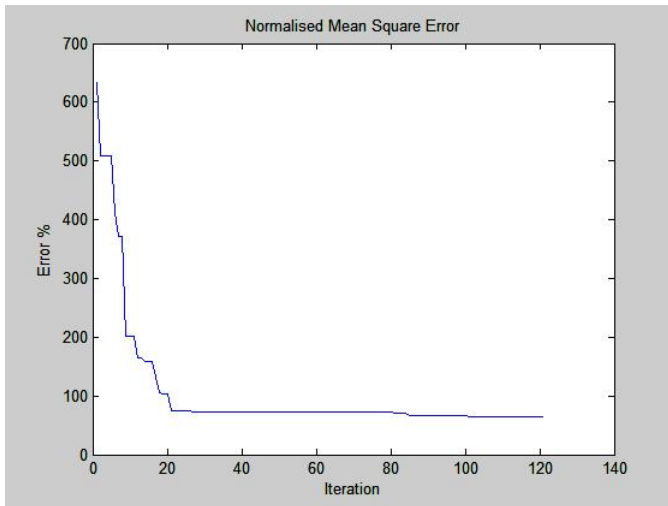
Τετραγωνικό Σφάλμα

Χρησιμοποιήσαμε ως datarecord τα δεδομένα από 370-520. Οι τιμές της γλυκόζης για αυτή την περίοδο φαίνεται στο σχήμα (6.1). Για την αρχικοποίηση των συντελεστών του disturbance χρησιμοποιήθηκαν τα πρώτα 25 δείγματα και τα υπόλοιπα χρησιμοποιήθηκαν σαν data για την εκπαίδευση.

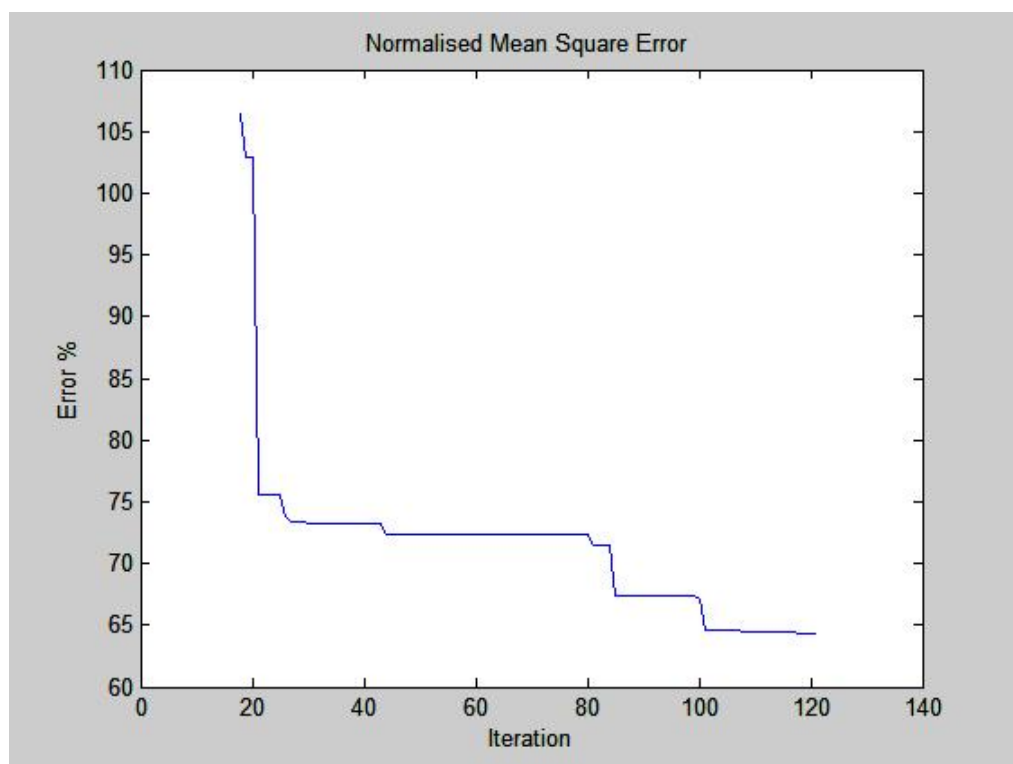
Για την εκτίμηση χρησιμοποιήσαμε τόσο την μέθοδο Levenberg-Marquardt όσο και το φίλτρο Kalman τα αποτελέσματα των οποίων θα συγκρίνουμε. Όσον αφορά το φίλτρο Kalman η τελική τιμή του κανονικοποιημένου μέσου τετραγωνικού σφάλματος μετά το πρώτο πέρασμα του datarecord ήταν 64%. Στο σχήμα 6.2 βλέπουμε την πορεία ελάττωσης του σφάλματος μέχρι να καταλήξει στην τελική τιμή του. Επειδή η τιμή αυτή δεν είναι ευδιάκριτη εξαιτίας των αρχικών τεράστιων τετραγωνικών σφαλμάτων, στο σχήμα 6.3 δίνουμε τα αποτελέσματα από την δέκατη όγδοο επανάληψη και έπειτα. Αν και μέχρι στιγμής δεν έχουμε παρουσιάσει διαφορετικά μεγέθη μοντέλων για να μπορέσει να γίνει η σύγκριση ωστόσο πρέπει να σημειώσουμε εδώ ότι φτάσαμε στην τιμή αυτή μετά από ένα μόνο πέρασμα του Datarecord και μόνο. Επιχειρήσαμε ένα ακόμη πέρασμα του Datarecord και καταλήξαμε σε σφάλμα 49.83%. Με την μέθοδο Levenberg-Marquardt καταλήξαμε σε σφάλμα 30,52%.



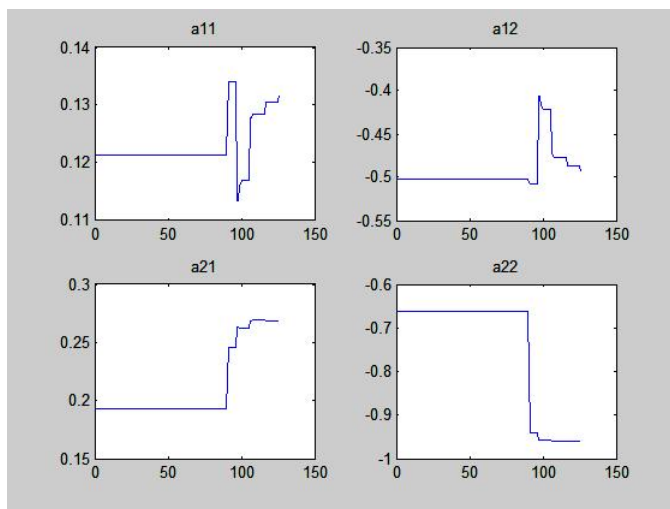
Σχήμα 6.1: Τιμές Γλυκόζης 370-520



Σχήμα 6.2: Τιμές Γλυκόζης 370-520



Σχήμα 6.3: Κανονικοποιημένο Σφάλμα

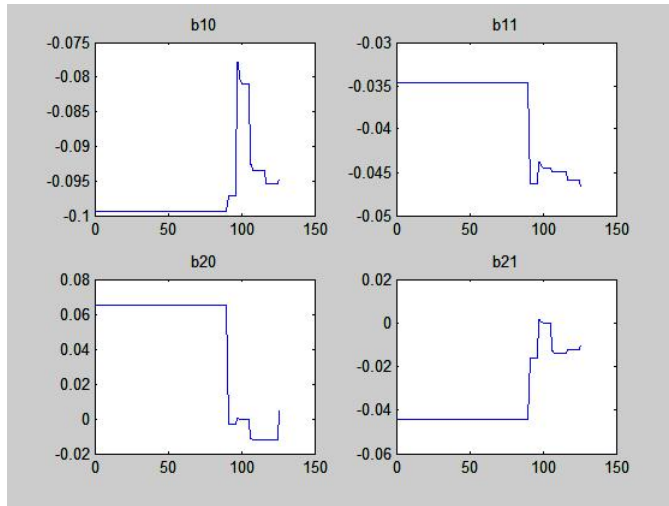


Σχήμα 6.4: Παράμετροι a

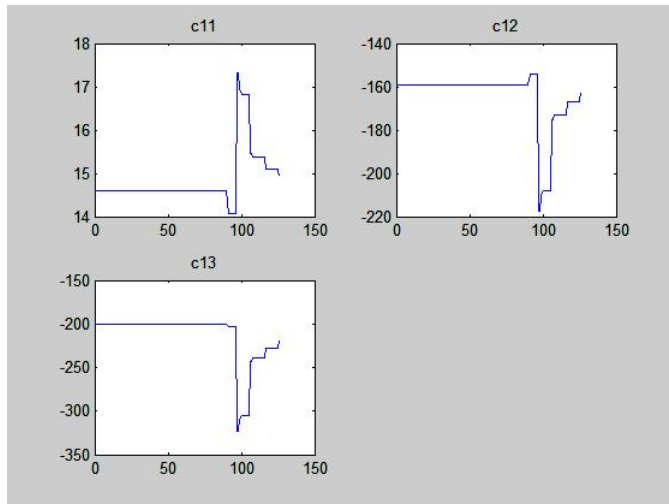
Τιμές Παραμέτρων

Τα διαγράμματα (6.4,6.5,6.6,6.7) απεικονίζουν την μεταβολή των τιμών των παραμέτρων του μοντέλου ινσουλίνης-γλυκόζης κατά το πρώτο πέρασμα του datarecord (φίλτρο Kalman). Επιπλέον, υπάρχει κι συνολικός πίνακας με τις αρχικές και τελικές τιμές κάθε μιας εκ των παραμέτρων του μοντέλου κι της διαταραχής τόσο με την μέθοδο Levenberg-Marquardt όσο και με το φίλτρο Kalman.

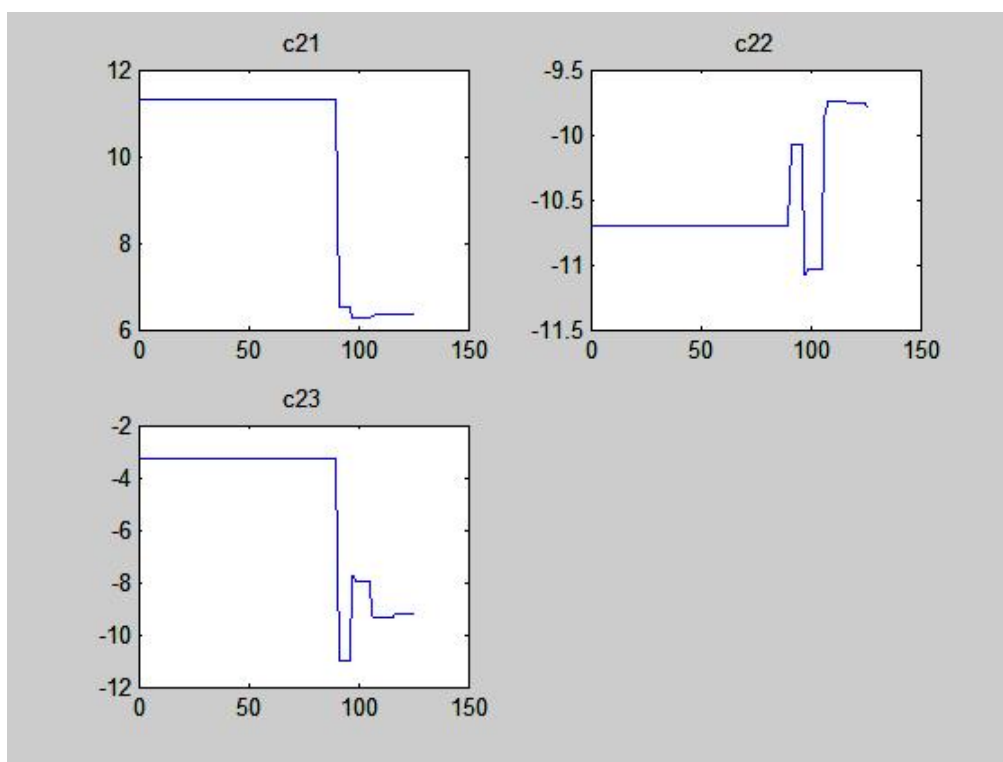
Αυτό που θα πρέπει να παρατηρήσουμε όσον αφορά τα σχήματα που δείχνουν την μεταβολή των παραμέτρων κατά το πρώτο πέρασμα του datarecord, είναι ότι οι τιμές μέχρι τα πρώτα περίπου 80 δείγματα παραμένουν σταθερές. Ο λόγος για το γεγονός αυτό είναι ότι μέχρι περίπου το δείγμα 35 δεν έχουμε καμία ένεση ινσουλίνης οπότε και δεν αλλάζουν οι συντελεστές αφού το μοντέλο δεν επηρεάζει την έξοδο και για την επόμενη πρώτη ένεση στο datarecord δεν πραγματοποιείται καμία αλλαγή στους συντελεστές του μοντέλου.



Σχήμα 6.5: Παράμετροι b



Σχήμα 6.6: Παράμετροι c_1



Σχήμα 6.7: Παράμετροι c_2

Όνομα Παραμέτρου	a_{11}	a_{12}	a_{21}	a_{22}
Αρχική τιμή	0,1214	-0.5023	0,193	-0,6632
Kalman Filter	0,1899	-0.7397	0,187	-0,5873
Levenberg-Mar	0.1337	-0.9088	0.1341	-1.0203

Όνομα Παραμέτρου	b_{10}	b_{11}	b_{20}	b_{21}
Αρχική τιμή	-0,0994	-0,0347	0,0651	-0,0442
Kalman Filter	-0,1360	-0,0781	0,0386	0,0123
Levenberg-Mar	-0.0805	-0.0335	0.0839	-0.0509

Όνομα Παραμέτρου	c_{11}	c_{12}	c_{13}	c_{21}	c_{22}	c_{23}
Αρχική τιμή	14,5938	-159,2303	-199,9099	11,3137	-10,6948	-3,2936
Kalman Filter	9,6895	-59,5355	-38,5237	10,2110	-14,1146	7,2823
Levenberg-Mar	28.4542	-114.6196	-146.3447	15.0000	-10.0000	-5.0000

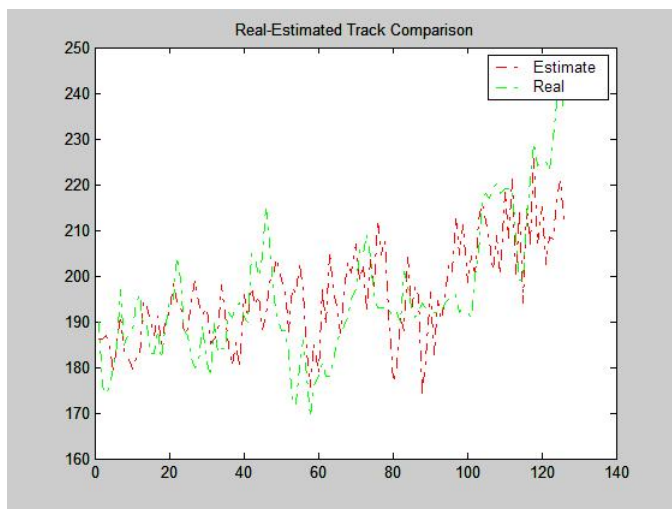
Όνομα Παραμέτρου	k_1	k_2	k_3	k_4	k_5
Αρχική τιμή	0,422	-0,772	-0,213	0,647	-0,7179
Kalman Filter	-0.1773	0.536	-0.2711	0.6943	-0.1316
Levenberg-Mar	0.3858	-0.929	-0.184	0.475	-0.6724

Όνομα Παραμέτρου	k_6	k_7	k_8	k_9	k_{10}
Αρχική τιμή	0,6796	0,4324	-0,2926	0,9249	-0,0735
Kalman Filter	-0.4806	0.8215	-0.2225	-0.1275	0.2967
Levenberg-Mar	0.4590	0.3347	-0.4484	0.6938	-0.1063

Όνομα Παραμέτρου	k_{11}	k_{12}	k_{13}	k_{14}	k_{15}
Αρχική τιμή	-0,2828	0.3234	0.1213	-0.4401	0.610
Kalman Filter	-0.6021	-0.0637	0.6776	-0.1157	0.0593
Levenberg-Mar	-0.2483	0.2766	0.0954	-0.5002	0.6479

Όνομα Παραμέτρου	k_{16}	k_{17}	k_{18}	k_{19}	k_{20}
Αρχική τιμή	-0.5317	-0.269	0.266	-0.3479	0.2594
Kalman Filter	-0.0118	-0.7806	0.1751	0.2739	-0.0841
Levenberg-Mar	-0.5408	-0.2372	0.3930	-0.6930	0.2011

Όνομα Παραμέτρου	k_{21}	k_{22}	k_{23}	k_{24}	k_{25}
Αρχική τιμή	0.3099	-0.2147	0.5936	-0.0647	0.6813
Kalman Filter	0.3607	-0.1347	0.2250	0.2955	-0.2016
Levenberg-Mar	0.3057	-0.3011	0.9837	-0.0984	0.7508

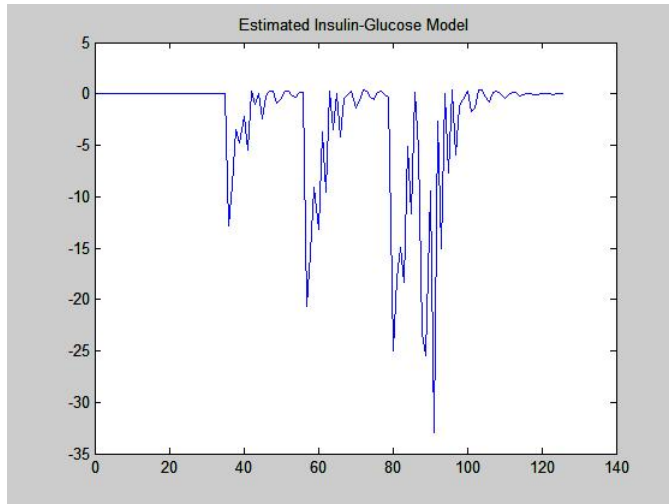


Σχήμα 6.8: Συγκριτικό Διάγραμμα

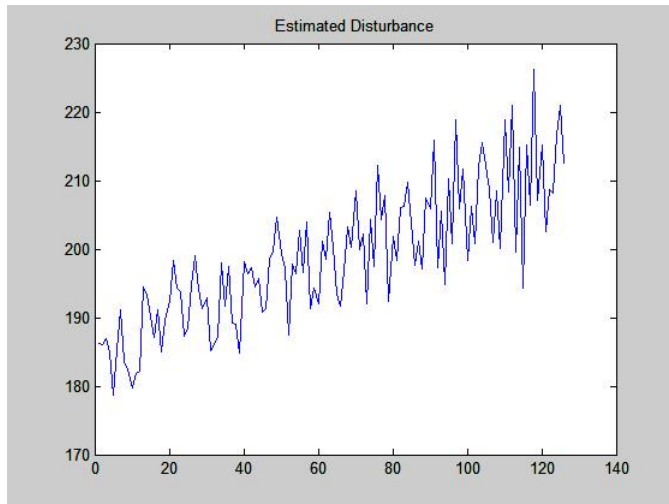
Εκτιμώμενη καμπύλη μοντέλου κι Disturbance

Στην ενότητα αυτή παραθέτουμε συγκριτικά διαγράμματα της καμπύλης που προκύπτει από τις εκτιμώμενες τιμές παραμέτρων σε σχέση με την πραγματική καμπύλη για κάθε μία εκ των δύο μεθόδων που χρησιμοποιήσαμε. Επιπλέον για κάθε μία εκ των δύο μεθόδων δίνουμε το μέρος της καμπύλης που οφείλεται στο disturbance και το μέρος της καμπύλης που οφείλεται στο μοντέλο ινσουλίνης γλυκόζης.

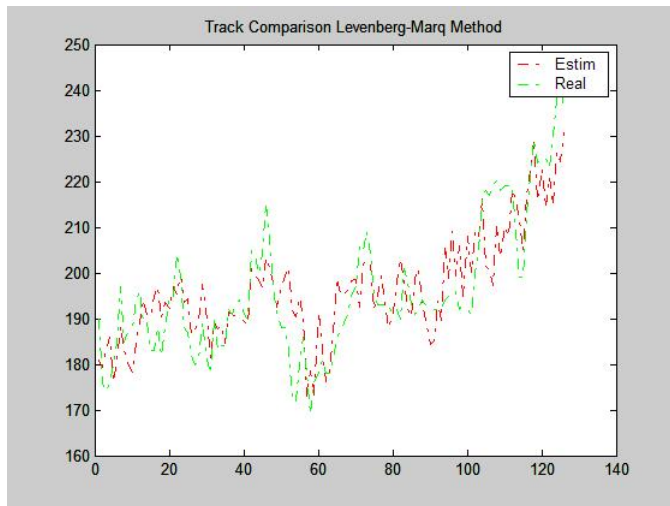
Αυτό που θα πρέπει να παρατηρήσουμε είναι ότι οι δύο μέθοδοι καταλήγουν σε αποτελέσματα για το μοντέλο ινσουλίνης-γλυκόζης τα οποία έχουν πολύ διαφορετική ερμηνεία κι διαφέρουν μεταξύ τους ουσιαστικά. Το μοντέλο που περιγράφεται με το φίλτρο Kalman δίνει μόνο αρνητικές τιμές ενώ το αντίστοιχο με την μέθοδο Levenberg-Marquardt δίνει και θετικές τιμές. Αυτό σημαίνει ότι στην πρώτη περίπτωση μια πιθανή έγχυση ινσουλίνης έχει αποκλειστικά ως αποτέλεσμα μείωση της γλυκόζης στο αίμα ενώ στην δεύτερη περίπτωση μια ένεση ινσουλίνης θα μπορούσε προς στιγμή να συντελέσει σε αύξηση της γλυκόζης. Βέβαια αυτή η αύξηση θα είναι πολύ μικρή στην συγκεκριμένη περίπτωση. Θα επανέλθουμε όμως στο ζήτημα αυτό σε επόμενο μοντέλο όπου θα είναι έντονη αυτή η διαφορά.



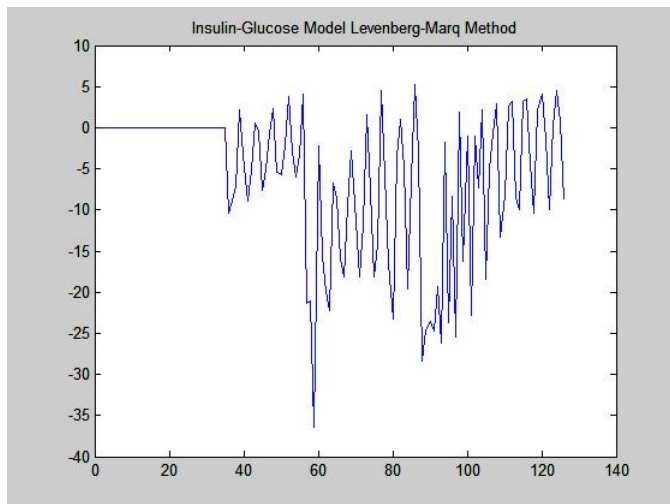
Σχήμα 6.9: Μοντέλο Ινσουλίνης-Γλυκόζης



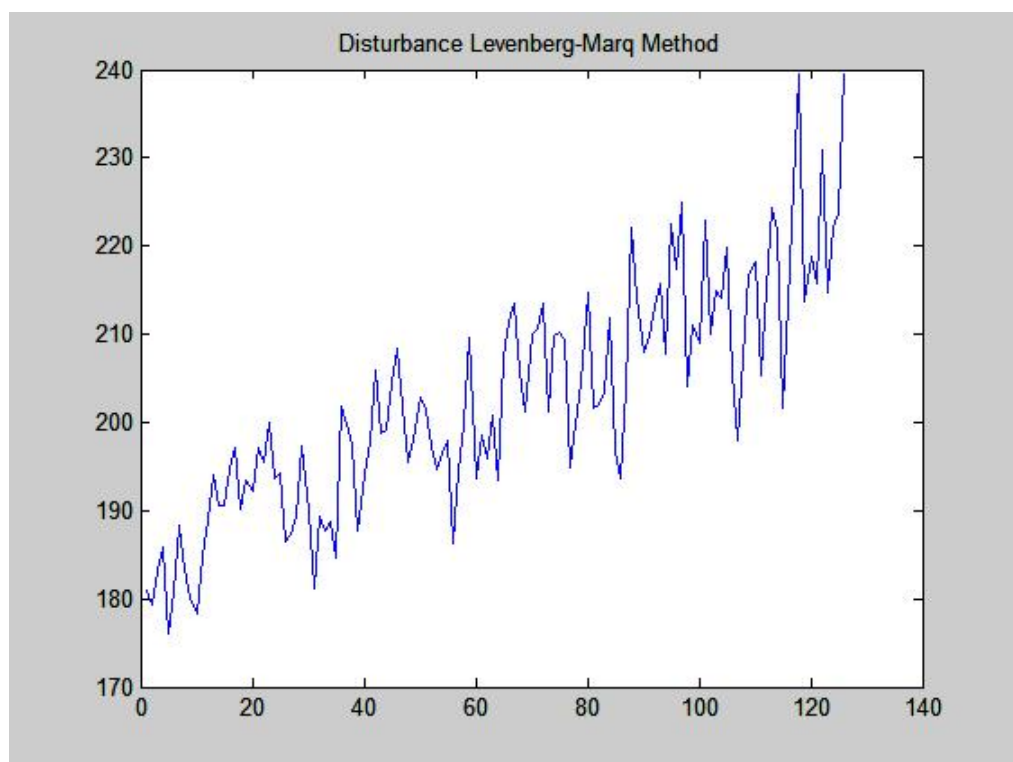
Σχήμα 6.10: Disturbance



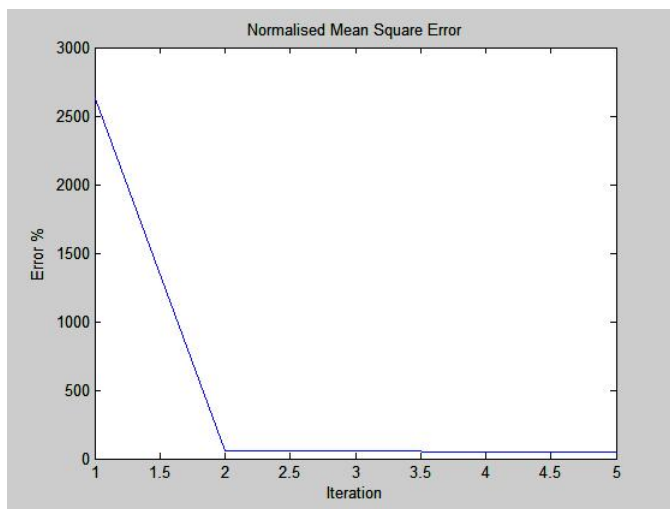
Σχήμα 6.11: Συγκριτικό Διάγραμμα Levenberg-Marquardt



Σχήμα 6.12: Μοντέλο Ινσουλίνης-Γλυκόζης Levenberg-Marquardt



Σχήμα 6.13: Disturbance Levenberg-Marquardt



Σχήμα 6.14: Κανονικοποιημένο Σφάλμα

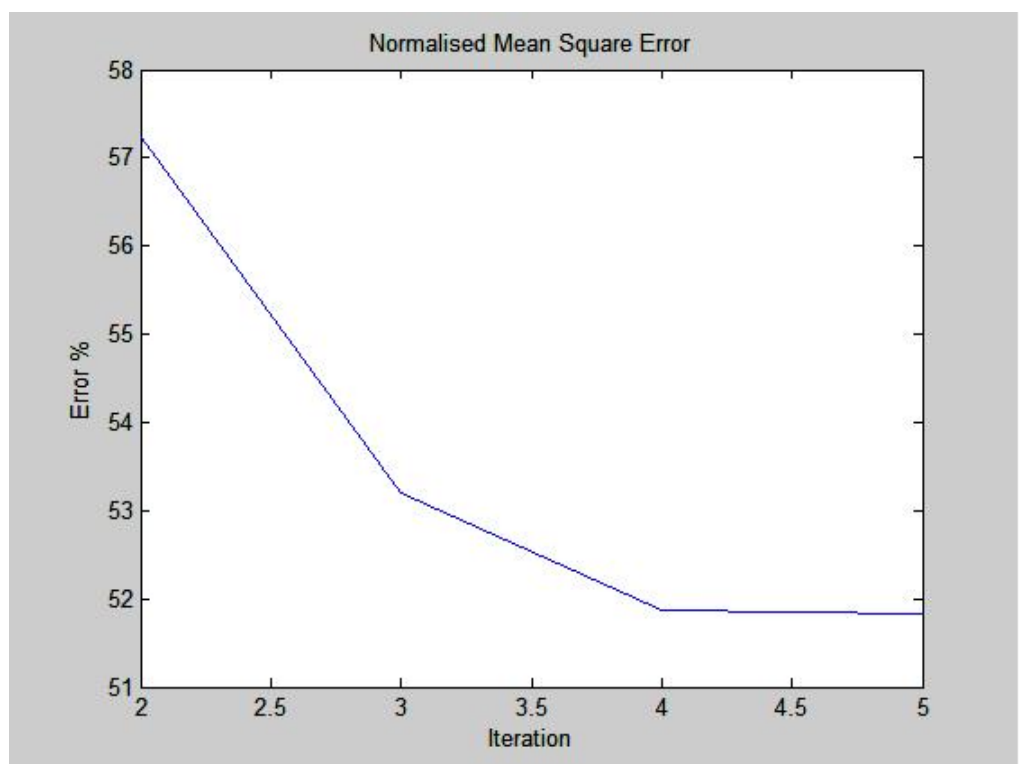
6.1.2 Τάξη Μοντέλου 20

Τετραγωνικό Σφάλμα

Αφού καταλήξαμε σε τιμή κανονικοποιημένου μέσου τετραγωνικού σφάλματος μοντελοποιώντας την διαταραχή με AR μοντέλο τάξης 25 προσπαθούμε να ελαττώσουμε την τάξη του μοντέλου. Θα χρησιμοποιήσουμε τάξη μοντέλου 20 το οποίο σημαίνει ότι από τα διαθέσιμα 150 δείγματα θα χρησιμοποιήσουμε τα 20 πρώτα για την αρχικοποίηση και τα υπόλοιπα 130 σαν datarecord. Τα ακόλουθα σε αυτή την ενότητα αφορούν αποκλειστικά το φίλτρο Kalman

Η τελική τιμή του κανονικοποιημένου τετραγωνικού σφάλματος μετά το πρώτο πέρασμα του datarecord ήταν 2627,7%(Φίλτρο Kalman) όπως φαίνεται και στο σχήμα (6.14). Ωστόσο στην συγκεκριμένη περίπτωση πρέπει να επισημάνουμε ότι το διάνυμα το οποίο χρησιμοποιήσαμε ως αρχικό διάνυμα για την διαδικασία έδινε τεράστιες τιμές κανονικοποιημένου μέσου τετραγωνικού σφάλματος (όπως μπορεί κανείς να καταλάβει κι από το αποτέλεσμα στο οποίο καταλήξαμε) εξαιτίας του ότι οδηγούσε σε πολύ μεγάλες τιμές την διαταραχή. Μετά όμως από το δεύτερο πέρασμα του datarecord οι διαδοχικές τιμές του κανονικοποιημένου μέσου τετραγωνικού σφάλματος ελαττώθηκαν όπως φαίνεται στο σχήμα (6.14). Μετά από πέντε διαδοχικά περάσματα του datarecord φτάσαμε σε κανονικοποιημένο τετραγωνικό σφάλμα της τάξης του 50%. Οι τιμές του κανονικοποιημένου μέσου σφάλματος από την δεύτερη επανάληψη κι έπειτα φαίνονται στο σχήμα (6.15). Με την μέθοδο Levenberg-Marquardt καταλήξαμε σε σφάλμα 29.83% ¹

¹Περάσαμε 5 φορές το datarecord



Σχήμα 6.15: Κανονικοποιημένο Σφάλμα

Τιμές Παραμέτρων

Οι τιμές των παραμέτρων στις οποίες καταλήξαμε μετά το πέρας των 6 περασιμάτων του datarecord και οι αντίστοιχες τιμές που προκύπτουν έπειτα την εφαρμογή της μεθόδου Levenberg-Marquardt δίνονται στον ακόλουθο πίνακα. Αυτό που πρέπει να σημειώσουμε και το οποίο φαίνεται και στους ακόλουθους πίνακες είναι ότι ξεκινήσαμε από τις ίδιες αρχικές συνθήκες και καταλήξαμε σε αρκετά διαφορετικές τιμές. Το γεγονός αυτό θα το επισημάνουμε και στην επόμενη υποενότητα αναλυτικότερα.

Όνομα Παραμέτρου	a_{11}	a_{12}	a_{21}	a_{22}
Αρχική τιμή	-0.0252	0.0062	0.1603	0.5899
Kalman Filter	0.0076	-0.0020	0.1361	0.4983
Levenberg-Mar	-0.0300	0.0060	0.1306	0.7290

Όνομα Παραμέτρου	b_{10}	b_{11}	b_{20}	b_{21}
Αρχική τιμή	-0.8186	-0.6686	-0.1200	-0.2692
Kalman Filter	0.0088	-0.0253	0.0156	-0.2271
Levenberg-Mar	-0.6116	-0.6260	-0.1379	-0.2238

Όνομα Παραμέτρου	c_{11}	c_{12}	c_{13}	c_{21}	c_{22}	c_{23}
Αρχική τιμή	-104.3262	-67.5760	-14.5697	22.0701	-2.1506	-5.1071
Kalman Filter	-214.7708	-59.6216	25.4806	-2.7071	-7.3168	-240.2220
Levenberg-Mar	-99.6838	-71.5180	-12.1967	19.6824	-1.6609	-3.5056

Όνομα Παραμέτρου	k_1	k_2	k_3	k_4	k_5
Αρχική τιμή	-0.4324	0.0579	-0.0011	-0.3266	0.2159
Kalman Filter	-0.0041	0.1629	-0.0506	-0.2114	0.0541
Levenberg-Mar	-0.4228	0.0419	-0.0017	-0.2207	0.2799

Όνομα Παραμέτρου	k_6	k_7	k_8	k_9	k_{10}
Αρχική τιμή	0.0638	0.7834	-0.2540	-0.0313	-0.0564
Kalman Filter	0.2274	0.4305	-0.0361	-0.0821	-0.0147
Levenberg-Mar	0.0848	0.5695	-0.2053	-0.0491	-0.1102

Όνομα Παραμέτρου	k_{11}	k_{12}	k_{13}	k_{14}	k_{15}
Αρχική τιμή	-0.4520	-0.2770	0.5144	0.1118	0.1727
Kalman Filter	-0.2865	0.1283	0.1356	-0.1863	-0.0543
Levenberg-Mar	-0.3243	-0.3962	0.6441	0.0962	0.2202

Όνομα Παραμέτρου	k_{16}	k_{17}	k_{18}	k_{19}	k_{20}
Αρχική τιμή	0.3822	0.3045	0.5442	-0.0543	-0.3955
Kalman Filter	0.5510	0.0606	0.3419	-0.0446	-0.1094
Levenberg-Mar	0.3943	0.5391	0.3657	-0.0502	-0.4282

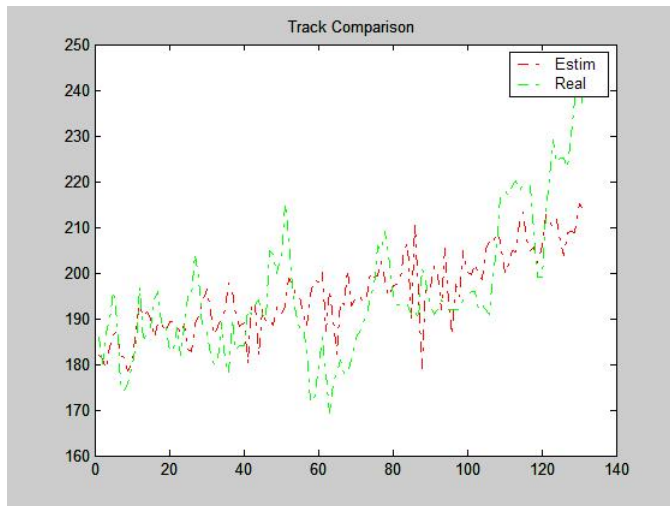
Εκτιμώμενη καμπύλη μοντέλου κι Disturbance

Τέλος παραθέτουμε συγκριτικά διαγράμματα της καμπύλης των πραγματικών δεδομένων με την καμπύλη που προκύπτει από τις εκτιμώμενες τιμές των παραμέτρων με τους δύο τρόπους εκτιμήσεων που αναφέραμε προηγουμένως δηλαδή με το Kalman Filter και με την μέθοδο Levenberg-Marquardt. Επιπλέον δίνουμε σε ξεχωριστά διαγράμματα την καμπύλη που περιγράφει το μέρος της εκτιμώμενης καμπύλης που οφείλεται στο disturbance και το μέρος που οφείλεται στο μοντέλο ινσουλίνης-γλυκόζης και για τις δύο μεθόδους εκτίμησης.

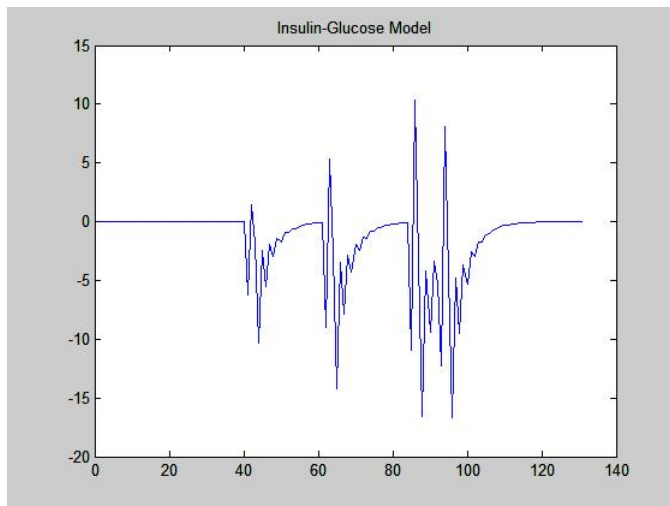
Στην προηγούμενη ενότητα είδαμε ότι παρότι είχαμε ξεκινήσει από ίδιες αρχικές τιμές παραμέτρων είχαμε καταλήξει σε διαφορετικές τιμές. Έίχαμε αναφέρει όταν εξετάζαμε τάξη μοντέλου 25 ότι οι διαφορετικές τιμές στις οποίες συνέκλιναν οι δύο διαφορετικές μέθοδοι που είχαμε χρησιμοποιήσει για την εκτίμηση των παραμέτρων κατεδύκνιαν διαφορετική αντίδραση του συστήματος(οργανισμού) στις ενέσεις ινσουλίνης. Το φαινόμενο αυτό δεν ήταν τόσο έντονο σε εκείνη την περίπτωση και δεν είχαμε δώσει μεγάλη έκταση σε εκείνο το σημείο. Στην συγκεκριμένη περίπτωση όμως οι δύο εκτιμήσεις δίνουν συστήματα τα οποία παρουσιάζουν πολύ μεγάλες διαφορές.

Αναλυτικότερα συγκρίνοντας τα διαγράμματα 6.17,6.20 βλέπουμε ότι οι δύο μέθοδοι δίνουν μοντέλα ινσουλίνης-γλυκόζης τα οποία συμπεριφέρονται πολύ διαφορετικά. Συγκεκριμένα το μοντέλο που προκύπτει με το φίλτρο Kalman δίνει τόσο θετικές όσο κι αρνητικές τιμές ενώ το αντίστοιχο με την μέθοδο Levenberg-Marquardt δίνει μόνο αρνητικές τιμές. Επιπλέον το εκτιμώμενο με την μέθοδο Levenberg-Marquardt μοντέλο επιτρέπει πολύ μεγαλύτερες αρνητικές τιμές του μοντέλου. Χαρακτηριστικά παρατηρούμε στο διάγραμμα ότι έχουμε ελάττωση μέχρι και 35mg/dl οφειλόμενη στην εγγεόμενη ινσουλίνη ενώ με το φίλτρο Kalman δεν υπάρχει μεγαλύτερη ελάττωση από 17,5mg/dl.

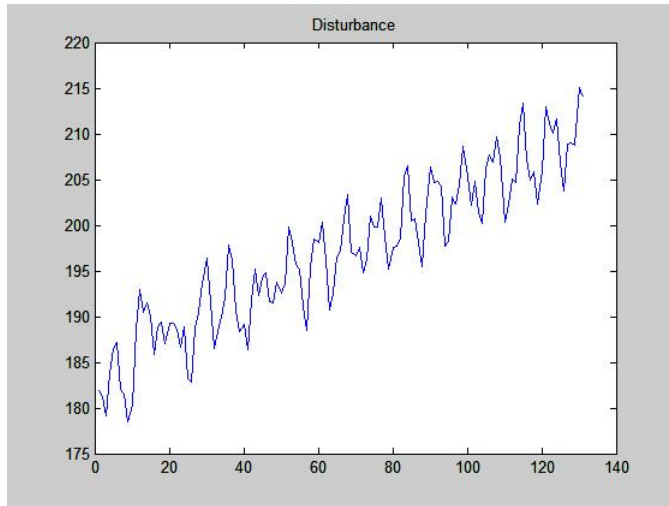
Εν κατακλείδι τα εκτιμώμενα διανύσματα τα οποία προκύπτουν με τις δύο διαφορετικές μεθόδους δίνουν πολύ διαφορετική απόκριση του μοντέλου ινσουλίνης γλυκόζης. Αυτό σημαίνει ότι το κάθε διάνυσμα περιγράφει διαφορετικά τον τρόπο που θα αντιδράσει ο οργανισμός σε πιθανή ένεση ινσουλίνης. Βλέπουμε λοιπόν ότι οι επιμέρους καμπύλες που παίρνουμε από τις δύο μεθόδους τόσο για το μοντέλο ινσουλίνης γλυκόζης όσο και για το disturbance είναι πολύ διαφορετικές μεταξύ τους(περιγράφουν διαφορετική αντίδραση οργανισμού), αν κι αντιστοιχούν σε κανονικοποιημένα μέσα τετραγωνικά σφάλματα τα οποία έχουν διαφορά 20% (30% για την μέθοδο Levenberg-Marquardt και 50% για το φίλτρο Kalman) .



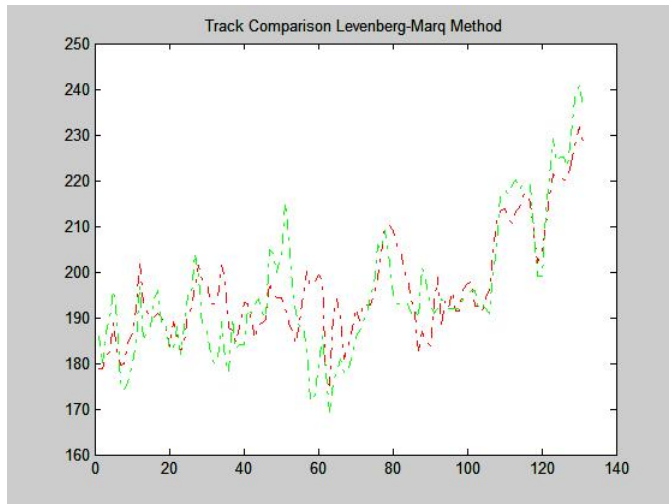
Σχήμα 6.16: Συγκριτικό Διάγραμμα



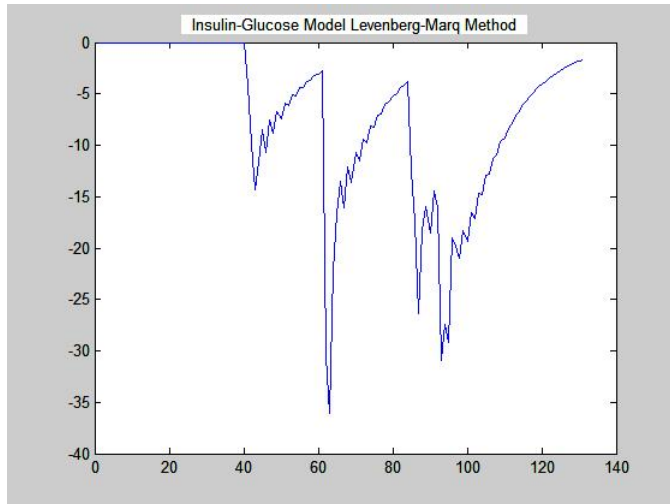
Σχήμα 6.17: Μοντέλο Ινσουλίνης-Γλυκόζης



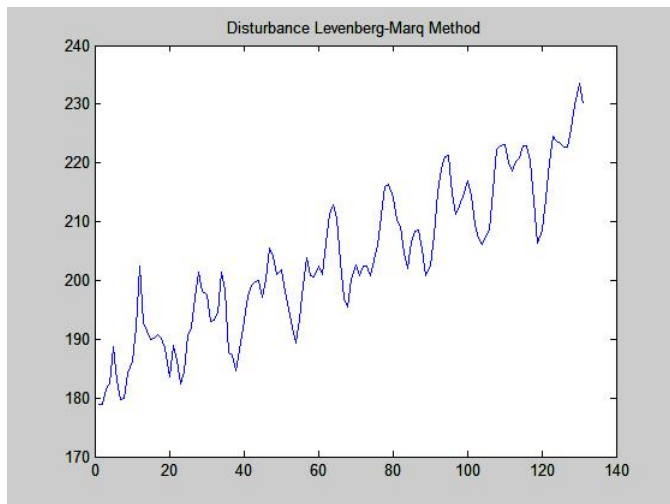
Σχήμα 6.18: Disturbance



Σχήμα 6.19: Συγκριτικό Διάγραμμα Levenberg-Marquardt



Σχήμα 6.20: Μοντέλο Ινσουλίνης-Γλυκόζης Levenberg Marquardt



Σχήμα 6.21: Disturbance Levenberg-Marquardt

6.1.3 Σύγκριση της Τάξης των 2 μοντέλων

Αφού παρουσιάσαμε τα αποτελέσματα και για τις δύο διαφορετικές τάξεις AR μοντέλων τίθεται το ερώτημα, το ποια τάξη θα ήταν προτιμότερη. Είναι σίγουρο όπως είχα τονίσει και στην αρχή του κεφαλαίου το ότι μεγαλύτερη τάξη μοντέλου θα δίνει τουλάχιστον εξίσου καλά αποτελέσματα με οποιαδήποτε μικρότερη τάξη.

Από εκεί κι έπειτα όμως αυτό που θα πρέπει να δούμε είναι αν η χρησιμοποίηση μοντέλου τάξης 25 έναντι 20 προσφέρει σημαντική μείωση του σφάλματος. Για τις συγκεκριμένες εκτιμήσεις παραμέτρων το φίλτρο Kalman για μοντέλο τάξης 25 δίνει σφάλμα 49.83% ενώ για τάξης 20 51,85% ενώ για αντίστοιχες τάξεις μοντέλων με την μέθοδο Levenberg-Marquardt παίρνουμε σφάλματα 30.52% και 29.83% αντίστοιχα. Θα τονίσουμε κι πάλι εδώ ότι δεν τίθεται ζήτημα για το γεγονός ότι με μικρότερης τάξης μοντέλου παίρνουμε μικρότερο σφάλμα. Τα δύο σφάλματα εκτιμήθηκαν σε διαφορετικού μεγέθους datarecord όπως ανέφερα και στην αρχή της ενότητας. Για το ίδιο datarecord τα αποτελέσματα που παίρνουμε είναι :

Τάξη μοντέλου	25	20
Kalman Filter	49.83	50.73
Levenberg-Mar	30.52	31.54

Απο τα προηγούμενα λοιπόν βλέπουμε ότι τέτοιες διαφορές λάθους δεν δικαιολογούν την επιλογή μεγαλύτερης τάξης μοντέλου κι επομένως θα ήταν προτιμότερη η επιλογή μοντέλου τάξης 20.

6.2 Δείγματα 1-150

6.2.1 Τάξη Μοντέλου 25

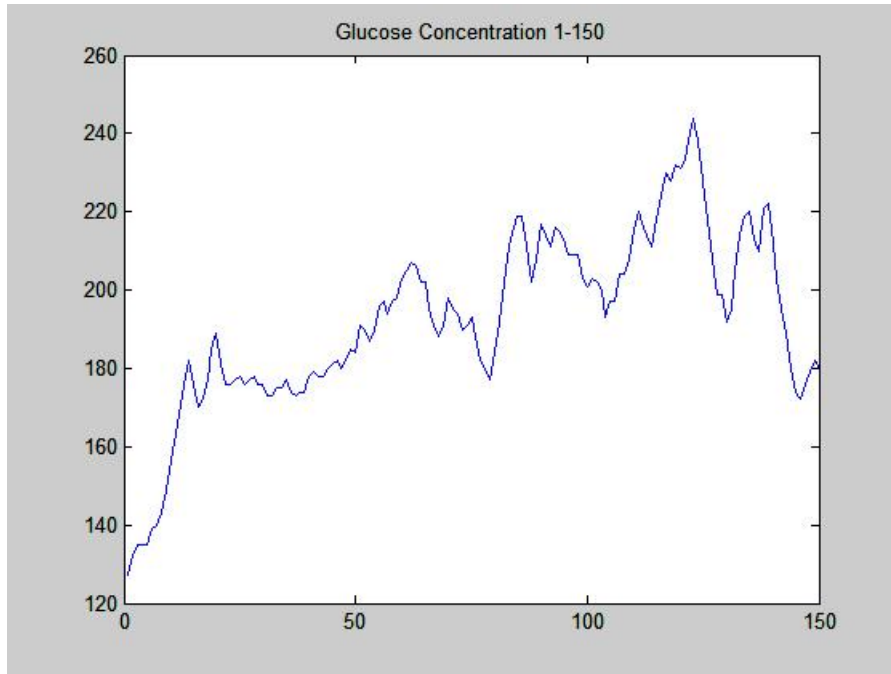
Τετραγωνικό Σφάλμα

Χρησιμοποιήσαμε ως datarecord τα δεδομένα από 1-150. Οι τιμές της γλυκόζης για αυτή την περίοδο φαίνεται στο σχήμα (6.22). Για την αρχικοποίηση των συντελεστών του disturbance χρησιμοποιήθηκαν τα πρώτα 25 δείγματα και τα υπόλοιπα χρησιμοποιήθηκαν σαν data για την εκπαίδευση.

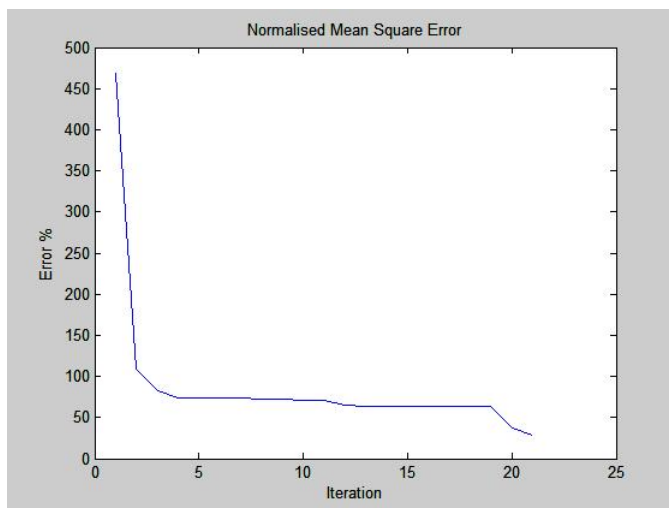
Η τελική τιμή του κανονικοποιημένου τετραγωνικού σφάλματος μετά το πρώτο πέρασμα του datarecord ήταν 469.2196 % όπως φαίνεται και στο σχήμα (6.23) Όπως και πριν προχωρήσαμε στη συνέχεια εκτελώντας διαδοχικά περάσματα του datarecord ώστε να πετύχουμε περαιτέρω μείωση στο σφάλμα. Μετά από 21 περάσματα του datarecord φτάσαμε σε τιμή λάθους 28.4530%. Η πορεία ελάττωσης του σφάλματος φαίνεται αναλυτικά στα σχήματα (6.23,6.24). Στο σχήμα 6.24 δίνουμε τις τιμές από την τρίτη επανάληψη κι έπειτα οι οποίες δεν φαίνονται καθαρά στο προηγούμενο σχήμα.

Η τιμή κανονικοποιημένου μέσου τετραγωνικού σφάλματος στην οποία καταλήξαμε με την μέθοδο Levenberg-Marquardt ήταν 34.52 % .²

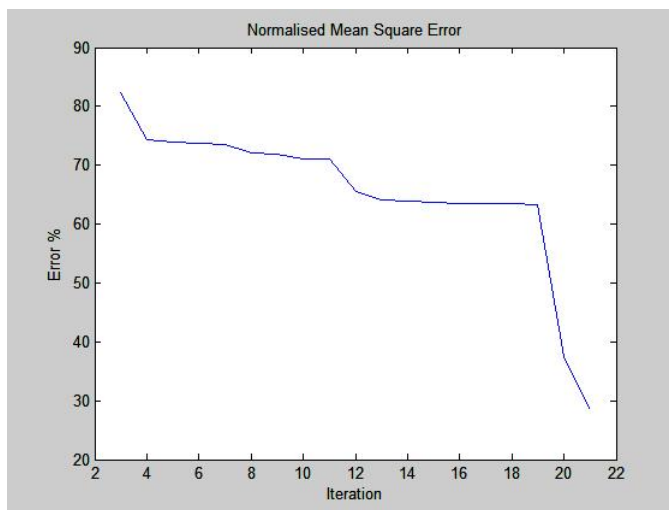
²Περάσαμε 20 φορές το datarecord



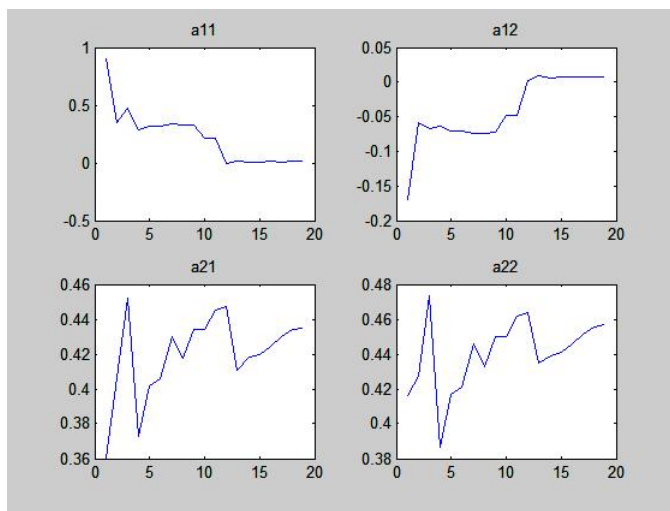
Σχήμα 6.22: Τιμές Γλυκόζης 1-150



Σχήμα 6.23: Κανονικοποιημένο Σφάλμα



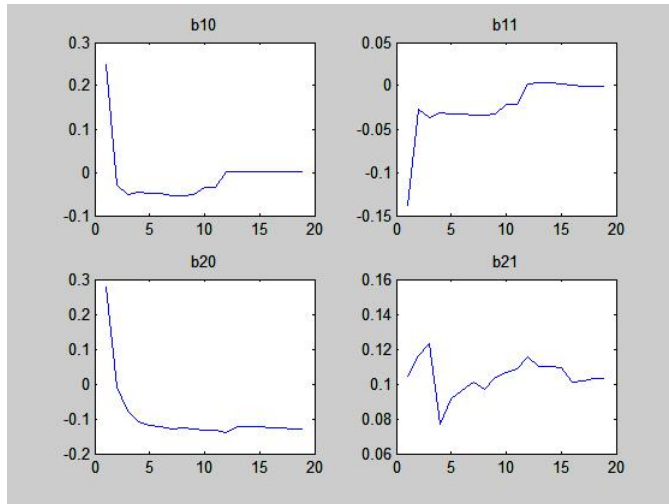
Σχήμα 6.24: Κανονικοποιημένο Σφάλμα



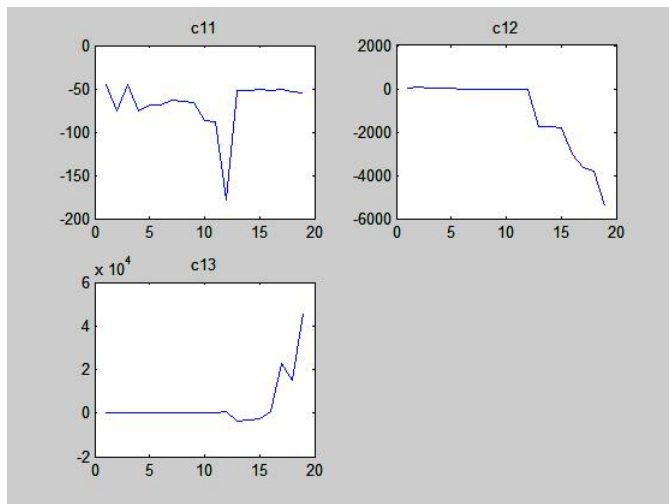
Σχήμα 6.25: Παράμετροι a

Τιμές Παραμέτρων

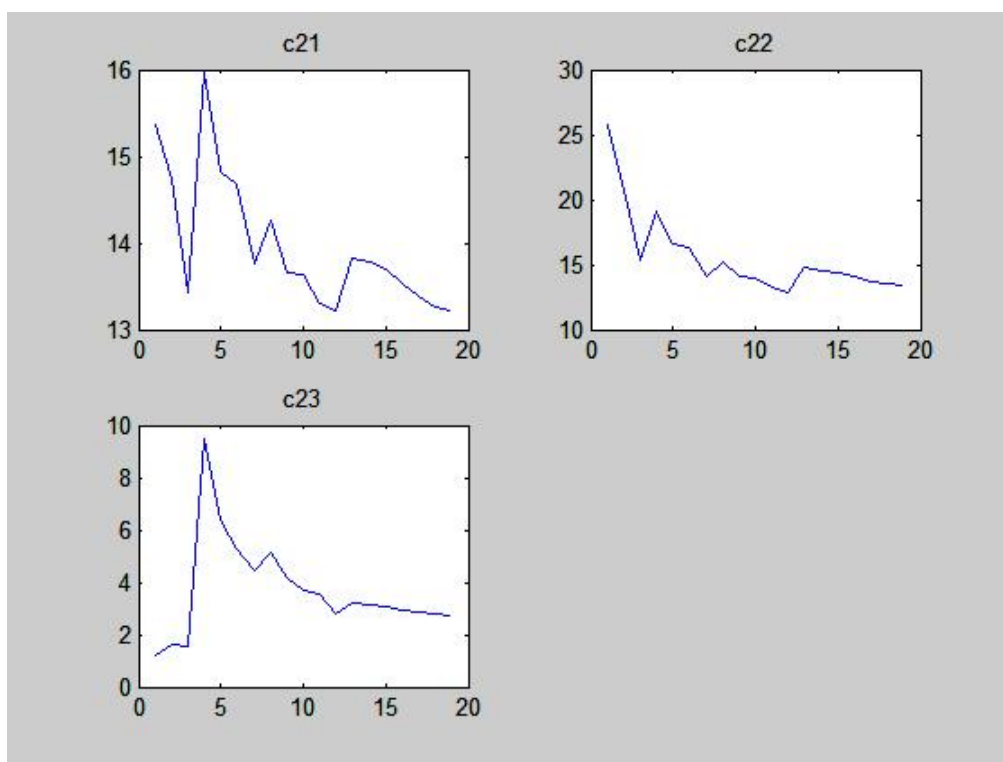
Τα διαγράμματα (6.25,6.26,6.27,6.28) απεικονίζουν την μεταβολή των τιμών των παραμέτρων του μοντέλου ινσουλίνης-γλυκόζης κατά το πρώτο πέρασμα του datarecord. Στα υπόλοιπα πέρασματα του datarecord όπως προείπα η μείωση του σφάλματος ήταν πολύ μικρή σε σχέση με την αντίστοιχη μείωση από το πρώτο πέρασμα και αντίστοιχα δεν είχαμε σημαντικές διακυμάνσεις στις τιμές των περισσοτέρων παραμέτρων, οπότε δεν είχε πρακτική σημασία να δώσουμε τα αντίστοιχα διαγράμματα. Επιπλέον, υπάρχει κι συνολικός πίνακας με τις αρχικές και τελικές τιμές κάθε μιας εκ των παραμέτρων του μοντέλου κι της διαταραχής.



Σχήμα 6.26: Παράμετροι b



Σχήμα 6.27: Παράμετροι c_1



Σχήμα 6.28: Παράμετροι c_2

Όνομα Παραμέτρου	a_{11}	a_{12}	a_{21}	a_{22}
Αρχική τιμή	0.1190	0.0528	0.3710	0.3954
Kalman Filter	0.0201	0.0081	0.4351	0.4567
Levenberg-Mar	0,8964	0,0706	0,7447	0,2138

Όνομα Παραμέτρου	b_{10}	b_{11}	b_{20}	b_{21}
Αρχική τιμή	0.2320	0.0422	0.2939	0.1031
Kalman Filter	0.001	-1.4948e-4	-0.1287	0.1031
Levenberg-Mar	0,1842	0,0240	0,2290	0,0645

Όνομα Παραμέτρου	c_{11}	c_{12}	c_{13}	c_{21}	c_{22}	c_{23}
Αρχική τιμή	-52.183	-4.6431	-3.1149	15.824	25.577	2.843
Kalman Filter	-55.492	-5.4138e+003	4.6127e+004	13.215	13.395	2.741
Levenberg-Mar	-59,970	-2,3662	-16,2308	15,000	23,437	4,499

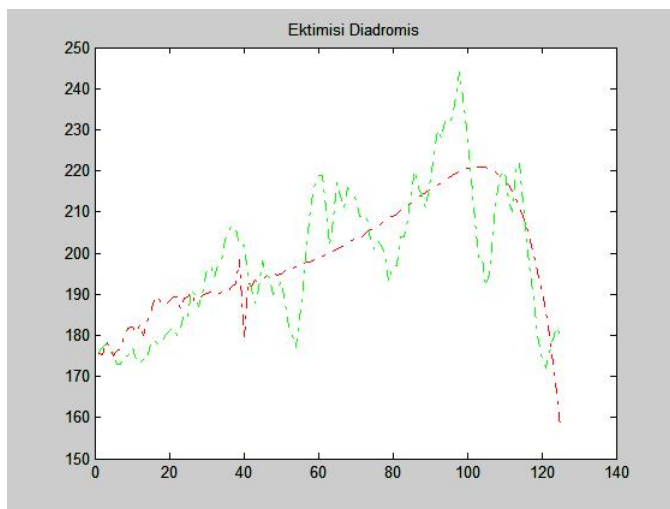
Όνομα Παραμέτρου	k_1	k_2	k_3	k_4	k_5
Αρχική τιμή	0.1329	0.0267	0.2416	0.3746	0.0592
Kalman Filter	1.2386	-0.3581	0.2902	0.1370	-0.5036
Levenberg-Mar	0,1791	0.2064	0.2217	0.2239	0.2037

Όνομα Παραμέτρου	k_6	k_7	k_8	k_9	k_{10}
Αρχική τιμή	0.0532	0.0172	0.1910	0.0242	0.0942
Kalman Filter	0.6137	-0.8028	0.9016	-0.6353	0.1562
Levenberg-Mar	0.1552	0.1262	0.1159	0.0904	0.0532

Όνομα Παραμέτρου	k_{11}	k_{12}	k_{13}	k_{14}	k_{15}
Αρχική τιμή	-0.0241	-0.1951	-0.0304	-0.0121	-0.0822
Kalman Filter	0.1361	-0.5870	0.6349	-0.3925	0.1557
Levenberg-Mar	-0.0234	-0.1012	-0.1260	-0.1253	-0.1081

Όνομα Παραμέτρου	k_{16}	k_{17}	k_{18}	k_{19}	k_{20}
Αρχική τιμή	-0.0568	-0.0396	-0.0175	0.0103	0.0693
Kalman Filter	-0.0070	-0.0888	0.2276	-0.2597	0.3108
Levenberg-Mar	-0.0208	-0.0689	-0.0494	0.0137	0.0382

Όνομα Παραμέτρου	k_{21}	k_{22}	k_{23}	k_{24}	k_{25}
Αρχική τιμή	0.0989	0.1615	0.0106	-0.0683	-0.1163
Kalman Filter	-0.1377	0.1193	-0.0499	-0.2509	0.1519
Levenberg-Mar	0.0738	0.0840	0.0326	-0.0456	-0.1554



Σχήμα 6.29: Συγκριτικό Διάγραμμα

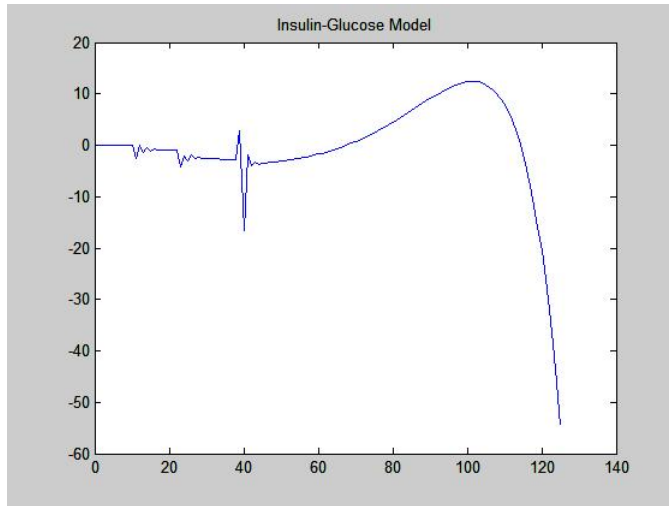
Εκτιμώμενη καμπύλη μοντέλου κι Disturbance

Στην υποενότητα αυτή παραθέτουμε συγκριτικά διαγράμματα της καμπύλης των πραγματικών δεδομένων με την καμπύλη που προκύπτει από τις εκτιμώμενες τιμές των παραμέτρων με τους δύο τρόπους εκτιμήσεων που αναφέραμε προηγουμένως δηλαδή με το Kalman Filter και με την μέθοδο Levenberg-Marquardt. Επιπλέον δίνουμε σε ξεχωριστά διαγράμματα την καμπύλη που περιγράφει το μέρος της εκτιμώμενης καμπύλης που οφείλεται στο disturbance και το μέρος που οφείλεται στο μοντέλο ινσουλίνης-γλυκόζης και για τις δύο μεθόδους εκτίμησης.

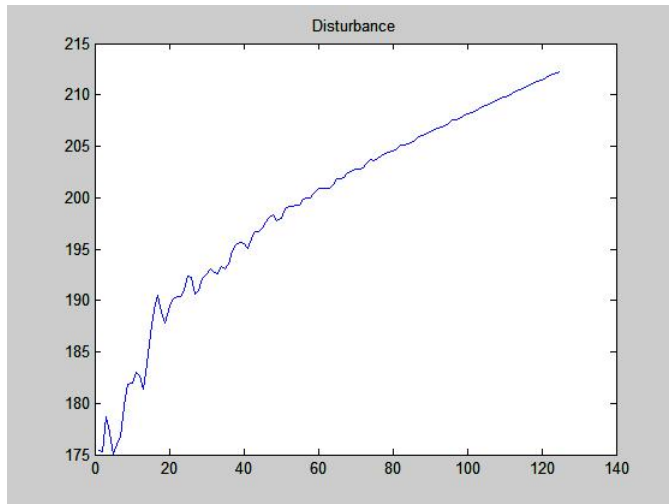
Και σε αυτή την περίπτωση παρατηρούμε ωστόσο ότι οι τιμές στις οποίες καταλήγουμε με τους δύο διαφορετικούς τρόπους εκτίμησης καταδικνείουν συστήματα (οργανισμό) τα οποία αντιδρούν με διαφορετικό τρόπο στις ενέσεις ινσουλίνης.

Αναλυτικότερα συγκρίνοντας τα διαγράμματα 6.30,6.33 βλέπουμε ότι οι δύο μέθοδοι δίνουν μοντέλα ινσουλίνης-γλυκόζης τα οποία συμπεριφέρονται πολύ διαφορετικά. Βλέπουμε καταρχήν ότι τα δύο μοντέλα παρουσιάζουν εντελώς διαφορετική συμπεριφορά σε πολλά διαστήματα. Χαρακτηριστικά για τιμές μετά το 98 δείγμα το μοντέλο με το φίλτρο Kalman δίνει φθίνουσα καμπύλη ενώ το αντίστοιχο με την μέθοδο Levenberg-Marquardt δίνει αύξουσα καμπύλη. Επιπλέον σημαντικό είναι και το γεγονός ότι το μοντέλο με το φίλτρο Kalman δίνει τόσο θετικές όσο κι αρνητικές τιμές οι οποίες φτάνουν μέχρι κι -62mg/dl ενώ το αντίστοιχο με την μέθοδο Levenberg-Marquardt δίνει μόνο αρνητικές τιμές με ελάχιστη τιμή -32mg/dl .

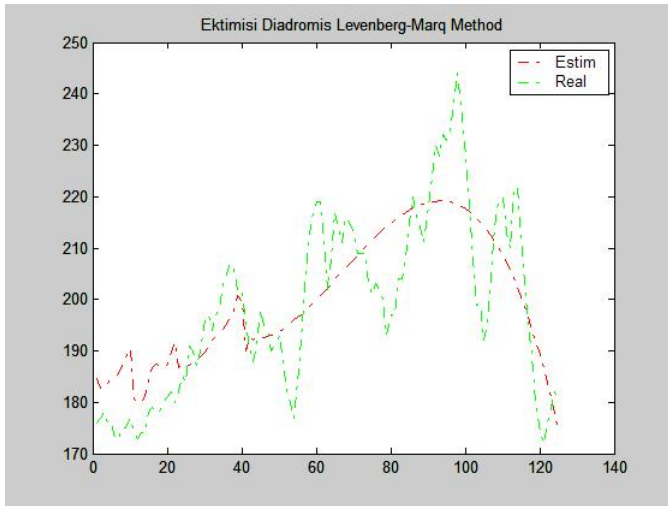
Εν κατακλείδι τα εκτιμώμενα διανύσματα τα οποία προκύπτουν με τις δύο διαφορετικές μεθόδους δίνουν πολύ διαφορετική απόκριση του μοντέλου ινσουλίνης γλυκόζης. Αυτό σημαίνει ότι το κάθε διάνυσμα περιγράφει διαφορετικά τον τρόπο που θα αντιδράσει ο οργανισμός σε πιθανή ένεση ινσουλίνης. Παρόλα αυτά αν κι οι επιμέρους καμπύλες που παίρνουμε από τις δύο μεθόδους τόσο για το μοντέλο ινσουλίνης γλυκόζης όσο και για το disturbance είναι πολύ διαφορετικές μεταξύ τους αντιστοιχούν σε κανονικοποιημένα τετραγωνικά σφάλματα τα οποία είναι πολύ κοντά μεταξύ τους (34.25% για την μέθοδο Levenberg-Marquardt και 28,453% για το φίλτρο Kalman).



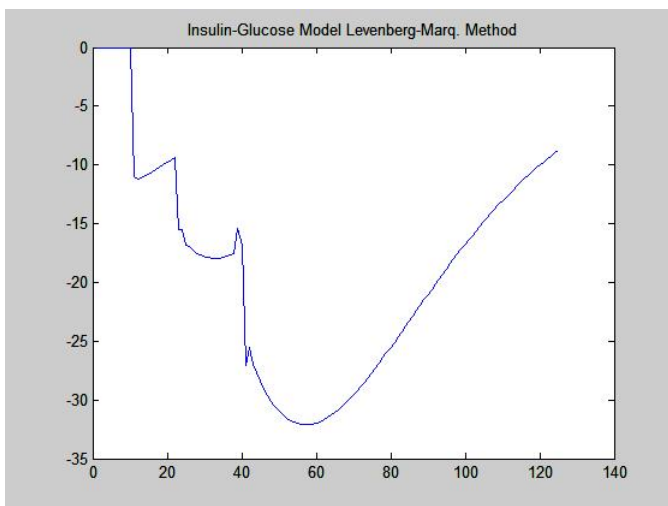
Σχήμα 6.30: Μοντέλο Ινσουλίνης-Γλυκόζης



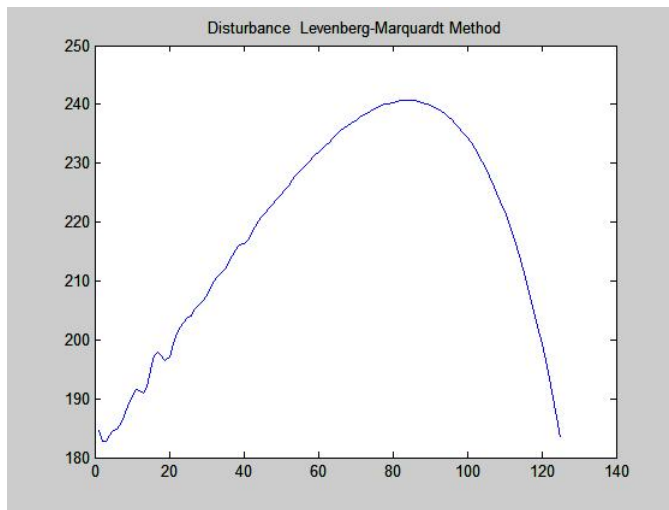
Σχήμα 6.31: Disturbance



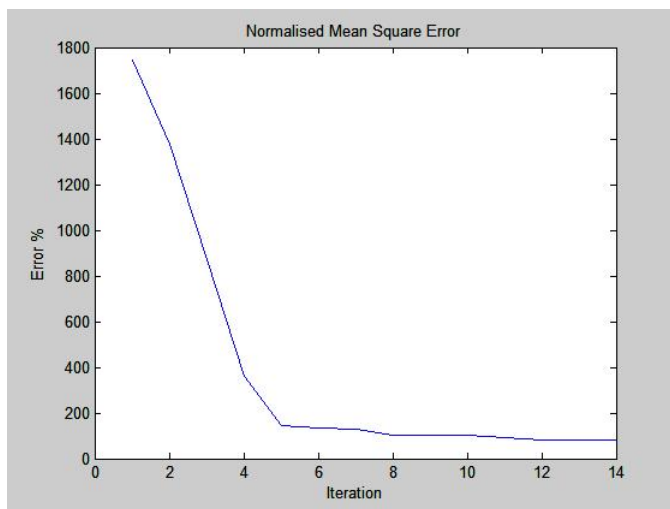
Σχήμα 6.32: Συγκριτικό Διάγραμμα Levenberg-Marquardt



Σχήμα 6.33: Μοντέλο Ινσουλίνης-Γλυκόζης Levenberg-Marquardt



Σχήμα 6.34: Disturbance Levenberg-Marquardt



Σχήμα 6.35: Κανονικοποιημένο Σφάλμα

6.2.2 Τάξη Μοντέλου 20

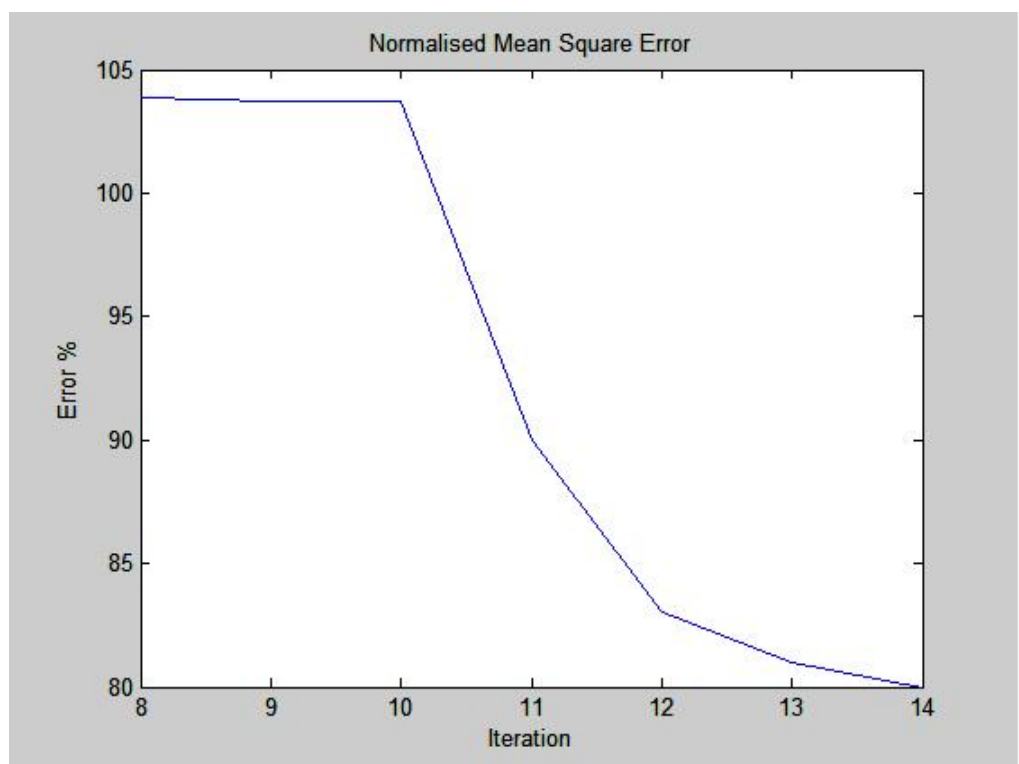
Τετραγωνικό Σφάλμα

Προσπαθούμε όπως και πριν να μειώσουμε την τάξη του μοντέλου σε 20. Έτσι για την αρχικοποίηση των συντελεστών του disturbance χρησιμοποιήθηκαν τα πρώτα 20 δείγματα και τα υπόλοιπα 130 χρησιμοποιήθηκαν σαν data για την εκπαίδευση.

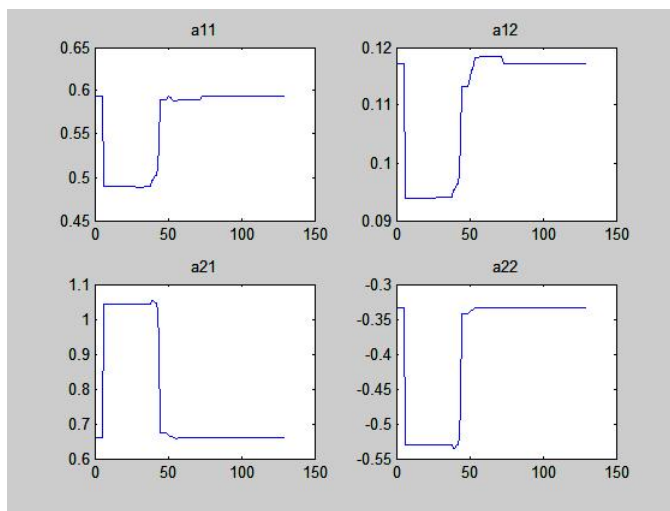
Η τελική τιμή του κανονικοποιημένου μέσου τετραγωνικού σφάλματος μετά το πρώτο πέρασμα του datarecord ήταν 1746,3% όπως φαίνεται και στο σχήμα (6.35). Ωστόσο στην συγκεκριμένη περίπτωση πρέπει να παρατηρήσουμε ότι το διάλυμα το οποίο χρησιμοποιήσαμε ως αρχικό διάνυσμα για την διαδικασία έδινε μεγάλες τιμές κανονικοποιημένου τετραγωνικού σφάλματος εξαιτίας του ότι οδηγούσε σε πολύ μεγάλες τιμές την διαταραχή. Για να πέσει το σφάλμα σε τιμή κοντά στο 100 % χρειάστηκε να φτάσουμε στην όγδοη επανάληψη (103,8%). Οι τιμές του σφάλματος από την επανάληψη αυτή κι έπειτα φαίνονται στο διάγραμμα (6.36)

Όπως και πριν προχωρήσαμε στη συνέχεια εκτελώντας διαδοχικά περάσματα του datarecord ώστε να πετύχουμε περαιτέρω μείωση στο σφάλμα. Η τιμή του σφάλματος στην οποία καταλήξαμε ήταν 80%. Η πορεία ελάττωσης του σφάλματος φαίνεται αναλυτικά στο σχήμα (6.35).³

³Περάσαμε 14 φορές το datarecord



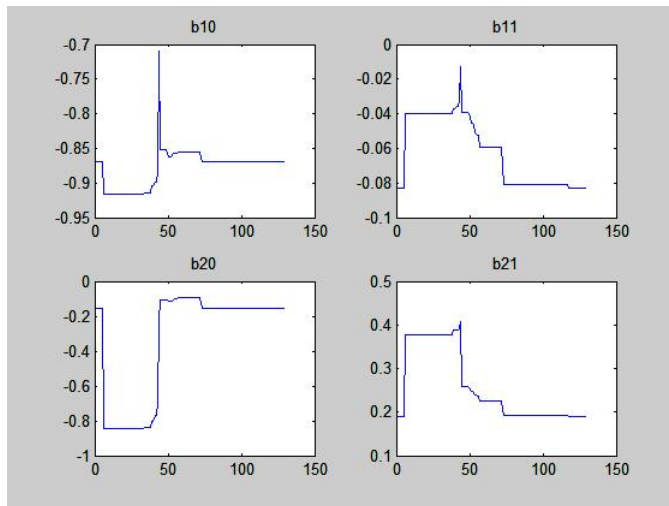
Σχήμα 6.36: Κανονικοποιημένο Σφάλμα



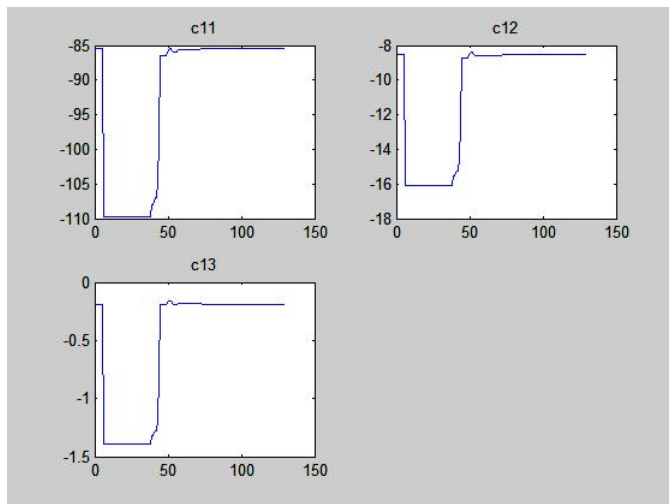
Σχήμα 6.37: Παράμετροι a

Τιμές Παραμέτρων

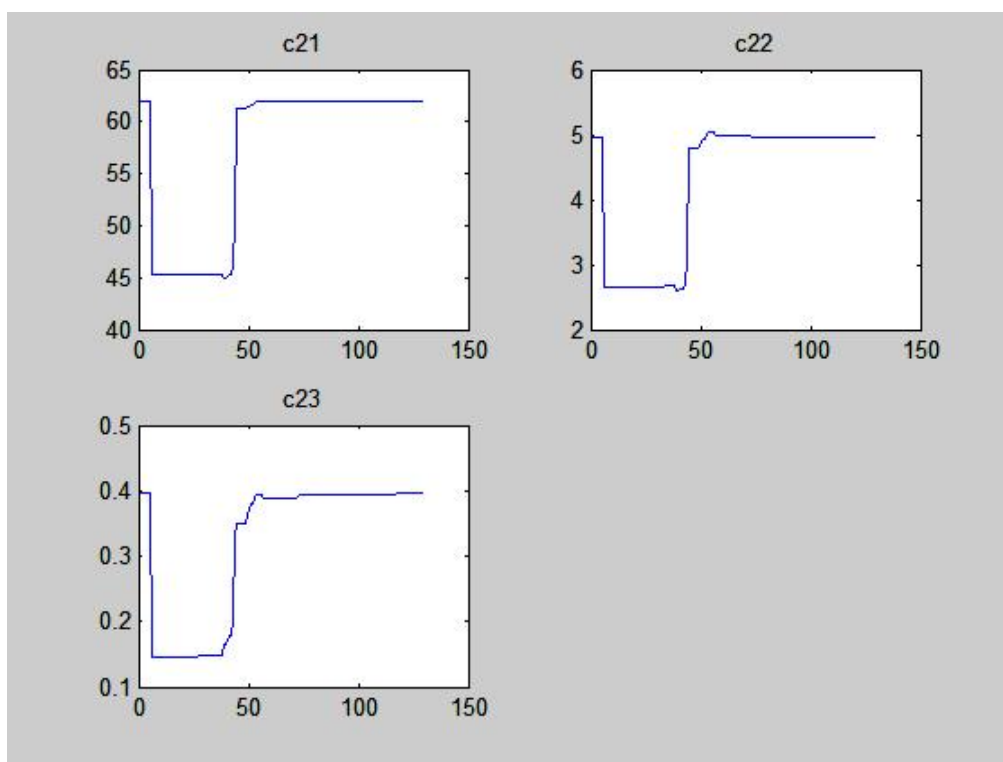
Τα διαγράμματα (6.37,6.38,6.39,6.40) απεικονίζουν την μεταβολή των τιμών των παραμέτρων του μοντέλου ινσουλίνης-γλυκόζης κατά το πρώτο πέρασμα του datarecord. Ο λόγος για τον οποίο παραθέτουμε μόνο διαγράμματα για τις μεταβολές των παραμέτρων κατά το πρώτο πέρασμα του datarecord είναι ο ίδιος με πριν. Επιπλέον, υπάρχει κι συνολικός πίνακας με τις αρχικές και τελικές τιμές κάθε μιας εκ των παραμέτρων του μοντέλου κι της διαταραχής.



Σχήμα 6.38: Παράμετροι b



Σχήμα 6.39: Παράμετροι c_1



Σχήμα 6.40: Παράμετροι c_2

Όνομα Παραμέτρου	a_{11}	a_{12}	a_{21}	a_{22}
Αρχική τιμή	0.4550	0.0273	1.0532	-0.4927
Kalman Filter	0.1444	0.0397	0.7059	-0.3457
Levenberg-Mar	0.9552	0.0211	1.2882	-0.3199

Όνομα Παραμέτρου	b_{10}	b_{11}	b_{20}	b_{21}
Αρχική τιμή	-0.9155	-0.0401	-0.8423	0.3778
Kalman Filter	0.0191	-0.0257	0.1150	-0.0161
Levenberg-Mar	-0.5982	-0.0351	-0.8467	0.2362

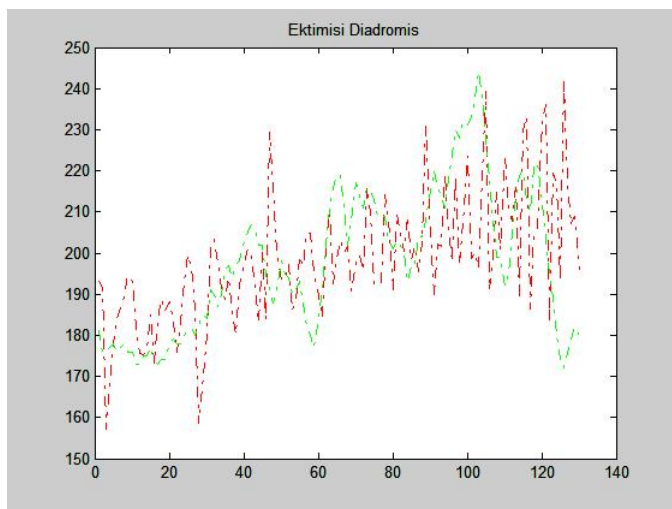
Όνομα Παραμέτρου	c_{11}	c_{12}	c_{13}	c_{21}	c_{22}	c_{23}
Αρχική τιμή	-109.7597	-16.0771	-1.3939	45.3451	2.6624	0.1468
Kalman Filter	-5.3989	821.4134	189.0128	-95.7165	74.1914	-22.5680
Levenberg-Mar	-74.0397	-13.2064	-0.9438	36.5586	3.4896	0.2144

Όνομα Παραμέτρου	k_1	k_2	k_3	k_4	k_5
Αρχική τιμή	-0.3666	0.5565	0.3111	-0.0547	0.2186
Kalman Filter	1.0379	0.1283	-0.3147	-0.0914	0.3179
Levenberg-Mar	-0.4812	0.5294	0.2321	-0.0766	0.1450

Όνομα Παραμέτρου	k_6	k_7	k_8	k_9	k_{10}
Αρχική τιμή	0.4106	0.5805	0.4665	-0.3240	0.0250
Kalman Filter	0.1045	0.2123	-0.0975	-0.5748	0.7049
Levenberg-Mar	0.3726	0.7206	0.3577	-0.3028	0.0272

Όνομα Παραμέτρου	k_{11}	k_{12}	k_{13}	k_{14}	k_{15}
Αρχική τιμή	0.1322	-0.5257	-0.7525	-0.6630	-0.0808
Kalman Filter	-0.3086	-0.4008	-0.0346	0.0018	0.4277
Levenberg-Mar	0.1023	-0.4448	-0.5217	-0.3934	-0.0781

Όνομα Παραμέτρου	k_{16}	k_{17}	k_{18}	k_{19}	k_{20}
Αρχική τιμή	0.5765	-0.1088	-0.9650	0.3196	0.4496
Kalman Filter	0.3542	-0.7905	-0.3117	1.3319	-0.6960
Levenberg-Mar	0.3672	-0.2137	-0.5612	0.4864	0.7153



Σχήμα 6.41: Συγκριτικό Διάγραμμα

Εκτιμώμενη καμπύλη μοντέλου κι Disturbance

Όπως και προηγούμενα στην υποενότητα αυτή παραθέτουμε συγκριτικά διαγράμματα της καμπύλης των πραγματικών δεδομένων με την καμπύλη που προκύπτει από τις εκτιμώμενες τιμές των παραμέτρων με τους δύο τρόπους εκτιμήσεων που αναφέραμε προηγουμένως δηλαδή με το Kalman Filter και με την μέθοδο Levenberg-Marquardt. Επιπλέον δίνουμε σε ξεχωριστά διαγράμματα την καμπύλη που περιγράφει το μέρος της εκτιμώμενης καμπύλης που οφείλεται στο disturbance και το μέρος που οφείλεται στο μοντέλο ινσουλίνης-γλυκόζης και για τις δύο μεθόδους εκτίμησης.

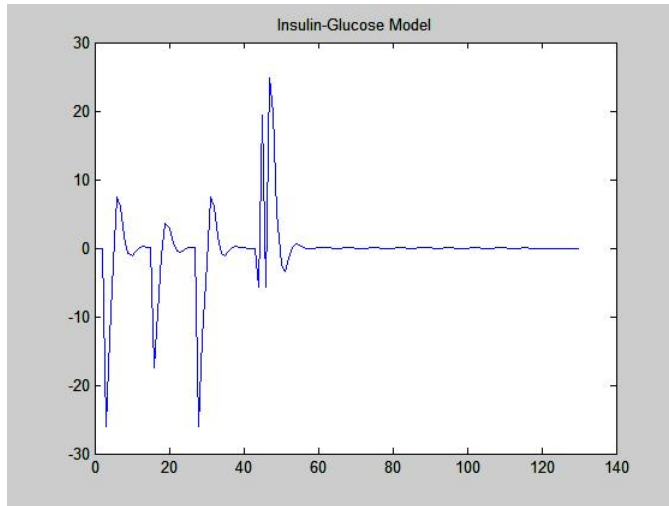
Και σε αυτή την περίπτωση παρατηρούμε ωστόσο ότι οι τιμές στις οποίες καταλήγουμε με τους δύο διαφορετικούς τρόπους εκτίμησης καταδυναμίζουν συστήματα (οργανισμό) τα οποία αντιδρούν με διαφορετικό τρόπο στις ενέσεις ινσουλίνης.

Αναλυτικότερα συγκρίνοντας τα διαγράμματα 6.42, 6.45 βλέπουμε ότι οι δύο μέθοδοι δίνουν μοντέλα ινσουλίνης-γλυκόζης τα οποία συμπεριφέρονται πολύ διαφορετικά. Αυτό που πρέπει να επισημάνουμε είναι ότι το μοντέλο που προκύπτει με την μέθοδο Levenberg-Marquardt δίνει αύξηση της εκρινόμενης γλυκόζης με τις ενέσεις ινσουλίνης.

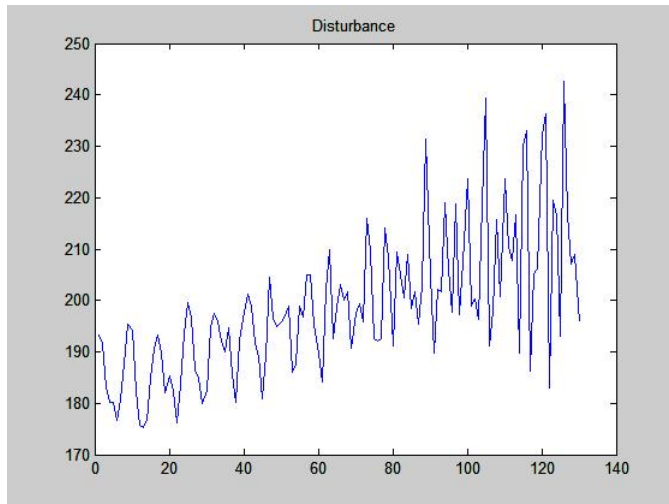
6.2.3 Σύγκριση της Τάξης των 2 μοντέλων

Ομοίως με πριν αφού παρουσιάσαμε τα αποτελέσματα βρισκόμαστε μπροστά στο ερώτημα ως προς ποια τάξη μοντέλου θα ήταν προτιμότερο να επιλέξουμε.

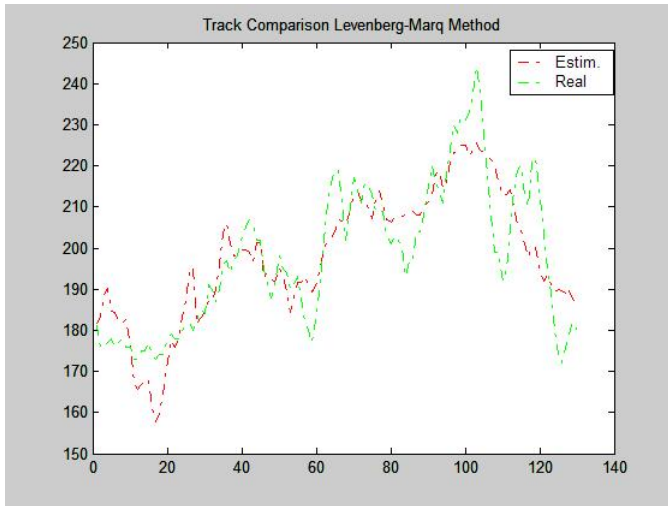
Αυτό που θα χρησιμοποιήσουμε ως κριτήριο κι σε αυτήν την περίπτωση είναι αν η χρησιμοποίηση μοντέλου τάξης 25 έναντι 20 προσφέρει σημαντική μείωση του σφάλματος. Για τις συγκεκριμένες εκτιμήσεις παραμέτρων το φίλτρο Kalman για μοντέλο τάξης 25 δίνει σφάλμα 28.453% ενώ για τάξης 20 80% ενώ για αντίστοιχες τάξεις μοντέλων με την μέθοδο Levenberg-Marquardt παίρνουμε σφάλματα 34.25% κι 26.52% αντίστοιχα. Θα τονίσουμε κι πάλι εδώ ότι δεν τίθεται ζήτημα για το γεγονός ότι με μικρότερης τάξης μοντέλου παίρνουμε μικρότερο σφάλμα. Τα δύο σφάλματα εκτιμήθηκαν σε διαφορετικού μεγέθους data record όπως ανέφερα και στην αρχή της



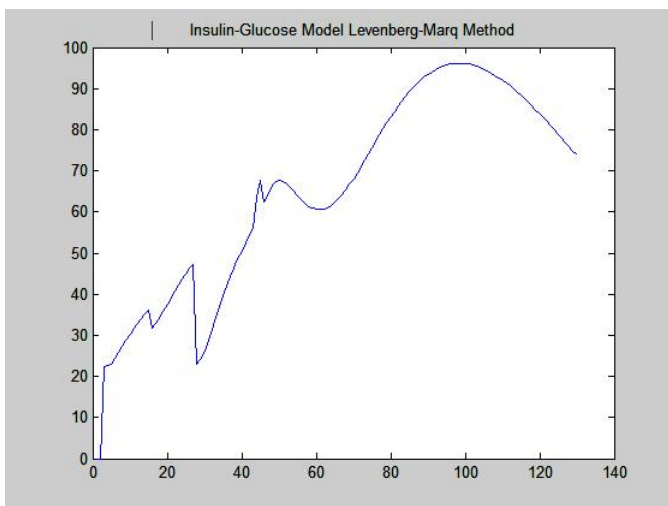
Σχήμα 6.42: Μοντέλο Ινσουλίνης-Γλυκόζης



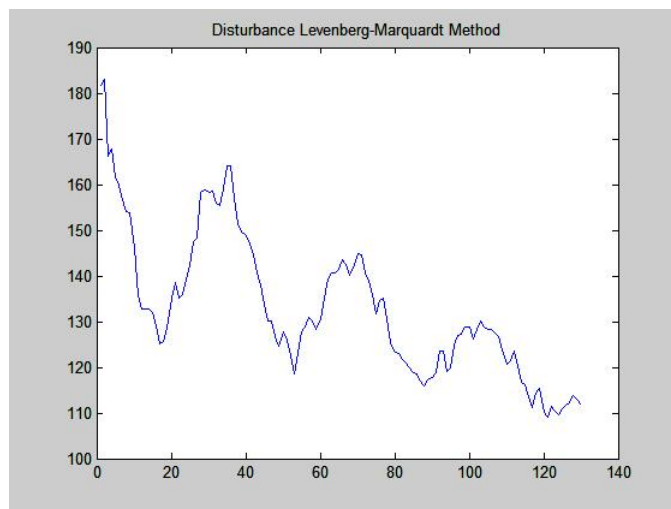
Σχήμα 6.43: Disturbance



Σχήμα 6.44: Συγκριτικό Διάγραμμα Levenberg-Marquardt



Σχήμα 6.45: Μοντέλο Ινσουλίνης-Γλυκόζης Levenberg-Marquardt



Σχήμα 6.46: Disturbance Levenberg-Marquardt

ενότητας .Για το ίδιο datarecord τα αποτελέσματα που παίρνουμε είναι :

Τάξη μοντέλου	25	20
Kalman Filter	28,45	82.34%
Levenberg-Mar	34.25	37.42

Σε αυτή την περίπτωση βέβαια τα συμπεράσματά μας είναι διαφορετικά με πριν. Για την μέθοδο Levenberg-Marquardt ελάττωση του σφάλματος κατά 3% δεν δικαιολογεί επιλογή μοντέλου αυξημένου κατά πέντε τάξεις. Για το Kalman Filter όμως δεν ισχύει το ίδιο καθώς η διαφορά στο σφάλμα είναι κοντά στο 60% οπότε θα πρέπει να επιλέξουμε μοντέλο τάξης 25 εάν θα θέλαμε να χρησιμοποιήσουμε αυτή την μέθοδο.

Bibliography

- [1] Dimitri P. Bertsekas *Nonlinear Programming* Athena Scientific 1995
- [2] David G. Luenberger *Linear and Nonlinear Programming* (2nd Edition) Addison-Wesley 1984
- [3] Ν.Γ.Μαράτος *Τεχνικές Βελτιστοποιήσεως(Σημειώσεις)* Εκδόσεις ΕΜΠ
- [4] M.S. Bazaraa, H.D. Sherali, C.M. Shetty *Nonlinear Programming Theory and Algorithms* John Wiley and Sons, 1973
- [5] Harold W. Sorenson *Parameter Estimation Principles and Problems* Control and Systems Theory Vol.9 Marcel Dekker 1980
- [6] Vasilis Z. Marmarelis *Nonlinear Dynamic Modeling of Physiological Systems* EMB, IEEE Press, Wiley-Interscience, 2004
- [7] M. G. Markakis, G. P. Papavassilopoulos, G. D. Mitsis, V. Z. Marmarelis *Computational Study of Model-Based Nonlinear Control of Blood Glucose* IEEE 2006
- [8] G. Bishop, G. Welch *An Introduction to the Kalman Filter* University of North Carolina at Chapel Hill, 2004
- [9] Simon Haykin *Kalman Filter and Neural Networks* John Wiley and Sons, 2001
- [10] Arthur Gelb *Applied Optimal Estimation* The Analytic Science Corporation, 16th Printing 2001
- [11] Guyton and Hall *Φυσιολογία του Ανθρώπου και Μηχανισμοί των νόσων* (Μεταφρ) Εκδόσεις Παρισιανού, 2001
- [12] Cecil *Βασική Παθολογία* Ιατρικές εκδόσεις Αίτσα, Τόμος 2, 2000
- [13] Μιχάλης Μαρκάκης *Μοντελοποίηση και Έλεγχος του Συστήματος Ρύθμισης της Γλυκόζης στο Αίμα* Πτυχιακή Εργασία, ΕΜΠ, 2005
- [14] Φρέρης Νικόλαος *Εκτίμηση των Παραμέτρων ενός νέου μοντέλου Ινσουλίνης-Γλυκόζης* Πτυχιακή Εργασία, ΕΜΠ, 2005
- [15] Tresp V., Briegel T. and Moody J. *Neural - Network Models for the Blood Glucose Metabolism of a Diabetic.* IEEE Transactions on Neural Networks, Vol. 10, No. 5., (1999)
- [16] Vicini P., Caumo A. and Cobelli C. *The Hot IVGTT Two δ Compartment Minimal Model: Indexes of Glucose Effectiveness and Insulin Sensitivity* Am. J. Physiol. 273: E1024 δ E1032, 1997

- [17] A. Caumo, P. Vicini, J. Zachwieja, A. Avogaro, K. Yarasheski, D. Bier, C. Cobelli *Undermodeling affects minimal model indexes: insights from a two-compartment model* 1999
- [18] Cobelli C., Finkelstein L., Carson E.R. *Mathematical modelling of endocrine and metabolic systems: model formulation, identification and validation* Mathematics and Computers in Simulation XXIV, 442-451, 1982
- [19] Carson E.R., Cobelli C., Finkelstein L. *Modeling and identification of metabolic systems* Am. J. Physiol. 240: R120-R129, 1981.