

Πολυτεχνείο Κρήτης
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών
Υπολογιστών (ΗΜΜΥ)

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Υπολογισμός παραμέτρων σε συστήματα που περιγράφονται από
Συνήθεις Διαφορικές Εξισώσεις με χρήση MATLAB



Επιμέλεια: Τζανάκης Σάββας

Επιβλέπων: Καθηγητής Ζερβάκης Μιχάλης

Μέλη επιτροπής: Καθηγητής Λιάβας Αθανάσιος,
Καθηγητής Καλογεράκης Νικόλαος

Πρόλογος

Για την εκπόνηση της παρούσας διπλωματικής εργασίας, θα ήθελα να ευχαριστήσω τους καθηγητές στο τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Πολυτεχνείου Κρήτης, τον κύριο Μιχάλη Ζερβάκη που δέχθηκε να είναι ο επιβλέπων της εργασίας αυτής και τον κύριο Αθανάσιο Λιάβα που δέχθηκε να είναι μέλος της επιτροπής. Επίσης, θα ήθελα να ευχαριστήσω θερμά τον κ. Νικόλαο Καλογέρακη, τρίτο μέλος της επιτροπής, καθηγητή του τμήματος Μηχανικών Περιβάλλοντος του Πολυτεχνείου Κρήτης, ο οποίος δέχθηκε να αναλάβει την επίβλεψη της εργασίας μου και στάθηκε δίπλα μου καθοδηγώντας με σε όλη τη διάρκεια της υλοποίησης της. Ακόμη, θα ήθελα να ευχαριστήσω όλο το διδακτικό και διοικητικό προσωπικό του τμήματος ΗΜΜΥ και του Πολυτεχνείου Κρήτης γενικότερα για το επίπεδο της εκπαίδευσης που προσφέρουν και τις ευκαιρίες που δίνουν στους φοιτητές. Τέλος, θα ήθελα να ευχαριστήσω την οικογένεια μου για την αμέριστη στήριξη τους καθ'όλη τη διάρκεια των σπουδών μου, οι οποίες χωρίς αυτούς δεν θα ήταν εφικτές.

Περίληψη

Ο υπολογισμός αγνώστων παραμέτρων σε συστήματα που περιγράφονται από Συνήθεις Διαφορικές Εξισώσεις (Ordinary Differential Equations, ODEs) είναι ένα θέμα ιδιαίτερα σημαντικό όσον αφορά τους τομείς της Μηχανικής, καθώς συναντάται συχνά σε αντίστοιχα συστήματα. Στην παρούσα εργασία, παρουσιάζονται το θεωρητικό υπόβαθρο σχετικά με δυναμικά μοντέλα από συνήθεις διαφορικές εξισώσεις, η μέθοδος Gauss-Newton που χρησιμοποιείται για τον υπολογισμό των παραμέτρων και ο αλγόριθμος που προκύπτει, μαζί με τις απαραίτητες τροποποιήσεις για τη βελτιστοποίηση του. Έπειτα, το μαθηματικό μοντέλο και η επίλυση του προβλήματος υπολογισμού των παραμέτρων που ήταν ανεπτυγμένο σε FORTRAN μετατρέπεται σε αντίστοιχο που εκτελείται στο περιβάλλον της MATLAB. Στο μοντέλο προστίθεται κώδικας, ο οποίος δημιουργεί γραφικά που δείχνουν την πορεία της σύγκλισης του εκτιμώμενου μοντέλου με τα πειραματικά δεδομένα και κώδικας που υπολογίζει βασικές στατιστικές υπολογιστικές εκφράσεις για τις παραμέτρους. Τέλος, δύο συγκεκριμένες εφαρμογές με πραγματικά πειραματικά δεδομένα (Πυρολυτική Αφυδρογόνωση του Βενζολίου σε Διφαινύλιο και Τριφαινύλιο & Καταλυτική Υδρογόνωση του 3-Hydroxypropanal (HPA) σε 1,3-Propanediol (PD)) εκτελούνται και αναλύονται τα αποτελέσματά τους, αριθμητικά και γραφικά.

Abstract

Estimation of unknown parameters in models described by Ordinary Differential Equations (ODEs) is a particularly important issue in the fields of engineering, as it is often encountered in corresponding systems. In this thesis, the theoretical background about ODEs, Gauss-Newton method that is used for the parameter estimation and the algorithm that results, along with the necessary modifications for the algorithm's optimization, are presented. Then, a mathematical model developed in FORTRAN is converted into a corresponding one that is executed in the environment of MATLAB. Code is added to the model, which creates graphs that show the course of convergence of the estimated model with the experimental data. Also added is code that calculates basic statistical computational expressions for the parameters. Finally, two specific applications with real experimental data (Pyrolytic Dehydrogenation of Benzene to Diphenyl and Triphenyl & Catalytic Hydrogenation of 3-Hydroxypropanal (HPA) to 1,3-Propanediol (PD)) are executed and their results are analyzed, numerically and graphically.

Περιεχόμενα

1. Εισαγωγή
 1. Κίνητρο
 2. Περιγραφή της εργασίας
2. Υπολογισμός παραμέτρων σε συστήματα από ODEs
 1. Περιγραφή των ODEs
 2. Επιλογή Αντικειμενικής Συνάρτησης και υπολογισμός
 3. Μέθοδος Gauss-Newton για ODEs
 4. Αλγόριθμος Gauss-Newton για συστήματα από ODEs
 1. Χρήση Λύτη Διαφορικών Εξισώσεων
 2. Χρήση του Κανόνα Διχοτόμησης και Υπέρβαση
 3. Ill-conditioning και Μέσα Αντιμέτωπισης
 5. Στατιστικός Λογισμός
3. Κύριες αλλαγές σε περιβάλλον MATLAB
4. Εφαρμογή
 1. Εφαρμογή 1: Πυρολυτική Αφυδρογόνωση του Βενζολίου σε Διφαινύλιο και Τριφαινύλιο
 1. Περιγραφή
 2. Αποτελέσματα
 2. Εφαρμογή 2: Καταλυτική Υδρογόνωση του 3-Hydroxypropanal (HPA) σε 1,3-Propanediol (PD)
 1. Περιγραφή
 2. Αποτελέσματα
5. Συμπεράσματα

Βιβλιογραφία

Λίστα Εικόνων

Εικόνα 3.1. Γενικό Διάγραμμα Ροής του Αλγορίθμου

Εικόνα 3.2. Παράδειγμα γραφικής παράστασης που δημιουργείται για συγκεκριμένο iteration

Εικόνα 4.1. Τελικό αποτέλεσμα του αλγορίθμου στην κονσόλα της πλατφόρμας MATLAB (Εφαρμογή 1)

Εικόνα 4.2. Τιμές των εκτιμώμενων παραμέτρων k_1 , k_2 ανά iteration. Την k_1 απεικονίζει η μπλε γραμμή και την k_2 απεικονίζει η κόκκινη γραμμή.

Εικόνα 4.3. Πορεία σύγκλισης του μοντέλου σε σχέση με τα πειραματικά δεδομένα. Οι συνεχείς γραμμές απεικονίζουν τα εκτιμώμενα μοντέλα x_1 και x_2 και οι κύκλοι τα πειραματικά δεδομένα τους.

Εικόνα 4.4. Τελικό αποτέλεσμα του αλγορίθμου στην κονσόλα της πλατφόρμας MATLAB (Εφαρμογή 2)

Εικόνα 4.5. Γραφική απεικόνιση για το δεδομένο σετ

Εικόνα 4.6. Γραφική απεικόνιση για δεδομένο σετ και $k(2) \times 100$

Εικόνα 4.7. Γραφική απεικόνιση για δεδομένο σετ και $k(4) \times 100$

Εικόνα 4.8. Γραφική απεικόνιση για δεδομένο σετ και $k(1) \times 100$

Λίστα Πινάκων

Πίνακας 4.1. Δεδομένα για την Πυρολυτική Αφυδρογόνωση του Βενζολίου

Πίνακας 4.2. Πυρολυτική αφυδρογόνωση του βενζολίου: Πειραματικά δεδομένα (data) και εκτιμώμενες παράμετροι του μοντέλου (model)

Πίνακας 4.3. Πυρολυτική αφυδρογόνωση του βενζολίου: Ανά προσέγγιση (iteration), τιμή της αντικειμενικής συνάρτησης και των εκτιμώμενων παραμέτρων k_1 , k_2

Πίνακας 4.4. Υδρογόνωση του HPA: Πειραματικά δεδομένα συλλεγμένα σε θερμοκρασία 318K και πιέσεις 2.6, 4.0 και 5.15 MPa

Πίνακας 4.5. Αριθμός κατάστασης του A ($\text{cond}(A)$), ανάλογα από τον αριθμό των εκτιμώμενων παραμέτρων και την ύπαρξη τροποποίησης Marquardt ή όχι.

Ακρωνύμια – Συντομεύσεις

cond	condition number, Αριθμός κατάστασης
COV	Covariance, Συνδιασπορά
d.f.	degrees of freedom, Βαθμοί ελευθερίας
GLS	Generalized Least Squares, Γενικευμένα Ελάχιστα Τετράγωνα
HPA	Hydroxypropanal
LS	Least Squares, Ελάχιστα Τετράγωνα
ML	Maximum Likelihood, Μέγιστη Πιθανοφάνεια
ODE	Ordinary Differential Equation, Συνήθης Διαφορική Εξίσωση
PD	Propanediol
RSE	Relative Standard Error, Σχετικό Τυπικό Σφάλμα
SSE	Sum of Square of Errors, Άθροισμα Ελαχίστων Τετραγώνων σφάλματος
WLS	Weighted Least Squares, Σταθμισμένα Ελάχιστα Τετράγωνα

1. Εισαγωγή

1.1. Κίνητρο

Η ανάγκη για τον υπολογισμό αγνώστων παραμέτρων σε συστήματα που περιγράφονται από ODEs συναντάται συχνά σε συστήματα μηχανικής και έχουν αναπτυχθεί πολύ ισχυροί και υπολογιστικά αποδοτικοί αλγόριθμοι για τη λύση του προβλήματος αυτού. Σκοπός της παρούσας διπλωματικής εργασίας είναι, δεδομένο μαθηματικό μοντέλο υπολογισμού παραμέτρων που εκτελείται σε FORTRAN να μετατραπεί σε αντίστοιχο που εκτελείται στο περιβάλλον της MATLAB και να τροποποιηθεί και βελτιστοποιηθεί. Στο νέο μοντέλο προστίθεται κώδικας, ο οποίος δημιουργεί γραφικά που δείχνουν την πορεία της σύγκλισης του εκτιμώμενου μοντέλου με τα πειραματικά δεδομένα και κώδικας που υπολογίζει βασικές στατιστικές υπολογιστικές εκφράσεις για τις παραμέτρους. Η λειτουργία του νέου μοντέλου ελέγχεται, εκτελώντας δύο συγκεκριμένες εφαρμογές με πραγματικά πειραματικά δεδομένα (Πυρολυτική Αφυδρογόνωση του Βενζολίου σε Διφαινύλιο και Τριφαινύλιο & Καταλυτική Υδρογόνωση του 3-Hydroxypropanal (HPA) σε 1,3-Propanediol (PD)).

1.2. Περιγραφή της εργασίας

Το παρόν κεφάλαιο, 1, αποτελεί την εισαγωγή της εργασίας.

Στο κεφάλαιο 2:

- περιγράφεται αναλυτικά η έννοια μιας συνήθους διαφορικής εξίσωσης,
- αναλύεται ο τρόπος που επιλέγεται και υπολογίζεται η Αντικειμενική Συνάρτηση (Objective Function),
- εξηγείται η μέθοδος Gauss-Newton για ODEs και παρουσιάζεται ο αλγόριθμος που την υλοποιεί, μαζί με τις τροποποιήσεις του (Χρήση Λύτη Διαφορικών Εξισώσεων, Χρήση του Κανόνα Διχοτόμησης, Ill-conditioning και Μέσα Αντιμετώπισης του) για την βελτίωση της αποτελεσματικότητας του υπολογισμού παραμέτρων και
- περιγράφεται ο στατιστικός λογισμός του αλγορίθμου.

Στο κεφάλαιο 3, περιγράφεται η λειτουργία του αλγορίθμου που πλέον έχει δομηθεί στη MATLAB και παρουσιάζονται αναλυτικότερα οι βασικές αλλαγές που υλοποιούνται. Στο κεφάλαιο 4, το θεωρητικό υπόβαθρο του κεφαλαίου 2 και η υλοποίηση υπολογιστικού προγράμματος του κεφαλαίου 3 συνδέονται με τα δεδομένα της εργασίας, δηλαδή μαθηματικά μοντέλα που έχουν αναπτυχθεί και πειραματικά δεδομένα που έχουν συλλεχθεί τα τελευταία χρόνια στο Εργαστήριο Βιοχημικής Μηχανικής και Περιβαλλοντικής Βιοτεχνολογίας του Πολυτεχνείου Κρήτης. Περιγράφονται δύο εφαρμογές, η Πυρολυτική Αφυδρογόνωση του

Βενζολίου σε Διφαινύλιο και Τριφαινύλιο (Εφαρμογή 1) και η Καταλυτική Υδρογόνωση του 3-Hydroxypropanal (HPA) σε 1,3-Propanediol (PD) (Εφαρμογή 2) και παρουσιάζονται τα αποτελέσματα τους, μετά την εκτέλεση του αλγορίθμου που αναπτύχθηκε στο περιβάλλον της MATLAB.

Τέλος, το κεφάλαιο 5 αποτελεί τα συνολικά συμπεράσματα της παρούσας διπλωματικής εργασίας.

2. Υπολογισμός παραμέτρων σε συστήματα από ODEs

2.1. Περιγραφή των ODEs

Σε ορισμένες περιπτώσεις συστημάτων που περιγράφονται από ένα σετ από συνήθεις διαφορικές εξισώσεις (Ordinary Differential Equations, ODEs), είναι δυνατή η αναλυτική επίλυση των διαφορικών εξισώσεων. Αυτό έχει ως αποτέλεσμα ο υπολογισμός των παραμέτρων να μπορεί να περιγραφεί αλγεβρικά. Εάν όμως, οι ODEs δε μπορούν να λυθούν αναλυτικά, το μαθηματικό μοντέλο είναι πιο περίπλοκο και η δομή του γενικά είναι η παρακάτω:

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{f}(\mathbf{x}(t), \mathbf{u}, \mathbf{k}), \mathbf{x}(t_0) = \mathbf{x}_0 \quad (2.1)$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) \quad (2.2)$$

όπου

$\mathbf{k} = [k_1, k_2, \dots, k_p]^T$ είναι διάνυσμα διάστασης p , το οποίο περιέχει παραμέτρους των οποίων η αριθμητική αξία είναι άγνωστη,

$\mathbf{x} = [x_1, x_2, \dots, x_n]^T$ είναι διάνυσμα διάστασης n , το οποίο περιέχει μεταβλητές κατάστασης (state variables),

$\mathbf{x}_0 = [x_{10}, x_{20}, \dots, x_{n0}]^T$ είναι διάνυσμα διάστασης n , το οποίο περιέχει αρχικές συνθήκες για τις μεταβλητές κατάστασης, γνωστές επ'ακριβώς,

$\mathbf{u} = [u_1, u_2, \dots, u_r]^T$ είναι διάνυσμα διάστασης r , το οποίο περιέχει μεταβλητές που καθορίζονται από τον χρήστη και η αριθμητική τους αξία είναι γνωστή ή έχουν μετρηθεί πειραματικά,

$\mathbf{f} = [f_1, f_2, \dots, f_n]^T$ είναι διάνυσμα διάστασης n , το οποίο περιέχει λειτουργίες / συναρτήσεις γνωστής μορφής,

$\mathbf{y} = [y_1, y_2, \dots, y_m]^T$ είναι διάνυσμα εξόδου διάστασης m , το οποίο περιέχει τις μεταβλητές που υπολογίζονται πειραματικά,

και

\mathbf{C} είναι πίνακας (παρατήρησης) διαστάσεων $m \times n$, ο οποίος δηλώνει ποιες από τις μεταβλητές κατάστασης (ή γραμμικούς συνδυασμούς μεταβλητών κατάστασης) υπολογίζονται πειραματικά.

Για να περιγραφεί πλήρως μια κατάσταση του συστήματος, στην πραγματικότητα η γνώση πολλών μεταβλητών είναι απαραίτητη. Οι μεταβλητές κατάστασης είναι το ελάχιστο σύνολο εξαρτημένων μεταβλητών που απαιτούνται για την πλήρη περιγραφή της κατάστασης του συστήματος. Το διάνυσμα εξόδου αντιπροσωπεύει κανονικά ένα υποσύνολο των μεταβλητών κατάστασης ή συνδυασμούς αυτών που μετρώνται. Επομένως, ο πίνακας παρατήρησης \mathbf{C} θα έχει μηδενικά παντού, εκτός από λίγες τοποθεσίες όπου θα έχει 1, τα οποία θα δείχνουν τις μεταβλητές κατάστασης που μετρώνται. Παράλληλα, η αρχική συνθήκη \mathbf{x}_0 είναι πολύ σημαντική, επιλέγεται από το χρήστη και είναι γνωστή επακριβώς. Τέλος, πρέπει να αναφερθεί ότι ως σφάλμα, παρουσιάζεται το σφάλμα στις μετρήσεις, οι ανεπάρκειες του μοντέλου, και εφόσον το μοντέλο αποτελείται από δυναμικά συστήματα, οι αυτόματες συσχετίσεις μεταξύ των διάφορων όρων σφάλματος.

2.2. Επιλογή Αντικειμενικής Συνάρτησης και υπολογισμός

Προτού ξεκινήσει η αναζήτηση για τις τιμές των παραμέτρων, πρέπει να επιλεγθεί η Αντικειμενική Συνάρτηση (objective function). Η συνάρτηση αυτή αναφέρεται ως $S(\mathbf{k})$ και πρόκειται για μια βαθμιαία συνάρτηση, η ελαχιστοποίηση της οποίας οδηγεί στην εύρεση του διάνυσματος \mathbf{k} , του ζητούμενου της παρούσας εργασίας. Πρόκειται για μια κατάλληλη μέτρηση της συνολικής απόκλισης των υπολογιζόμενων από το μοντέλο τιμών σε σχέση με τις δεδομένες παραμέτρους και τις τιμές τους. Η επιλογή της αντικειμενικής συνάρτησης είναι πολύ σημαντική, καθώς επηρεάζει τις τιμές των παραμέτρων αλλά και τις στατιστικές τους ιδιότητες.

Στην παρούσα εργασία, η αντικειμενική συνάρτηση είναι η μέθοδος των Ελάχιστων Τετραγώνων (Least Squares, LS) (Englezos and Kalogerakis, 1993 & 2001, Gill and Murray, 1975) και οι παραλλαγές της και παρουσιάζεται ως εξής:

$$S_{LS}(k) = \sum_{i=1}^N \mathbf{e}_i^T \mathbf{Q}_i \mathbf{e}_i \quad (2.4)$$

όπου

\mathbf{e}_i είναι το διάνυσμα των ‘υπολειμμάτων’ (δηλαδή η απόκλιση κάθε φορά από τις υπολογιζόμενες από το μοντέλο τιμές) από το i -οστό πείραμα και \mathbf{Q}_i είναι πίνακας στάθμισης (weighting matrix) διαστάσεων $m \times X$ που δίνεται από το χρήστη.

Πιο συγκεκριμένα, για συστήματα που περιγράφονται από ODEs η αντικειμενική συνάρτηση LS γίνεται:

$$S_{LS}(k) = \sum_{i=1}^N [\hat{\mathbf{y}}_i - \mathbf{y}(t_i, k)]^T \mathbf{Q}_i [\hat{\mathbf{y}}_i - \mathbf{y}(t_i, k)] \quad (2.5)$$

και αν πρόκειται για μοντέλο με μετρήσεις σε περισσότερα από ένα πειράματα (runs), κάτι σύχνα απαραίτητο:

$$S_{LS}(k) = \sum_{j=1}^{N_R} \left[\sum_{i=1}^N [\hat{\mathbf{y}}_i - \mathbf{y}(t_i, k)]^T \mathbf{Q}_i [\hat{\mathbf{y}}_i - \mathbf{y}(t_i, k)] \right]_j \quad (2.6)$$

όπου N_R είναι ο αριθμός των runs.

Ανάλογα από τον πίνακα στάθμισης \mathbf{Q}_i , η αντικειμενική συνάρτηση εμπίπτει στις παρακάτω περιπτώσεις:

1. Απλά ή Αβαρή Ελάχιστα Τετράγωνα (Simple or Unweighted LS)

Σε αυτή την περίπτωση, δεν υπάρχει παράγοντας στάθμισης για το άθροισμα των τετραγώνων των σφαλμάτων (sum of squares of errors, SSE), επομένως $\mathbf{Q}_i = \mathbf{I}$ και:

$$S_{LS}(k) = \sum_{i=1}^N \mathbf{e}_i^T \mathbf{e}_i \quad (2.7)$$

2. Σταθμισμένα Ελάχιστα Τετράγωνα (Weighted LS, WLS)

Σε αυτή την περίπτωση, ο πίνακας στάθμισης \mathbf{Q}_i είναι ίδιος για όλα τα πειράματα, επομένως $\mathbf{Q}_i = \mathbf{Q}$ για $i=1,2,\dots,N$.

3. Γενικευμένα Ελάχιστα Τετράγωνα (Generalized Least Squares, GLS)

Σε αυτή την περίπτωση, ο πίνακας στάθμισης \mathbf{Q}_i είναι διαφορετικός για κάθε πείραμα. Για να επιλεχθεί, λαμβάνονται υπόψη η μέθοδος της Μέγιστης Πιθανοφάνειας (Maximum Likelihood, ML, Ζαΐρης, 2010) και το Κριτήριο της Ορίζουσας (The Determinant Criterion).

Στατιστικά, ο πίνακας στάθμισης \mathbf{Q}_i πρέπει να είναι ίσος με τον αντίστροφο του $\mathbf{COV}(\mathbf{e}_i)$ (covariance, συνδιασπορά). Ωστόσο, σε πολλές περιπτώσεις οι πληροφορίες για το σφάλμα των μετρήσεων δεν αρκούν και είναι ιδιαίτερα χρήσιμο ο \mathbf{Q}_i να χρησιμοποιείται σαν πίνακας ομαλοποίησης (normalization matrix) για τις εκτιμώμενες αποκρίσεις (-μετρήσεις) ίδιας τάξης μεγέθους. Αν οι τιμές του μοντέλου δεν αλλάζουν σημαντικά από σημείο δεδομένων σε σημείο

δεδομένων, μπορεί να χρησιμοποιηθεί σταθερός \mathbf{Q} και η απλούστερη δεκτή μορφή αυτού είναι η εξής:

$$\mathbf{Q}_{jj} = \frac{1}{\left(\frac{1}{N} \sum_{i=1}^N \hat{\mathbf{y}}_{j,i} \right)^2} \quad (2.8)$$

δηλαδή διαγώνιος πίνακας του οποίου το j -οστό στοιχείο στη διαγώνιο είναι το αντίστροφο της τετραγωνισμένης μέσης απόκρισης της j -οστής μεταβλητής. Αυτό είναι αντίστοιχο με την παραδοχή συνεχούς τυπικού σφάλματος στην μέτρηση της j -οστής μεταβλητής και παράλληλα τα τυπικά σφάλματα των διαφορετικών μεταβλητών απόκρισης (δηλαδή των μετρήσεων του μοντέλου) είναι ανάλογα με την μέση τιμή των μεταβλητών. Η επιλογή αυτή του πίνακα \mathbf{Q} είναι ασφαλής όταν δεν υπάρχουν άλλες διαθέσιμες πληροφορίες και ο υπολογισμός των ελάχιστων τετραγώνων δίνει την ίδια βάση σε κάθε σφάλμα από διαφορετικές μεταβλητές απόκρισης. Αν οι μετρήσεις των μεταβλητών απόκρισης έχουν διαφορές μεγέθους αρκετών τάξεων, τότε χρησιμοποιείται η εξής μορφή του \mathbf{Q} :

$$\mathbf{Q}_{i,jj} = \frac{1}{\left(\hat{\mathbf{y}}_{j,i} \right)^2} \quad (2.9)$$

Αυτό είναι αντίστοιχο με την παραδοχή ότι το τυπικό σφάλμα στη i -οστή μέτρηση της j -οστής μεταβλητής απόκρισης είναι ανάλογο της τιμής της, ξανά μια ασφαλής επιλογή, αφού ο υπολογισμός των ελάχιστων τετραγώνων δίνει την ίδια βάση σε όλα τα σημεία δεδομένων.

2.3. Μέθοδος Gauss-Newton για ODEs

Η ελαχιστοποίηση της εξίσωσης (2.5), μπορεί να επιτευχθεί χρησιμοποιώντας διαφορετικές τεχνικές βελτιστοποίησης. Η μέθοδος Gauss-Newton είναι η καταλληλότερη για τα μοντέλα ODEs (Bard, 1970, 1974, Kalogerakis, 1983) και επιλέγεται στην εργασία αυτή. Αναλύεται παρακάτω και παρουσιάζονται και οι λόγοι που έγιναν συγκεκριμένες επιλογές και τροποποιήσεις στον αλγόριθμο.

Έστω ότι $\mathbf{k}^{(j)}$ η εκτίμηση (ορίζεται και ως «εκτιμήτρια», Ζαΐρης, 2010) των αγνώστων παραμέτρων στην j -οστή προσέγγιση (iteration). Η γραμμικοποίηση του διανύσματος εξόδου και η διατήρηση των όρων πρώτης τάξης αποδίδει:

$$\mathbf{y}(t_i, \mathbf{k}^{(j+1)}) = \mathbf{y}(t_i, \mathbf{k}^{(j)}) + \left(\frac{\partial \mathbf{y}^T}{\partial \mathbf{k}} \right)_i^T \Delta \mathbf{k}^{(j+1)} \quad (2.10)$$

Έστω ότι μεταξύ του διανύσματος εξόδου και των μεταβλητών κατάστασης υπάρχει γραμμική σχέση ($\mathbf{y}=\mathbf{C}\mathbf{x}$), τότε:

$$\mathbf{y}(t_i, \mathbf{k}^{(j+1)}) = \mathbf{C}\mathbf{x}(t_i, \mathbf{k}^{(j)}) + \mathbf{C} \left(\frac{\partial \mathbf{y}^T}{\partial \mathbf{k}} \right)_i^T \Delta \mathbf{k}^{(j+1)} \quad (2.11)$$

Όσον αφορά τα μοντέλα ODEs, ο πίνακας ευαισθησίας $\mathbf{G}(t_i) \equiv (\partial \mathbf{x}^T / \partial \mathbf{k})^T$ δε μπορεί να αποκτηθεί από μία απλή παραγωγή. Είναι όμως δυνατό να βρεθεί διαφορική εξίσωση, την οποία ικανοποιεί ο $\mathbf{G}(t)$, ο οποίος πλέον θεωρείται συνάρτηση του χρόνου και λύνεται ταυτόχρονα με σκετ άλλων διαφορικών εξισώσεων. Το σκετ αυτό αποκτάται παραγωγίζοντας την (2.1) ως προς το \mathbf{k} :

$$\frac{\partial}{\partial \mathbf{k}} \left(\frac{d\mathbf{x}}{dt} \right) = \frac{\partial}{\partial \mathbf{k}} (\mathbf{f}(\mathbf{x}, \mathbf{u}, \mathbf{k})) \quad (2.12)$$

Αντιστρέφοντας τη σειρά της παραγωγίσης στο αριστερό μέρος της (2.11) και εκτελώντας πεπλεγμένη παραγωγή στο δεξί μέρος:

$$\frac{d}{dt} \left[\left(\frac{\partial \mathbf{x}^T}{\partial \mathbf{k}} \right)^T \right] = \left(\frac{\partial \mathbf{f}^T}{\partial \mathbf{x}} \right)^T \left(\frac{\partial \mathbf{x}^T}{\partial \mathbf{k}} \right)^T + \left(\frac{\partial \mathbf{f}^T}{\partial \mathbf{k}} \right)^T \quad (2.13)$$

Η αλλιώς:

$$\frac{d\mathbf{G}(t)}{dt} = \left(\frac{\partial \mathbf{f}^T}{\partial \mathbf{x}} \right)^T \mathbf{G}(t) + \left(\frac{\partial \mathbf{f}^T}{\partial \mathbf{k}} \right)^T \quad (2.14)$$

Η αρχική συνθήκη $\mathbf{G}(t_0)$ αποκτάται παραγωγίζοντας την αρχική συνθήκη $\mathbf{x}(t_0)=\mathbf{x}_0$ ως προς \mathbf{k} και εφόσον η αρχική κατάσταση είναι ανεξάρτητη των παραμέτρων:

$$\mathbf{G}(t_0)=\mathbf{0} \quad (2.15)$$

Η (2.11) είναι πίνακας διαφορικών εξισώσεων και αντιπροσωπεύει ένα set από $n \times p$ ODEs. Μόλις αποκτηθούν οι συντελεστές ευαισθησίας (sensitivity coefficients), λύνοντας τις ODEs αυτές, τότε το διάνυσμα εξόδου, $\mathbf{y}(t_i, \mathbf{k}^{(j+1)})$, μπορεί να υπολογιστεί. Αντικαθιστώντας το διάνυσμα στην αντικειμενική συνάρτηση και χρησιμοποιώντας τη στατική συνθήκη $\partial S(\mathbf{k}^{(j+1)}) / \partial \mathbf{k}^{(j+1)} = \mathbf{0}$:

$$\mathbf{A} \Delta \mathbf{k}^{(j+1)} = \mathbf{b} \quad (2.16)$$

όπου

$$\mathbf{A} = \sum_{i=1}^N \mathbf{G}^T(t_i) \mathbf{C}^T \mathbf{Q}_i \mathbf{C} \mathbf{G}(t_i) \quad (2.17)$$

και

$$\mathbf{b} = \sum_{i=1}^N \mathbf{G}^T(t_i) \mathbf{C}^T \mathbf{Q}_i [\hat{\mathbf{y}}_i - \mathbf{C}\mathbf{x}(t_i, \mathbf{k}^{(j)})] \quad (2.18)$$

Η λύση της (2.14) δίνει το $\Delta \mathbf{k}^{(j+1)}$ και επομένως, το $\mathbf{k}^{(j+1)}$ αποκτάται από:

$$\mathbf{k}^{(j+1)} = \mathbf{k}^{(j)} + \mu \Delta \mathbf{k}^{(j+1)} \quad (2.19)$$

Το μ είναι παράμετρος βήματος ($0 < \mu \leq 1$), η οποία αποφασίζεται από τον κανόνα της διχοτόμησης, που θα αναλυθεί παρακάτω.

Συμπερασματικά, σε κάθε iteration, δεδομένων των τωρινών εκτιμήσεων για τις παραμέτρους ($\mathbf{k}^{(j)}$), αποκτούνται τα $\mathbf{x}(t)$ και $\mathbf{G}(t)$, ολοκληρώνοντας τις διαφορικές εξισώσεις κατάστασης και ευαισθησίας. Χρησιμοποιώντας αυτές τις τιμές υπολογίζεται η έξοδος του μοντέλου $\mathbf{y}(t_i, \mathbf{k}^{(j)})$ και οι συντελεστές ευαισθησίας $\mathbf{G}(t_i)$ για κάθε σημείο δεδομένων $i=1,2,\dots,N$. Έτσι, ρυθμίζονται ο πίνακας \mathbf{A} και το διάνυσμα \mathbf{b} και η εξίσωση (2.16) λύνεται, δίνοντας το $\Delta \mathbf{k}^{(j+1)}$ και επομένως, το $\mathbf{k}^{(j+1)}$.

Η διαδικασία αυτή παράγει μια ακολουθία από εκτιμήσεις παραμέτρων, $\mathbf{k}^{(1)}, \mathbf{k}^{(2)}, \dots$, που συχνά συγκλίνει στο βέλτιστο, \mathbf{k}^* , αν η αρχική εκτίμηση (initial guess) $\mathbf{k}^{(0)}$ ήταν ικανοποιητικά κοντά. Οι τιμές από τις παραμέτρους που συγκλίνουν αντιπροσωπεύουν τις εκτιμήσεις των Least Squares (LS), Weighted Least Squares (WLS) ή Generalized Least Squares (GLS), ανάλογα από την επιλογή των πινάκων στάθμισης \mathbf{Q}_i .

Η μέθοδος περιγράφεται αλγοριθμικά στο παρακάτω υποκεφάλαιο.

2.4. Αλγόριθμος Gauss-Newton για συστήματα από ODEs

Βηματική περιγραφή του αλγορίθμου:

1. Είσαγε την αρχική εκτίμηση για τις παραμέτρους $\mathbf{k}^{(0)}$.
2. Για $j = 0, 1, 2, \dots$, επανάλαβε.
3. Ολοκλήρωσε τις εξισώσεις κατάστασης και ευαισθησίας για να λάβεις τα $\mathbf{x}(t)$ και $\mathbf{G}(t)$. Για κάθε iteration t_i , $i=1,2,\dots,N$, υπολόγισε τα $\mathbf{y}(t_i, \mathbf{k}^{(j)})$ και $\mathbf{G}(t_i)$ και ρύθμισε τον πίνακα \mathbf{A} και το διάνυσμα \mathbf{b} .
4. Λύσε τη γραμμική εξίσωση $\mathbf{A} \Delta \mathbf{k}^{(j+1)} = \mathbf{b}$ και πάρε το $\Delta \mathbf{k}^{(j+1)}$.
5. Προσδιόρισε το μ χρησιμοποιώντας τον κανόνα της διχοτόμησης και λάβε το $\mathbf{k}^{(j+1)} = \mathbf{k}^{(j)} + \mu \Delta \mathbf{k}^{(j+1)}$.
6. Συνέχισε έως ότου επιτευχθεί ο μέγιστος αριθμός iterations ή επιτευχθεί η σύγκλιση.
7. Υπολόγισε τις στατιστικές ιδιότητες των εκτιμήσεων των παραμέτρων.

Όσον αφορά το βήμα 1, αξίζει να αναφερθεί ότι η επιλογή της αρχικής εκτίμησης, μαζί με την οπτική επιθεώρηση των δεδομένων στο χέρι, είναι βασικές ενέργειες του χρήστη για την βέλτιστη ανάπτυξη του αλγορίθμου. Μέσα από την οπτική επιθεώρηση, εντοπίζονται πιθανές

ακραίες τιμές (outliers), που οι Barnett et al. (1994) ορίζουν ως τις παρατηρήσεις σε ένα δείγμα που φαίνεται να είναι ασυνεπείς με το υπόλοιπο του δείγματος, για λόγους όπως ένα λάθος στους υπολογισμούς ή στην κωδικοποίηση των δεδομένων. Ένα δυνητικό outlier πρέπει πάντα να εξετάζεται προσεκτικά με αρχικό έλεγχο για λάθη κατά την εγγραφή ή τον πρώιμο χειρισμό των δεδομένων. Έπειτα, πρέπει να συμβαίνει προσεκτική εξέταση των πειραματικών συνθηκών κατά τη διάρκεια του συγκεκριμένου πειράματος και η απόρριψη του να συμβαίνει μόνο αν υπάρχουν ισχυροί μη στατιστικοί λόγοι. Τέλος, πρέπει να δίνεται σημαντική βάση στην επιλογή της αρχικής εκτίμησης ($\mathbf{k}^{(0)}$), καθώς μια καλή επιλογή αυτής οδηγεί σε γρήγορη σύγκλιση του μοντέλου στις βέλτιστες τιμές των παραμέτρων. Η αρχική εκτίμηση επιλέγεται από το χρήστη μελετώντας τη φύση, τη δομή και τη συμπεριφορά του κάθε μοντέλου.

2.4.1. Χρήση Λύτη Διαφορικών Εξισώσεων

Η ταυτόχρονη ολοκλήρωση των εξισώσεων κατάστασης και ευαισθησίας είναι η βέλτιστη προσέγγιση για εφαρμογή της μεθόδου Gauss-Newton, χωρίς να υπάρξει ανάγκη για αποθήκευση του $\mathbf{x}(t)$ συναρτήσει του χρόνου. Το $\mathbf{x}(t)$ είναι απαραίτητο για την λύση των Ιακωβιανών (Jacobian) συναρτήσεων στην (2.12), κατά τη λύση της διαφορικής εξίσωσης για να αποκτηθεί το $\mathbf{G}(t)$. Για να συμβεί αυτή η ολοκλήρωση, χρησιμοποιείται Λύσης διαφορικών εξισώσεων.

Ισχύει ότι:

$$\mathbf{G}(t) \equiv \left(\frac{\partial \mathbf{x}^T}{\partial \mathbf{k}} \right)^T = \left[\left(\frac{\partial \mathbf{x}}{\partial k_1} \right), \left(\frac{\partial \mathbf{x}}{\partial k_2} \right), \dots, \left(\frac{\partial \mathbf{x}}{\partial k_p} \right) \right] = [\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_p] \quad (2.20)$$

Τα διανύσματα $\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_p$ είναι διάστασης n , αντιπροσωπεύουν τους αντίστοιχους συντελεστές ευαισθησίας $\mathbf{k}_1, \mathbf{k}_2, \dots, \mathbf{k}_p$ και ικανοποιούν τις εξής ODEs:

$$\frac{d\mathbf{g}_1(t)}{dt} = \left(\frac{\partial \mathbf{f}^T}{\partial \mathbf{x}} \right)^T \mathbf{g}_1(t) + \left(\frac{\partial \mathbf{f}}{\partial k_1} \right); \quad \mathbf{g}_1(t_0) = 0 \quad (2.21a)$$

$$\frac{d\mathbf{g}_2(t)}{dt} = \left(\frac{\partial \mathbf{f}^T}{\partial \mathbf{x}} \right)^T \mathbf{g}_2(t) + \left(\frac{\partial \mathbf{f}}{\partial k_2} \right); \quad \mathbf{g}_2(t_0) = 0 \quad (2.21b)$$

·
·

$$\frac{d\mathbf{g}_p(t)}{dt} = \left(\frac{\partial \mathbf{f}^T}{\partial \mathbf{x}} \right)^T \mathbf{g}_p(t) + \left(\frac{\partial \mathbf{f}}{\partial k_p} \right); \quad g_p(t_0) = 0 \quad (2.21c)$$

Ο λύτης απαιτεί οι εξισώσεις να είναι στην εξής μορφή:

$$\frac{d\mathbf{z}}{dt} = \boldsymbol{\varphi}(\mathbf{z}) \quad ; \quad \mathbf{z}(t_0) \text{ δεδομένο} \quad (2.22)$$

Επομένως, παραγέται το διάνυσμα \mathbf{z} διάστασης $n \times (p+1)$:

$$\mathbf{z} = \begin{bmatrix} \mathbf{x}(t) \\ \hline \left(\frac{\partial \mathbf{x}}{\partial k_1} \right) \\ \hline \left(\frac{\partial \mathbf{x}}{\partial k_2} \right) \\ \hline \vdots \\ \hline \left(\frac{\partial \mathbf{x}}{\partial k_p} \right) \end{bmatrix} = \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{g}_1(t) \\ \mathbf{g}_2(t) \\ \vdots \\ \mathbf{g}_p(t) \end{bmatrix} \quad (2.23)$$

και το διάνυσμα συνάρτησης $\boldsymbol{\varphi}(\mathbf{z})$ διάστασης $n \times (p+1)$:

$$\boldsymbol{\varphi}(\mathbf{z}) = \begin{bmatrix} \mathbf{f}(\mathbf{x}, \mathbf{u}, \mathbf{k}) \\ \hline \left(\frac{\partial \mathbf{f}^T}{\partial \mathbf{x}} \right)^T \mathbf{g}_1(t) + \left(\frac{\partial \mathbf{f}}{\partial k_1} \right) \\ \hline \left(\frac{\partial \mathbf{f}^T}{\partial \mathbf{x}} \right)^T \mathbf{g}_2(t) + \left(\frac{\partial \mathbf{f}}{\partial k_2} \right) \\ \hline \vdots \\ \hline \left(\frac{\partial \mathbf{f}^T}{\partial \mathbf{x}} \right)^T \mathbf{g}_p(t) + \left(\frac{\partial \mathbf{f}}{\partial k_p} \right) \end{bmatrix} \quad (2.24)$$

με Ιακωβιανή (Jacobian):

$$\boldsymbol{\Phi}_{\mathbf{z}} = \left(\frac{\partial \boldsymbol{\varphi}^T}{\partial \mathbf{z}} \right)^T = \begin{bmatrix} \left(\frac{\partial \mathbf{f}^T}{\partial \mathbf{x}} \right)^T & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ * & \left(\frac{\partial \mathbf{f}^T}{\partial \mathbf{x}} \right)^T & \mathbf{0} & \dots & \mathbf{0} \\ * & \mathbf{0} & \left(\frac{\partial \mathbf{f}^T}{\partial \mathbf{x}} \right)^T & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ * & \mathbf{0} & \mathbf{0} & \dots & \left(\frac{\partial \mathbf{f}^T}{\partial \mathbf{x}} \right)^T \end{bmatrix} \quad (2.25)$$

Οι αστερίσκοι (*) στην πρώτη στήλη αντιπροσωπεύουν όρους που έχουν παραγώγους δεύτερης τάξης του \mathbf{f} ως προς το \mathbf{x} και επομένως μπορούν πρακτικά να αγνοηθούν. Επομένως, η Jacobian μπορεί να θεωρηθεί σαν διαγώνιος πίνακας (block diagonal matrix), όσον αφορά τον λύτη ODE.

Η χρήση του λύτη διευκολύνει τη διαδικασία της μεθόδου Gauss-Newton και έχει ως αποτέλεσμα την μείωση των απαιτήσεων μνήμης του προγράμματος και την αύξηση της ευρωστίας (robustness) της ολοκλήρωσης των εξισώσεων.

2.4.2. Χρήση του Κανόνα Διχοτόμησης και Υπέρβαση

Είναι συχνό η παράμετρος βήματος μ να ξεκινάει με την τιμή $\mu=1$ και να υποδιπλασιάζεται, όσο η αντικειμενική συνάρτηση είναι μικρότερη από αυτήν του προηγούμενου iteration. Η πρώτη τιμή που ικανοποιεί την παρακάτω ανισότητα ικανοποιεί τον κανόνα διχοτόμησης (bisection rule) και είναι δεκτή:

$$S(\mathbf{k}^{(j)} + \mu\Delta\mathbf{k}^{(j+1)}) < S(\mathbf{k}^{(j)}) \quad (2.26)$$

Στην περίπτωση των μοντέλων ODEs, η αντικειμενική συνάρτηση $S(\mathbf{k}^{(j)} + \mu\Delta\mathbf{k}^{(j+1)})$ για συγκεκριμένη τιμή του μ υποδηλώνει την ολοκλήρωση των εξισώσεων κατάστασης. Όταν η συνάρτηση αυτή γίνεται μεγαλύτερη από το $S(\mathbf{k}^{(j)})$, επιλέγεται μικρότερη τιμή για το μ και υπολογίζονται οι εξισώσεις κατάστασης. Η διαδικασία αυτή εξοικονομεί χρόνο υπολογισμού και αποφεύγεται η αριθμητική αστάθεια, αφού η αντικειμενική συνάρτηση μεγαλώνει γρήγορα και το πρόγραμμα σταματάει πριν συμβεί υπολογιστική υπερχείλιση (overflow) (Kalogerakis and Luus, 1983a).

Κάποιες φορές, κατά την επιλογή κατάλληλου μ , το $\Delta\mathbf{k}^{(j+1)}$ είναι υπερβολικά μεγάλο, με αποτέλεσμα να προκαλείται υπέρβαση (overstepping). Αυτό οδηγεί σε αριθμητική αστάθεια και υπολογιστικό overflow, προκαλώντας αδυναμία υπολογισμού για το διανύσμα εξόδου του πρώτου σημείου δεδομένων και για τον έλεγχο της ανισότητας (2.26). Για να αποφευχθεί αυτό, είναι σημαντική η επιλογή καλής ρουτίνας ολοκλήρωσης, η οποία θα ελέγχει τις προϋποθέσεις ανοχής, θα σταματάει την ολοκλήρωση, θα υποδιπλασιάζει το μ και θα συνεχίζει την ολοκλήρωση, προτού υπάρξει υπέρβαση. Ίσως χρειαστούν αρκετές επανεκκινήσεις της ρουτίνας προτού βρεθεί και επιλεγθεί κατάλληλο μ .

2.4.3. Ill-conditioning και Μέσα Αντιμετώπισης

Κακή κατάσταση (ill-conditioning) σε ένα πρόβλημα υπολογισμού παραμέτρων, έγκειται όταν ο πίνακας \mathbf{A} γίνεται μη αντιστρέψιμος (singular) ή σχεδόν μη αντιστρέψιμος και βρίσκεται σε κακή κατάσταση (ill-conditioned). Αυτό συμβαίνει όταν δύο ή περισσότερες από τις άγνωστες παραμέτρους συσχετίζονται σε μεγάλο βαθμό ή μία από τις παραμέτρους δεν έχει μετρήσιμη επίδραση στις μεταβλήτες απόκρισης.

Ο βαθμός ill-conditioning ενός αντιστρέψιμου πίνακα \mathbf{A} είναι ο αριθμός κατάστασης (condition number) και ορίζεται ως:

$$\text{cond}(\mathbf{A}) = \|\mathbf{A}\| \|\mathbf{A}^{-1}\| \quad (2.27)$$

Ο αριθμός κατάστασης είναι πάντα μεγαλύτερος του ενός και αντιπροσωπεύει τη μέγιστη διεύρυνση των σφαλμάτων στη δεξιά πλευρά του διανύσματος λύσης. Είναι επίσης ίσος με την τετραγωνική ρίζα της αναλογίας μεγαλύτερης και μικρότερης μοναδικής (singular) τιμής του \mathbf{A} . Στα προγράμματα υπολογισμού παραμέτρων, ο \mathbf{A} είναι θετικά ορισμένος και συμμετρικός,

επομένως το $\text{cond}(\mathbf{A})$ είναι ίσο και με την αναλογία μεγαλύτερης και μικρότερης ιδιοτιμής του \mathbf{A} :

$$\text{cond}(\mathbf{A}) = \frac{\lambda_{\max}}{\lambda_{\min}} \quad (2.28)$$

Γενικά, αν ο αριθμός κατάστασης είναι μικρότερος από 10^3 , τότε το πρόβλημα είναι καλά δομημένο (well-posed problem), αν είναι μεγαλύτερος από 10^{10} , τότε το **πρόγραμμα** είναι σε σχετικά κακή κατάσταση και αν είναι μεγαλύτερος από 10^{30} , τότε το **πρόγραμμα** βρίσκεται σε πολύ κακή κατάσταση και ίσως υπάρξουν προβλήματα υπολογιστικού overflow. Ο αριθμός κατάστασης πίνακα \mathbf{A} είναι στενά συνδεδεμένος με την ευαισθησία της λύσης του γραμμικού συστήματος εξισώσεων $\mathbf{Ax}=\mathbf{b}$. Κατά την επίλυση αυτής της εξίσωσης, το σφάλμα στη λύση μπορεί να μεγεθυνθεί κατά ένα ποσό τόσο μεγάλο όσο το $\text{cond}(\mathbf{A})$ επί τη νόρμα του σφάλματος στα \mathbf{A} και \mathbf{b} εξαιτίας της ύπαρξης σφάλματος στα δεδομένα.

$$\frac{\|\Delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq \text{cond}(\mathbf{A}) \left(\frac{\|\Delta \mathbf{b}\|}{\|\mathbf{b}\|} + \frac{\|\Delta \mathbf{A}\|}{\|\mathbf{A}\|} \right) \quad (2.29)$$

Επομένως, σε ένα πρόβλημα σε κακή κατάσταση (ill-conditioned) το σφάλμα στο διάνυσμα λύσης θα είναι μεγάλο, ενώ σε ένα πρόβλημα σε καλή κατάσταση (well-conditioned) θα είναι μικρό. Στον υπολογισμό παραμέτρων, το διάνυσμα \mathbf{b} αποτελείται από γραμμικό συνδυασμό των μεταβλητών απόκρισης (μετρήσεις), οι οποίες περιέχουν σφάλμα. Αντιθέτως, ο πίνακας \mathbf{A} δεν εξαρτάται ρητά από τις μεταβλητές απόκρισης. Εξαρτάται από τους συντελεστές ευαισθησίας των παραμέτρων, που με τη σειρά τους εξαρτώνται από τις ανεξάρτητες μεταβλητές (οι οποίες είναι γνωστές επ'ακριβώς) και από το διάνυσμα εκτιμώμενων παραμέτρων \mathbf{k} , το οποίο ενσωματώνει την αβεβαιότητα στα δεδομένα. Επομένως, το μεγαλύτερο μέρος αβεβαιότητας στην (2.29) θα βρίσκεται στο $\Delta \mathbf{b}$.

Αν ο πίνακας \mathbf{A} είναι σε κακή κατάσταση στην βέλτιστη περίπτωση (δηλαδή $\mathbf{k}=\mathbf{k}^*$, όπου $*$ =optimal), τότε το πρόβλημα είναι σε «αληθινά» κακή κατάσταση και οι εκτιμώμενες παράμετροι θα έχουν πολύ αμφισβητήσιμες τιμές με απαράδεκτα μεγάλες διακυμάνσεις. Η μόνη βελτίωση σε τέτοιο πρόβλημα είναι η επανεξέταση και επανασχεδίαση του μοντέλου και της ρουτίνας.

Αντιθέτως, αν ο πίνακας \mathbf{A} είναι σε καλή κατάσταση στην βέλτιστη περίπτωση, μπορεί εύκολα να είναι ill-conditioned όταν οι παράμετροι βρίσκονται μακριά από τις βέλτιστες τιμές τους. Αυτό συμβαίνει συχνά κατά τον υπολογισμό παραμέτρων και αντιμετωπίζεται βρίσκοντας μέσο μετακίνησης των εκτιμήσεων των παραμέτρων από την αρχική επιλογή στη βέλτιστη, ακόμα και αν ο αριθμός κατάστασης του \mathbf{A} είναι υπερβολικά μεγάλος για τα αρχικά iterations. Στην παρούσα εργασία, τα μέσα αυτά είναι (α) η τροποποίηση του Marquardt, (β) η χρήση ψευδοαντίστροφων (pseudoinverse) πινάκων και (γ) η αλλαγή κλίμακας (scaling) του πίνακα \mathbf{A} :

(α) Ο Levenberg το 1944 και αργότερα ο Marquardt (1963) πρότειναν την τροποποίηση των κανονικών εξισώσεων προσθέτοντας έναν μικρό θετικό αριθμό, γ^2 , στα διαγώνια στοιχεία του \mathbf{A} .

Σε κάθε iteration η αύξηση στο διάνυσμα παραμέτρων αποκτάται επιλύοντας την ακόλουθη εξίσωση:

$$(\mathbf{A} + \gamma^2 \mathbf{I}) \Delta \mathbf{k}^{(j+1)} = \mathbf{b} \quad (2.30)$$

Η τροποποίηση του Levenberg-Marquardt αποκτάται και εξετάζοντας τις ιδιοτιμές του τροποποιημένου πίνακα $(\mathbf{A} + \gamma^2 \mathbf{I})$. Ισχύει ότι η αποσύνθεση του \mathbf{A} είναι $\mathbf{V}^T \mathbf{A} \mathbf{V}$ και:

$$\mathbf{A} + \gamma^2 \mathbf{I} = \mathbf{V}^T \mathbf{A} \mathbf{V} + \gamma^2 \mathbf{V}^T \mathbf{V} = \mathbf{V}^T (\mathbf{A} + \gamma^2 \mathbf{I}) \mathbf{V} = \mathbf{V}^T \mathbf{\Lambda}^M \mathbf{V} \quad (2.31)$$

όπου

$$\mathbf{\Lambda}^M = \text{diag}(\lambda_1 + \gamma^2, \lambda_2 + \gamma^2, \dots, \lambda_p + \gamma^2) \quad (2.32)$$

Όλες οι ιδιοτιμές του πίνακα \mathbf{A} αυξάνονται κατά γ^2 , επομένως οι μεγάλες ιδιοτιμές θα παραμένουν πρακτικά ίδιές και οι μικρές, που είναι πολύ μικρότερες από το γ^2 , θα ισούνται πρακτικά με γ^2 . Επομένως:

$$\text{cond}(\mathbf{A}) = \frac{\lambda_{\max} + \gamma^2}{\lambda_{\min} + \gamma^2} \approx \frac{\lambda_{\max}}{\gamma^2} \quad (2.33)$$

Η μέθοδος αυτή πρέπει να χρησιμοποιείται μόνο σε προβληματικές περιπτώσεις, καθώς ο χρήστης καλείται να αποφασίσει πως θα χειριστεί το γ^2 του Marquardt αλλά και το μέγεθος του βήματος λ . Βασικό είναι να χρησιμοποιείται γ^2 μεγαλύτερο από τις πιο μικρές ιδιοτιμές του \mathbf{A} που παρατηρούνται, έτσι ώστε να υπολογιστεί το $\|\Delta \mathbf{k}^{(j+1)}\|$ και μέσω του κανόνα διχοτόμησης να βρεθεί δεκτή παράμετρος βήματος.

(β) Η αποσύνθεση θετικά ορισμένου και συμμετρικού πίνακα \mathbf{A} είναι $\mathbf{V}^T \mathbf{A} \mathbf{V}$, όπου ο \mathbf{V} είναι ορθογωνικός πίνακας (δηλαδή $\mathbf{V}^{-1} = \mathbf{V}^T$) και $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_p)$ με ιδιοτιμές σε φθίνουσα σειρά ($\lambda_1 > \lambda_2 > \dots > \lambda_p$). Επομένως:

$$\mathbf{A}^{-1} = \mathbf{V}^T \mathbf{\Lambda}^{-1} \mathbf{V} \quad (2.34)$$

Αν ο πίνακας \mathbf{A} βρίσκεται σε καλή κατάσταση, η παραπάνω εξίσωση χρησιμοποιείται άμεσα. Αλλιώς, χρησιμοποιείται χωρίς περαιτέρω υπολογιστική προσπάθεια ο ψευδοαντίστροφος (pseudoinverse) του και στην παραπάνω εξίσωση αντι για $\mathbf{\Lambda}^{-1}$ χρησιμοποιείται πίνακας $\mathbf{\Lambda}^\#$. Αυτός αποκτάται, αντιστρέφοντας τις ιδιοτιμές μεγαλύτερες από έναν μικρό αριθμό δ που ορίζει ο χρήστης και μηδενίζοντας τον υπόλοιπο πίνακα, δηλαδή:

$$\mathbf{\Lambda}^\# = \text{diag}(1/\lambda_1, 1/\lambda_2, \dots, 1/\lambda_\pi, 0, \dots, 0) \quad (2.35)$$

όπου $\lambda_j \geq \delta$, $j=1,2,\dots,\pi$ και $\lambda_j < \delta$, $j=\pi+1,\pi+2,\dots,p$. Πρακτικά, έτσι αγνοούνται οι ιδιοτιμές πολύ κοντά στο μηδέν και βρίσκεται λύση στην εξίσωση (2.16) με $\|\Delta \mathbf{k}^{(j+1)}\|$ όσο πιο μικρό γίνεται.

(γ) Όταν οι παράμετροι διαφέρουν περισσότερο από μία τάξη μεγέθους, υπάρχει περίπτωση ο πίνακας \mathbf{A} να φαίνεται ότι βρίσκεται σε κακή κατάσταση, ακόμα κι αν το μοντέλο του υπολογισμού των παραμέτρων έχει ρυθμιστεί σωστά. Αυτό αντιμετωπίζεται με την εισαγωγή των μειωμένων συντελεστών ευαισθησίας:

$$\mathbf{G}_{R_{ij}} = \left(\frac{\partial y_i}{\partial k_j} \right) k_j \quad (2.36)$$

και

$$\mathbf{G}_R = \mathbf{G} \mathbf{K} \quad (2.37)$$

όπου

$$\mathbf{K} = \text{diag}(k_1, k_2, \dots, k_p) \quad (2.38)$$

Αυτή η αλλαγή κλίμακας (scaling), μετατρέπει την (2.16) σε:

$$\mathbf{A}_R \Delta \mathbf{k}_R^{(j+1)} = \mathbf{b}_R \quad (2.39)$$

όπου

$$\mathbf{A}_R = \mathbf{K} \left[\sum_{i=1}^N \mathbf{G}_i^T \mathbf{Q}_i \mathbf{G}_i \right] \mathbf{K} = \mathbf{K} \mathbf{A} \mathbf{K} \quad (2.40)$$

$$\mathbf{b}_R = \mathbf{K} \left[\sum_{i=1}^N \mathbf{G}_i^T \mathbf{Q}_i \left[\hat{\mathbf{y}}_i - f(\mathbf{x}_i, \mathbf{k}^{(j)}) \right] \right] = \mathbf{K} \mathbf{b} \quad (2.41)$$

και

$$\Delta \mathbf{k}_R^{(j+1)} = \mathbf{K}^{-1} \Delta \mathbf{k}^{(j+1)} \quad (2.42)$$

Οι εκτιμώμενες παράμετροι για το επόμενο iteration υπολογίζονται έτσι:

$$\mathbf{k}^{(j+1)} = \mathbf{k}^{(j)} + \mu \mathbf{K} \Delta \mathbf{k}_R^{(j+1)} \quad (2.43)$$

ή αντίστοιχα

$$\mathbf{k}_i^{(j+1)} = \mathbf{k}_i^{(j)} \left[1 + \mu \Delta \mathbf{k}_{R,i}^{(j+1)} \right]; \quad i = 1, \dots, p \quad (2.44)$$

Η τροποποίηση της αλλαγής κλίμακας του πίνακα \mathbf{A} οδηγεί σε σημαντική βελτίωση της κατάστασης του και ο αριθμός κατάστασης ($\text{cond}(\mathbf{A}_R)$) δίνει μια πολύ περισσότερο αξιόπιστη μέτρηση του βαθμού του ill-conditioning, όσον αφορά το πρόβλημα υπολογισμού παραμέτρων.

2.5. Στατιστικός Λογισμός

Αφού έχουν υπολογιστεί οι άγνωστες παράμετροι, εκτελούνται κάποιοι ακόμα υπολογισμοί για βασικές στατιστικές εκφράσεις για τις παραμέτρους. Το επιπλέον αυτό υπολογιστικό βήμα είναι ιδιαίτερα πολύτιμο για το χρήστη, καθώς τα αποτελέσματα του

αποτελούν μέτρο της ποιότητας του μοντέλου και της αξιοπιστίας των εκτιμήσεων των παραμέτρων.

Κατά τη χρήση της μεθόδου Gauss-Newton, οι εξισώσεις του μοντέλου γραμμικοποιούνται και σε κάθε iteration λύνεται το αντίστοιχο γραμμικό LS πρόβλημα. Επομένως, οι τιμές των παραμέτρων που υπολογίζονται έχουν ιδιότητες γραμμικού LS. Οι εκτιμήσεις είναι κανονικά κατανομημένες και αμερόληπτες και ο πίνακας συνδιασποράς (covariance, COV) δίνεται από την εξής εξίσωση:

$$COV(\mathbf{k}^*) = \sigma_\varepsilon^2 [\mathbf{A}^*]^{-1} \quad (2.45)$$

όπου \mathbf{A}^* είναι ο πίνακας \mathbf{A} για \mathbf{k}^* (βέλτιστη εκτίμηση) και σ_ε^2 είναι η διασπορά (variance), υπολογιζόμενη παρακάτω:

$$\hat{\sigma}_\varepsilon^2 = \frac{S(\mathbf{k}^*)}{(d.f.)} = \frac{S(\mathbf{k}^*)}{Nm - p} \quad (2.46)$$

όπου (d.f.)= N^*m-p και είναι οι βαθμοί ελευθερίας (degrees of freedom), δηλαδή ο συνολικός αριθμός μετρήσεων (N^*m) μείον τον αριθμό των αγνώστων παραμέτρων (p).

Το $(1-\alpha)100\%$ οριακό διάστημα εμπιστοσύνης για κάθε παράμετρο \mathbf{k}_i , $i=1,2,\dots,p$ είναι:

$$\mathbf{k}_i^* - t_{\alpha/2}^v \hat{\sigma}_{k_i} \leq \mathbf{k}_i \leq \mathbf{k}_i^* + t_{\alpha/2}^v \hat{\sigma}_{k_i} \quad (2.47)$$

όπου το $t_{\alpha/2}^v$ αποκτάται από τους πίνακες κατανομής του t κατά Student με $v=(N^*m-p)$ βαθμούς ελευθερίας και α επιλεγμένο από το χρήστη. Το τυπικό σφάλμα για την παράμετρο \mathbf{k}_i , $\hat{\sigma}_{k_i}$, αποκτάται από την τετραγωνική ρίζα της διασποράς ($\hat{\sigma}_\varepsilon^2$) επί την τετραγωνική ρίζα των αντίστοιχων διαγώνιων στοιχείων του αντίστροφου του πίνακα \mathbf{A}^* :

$$\hat{\sigma}_{k_i} = \hat{\sigma}_\varepsilon \sqrt{\left\{ [\mathbf{A}^*]^{-1} \right\}_{ii}} \quad (2.48)$$

Για $v \geq 30$, το $t_{\alpha/2}^v \approx z_{\alpha/2}$, όπου το $z_{\alpha/2}$ αποκτάται από τους πίνακες τυπικής κανονικής κατανομής. Πιο απλά, όταν υπάρχουν πολλά σημεία δεδομένων, το 95% διάστημα εμπιστοσύνης είναι 2 φορές το τυπικό σφάλμα $\hat{\sigma}_{k_i}$.

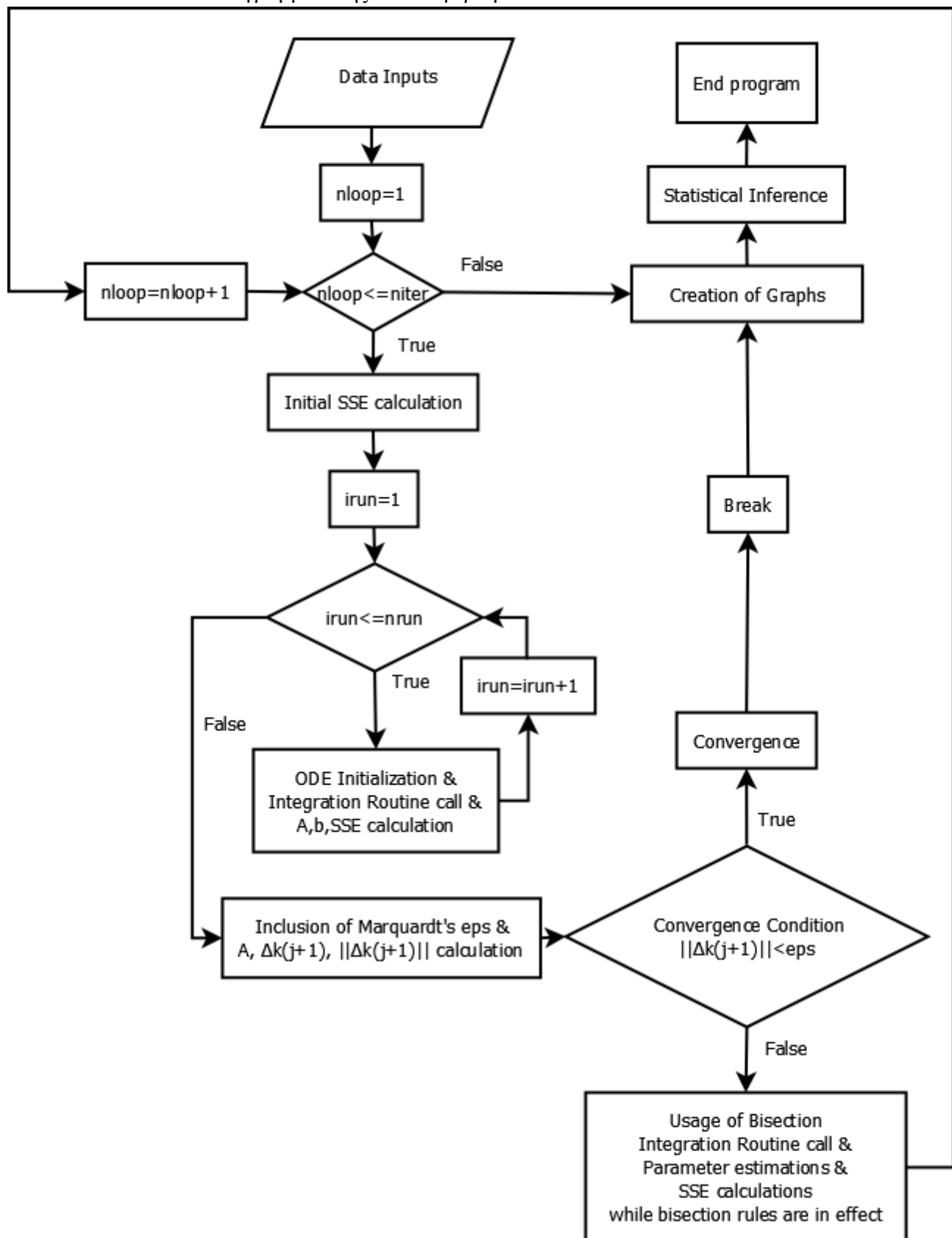
3. Κύριες αλλαγές σε περιβάλλον *MATLAB*

Ο αλγόριθμος της παρούσας εργασίας αναπτύσσεται σε περιβάλλον MATLAB και βασίζεται σε υπαρκτό αλγόριθμο ανεπτυγμένο σε FORTRAN. Αποτελείται από πέντε αρχεία:

- fmodel.m
- fmodelall.m
- Jmodel.m
- Jmodelall.m
- ODE.m

Στα πρώτα τέσσερα αρχεία αναπτύσσονται οι εξισώσεις του μοντέλου και στο πέμπτο αναπτύσσεται το πρόγραμμα που εκτελείται και χρησιμοποιεί τα παραπάνω. Στο παρόν κεφάλαιο παρουσιάζονται οι κύριες αλλαγές και προσθήκες που έγιναν στο πρόγραμμα για την επίλυση του προβλήματος υπολογισμού παραμέτρων και την παρουσίαση των αποτελεσμάτων γραφικά και αριθμητικά. Αρχικά, παρουσιάζεται ένα γενικό διάγραμμα ροής του αλγορίθμου:

Εικόνα 3.1. Γενικό Διάγραμμα Ροής του Αλγορίθμου



Ξεκινώντας τη λειτουργία του, ο αλγόριθμος δέχεται τις απαραίτητες εισόδους (data inputs) από τον χρήστη (αρχικές συνθήκες (x_0), μέγιστος αριθμός προσεγγίσεων (niter, number of iterations), αριθμός πειραμάτων (nrun), αριθμός εκτιμώμενων παραμέτρων (npar), αρχική εκτιμήση για τις παραμέτρους (\mathbf{p}_0 διάνυσμα στον αλγόριθμο). Έπειτα, για κάθε iteration, υπολογίζεται το αρχικό Άθροισμα Ελαχίστων Τετραγώνων σφάλματος (Sum of Square of Errors, SSE) και για κάθε πείραμα αρχικοποιείται το μοντέλο της ODE και καλείται η Ρουτίνα Ολοκλήρωσης. Υπολογίζεται ξανά το SSE, ο πίνακας \mathbf{A} και το διάνυσμα \mathbf{b} . Όταν τελειώσει αυτή η επανάληψη, προστίθεται η σταθερά του Marquardt (αν υπάρχει) και υπολογίζονται ο \mathbf{A} και τα $\Delta \mathbf{k}^{(j+1)}$, $\|\Delta \mathbf{k}^{(j+1)}\|$. Ελέγχεται αν ισχύει η συνθήκη σύγκλισης ($\|\Delta \mathbf{k}^{(j+1)}\| < \text{eps}$, όπου $\text{eps} = \text{eps} = 0.0005 * \text{npar}$). Αν ισχύει, η σύγκλιση έχει επιτευχθεί, η κύρια επανάληψη του προγράμματος (nloop \leq niter) σταμάταται (συμβαίνει “break”), παράγονται τα γραφικά και οι στατιστικές εκφράσεις και το πρόγραμμα σταματάει τη λειτουργία του. Αν η συνθήκη δεν ισχύει, χρησιμοποιείται η διχοτόμηση (bisection). Όσο οι κανόνες της διχοτόμησης δεν ικανοποιούνται (στον κώδικα, όσο ισχύει “while SSEnew > SSEtotal”), συμβαίνει διχοτόμηση (dpnew(i) = 0.50 * dpnew(i)), η ρουτίνα ολοκλήρωσης καλείται και το SSE υπολογίζεται. Όταν πλέον οι κανόνες αυτοί ικανοποιηθούν, ολοκληρώνεται μία επανάληψη του κύριου for-loop. Η διαδικασία αυτή επαναλαμβάνεται μέχρι είτε να υπάρξει σύγκλιση και να συμβεί break, είτε μέχρι να πραγματοποιηθεί ο μέγιστος αριθμός iterations. Τελικά, εμφανίζονται τα γραφικά και τα στατιστικά που υπολογίστηκαν και η λειτουργία του προγράμματος ολοκληρώνεται.

Ρουτίνα Ολοκλήρωσης

```
for irun=1:nrun
xinit=[x0(irun,:) zeros(1,neq-nx)]';
np=nobs(irun);
tinit=t0(irun);
tfinal=tp(irun,np);
tspan=[tinit tfinal];

%-----Call INTEGRATION ROUTINE
%press=pressure(irun);
options = odeset('Jacobian',@Jmodelall);
solution = ode15s(@fmodelall,tspan,xinit,options);
solutionstorage(nloop,:)=solution;
```

Για κάθε πείραμα (run) του μοντέλου, ρυθμίζονται τα πειραματικά δεδομένα (tspan για το χρόνο, xinit για τις τιμές). Έπειτα καθορίζονται οι επιλογές για τη λύση του μοντέλου (options, οι Ιακωβιανές εξισώσεις) και καλείται η συνάρτηση ode15s της MATLAB, αφού δεχθεί τα απαραίτητα ορίσματα. Τα αποτελέσματα της αποθηκεύονται στο solution και στο solutionstorage, μεταβλητή που αποθηκεύει τα αποτελέσματα για κάθε προσέγγιση (iteration).

Υπολογισμός της Αντικειμενικής Συνάρτησης

```
tdata=tp(irun,1:lengthDatafile(irun));
xmodel = deval(solution,tdata);
%Update Objective function and matrix A, b
%Select appropriate weighting factor (GLS vs WLS)
ydata_irun=squeeze(ydata(irun,:,1:np))';
ymodel=xmodel(1:ny,1:np)';
error=ydata_irun - ymodel;
%-----compute GLS or WLS weights
    if (GLS == 'on')
        for inp=1:np
            for j=1:ny
                qy(j,inp)= q(j)/(abs(ydata_irun(inp,j)) + epsmin);
                errorq(inp,j)=error(inp,j)*qy(j,inp);
            end
        end
    else
        for inp=1:np
            for j=1:ny
                qy(j,inp)= q(j);
            end
        end
        errorq=error*diag(q);
    end
SSEtotal=SSEtotal+norm(errorq);
```

Από το solution, προκύπτει το xmodel (μέσω της deval, που αξιολογεί τη λύση της διαφορικής εξίσωσης για τις τιμές του χρόνου που υπάρχουν μετρήσεις) και το ymodel. Υπολογίζεται το σφάλμα (error) για τη συγκεκριμένη επανάληψη και, ανάλογα από την αντικειμενική συνάρτηση που έχει επιλεχθεί (GLS ή WLS), σταθμίζεται από τον πίνακα στάθμισης qy. Προκύπτει το σταθμισμένο σφάλμα (errorq), του οποίου η νόρμα προστίθεται στην αντικειμενική συνάρτηση. Έτσι, υπολογίζεται η νέα τιμή της αντικειμενικής συνάρτησης (SSEtotal).

Δημιουργία γραφικών

Όπως αναφέρεται και παραπάνω, στο solution (και solutionstorage) αποθηκεύονται τα αποτελέσματα της λύσης του μαθηματικού μοντέλου. Γράφημα (plot) με δεδομένα τα solution.x(:) και solution.y(n,:), δημιουργεί συνεχή γραμμή με τις λύσεις του μοντέλου για το συγκεκριμένο iteration, για ένα x_n . Παράλληλα, plot με δεδομένα τα tdata και ydata_irun(:,n) δημιουργεί σημεία, τα οποία απεικονίζουν τα πειραματικά δεδομένα, για ένα x_n . Αυτά τα δύο plots, εμφανιζόμενα μαζί, δείχνουν την απόσταση των εκτιμήσεων του μοντέλου (συνεχής γραμμή) από τα πειραματικά δεδομένα. Αντίστοιχα, όταν τα plots που προκύπτουν για κάθε x, εμφανίζονται μαζί, τότε υπάρχει συνολική εικόνα για το μοντέλο για το συγκεκριμένο iteration. Τέλος, όταν υπάρξει

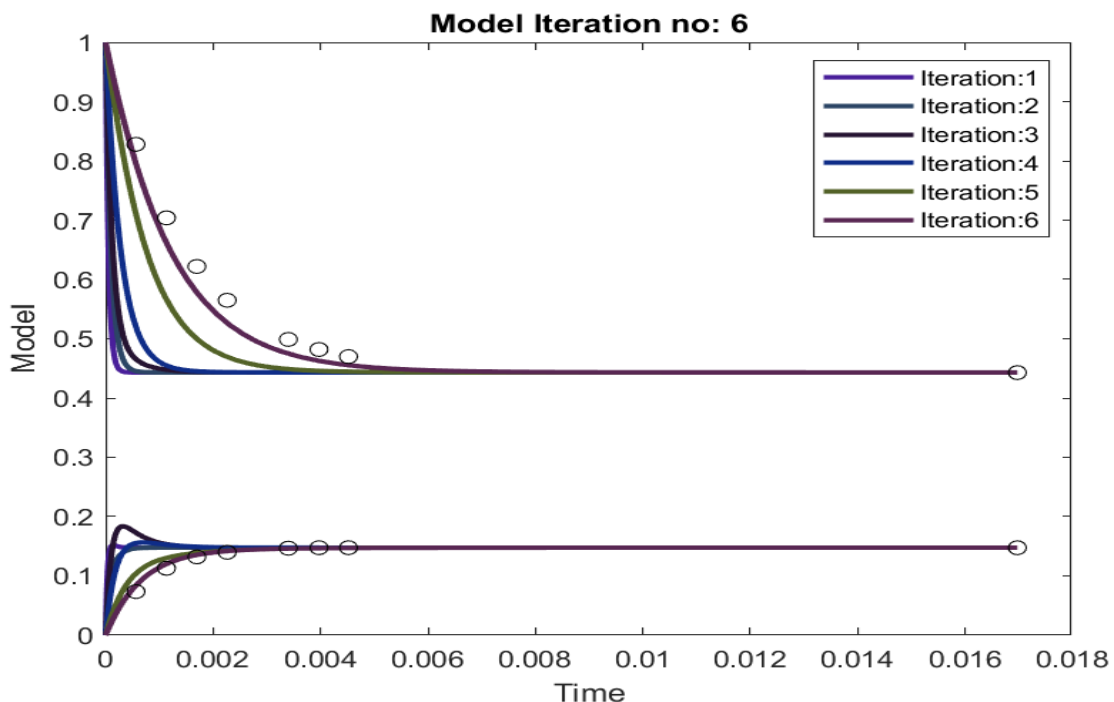
επανάληψη αυτής της διαδικασίας για κάθε iteration, χρησιμοποιώντας τα δεδομένα που έχουν αποθηκευτεί στο solutionstorage, φαίνεται η αλλαγή των εκτιμήσεων του μοντέλου ανά iteration και η μείωση της απόστασης, όσον αφορά τη σύγκλιση εκτιμώμενων και πειραματικών δεδομένων.

```

plotmake=ceil(sqrt(nloop));
for j=1:nloop
    subplot(plotmake,plotmake,j)
    for i=1:j
        solutiontemp=solutionstorage(i,:,:);
        solutiontemp2=squeeze(solutiontemp);
        plot(solutiontemp2.x(:),solutiontemp2.y(1,:), 'Color',[0 0.4470
0.7410], 'LineWidth',2)
        hold on
        plot(tdata,ydata_irun(:,1), 'o', 'Color',[0 0.4470
0.7410], 'LineWidth',2)
        hold on
        plot(solutiontemp2.x(:),solutiontemp2.y(2,:), 'Color',[0.6350 0.0780
0.1840], 'LineWidth',2)
        xlabel('Time'), ylabel('Model')
        hold on
        plot(tdata,ydata_irun(:,2), 'o', 'Color',[0.6350 0.0780
0.1840], 'LineWidth',2)
        end
        title(['Model Iteration no: ', num2str(j)])
    end
end

```

Εικόνα 3.2. Παράδειγμα γραφικής παράστασης που δημιουργείται για συγκεκριμένο iteration (συνολικό plot παρουσιάζεται στο επόμενο κεφάλαιο):



4. Εφαρμογή

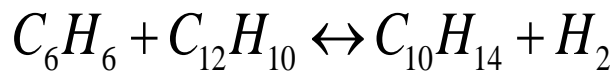
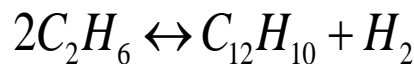
Στα παραπάνω κεφάλαια, αναλύεται η απαραίτητη θεωρία, όσον αφορά τις συνήθεις διαφορικές εξισώσεις και τη μέθοδο Gauss-Newton που χρησιμοποιείται για τον υπολογισμό άγνωστων παραμέτρων. Στην παρούσα εργασία, ως δεδομένο υπάρχουν μαθηματικά μοντέλα που έχουν αναπτυχθεί και πειραματικά δεδομένα που έχουν συλλεχθεί τα τελευταία χρόνια στο Εργαστήριο Βιοχημικής Μηχανικής και Περιβαλλοντικής Βιοτεχνολογίας του Πολυτεχνείου Κρήτης.

Όπως αναφέρεται και στην εισαγωγή, σκοπός της παρούσας εργασίας είναι η μετατροπή των μαθηματικών μοντέλων, ώστε να εκτελούνται στο περιβάλλον της MATLAB και όχι σε αντίστοιχο της FORTRAN. Παράλληλα, προστίθεται κώδικας, ο οποίος δημιουργεί γραφικά που δείχνουν την πορεία της σύγκλισης του εκτιμώμενου μοντέλου με τα πειραματικά δεδομένα. Τέλος, προστίθεται κώδικας που υπολογίζει βασικές στατιστικές υπολογιστικές εκφράσεις για τις παραμέτρους. Παρακάτω περιγράφονται οι δύο εφαρμογές που αναλύονται στην εργασία αυτή και τα αποτελέσματα που προκύπτουν από την εκτέλεση του αλγορίθμου που αναπτύχθηκε.

4.1. Εφαρμογή 1: Πυρολυτική Αφυδρογόνωση του Βενζολίου σε Διφαινύλιο και Τριφαινύλιο

4.1.1. Περιγραφή

Η πρώτη εφαρμογή του αλγορίθμου, ο οποίος αναπτύχθηκε στη MATLAB, αφορά την πυρολυτική αφυδρογόνωση του βενζολίου σε διφαινύλιο και τριφαινύλιο (Seinfeld and Gavalas, 1970, Hougen and Watson, 1948) και περιγράφεται ως εξής:



με μοντέλο

$$\frac{dx_1}{dt} = -r_1 - r_2 \quad (4.1a)$$

$$\frac{dx_2}{dt} = \frac{r_1}{2} - r_2 \quad (4.1b)$$

όπου

$$r_1 = k_1 \left[x_1^2 - x_2(2 - 2x_1 - x_2) / 3K_1 \right] \quad (4.2a)$$

$$r_2 = k_2 [x_1 x_2 - (1 - x_1 - 2x_2)(2 - 2x_1 - x_2) / 9K_2] \quad (4.2b)$$

όπου το x_1 απεικονίζει lb-mole βενζίνης ανά lb-mole αγνής τροφοδοτούμενης βενζίνης και το x_2 απεικονίζει lb-mole διφαινόλης ανά lb-mole αγνής τροφοδοτούμενης βενζίνης. Οι παράμετροι K_1 και K_2 είναι σταθερές ισορροπίας και οι παράμετροι k_1 και k_2 είναι άγνωστες σταθερές ρυθμού αντίδρασης. Η τροφοδοσία στον αντιδραστήρα ροής είναι αγνή βενζίνη και τα δεδομένα του πειράματος είναι οχτώ μετρήσεις των x_1 και x_2 που προκύπτουν από οχτώ τιμές ταχύτητας ανάστροφου διαστήματος (Reciprocal Space Velocity):

Πίνακας 4.1. Δεδομένα για την Πυρολυτική Αφυδρογόνωση του Βενζολίου

Reciprocal Space Velocity (t) × 10 ⁴	x_1	x_2
5.63	0.828	0.0737
11.32	0.704	0.113
16.97	0.622	0.1322
22.62	0.565	0.1400
34.0	0.499	0.1468
39.7	0.482	0.1477
45.2	0.470	0.1477
169.7	0.443	0.1476

Πηγή: Seinfeld and Gavalas (1970), Hougen and Watson (1948)

Μετρώνται και οι δύο μεταβλήτες κατάσταση, επομένως το διάνυσμα εξόδου είναι το ίδιο με το διάνυσμα κατάσταση, δηλαδή $y_1=x_1$ και $y_2=x_2$. Από την επανάληψη της μικρότερης ταχύτητας, προκύπτει ότι $K_1=0.242$ και $K_2=0.428$.

Επομένως:

$$\frac{dx_1}{dt} = f_1(x_1, x_2; k_1, k_2) \quad (4.3a)$$

$$\frac{dx_2}{dt} = f_2(x_1, x_2, k_1, k_2) \quad (4.3b)$$

Όπου $f_1 = (-r_1 - r_2)$ και $f_2 = r_1 / 2 - r_2$.

Ο πίνακας ευαισθησίας $\mathbf{G}(t)$ είναι πινάκας 2x2 με στοιχεία:

$$\mathbf{G}(t) = [\mathbf{g}_1(t), \mathbf{g}_2(t)] = \left[\frac{\partial x}{\partial k_1}, \frac{\partial x}{\partial k_2} \right] = \begin{bmatrix} \mathbf{G}_{11}(t) & \mathbf{G}_{12}(t) \\ \mathbf{G}_{21}(t) & \mathbf{G}_{22}(t) \end{bmatrix} = \begin{bmatrix} \frac{\partial x_1}{\partial k_1} & \frac{\partial x_1}{\partial k_2} \\ \frac{\partial x_2}{\partial k_1} & \frac{\partial x_2}{\partial k_2} \end{bmatrix} \quad (4.4)$$

Οι εξισώσεις (2.21) τότε γίνονται:

$$\frac{d\mathbf{g}_1(t)}{dt} = \left(\frac{\partial f^T}{\partial x} \right)^T \mathbf{g}_1(t) + \left(\frac{\partial f}{\partial k_1} \right); \quad \mathbf{g}_1(t_0)=0 \quad (4.5a)$$

$$\frac{d\mathbf{g}_2(t)}{dt} = \left(\frac{\partial f^T}{\partial x} \right)^T \mathbf{g}_2(t) + \left(\frac{\partial f}{\partial k_2} \right); \quad \mathbf{g}_2(t_0)=0 \quad (4.5b)$$

Επομένως:

$$\begin{bmatrix} \frac{d\mathbf{G}_{11}}{dt} \\ \frac{d\mathbf{G}_{21}}{dt} \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix} \begin{bmatrix} \mathbf{G}_{11} \\ \mathbf{G}_{21} \end{bmatrix} + \begin{bmatrix} \frac{\partial f_1}{\partial k_1} \\ \frac{\partial f_2}{\partial k_1} \end{bmatrix}; \quad \mathbf{G}_{11}(t_0)=0, \mathbf{G}_{21}(t_0)=0 \quad (4.6a)$$

$$\begin{bmatrix} \frac{d\mathbf{G}_{12}}{dt} \\ \frac{d\mathbf{G}_{22}}{dt} \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix} \begin{bmatrix} \mathbf{G}_{12} \\ \mathbf{G}_{22} \end{bmatrix} + \begin{bmatrix} \frac{\partial f_1}{\partial k_2} \\ \frac{\partial f_2}{\partial k_2} \end{bmatrix}; \quad \mathbf{G}_{21}(t_0)=0, \mathbf{G}_{22}(t_0)=0 \quad (4.6b)$$

Τέλος:

$$\frac{d\mathbf{G}_{11}}{dt} = \left(\frac{\partial f_1}{\partial x_1} \right) \mathbf{G}_{11} + \left(\frac{\partial f_1}{\partial x_2} \right) \mathbf{G}_{21} + \frac{\partial f_1}{\partial k_1}; \quad \mathbf{G}_{11}(0) = 0 \quad (4.7a)$$

$$\frac{d\mathbf{G}_{21}}{dt} = \left(\frac{\partial f_2}{\partial x_1} \right) \mathbf{G}_{11} + \left(\frac{\partial f_2}{\partial x_2} \right) \mathbf{G}_{21} + \frac{\partial f_2}{\partial k_1}; \quad \mathbf{G}_{21}(0) = 0 \quad (4.7b)$$

$$\frac{d\mathbf{G}_{12}}{dt} = \left(\frac{\partial f_1}{\partial x_1} \right) \mathbf{G}_{12} + \left(\frac{\partial f_1}{\partial x_2} \right) \mathbf{G}_{22} + \frac{\partial f_1}{\partial k_2}; \quad \mathbf{G}_{12}(0) = 0 \quad (4.7c)$$

$$\frac{d\mathbf{G}_{22}}{dt} = \left(\frac{\partial f_2}{\partial x_1} \right) \mathbf{G}_{12} + \left(\frac{\partial f_2}{\partial x_2} \right) \mathbf{G}_{22} + \frac{\partial f_2}{\partial k_2}; \quad \mathbf{G}_{22}(0) = 0 \quad (4.7d)$$

όπου

$$\left(\frac{\partial f_1}{\partial x_1} \right) = -k_1 \left(2x_1 + \frac{2x_2}{3K_1} \right) - k_2 \left(x_2 - \frac{1}{9K_2} (4x_1 - 4 + 5x_2) \right) \quad (4.8a)$$

$$\left(\frac{\partial f_1}{\partial x_2}\right) = -k_1 \left(\frac{1}{3K_1}(2x_2 + 2x_1 - 2)\right) - k_2 \left(x_1 - \frac{1}{9K_2}(5x_1 - 5 + 4x_2)\right) \quad (4.8b)$$

$$\left(\frac{\partial f_2}{\partial x_1}\right) = \frac{k_1}{2} \left(2x_1 + \frac{2x_2}{3K_1}\right) - k_2 \left(x_2 - \frac{1}{9K_2}(4x_1 - 4 + 5x_2)\right) \quad (4.8c)$$

$$\left(\frac{\partial f_2}{\partial x_2}\right) = \frac{k_1}{2} \left(\frac{1}{3K_1}(2x_2 + 2x_1 - 2)\right) - k_2 \left(x_1 - \frac{1}{9K_2}(5x_1 + 4x_2 - 5)\right) \quad (4.8d)$$

$$\left(\frac{\partial f_1}{\partial k_1}\right) = - \left[x_1^2 + \frac{1}{3K_1}(x_2^2 + 2x_1x_2 - 2x_2) \right] \quad (4.8e)$$

$$\left(\frac{\partial f_2}{\partial k_1}\right) = \frac{1}{2} \left[x_1^2 + \frac{1}{3K_1}(x_2^2 + 2x_1x_2 - 2x_2) \right] \quad (4.8f)$$

$$\left(\frac{\partial f_1}{\partial k_2}\right) = - \left[x_1x_2 - \frac{1}{9K_2}(2x_1^2 - 4x_1 + 5x_1x_2 - 5x_2 + 2x_2^2 + 2) \right] \quad (4.8g)$$

$$\left(\frac{\partial f_2}{\partial k_2}\right) = - \left[x_1x_2 - \frac{1}{9K_2}(2x_1^2 - 4x_1 + 5x_1x_2 - 5x_2 + 2x_2^2 + 2) \right] \quad (4.8h)$$

Οι τέσσερις εξισώσεις ευαισθησίας (4.7) λύνονται ταυτόχρονα με τις δύο εξισώσεις κατάστασης (4.3). Η ολοκλήρωση αυτών των έξι $[=n*(p+1)=2*(2+1)]$ εξισώσεων αποδίδουν τα $\mathbf{x}(t)$ και $\mathbf{G}(t)$, που χρησιμοποιούνται για τη ρύθμιση του πίνακα \mathbf{A} και του διανύσματος \mathbf{b} σε κάθε προσέγγιση (iteration) της μεθόδου Gauss-Newton.

Η ODE, η οποία ικανοποιείται για ένα συγκεκριμένο στοιχείο, \mathbf{G}_{ij} , του πίνακα ευαισθησίας ($n \times p$) διαστάσεων, μπορεί να γραφτεί άμεσα έτσι:

$$\frac{d\mathbf{G}_{ij}}{dt} \equiv \frac{d}{dt} \left(\frac{\partial x_i}{\partial k_j} \right) = \sum_{k=1}^n \left(\frac{\partial f_i}{\partial x_k} \right) \left(\frac{\partial x_k}{\partial k_j} \right) + \frac{\partial f_i}{\partial k_j} \equiv \sum_{k=1}^n \left(\frac{\partial f_i}{\partial x_k} \right) \mathbf{G}_{kj} + \frac{\partial f_i}{\partial k_j} \quad (4.9)$$

4.1.2. Αποτελέσματα

Σύμφωνα με τα παραπάνω και έπειτα από την εκτέλεση του αλγορίθμου στη Matlab, για $k_1=k_2=10000$, προκύπτουν τα εξής αποτελέσματα:

Η μέθοδος Gauss-Newton συγκλίνει σε 9 iterations και η παράμετρος του Marquardt παραμένει 0 κατά τη διάρκεια της εκτέλεσης. Στον πίνακα 4.2 είναι εμφανές ότι τα πειραματικά δεδομένα (data) είναι σχεδόν πανομοιότυπα με τα αποτελέσματα του μοντέλου (model), επόμενος ο αλγόριθμος είναι ακριβής. Στον πίνακα 4.3 φαίνεται η ελαχιστοποίηση της αντικειμενικής συνάρτησης και η μετατροπή των k_1 , k_2 ανά iteration. Τα αποτελέσματα συμφωνούν με τα ήδη υπάρχοντα αποτελέσματα των Seinfeld and Gavalas (1970), Hougen and Watson (1948) και Kalogerakis and Luus (1983), από τους οποίους χρησιμοποιήθηκαν διαφορετικές μέθοδοι και αλγόριθμοι σε άλλες πλατφόρμες.

Πίνακας 4.2. Πυρολυτική αφυδρογόνωση του βενζολίου: Πειραματικά δεδομένα (data) και εκτιμώμενες παράμετροι του μοντέλου (model)

<i>Reciprocal Space Velocity (t) × 10⁴</i>	$x_1(data)$	$x_1(model)$	$x_2(data)$	$x_2(model)$
5.63	0.828	0.8283	0.0737	0.0738
11.32	0.704	0.7054	0.113	0.1122
16.97	0.622	0.6215	0.1322	0.1313
22.62	0.565	0.5643	0.1400	0.1407
34.0	0.499	0.4990	0.1468	0.1473
39.7	0.482	0.4811	0.1477	0.1482
45.2	0.470	0.4693	0.1477	0.1485
169.7	0.443	0.4433	0.1476	0.1477

Πίνακας 4.3. Πυρολυτική αφυδρογόνωση του βενζολίου: Ανά προσέγγιση (iteration), τιμή της αντικειμενικής συνάρτησης και των εκτιμώμενων παραμέτρων k_1 , k_2

<i>Iteration</i>	<i>LS Objective Function</i>	k_1	k_2
1	0.5238	10000	10000
2	0.5219	4899.9	9515.2
3	0.5086	4233	1717.3
4	0.4644	1918.7	1663.5
5	0.2493	680.5209	1000.3
6	0.0865	435.4234	533.9674
7	0.0150	343.4356	373.0364
8	0.0025	354.3214	399.2941
9	0.0024	354.5886	400.3349

Εικόνα 4.1. Τελικό αποτέλεσμα του αλγορίθμου στην κονσόλα της πλατφόρμας MATLAB (Εφαρμογή 1)

```
The parameter estimation problem has converged successfully.

Standard Error for k(1): 0.033253
Relative Standard Error: 0.009378
Confidence interval for k(1):
354.518415<=k(1)<=354.661056
Interval Ta*Sk(1)=0.071320

Standard Error for k(2): 0.157469
Relative Standard Error: 0.039335
Confidence interval for k(2):
399.994147<=k(2)<=400.669621
Interval Ta*Sk(2)=0.337737

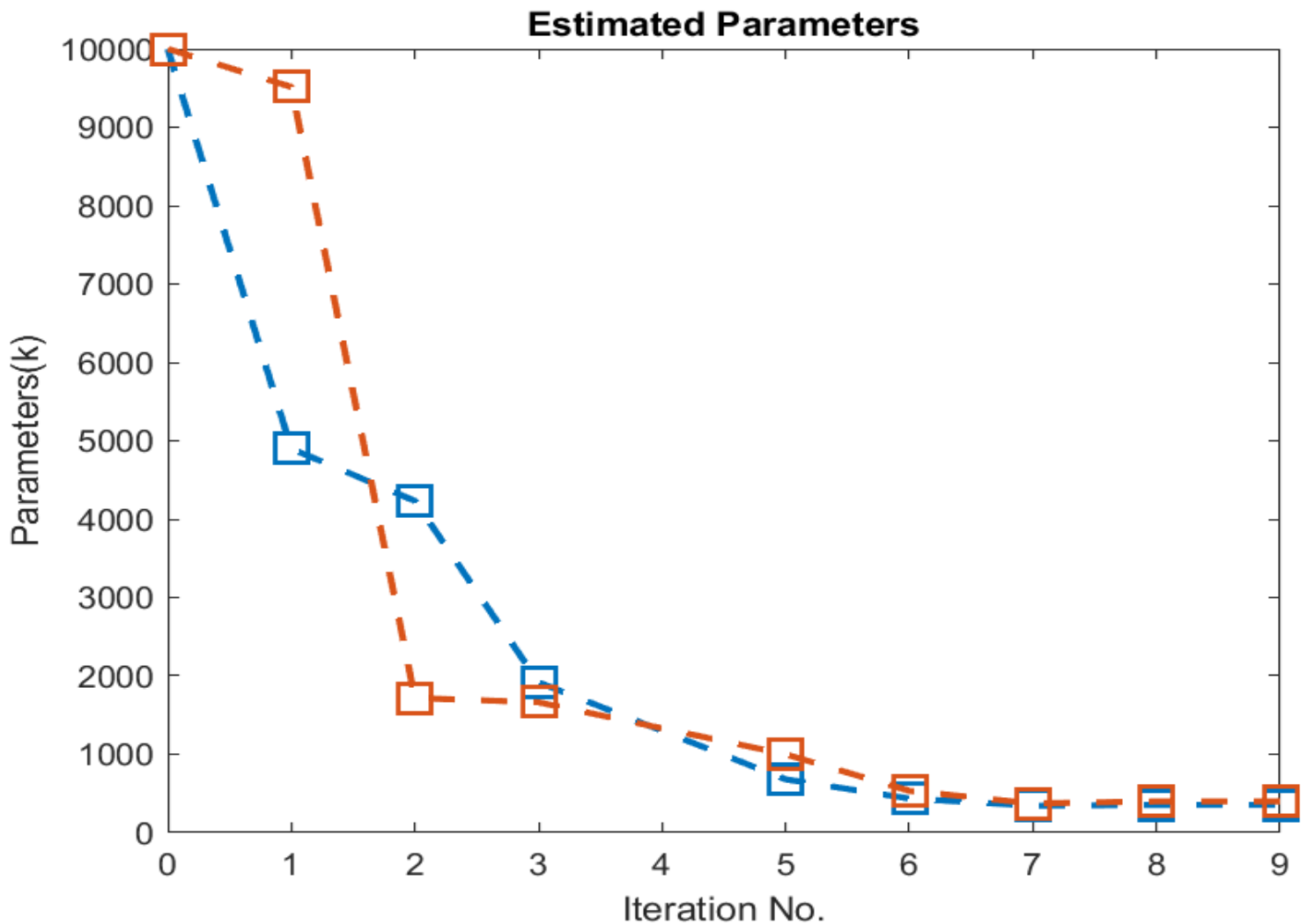
Variance Se^2: 0.000170

Condition number of matrix A: 23.449974
```

Όσον αφορά τα στατιστικά αποτελέσματα, όταν ο αλγόριθμος ολοκληρωθεί, παρουσιάζονται οι τυπικές αποκλίσεις και τα διαστήματα εμπιστοσύνης των παραμέτρων, το σχετικό τυπικό σφάλμα τους, ο αριθμός κατάστασης του πίνακα **A** και η διασπορά. Φαίνονται παραπάνω, στην εικόνα 4.1.

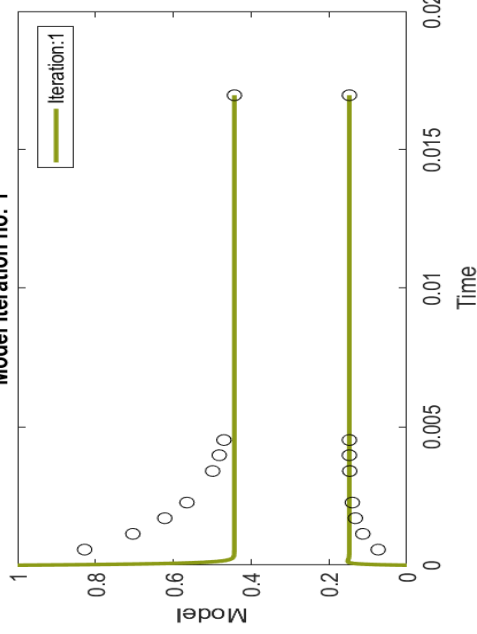
Παρακάτω παρουσιάζονται τα γραφικά, όπως προκύπτουν από τη MATLAB:

Εικόνα 4.2. Τιμές των εκτιμώμενων παραμέτρων k_1 , k_2 ανά iteration. Την k_1 απεικονίζει η μπλε γραμμή και την k_2 απεικονίζει η κόκκινη γραμμή.

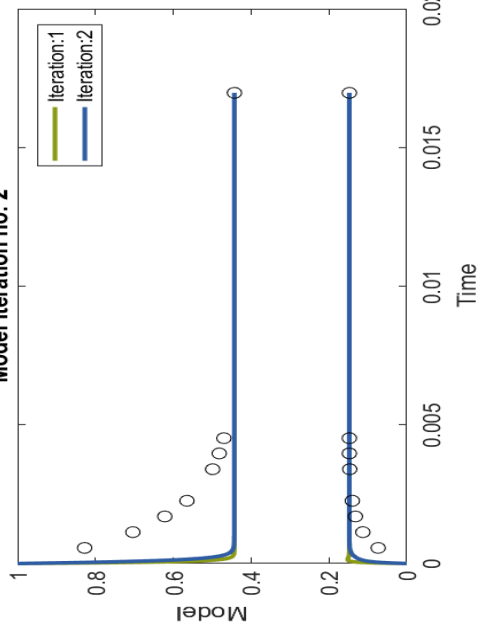


Εικόνα 4.3. Πορεία σύγκλισης του μοντέλου σε σχέση με τα πειραματικά δεδομένα. Οι συνεχείς γραμμές απεικονίζουν τα εκτιμώμενα μοντέλα x_1 και x_2 και οι κύκλοι τα πειραματικά δεδομένα τους.

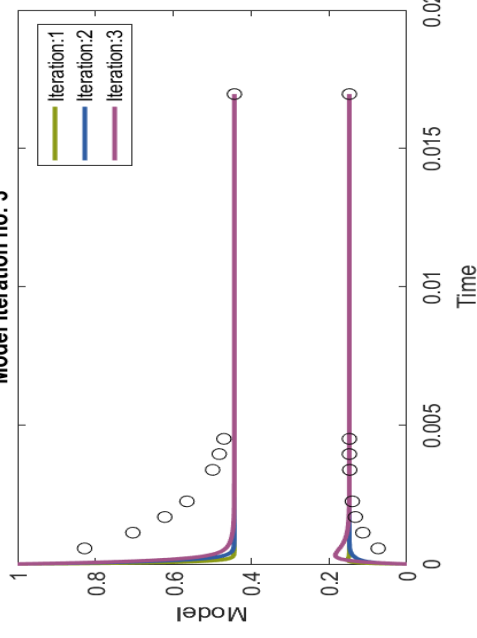
Model Iteration no: 1



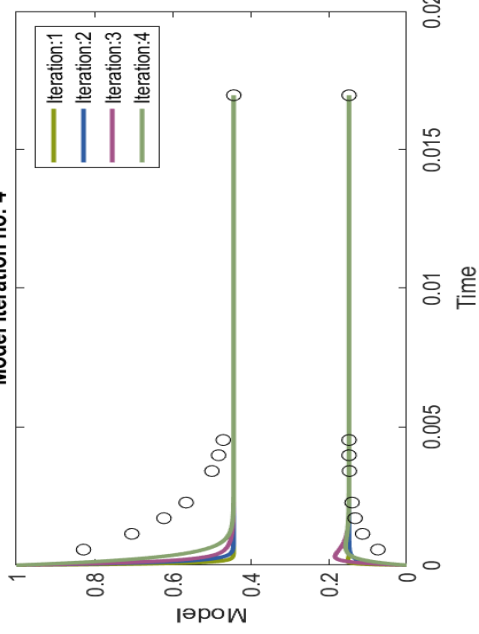
Model Iteration no: 2



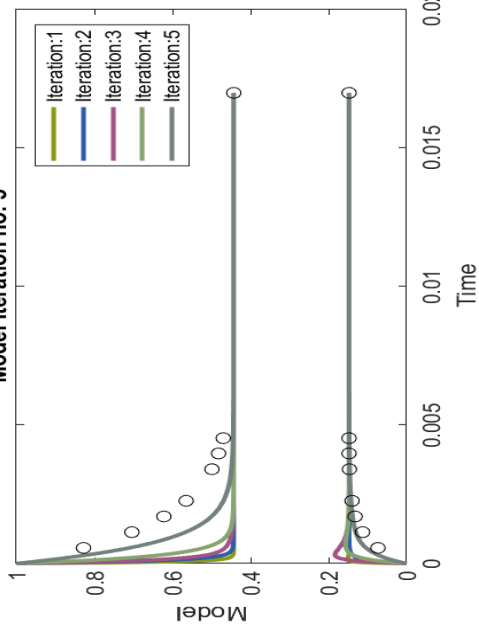
Model Iteration no: 3



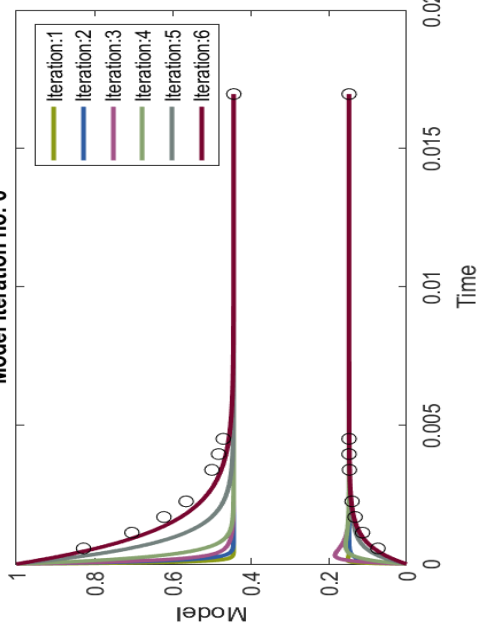
Model Iteration no: 4



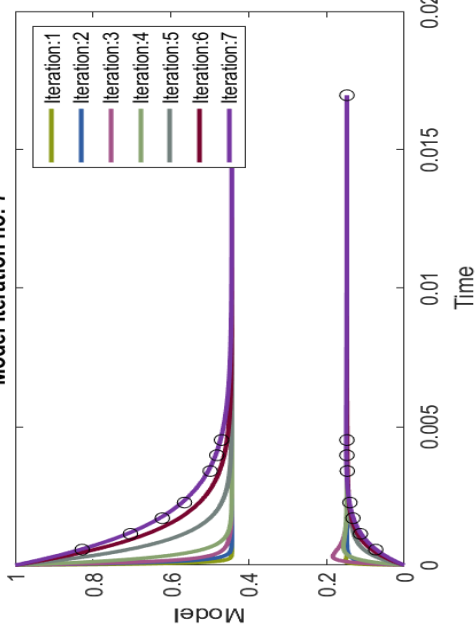
Model Iteration no: 5



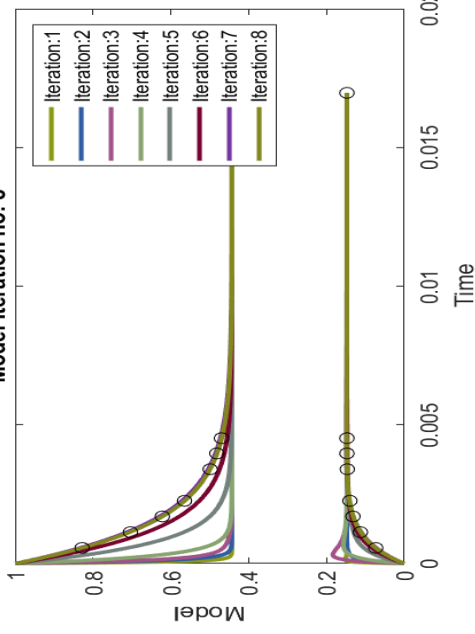
Model Iteration no: 6



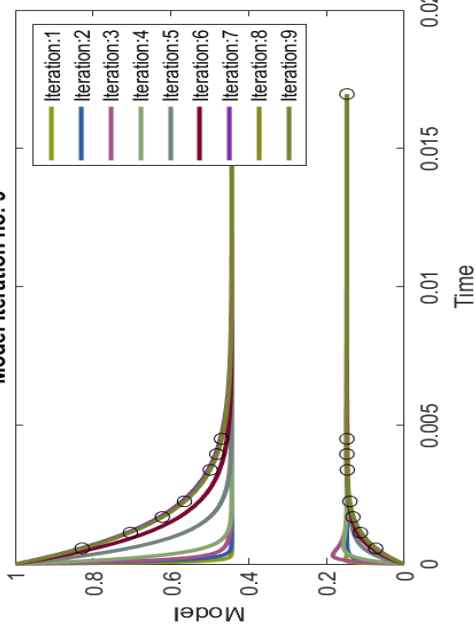
Model Iteration no: 7



Model Iteration no: 8



Model Iteration no: 9



4.2. Εφαρμογή 2: Καταλυτική Υδρογόνωση του 3-Hydroxypropanal (HPA) σε 1,3-Propanediol (PD)

4.2.1. Περιγραφή

Η υδρογόνωση της 3-υδροξυπροπανάλης (HPA) σε 1,3-προπανοδιόλη (PD) με σκόνη καταλύτη Ni / SiO₂ / Al₂O₃ μελετήθηκε από την ομάδα του καθηγητή Hoffman στο Πανεπιστήμιο Friedrich-Alexander στο Erlangen της Γερμανίας (Zhu et al., 1997) . Το PD είναι ένα δυνητικά ελκυστικό μονομερές για πολυμερή όπως το τερεφθαλικό πολυπροπυλένιο. Χρησιμοποιήθηκε αεροστεγής λέβητας βρασμού ανάδευσης κατά παρτίδες. Τα πειραματικά δεδομένα αποτελούνται από μετρήσεις της συγκέντρωσης HPA και PD (CHPA, CPD) σε σχέση με το χρόνο σε διάφορες θερμοκρασίες και πιέσεις. Η ίδια ομάδα πρότεινε και μαθηματικό μοντέλο που περιγράφει τους ρυθμούς κατανάλωσης HPA, το σχηματισμό PD, καθώς και το σχηματισμό ακρολεΐνης (Ac). Το μοντέλο είναι το εξής:

$$\frac{dC_{HPA}}{dt} = -[r_1 + r_2]C_k - [r_3 + r_4 - r_{-3}] \quad (4.10a)$$

$$\frac{dC_{PD}}{dt} = [r_1 - r_2]C_k \quad (4.10b)$$

$$\frac{dC_{Ac}}{dt} = r_3 - r_4 - r_{-3} \quad (4.10c)$$

όπου $C_k = 10 \text{ g/L}$, συγκέντρωση καταλύτη.

Ρυθμοί Αντίδρασης:

$$r_1 = \frac{k_1 P C_{HPA}}{H \left[1 + \left(\frac{K_1 P}{H} \right)^{0.5} + K_2 C_{HPA} \right]^3} \quad (4.11a)$$

$$r_2 = \frac{k_2 C_{PD} C_{HPA}}{1 + \left(\frac{K_1 P}{H} \right)^{0.5} + K_2 C_{HPA}} \quad (4.11b)$$

$$r_3 = k_3 C_{HPA} \quad (4.11c)$$

$$r_{-3} = k_{-3} C_{Ac} \quad (4.11d)$$

$$r_4 = k_4 C_{Ac} C_{HPA} \quad (4.11e)$$

$k_{1...4}$: σταθερές ρυθμού (L / (mol min g))

K_1, K_2 : σταθερές ισορροπίας προσρόφησης (L / mol) για τα H_2 και HPA αντίστοιχα

P: πίεση υδρογόνου (MPa) στον αντιδραστήρα

H: σταθερά του νόμου του Henry, $H=1379$ (L bar / mol) στα 298 K

Οι παράμετροι $k_1, k_2, k_3, k_{-3}, k_4, K_1, K_2$ προσδιορίζονται από τις συγκεντρώσεις των HPA και PD που μετρώνται.

Πίνακας 4.4. Υδρογόνωση του HPA: Πειραματικά δεδομένα συλλεγμένα σε θερμοκρασία 318K και πιέσεις 2.6, 4.0 και 5.15 MPa

Pressure (MPa)	Time (min)	C_{HPA} (mol/L)	C_{PD} (mol/L)
2.6	10	1.37395	0.0
	20	1.25821	0.0197109
	30	1.18707	0.0642576
	40	1.13292	0.136399
	50	1.03556	0.238633
	60	0.961339	0.304599
	80	0.734436	0.492378
	100	0.564551	0.732326
	120	0.374385	0.887254
	140	0.214799	1.04284
	160	0.100976	1.17306
	180	0.0364192	1.25769
	200	0.00530892	1.26032
4.0	10	1.3295	0.00262812
	20	1.31157	0.0525624
	30	1.22828	0.120736
	40	1.087	0.241393
	50	0.994539	0.384888

	60	0.811825	0.4682
	80	0.600962	0.773193
	100	0.386302	0.990802
	120	0.204222	1.14954
	140	0.0782304	1.28
	160	0.0277708	1.29
	180	0.00316296	1.30
	200	0.00210864	1.30
5.15	10	1.36324	0.00262812
	20	1.25882	0.0700394
	30	1.17918	0.184363
	40	0.972102	0.354008
	50	0.825203	0.469777
	60	0.697109	0.607359
	80	0.421451	0.852431
	100	0.232296	1.03535
	120	0.128095	1.16413
	140	0.0289817	1.30053
	160	0.00962368	1.31971

Προστίθενται τα παρακάτω διανύσματα:

$$\mathbf{x} = [x_1, x_2, x_3]^T = [C_{HPA}, C_{PD}, C_{Ac}]^T$$

$$\mathbf{k} = [k_1, k_2, k_3, k_4, k_5, k_6, k_7]^T = [k_1, k_2, k_3, k_{-3}, k_4, K_1, K_2]^T$$

$$\mathbf{y} = [y_1, y_2]^T = [C_{HPA}, C_{PD}]^T$$

και το μοντέλο ODE γίνεται:

$$\frac{dx_1}{dt} = f_1(x_1, x_2, x_3; k_1, k_2, \dots, k_7; u_1, u_2) \quad (4.12a)$$

$$\frac{dx_2}{dt} = f_2(x_1, x_2, x_3; k_1, k_2, \dots, k_7; u_1, u_2) \quad (4.12b)$$

$$\frac{dx_3}{dt} = f_3(x_1, x_2, x_3; k_1, k_2, \dots, k_7; u_1, u_2) \quad (4.12c)$$

$$\text{Με πίνακα παρατήρησης } \mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad (4.13)$$

Στις εξισώσεις (4.12), το u_1 απεικονίζει τη συγκεντρώση του καταλύτη που είναι παρών στον αντιδραστήρα (C_k) και το u_2 την πίεση του υδρογόνου (P). Όσον αφορά τον υπολογισμό των παραμέτρων, και οι δύο αυτές μεταβλητές είναι γνωστές επ'ακριβώς, επιλέγονται από το χρήστη και έχουν μεγάλη ισχύ, όσον αφορά την ποιότητα των υπολογισμών παραμέτρων. Το μοντέλο μετατρέπεται:

$$\frac{dx_1}{dt} = -u_1(r_1 + r_2) - (k_3x_1 + k_5x_3x_1 - k_4x_3) \quad (4.14a)$$

$$\frac{dx_2}{dt} = u_1(r_1 - r_2) \quad (4.14b)$$

$$\frac{dx_3}{dt} = (k_3x_1 - k_5x_3x_1 - k_4x_3) \quad (4.14c)$$

όπου

$$r_1 = \frac{k_1u_2x_1}{H \left[1 + \left(\frac{k_6u_2}{H} \right)^{0.5} + k_7x_1 \right]^3} \quad (4.15a)$$

$$r_2 = \frac{k_2x_2x_1}{1 + \left(\frac{k_6u_2}{H} \right)^{0.5} + k_7x_1} \quad (4.15b)$$

$$r_3 = k_3x_1 \quad (4.15c)$$

$$r_{-3} = k_4x_3 \quad (4.15d)$$

$$r_4 = k_5x_3x_1 \quad (4.15e)$$

Ο πίνακας ευαισθησίας, $\mathbf{G}(t)$, είναι διαστάσεων 3×7 με στοιχεία:

$$\mathbf{G}(t) = [\mathbf{g}_1(t), \mathbf{g}_2(t), \dots, \mathbf{g}_7(t)] = \left[\left(\frac{\partial \mathbf{x}}{\partial k_1} \right), \left(\frac{\partial \mathbf{x}}{\partial k_2} \right), \dots, \left(\frac{\partial \mathbf{x}}{\partial k_7} \right) \right] \quad (4.16a)$$

$$\mathbf{G}(t) = \begin{bmatrix} \mathbf{G}_{11}(t) & \cdot & \cdot & \cdot & \mathbf{G}_{17}(t) \\ \mathbf{G}_{21}(t) & \cdot & \cdot & \cdot & \mathbf{G}_{27}(t) \\ \mathbf{G}_{31}(t) & \cdot & \cdot & \cdot & \mathbf{G}_{37}(t) \end{bmatrix} = \begin{bmatrix} \left(\frac{\partial x_1}{\partial k_1} \right) & \cdot & \cdot & \cdot & \left(\frac{\partial x_1}{\partial k_7} \right) \\ \left(\frac{\partial x_2}{\partial k_1} \right) & \cdot & \cdot & \cdot & \left(\frac{\partial x_2}{\partial k_7} \right) \\ \left(\frac{\partial x_3}{\partial k_1} \right) & \cdot & \cdot & \cdot & \left(\frac{\partial x_3}{\partial k_7} \right) \end{bmatrix} \quad (4.16b)$$

και οι εξισώσεις (2.21) γίνονται:

$$\frac{d\mathbf{g}_1(t)}{dt} = \left(\frac{\partial \mathbf{f}^T}{\partial \mathbf{x}} \right)^T \mathbf{g}_1(t) + \left(\frac{\partial \mathbf{f}}{\partial k_1} \right); \quad \mathbf{g}_1(t_0)=0 \quad (4.17a)$$

$$\frac{d\mathbf{g}_2(t)}{dt} = \left(\frac{\partial \mathbf{f}^T}{\partial \mathbf{x}} \right)^T \mathbf{g}_2(t) + \left(\frac{\partial \mathbf{f}}{\partial k_2} \right); \quad \mathbf{g}_2(t_0)=0 \quad (4.17b)$$

·
·
·

$$\frac{d\mathbf{g}_7(t)}{dt} = \left(\frac{\partial \mathbf{f}^T}{\partial \mathbf{x}} \right)^T \mathbf{g}_7(t) + \left(\frac{\partial \mathbf{f}}{\partial k_7} \right); \quad \mathbf{g}_7(t_0)=0 \quad (4.17c)$$

όπου

$$\left(\frac{\partial \mathbf{f}^T}{\partial \mathbf{x}} \right)^T = \begin{bmatrix} \left(\frac{\partial f_1}{\partial x_1} \right) & \left(\frac{\partial f_1}{\partial x_2} \right) & \left(\frac{\partial f_1}{\partial x_3} \right) \\ \left(\frac{\partial f_2}{\partial x_1} \right) & \left(\frac{\partial f_2}{\partial x_2} \right) & \left(\frac{\partial f_2}{\partial x_3} \right) \\ \left(\frac{\partial f_3}{\partial x_1} \right) & \left(\frac{\partial f_3}{\partial x_2} \right) & \left(\frac{\partial f_3}{\partial x_3} \right) \end{bmatrix} \quad (4.18a)$$

και

$$\left(\frac{\partial \mathbf{f}}{\partial k_j}\right) = \begin{bmatrix} \left(\frac{\partial f_1}{\partial k_j}\right) \\ \left(\frac{\partial f_2}{\partial k_j}\right) \\ \left(\frac{\partial f_3}{\partial k_j}\right) \end{bmatrix} ; \quad j=1,2,\dots,7 \quad (4.18b)$$

Επομένως:

$$\frac{dG_{11}}{dt} = \left(\frac{\partial f_1}{\partial x_1}\right)G_{11} + \left(\frac{\partial f_1}{\partial x_2}\right)G_{21} + \left(\frac{\partial f_1}{\partial x_3}\right)G_{31} + \frac{\partial f_1}{\partial k_1} ; \quad G_{11}(0) = 0$$

$$\frac{dG_{21}}{dt} = \left(\frac{\partial f_2}{\partial x_1}\right)G_{11} + \left(\frac{\partial f_2}{\partial x_2}\right)G_{21} + \left(\frac{\partial f_2}{\partial x_3}\right)G_{31} + \frac{\partial f_2}{\partial k_1} ; \quad G_{21}(0) = 0$$

.

.

.

$$\frac{dG_{17}}{dt} = \left(\frac{\partial f_1}{\partial x_1}\right)G_{17} + \left(\frac{\partial f_1}{\partial x_2}\right)G_{27} + \left(\frac{\partial f_1}{\partial x_3}\right)G_{37} + \frac{\partial f_1}{\partial k_7} ; \quad G_{17}(0) = 0$$

$$\frac{dG_{27}}{dt} = \left(\frac{\partial f_2}{\partial x_1}\right)G_{17} + \left(\frac{\partial f_2}{\partial x_2}\right)G_{27} + \left(\frac{\partial f_2}{\partial x_3}\right)G_{37} + \frac{\partial f_2}{\partial k_7} ; \quad G_{27}(0) = 0$$

$$\frac{dG_{37}}{dt} = \left(\frac{\partial f_3}{\partial x_1}\right)G_{17} + \left(\frac{\partial f_3}{\partial x_2}\right)G_{27} + \left(\frac{\partial f_3}{\partial x_3}\right)G_{37} + \frac{\partial f_3}{\partial k_7} ; \quad G_{37}(0) = 0$$

Οι παραπάνω εξισώσεις αποτελούν συνολικά την (4.19).

Οι μερικές παράγωγοι σε σχέση με τις μεταβλητές κατάστασης στην εξίσωση (4.18a) που απαιτούνται στις παραπάνω ODE, δίδονται στη συνέχεια:

$$\left(\frac{\partial f_1}{\partial x_1}\right) = -u_1 \left(\frac{\partial r_1}{\partial x_1} + \frac{\partial r_2}{\partial x_1}\right) - k_3 - k_5 x_3 \quad (4.20a)$$

$$\left(\frac{\partial f_1}{\partial x_2}\right) = -u_1 \left(\frac{\partial r_2}{\partial x_2}\right) = -u_1 \frac{r_2}{x_2} \quad (4.20b)$$

$$\left(\frac{\partial f_1}{\partial x_3}\right) = k_4 - k_5 x_1 \quad (4.20c)$$

$$\left(\frac{\partial f_2}{\partial x_1}\right) = u_1 \left(\frac{\partial r_1}{\partial x_1} - \frac{\partial r_2}{\partial x_1}\right) \quad (4.20d)$$

$$\left(\frac{\partial f_2}{\partial x_2}\right) = -u_1 \left(\frac{\partial r_2}{\partial x_2}\right) = -u_1 \frac{r_2}{x_2} \quad (4.20e)$$

$$\left(\frac{\partial f_2}{\partial x_3}\right) = 0 \quad (4.20f)$$

$$\left(\frac{\partial f_3}{\partial x_1}\right) = k_3 - k_5 x_3 \quad (4.20g)$$

$$\left(\frac{\partial f_3}{\partial x_2}\right) = 0 \quad (4.20h)$$

$$\left(\frac{\partial f_3}{\partial x_3}\right) = -k_5 x_1 - k_4 \quad (4.20i)$$

Οι μερικές παράγωγοι σε σχέση με τις μεταβλητές κατάστασης στην εξίσωση (4.18b) που απαιτούνται στις παραπάνω ODE, δίδονται στη συνέχεια:

$$\left(\frac{\partial \mathbf{f}}{\partial k_1}\right) = \begin{bmatrix} \left(\frac{\partial f_1}{\partial k_1}\right) \\ \left(\frac{\partial f_2}{\partial k_1}\right) \\ \left(\frac{\partial f_3}{\partial k_1}\right) \end{bmatrix} = \begin{bmatrix} -u_1 \left(\frac{\partial r_1}{\partial k_1}\right) \\ u_1 \left(\frac{\partial r_1}{\partial k_1}\right) \\ 0 \end{bmatrix} \quad (4.21a)$$

$$\left(\frac{\partial \mathbf{f}}{\partial k_2}\right) = \begin{bmatrix} \left(\frac{\partial f_1}{\partial k_2}\right) \\ \left(\frac{\partial f_2}{\partial k_2}\right) \\ \left(\frac{\partial f_3}{\partial k_2}\right) \end{bmatrix} = \begin{bmatrix} -u_1 \left(\frac{\partial r_2}{\partial k_2}\right) \\ -u_1 \left(\frac{\partial r_2}{\partial k_2}\right) \\ 0 \end{bmatrix} \quad (4.21b)$$

$$\left(\frac{\partial \mathbf{f}}{\partial k_3}\right) = \begin{bmatrix} \left(\frac{\partial f_1}{\partial k_3}\right) \\ \left(\frac{\partial f_2}{\partial k_3}\right) \\ \left(\frac{\partial f_3}{\partial k_3}\right) \end{bmatrix} = \begin{bmatrix} -x_1 \\ 0 \\ x_1 \end{bmatrix} \quad (4.21c)$$

$$\left(\frac{\partial \mathbf{f}}{\partial k_4}\right) = \begin{bmatrix} \left(\frac{\partial f_1}{\partial k_4}\right) \\ \left(\frac{\partial f_2}{\partial k_4}\right) \\ \left(\frac{\partial f_3}{\partial k_4}\right) \end{bmatrix} = \begin{bmatrix} x_3 \\ 0 \\ -x_3 \end{bmatrix} \quad (4.21d)$$

$$\left(\frac{\partial \mathbf{f}}{\partial k_5}\right) = \begin{bmatrix} \left(\frac{\partial f_1}{\partial k_5}\right) \\ \left(\frac{\partial f_2}{\partial k_5}\right) \\ \left(\frac{\partial f_3}{\partial k_5}\right) \end{bmatrix} = \begin{bmatrix} -x_3 x_1 \\ 0 \\ -x_3 x_1 \end{bmatrix} \quad (4.21e)$$

$$\begin{pmatrix} \frac{\partial \mathbf{f}}{\partial k_6} \end{pmatrix} = \begin{pmatrix} \left(\frac{\partial f_1}{\partial k_6} \right) \\ \left(\frac{\partial f_2}{\partial k_6} \right) \\ \left(\frac{\partial f_3}{\partial k_6} \right) \end{pmatrix} = \begin{pmatrix} -u_1 \left(\frac{\partial r_1}{\partial k_6} + \frac{\partial r_2}{\partial k_6} \right) \\ u_1 \left(\frac{\partial r_1}{\partial k_6} - \frac{\partial r_2}{\partial k_6} \right) \\ 0 \end{pmatrix} \quad (4.21f)$$

$$\begin{pmatrix} \frac{\partial \mathbf{f}}{\partial k_7} \end{pmatrix} = \begin{pmatrix} \left(\frac{\partial f_1}{\partial k_7} \right) \\ \left(\frac{\partial f_2}{\partial k_7} \right) \\ \left(\frac{\partial f_3}{\partial k_7} \right) \end{pmatrix} = \begin{pmatrix} -u_1 \left(\frac{\partial r_1}{\partial k_7} + \frac{\partial r_2}{\partial k_7} \right) \\ u_1 \left(\frac{\partial r_1}{\partial k_7} - \frac{\partial r_2}{\partial k_7} \right) \\ 0 \end{pmatrix} \quad (4.21g)$$

Οι 21 εξισώσεις (που δίδονται ως Εξίσωση (4.19)) πρέπει να επιλυθούν ταυτόχρονα με τις τρεις εξισώσεις κατάστασης (4.15). Η ολοκλήρωση αυτών των 24 εξισώσεων αποδίδει τα $\mathbf{x}(t)$ και $\mathbf{G}(t)$ που χρησιμοποιούνται στη ρύθμιση του πίνακα \mathbf{A} και του διανύσματος \mathbf{b} σε κάθε iteration της μεθόδου Gauss-Newton.

4.2.2. Αποτελέσματα

Για την παραπάνω εφαρμογή, τα αποτελέσματα του αλγορίθμου είναι διαφορετικά σε δραστικό βαθμό. Αυτό συμβαίνει για δύο λόγους, το σύστημα των ODEs είναι δύσκαμπτο και σε κακή κατάσταση. Δυσκαμψία (stiffness, Chapra and Canale, 2018) είναι ένα ειδικό πρόβλημα που μπορεί να εμφανιστεί στη λύση μιας συνήθους διαφορικής εξίσωσης. Ένα δύσκαμπτο σύστημα είναι αυτό το οποίο περιλαμβάνει συνιστώσες (όρους) που μεταβάλλονται πολύ γρήγορα αλλά ταυτόχρονα και συνιστώσες οι οποίες μεταβάλλονται πολύ αργά. Σε πολλές περιπτώσεις οι ταχέως μεταβαλλόμενοι όροι είναι εφήμεροι και υφίστανται απόσβεση (δηλαδή γίνονται αμελητέοι) πολύ γρήγορα. Μετά την εξάλειψη τους, η λύση καθορίζεται από τους παράγοντες που μεταβάλλονται αργά. Αν και τα μεταβατικά φαινόμενα λαμβάνουν χώρα για ένα πολύ μικρό διάστημα της περιοχής ολοκλήρωσης, μπορούν σε μερικές περιπτώσεις να καθορίσουν το χρονικό βήμα της συνολικής λύσης.

Παράλληλα, χρησιμοποιώντας ως αρχικές εκτιμήσεις για τις παραμέτρους, σετ παραμέτρων υπολογισμένο με άλλο πρόγραμμα (Englezos and Kalogerakis, 2001, $\mathbf{k}_0=[2.6866$

0.236E-8 0.672E-3 0.126E-5 0.0273 35.56 2.57]), κοντά στο βέλτιστο, παρουσιάζονται τα εξής αποτελέσματα. Στον παρακάτω πίνακα, ο αλγόριθμος εκτελείται για διαφορετικό αριθμό εκτιμώμενων παραμέτρων (από 1 έως 7) με τροποποίηση του Marquardt ή και όχι, για 1 iteration. Οι παράμετροι που πρόκειται να υπολογιστούν πολλαπλασιάζονται με το 1.2, ενώ οι υπόλοιπες παραμένουν ίσες με το παραπάνω σετ που επιλέχθηκε.

Πίνακας 4.5. Αριθμός κατάστασης του \mathbf{A} ($\text{cond}(\mathbf{A})$), ανάλογα από τον αριθμό των εκτιμώμενων παραμέτρων και την ύπαρξη τροποποίησης Marquardt ή όχι.

Εκτιμώμενες Παράμετροι(k)	Marquardt's $\gamma = 0.0$	Marquardt's $\gamma = 10^4$
1	1	1
2	$6.1562 * 10^{12}$	$2.1123 * 10^5$
3	$7.1716 * 10^{13}$	$2.0173 * 10^5$
4	$1.9016 * 10^{14}$	$2.0174 * 10^5$
5	$1.9584 * 10^{14}$	$2.0006 * 10^5$
6	$1.8974 * 10^{17}$	$7.0324 * 10^7$
7	$2.2453 * 10^{17}$	$6.2803 * 10^7$

Όπως φαίνεται από τα αποτελέσματα, το σύστημα βρίσκεται σε κακή κατάσταση, ανεξαρτήτως από τον αριθμό των παραμέτρων που υπολογίζονται. Αν χρησιμοποιηθεί τροποποίηση του Marquardt, ο πίνακας \mathbf{A} φαίνεται ότι βρίσκεται σε καλή κατάσταση, αλλά κάτι τέτοιο δεν ισχύει. Όσον αφορά τα στατιστικά αποτελέσματα, παρουσιάζονται παρακάτω, χρησιμοποιώντας το επιλεγόμενο σετ εκτιμήσεων:

Εικόνα 4.4. Τελικό αποτέλεσμα του αλγορίθμου στην κονσόλα της πλατφόρμας MATLAB (Εφαρμογή 2)

Standard Error for k(1): 0.157389
 Relative Standard Error: 5.858311
 Confidence interval for k(1):
 $2.349033 \leq k(1) \leq 3.024167$
 Interval $Ta \cdot Sk(1) = 0.337567$

Standard Error for k(2): 11.445242
 Relative Standard Error: 484967893353.717773
 Confidence interval for k(2):
 $-24.547603 \leq k(2) \leq 24.547603$
 Interval $Ta \cdot Sk(2) = 24.547603$

Standard Error for k(3): 1.691711
 Relative Standard Error: 251742.737586
 Confidence interval for k(3):
 $-3.627688 \leq k(3) \leq 3.629032$
 Interval $Ta \cdot Sk(3) = 3.628360$

Standard Error for k(4): 11.445242
 Relative Standard Error: 908352515.391897
 Confidence interval for k(4):
 $-24.547601 \leq k(4) \leq 24.547603$
 Interval $Ta \cdot Sk(4) = 24.547602$

Standard Error for k(5): 9.454102
 Relative Standard Error: 34630.408874
 Confidence interval for k(5):
 $-20.249731 \leq k(5) \leq 20.304331$
 Interval $Ta \cdot Sk(5) = 20.277031$

Standard Error for k(6): 0.007268
 Relative Standard Error: 0.020439
 Confidence interval for k(6):
 $35.544411 \leq k(6) \leq 35.575589$
 Interval $Ta \cdot Sk(6) = 0.015589$

Standard Error for k(7): 0.188671
 Relative Standard Error: 7.341289
 Confidence interval for k(7):
 $2.165341 \leq k(7) \leq 2.974659$
 Interval $Ta \cdot Sk(7) = 0.404659$

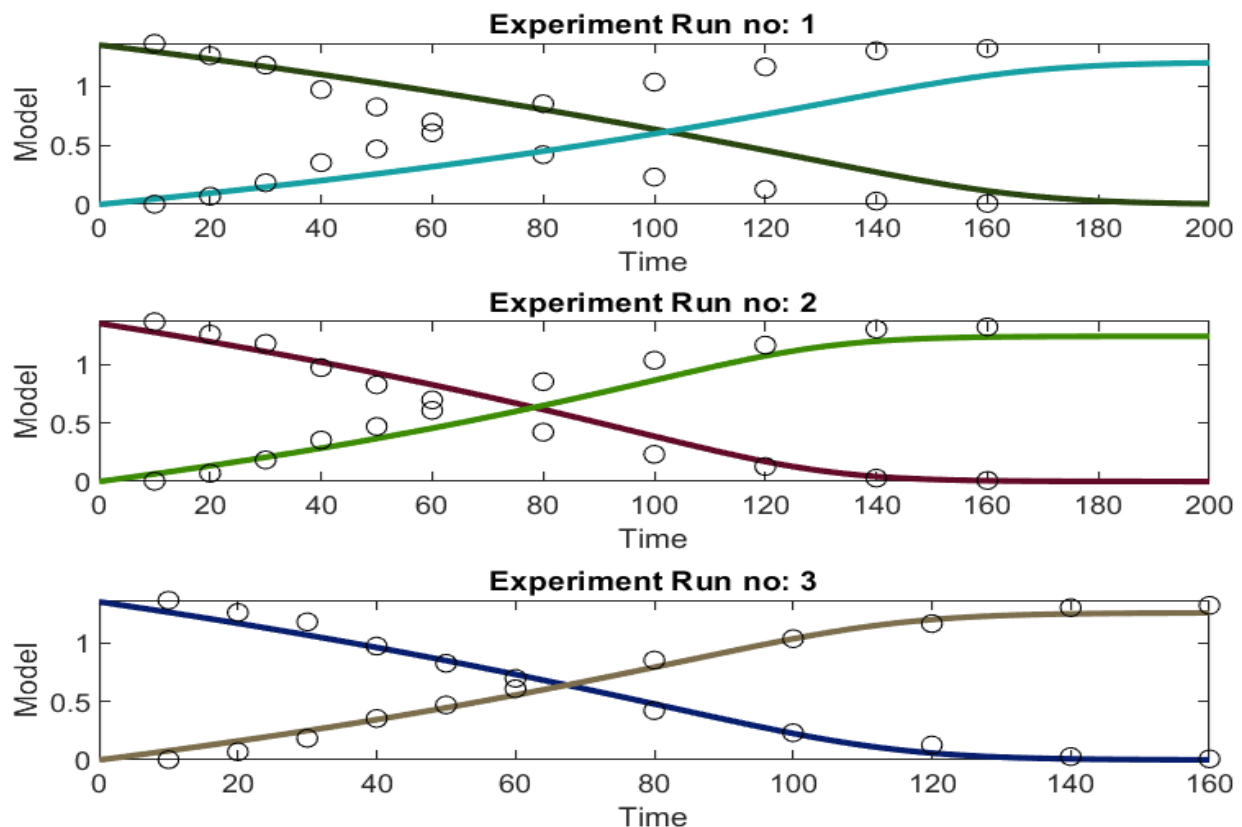
Variance Se^2 : 0.013099

Condition number of matrix A: 46240721.418457

SSE: 0.877657

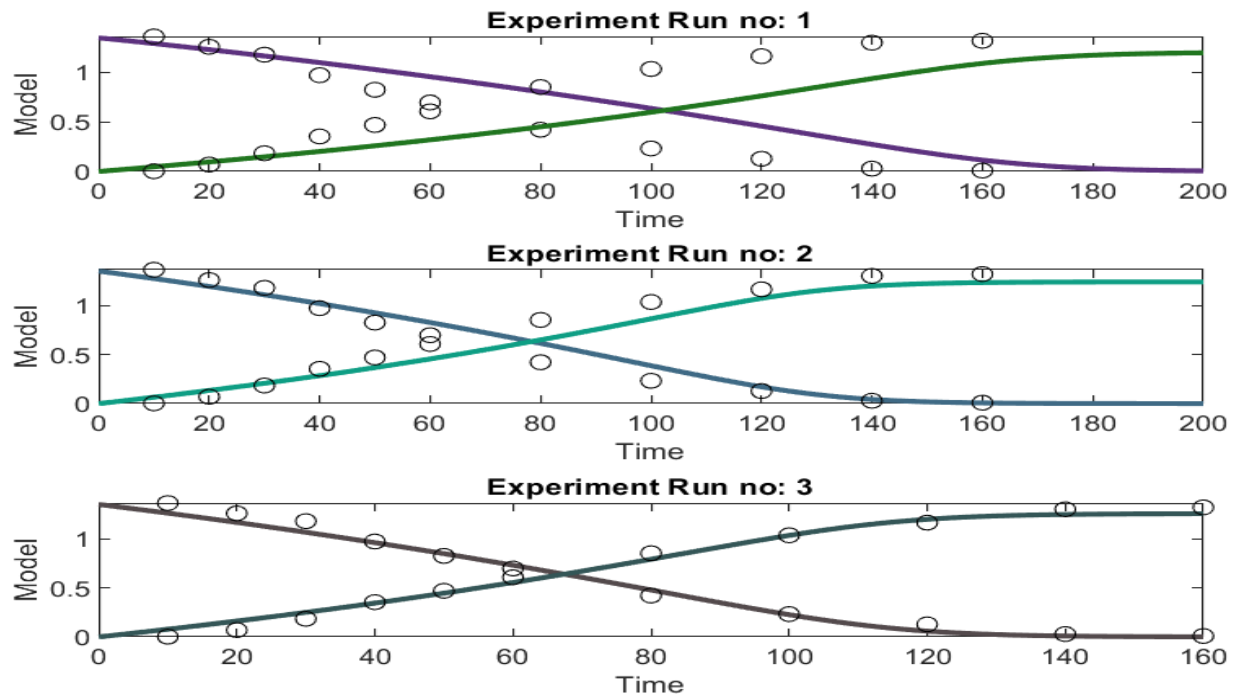
Όπως φαίνεται από τις στατιστικές εκφράσεις, κάποιες παράμετροι είναι ιδιαίτερα ασταθείς, όπως για παράδειγμα η $k(2)$ με ποσοστιαία τυπική απόκλιση / σχετικό τυπικό σφάλμα (Relative Standard Error, RSE) $48.5 \times 10^{10}\%$, η $k(4)$ με $9.08 \times 10^8\%$ και η $k(5)$ με $34.63 \times 10^3\%$. Παράλληλα, εμφανίζοντας τα παρακάτω γραφικά, είναι εμφανές ότι κάποιες παράμετροι επηρεάζουν το μοντέλο σε πολύ μικρό βαθμό, και αυτό φαίνεται πολλαπλασιάζοντας τις επί 100 με το μοντέλο να παραμένει πρακτικά ίδιο.

- Δεδομένο Σετ – Εικόνα 4.5. Γραφική απεικόνιση για το δεδομένο σετ:



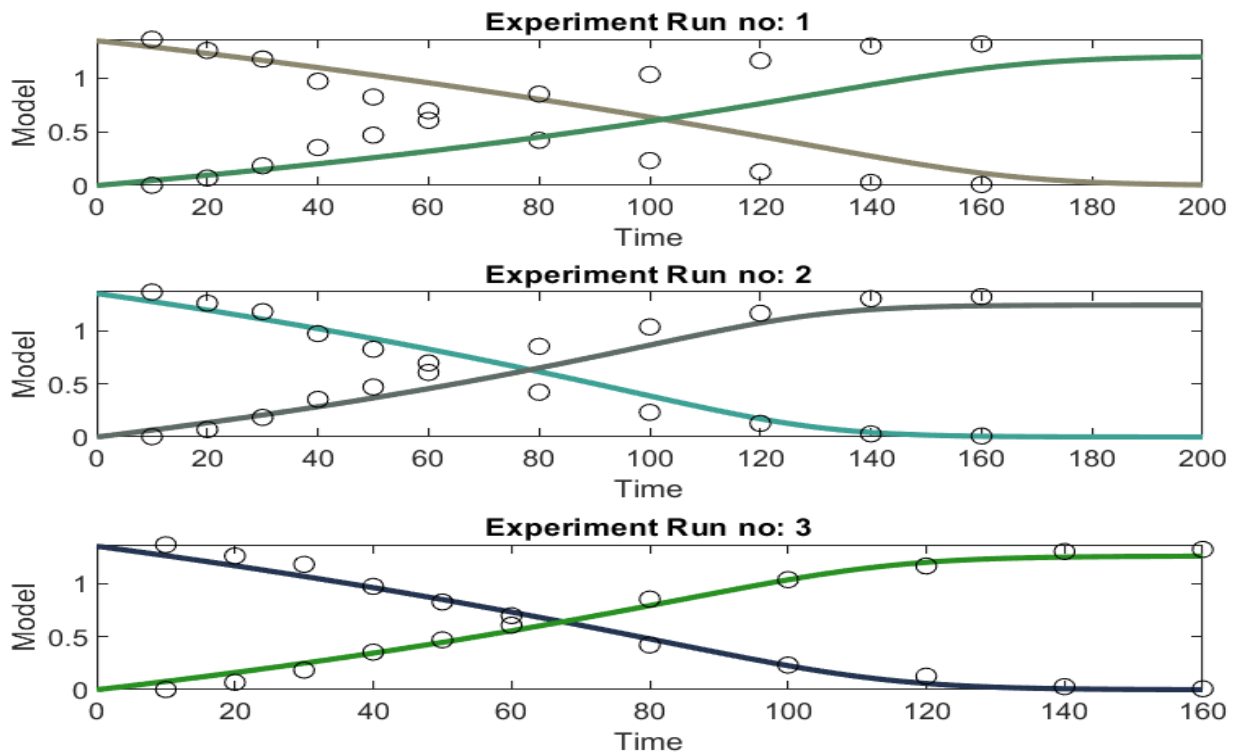
με αποτέλεσμα της LS συνάρτησης: 0.87766

- $k(2) \times 100$ – Εικόνα 4.6. Γραφική απεικόνιση για δεδομένο σεν και $k(2) \times 100$:



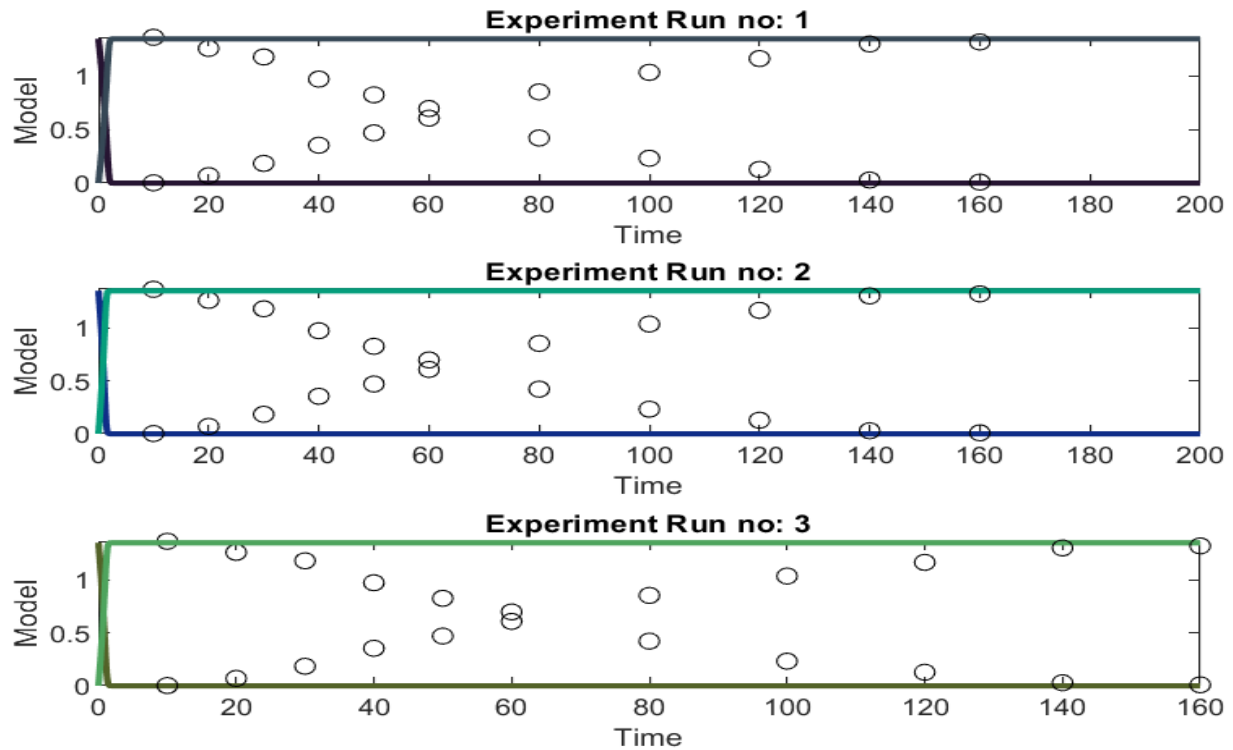
με αποτέλεσμα της LS συνάρτησης: 0.87777

- $k(4) \times 100$ – Εικόνα 4.7. Γραφική απεικόνιση για δεδομένο σεν και $k(4) \times 100$:



με αποτέλεσμα της LS συνάρτησης: 0.8779

- $k(1) \times 100$ – Εικόνα 4.8. Γραφική απεικόνιση για δεδομένο σετ και $k(1) \times 100$:



με αποτέλεσμα της LS συνάρτησης: 12.3701

Είναι εμφανές ότι, οι παράμετροι με μεγάλο RSE δεν επηρεάζουν το μοντέλο. Αντίθετως, όπως φαίνεται και για την $k(1)$, οι παράμετροι με μικρό RSE επηρεάζουν κατά πολύ το μοντέλο και το αποτέλεσμα της αντικειμενικής συνάρτησης.

Συμπερασματικά, το σύστημα της εφαρμογής είναι προβληματικό από τη φύση του, καθώς βρίσκεται σε κακή κατάσταση, είναι δύσκαμπτο και δεν συγκλίνει, ακόμα και με αρχικές εκτιμήσεις παραμέτρων ιδιαίτερα ευνοϊκές για τη λύση του. Επιλέγεται 1 iteration για την εκτέλεση του προγράμματος, καθώς δεν υπάρχει σύγκλιση και το πρόγραμμα δεν τερματίζει ή δεν εμφανίζει βασικές μεταβολές, ανεξάρτητα από τον αριθμό των iterations.

5. Συμπεράσματα

Στην παρούσα διπλωματική εργασία, σκοπός ήταν η ανάπτυξη προγράμματος στο περιβάλλον της MATLAB, το οποίο υπολογίζει άγνωστες παραμέτρους σε συστήματα που περιγράφονται από ODEs. Παράλληλα, ζητούμενο ήταν να υπολογίζονται βασικές στατιστικές εκφράσεις για τις παραμέτρους που υπολογίζονται και να δημιουργούνται γραφήματα που απεικονίζουν τις εκτιμώμενες τιμές και γραφικές παραστάσεις του μοντέλου και τα πειραματικά δεδομένα. Τα δύο επιπλέον αυτά χαρακτηριστικά, βοηθούν το χρήστη: 1) να αντιληφθεί ευκολότερα το αν το σύστημα που ανέπτυξε και περιγράφει είναι ποιοτικό και ορθά δομημένο και 2) να ελέγξει οπτικά και γρήγορα αν υπάρχει σύγκλιση, την απόσταση μεταξύ εκτιμώμενων και πειραματικών σημείων δεδομένων και τη συνολική δομή του συστήματος.

Όπως φαίνεται κατά την εκτέλεση και επίλυση της εφαρμογής 1, τα ζητούμενα ικανοποιήθηκαν. Το πρόγραμμα δημιουργεί μοντέλο που συγκλίνει σχεδόν τέλεια με τα πειραματικά δεδομένα και τα γραφήματα και τα στατιστικά αποτελέσματα υπολογίζονται και εμφανίζονται απρόσκοπτα. Ωστόσο, το παρόν πρόγραμμα έχει και τους περιορισμούς του. Ήταν αδύνατο να παράξει μοντέλο που να επιφέρει σύγκλιση για την εφαρμογή 2 χωρίς πολύ σημαντικές αλλαγές που πλέον βγαίνουν εκτός θέματος, ή μάλλον επεκτείνουν σε υπερβολικά μεγάλο βαθμό το θέμα. Παρ'όλα αυτά, ακόμα και σε αυτήν την περίπτωση, το πρόγραμμα παράγει χρήσιμες πληροφορίες που μπορούν να αξιοποιηθούν από το χρήστη. Αξίζει επίσης να αναφερθεί ότι το σύστημα της εφαρμογής 2 είναι εκ φύσεως ιδιότροπο και σε κακή κατάσταση.

Όσον αφορά μελλοντικές κατευθύνσεις, η παρούσα εργασία εστιάζει σε έναν τομέα βασικότατο και ιδιαίτερα εκτενή για τις θετικές επιστήμες, αφού πλέον διαφορικές εξισώσεις χρησιμοποιούνται σε οποιαδήποτε σχεδόν επιστήμη χρησιμοποιούνται μαθηματικά, στατιστική κλπ. Επομένως, οι κατευθύνσεις που μπορεί να πάρει κάποιος είναι πάρα πολλές. Μία επόμενη υλοποίηση βασισμένη πάνω στην παρούσα εργασία θα μπορούσε να είναι η περαιτέρω ανάπτυξη του αλγορίθμου, ώστε να μπορεί να επιλυεί και προβλήματα δύσκαμπτα ή/και σε πολύ κακή κατάσταση, μέσα από επιπλέον τροποποιήσεις και βελτιστοποιήσεις του αντίστοιχου συστήματος.

Βιβλιογραφία

- Ζαΐρης, Π., "Στατιστική Μεθοδολογία ", ΚΡΙΤΙΚΗ, ISBN13 9789602186923, 2010.
- Καλογεροπούλου, Α., Γκίκας, Μ., Καραγιαννάκης, Δ. και Μ. Λάμπρου, "Αγγλοελληνικό Λεξικό Μαθηματικών Όρων", ΤΡΟΧΑΛΙΑ, ISBN 960-7022-33-5, 2012.
- Bard, Y., "Comparison of Gradient Methods for the Solution of Nonlinear Parameter Estimation Problems", *SIAM J. Numer. Anal.*, 7, 157-186 (1970).
- Bard, Y., *Nonlinear Parameter Estimation*, Academic Press, New York, NY, 1974.
- Barnett, V., T. Lewis, and V. Rothamsted, *Outliers in Statistical Data*, 3rd edition, Wiley, New York, NY, 1994.
- Chapra, S. and R. Canale, "Αριθμητικές Μέθοδοι για Μηχανικούς", ΕΚΔΟΣΕΙΣ ΤΖΙΟΛΑ, ISBN 978-960-418-763-8, 2018.
- Englezos, P. and N. Kalogerakis, "Constrained Least Squares Estimation of Binary Interaction Parameters in Equations of State", *Computers Chem. Eng.*, 17, 117-121 (1993).
- Englezos, P., N. Kalogerakis and P.R. Bishnoi, "A Systematic Approach for the Efficient Estimation of Interaction Parameters in Equations of State Using Binary VLE Data", *Can. J. Chem. Eng.*, 71, 322-326 (1993).
- Englezos, P. and N. Kalogerakis, "Applied Parameter Estimation for Chemical Engineers", *Marcel Dekker, Inc.*, QA274.4 E64, 2001.
- Hougen, O., and K.M. Watson, *Chemical Process Principles*, Vol. 3, J. Wiley, New York, NY, 1948.
- Kalogerakis, N. and R. Luus, "Increasing the Size of the Region of Convergence in Parameter Estimation", *proc. American Control Conference*, Arlington, Virginia, 1, 358-364 (1982).
- Kalogerakis, N., and R. Luus, "Simplification of Quasilinearization Method for Parameter Estimation", *AIChE J.*, 29, 858-864 (1983a).
- Kalogerakis, N., and R. Luus, "Improvement of Gauss-Newton Method for Parameter Estimation through the Use of Information Index", *Ind. Eng. Chem. Fundam.*, 22, 436-445 (1983b).

- Kalogerakis, N., "Parameter Estimation of Systems Described by Ordinary Differential Equations", *Ph.D. thesis*, Dept. of Chemical Engineering and Applied Chemistry, University of Toronto, ON, Canada, 1983.
- Levenberg, K., "A Method for the Solution of Certain Non-linear Problems in Least Squares", *Quart. Appl. Math.*, II(2), 164-168 (1944).
- Marquardt, D.W., "An Algorithm for Least-Squares Estimation of Nonlinear Parameters", *J. Soc. Indust. Appl. Math.*, 11(2), 431-441 (1963).
- Seinfeld, J.H., and G.R. Gavalas, "Analysis of Kinetic Parameters from Batch and Integral Reaction Experiments", *AIChE J.*, 16, 644-647 (1970).
- Wang, B.C. and R. Luus, "Increasing the Size of the Region of Convergence for Parameter Estimation Through the Use of Shorter Data-Length", *Int. J. Control*, 31, 947-957 (1980).
- Zhu, X. D., G. Valerius, and H. Hofmann, "Intrinsic Kinetics of 3-Hydroxy-propanal Hydrogenation over Ni/SiO₂/Al₂O₃ Catalyst", *Ind. Eng. Chem. Res.*, 36, 3897-2902 (1997).