

Προσομοίωση και Μοντελοποίηση Δυναμικών Συστημάτων Τελική Εργασία

Αριστείδης Δασκαλόπουλος (ΑΕΜ: 10640)

6 Ιουλίου 2025

Περιεχόμενα

1	Θέμα 1	2
1.1	Ερώτημα Α	3
1.1.1	Μαθηματική Ανάλυση Μεθόδου Lyapunov Μ δομής	3
1.1.2	Αλγόριθμος Lyapunov με Προβολή	4
1.1.3	Προσομοίωση στο MATLAB	7
1.1.4	Μελέτη Ευστάθειας Συστήματος Εκτίμησης & Επιλογή Εισόδου $u(t)$. . .	7
1.1.5	Αποτελέσματα MATLAB και Συμπεράσματα	8
1.2	Ερώτημα Β	11
1.2.1	Γραμμικά Παραμετροποιημένη Μορφή	11
1.2.2	Εισαγωγή Σφάλματος Πόλωσης στο Σύστημα	13
1.2.3	Μοντελοποίηση ω στο MATLAB	14
1.2.4	Αποτελέσματα MATLAB και Συμπεράσματα	15
2	Θέμα 2	19
2.1	Ανάλυση Βημάτων	19
2.1.1	Διαδικασία Επιλογής Δομής	19
2.1.2	Εκτίμηση Παραμέτρων σε Πραγματικό Χρόνο	20
2.1.3	Αξιολόγηση και Σύγκριση Μοντέλων	21
2.2	Αποτελέσματα	21

Θέμα 1

Το δοθέν γραμμικό σύστημα που καλούμαστε να μελετήσουμε είναι το εξής:

$$\dot{x}(t) = Ax(t) + Bu(t), \quad (1-1)$$

όπου:

$$x(t) = \begin{bmatrix} x_1(t) & x_2(t) \end{bmatrix}^T \in \mathbb{R}^2 \text{ η κατάσταση του συστήματος, } u(t) \in \mathbb{R} \text{ η είσοδος,}$$
$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \text{και} \quad B = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix},$$

με τους A και B να θεωρούνται σταθεροί αλλά άγνωστοι/προς εκτίμηση πίνακες, για τα στοιχεία των οποίων γνωρίζουμε ότι ισχύουν οι παρακάτω περιορισμοί:

$$\boxed{a_{11} \in [-3, -1] \text{ \& } b_2 \in [1, +\infty)}$$

Στα πλαίσια της ανάλυσής μας θεωρούνται μετρήσιμες οι καταστάσεις του συστήματος, $x_1(t)$ και $x_2(t)$, όπως και η είσοδος $u(t)$. Επίσης οι πραγματικές τιμές των παραμέτρων προς εκτίμηση είναι:

$$\begin{array}{cc|c} A & & B \\ a_{11} = -2.15 & a_{12} = 0.25 & b_1 = 0 \\ a_{21} = -0.75 & a_{22} = -2 & b_2 = 1.5 \end{array}$$

Μπορούμε εύκολα να δούμε ότι οι ιδιοτιμές του 2×2 πίνακα, A , είναι συζυγείς μιγαδικές και βρίσκονται στο ανοικτό αριστερό ημιεπίπεδο.

Στην συνέχεια θα παρουσιάσουμε έναν αλγόριθμο πραγματικού χρόνου (online) για την εκτίμηση των αγνώστων πινάκων, θα μελετήσετε την ενστάθεια του συστήματος εκτίμησης που σχεδιάσαμε και, τέλος, θα εξάγουμε συμπεράσματα για την ακρίβεια των αποτελεσμάτων μας μέσω των ζητούμενων γραφικών παραστάσεων.

[1.1] Ερώτημα Α

Ο αλγόριθμος πραγματικού χρόνου για την εκτίμηση των \hat{A} και \hat{B} που θα παρουσιάσουμε στην παρακάτω ανάλυση εφαρμόζει την μέθοδο Lyapunov με προβολή—μιας και θέλουμε να κινούμαστε στον χώρο των λύσεων που ικανοποιούν τους περιορισμούς που αναφέραμε. Η δε δομή του συστήματος αναγνώρισης που επιλέγουμε να παρουσιάσουμε είναι η μεικτή (M) δομή (series-parallel configuration).

[1.1.1] Μαθηματική Ανάλυση Μεθόδου Lyapunov M δομής

Στη M δομή, όπως γνωρίζουμε, χρησιμοποιούνται οι πραγματικές τιμές των $x_1(t)$ και $x_2(t)$ στο μοντέλο, οπότε έχουμε:

$$\dot{\hat{x}} = \hat{A}x + \hat{B}u + C(x - \hat{x}), \quad (1-2)$$

όπου \hat{A}, \hat{B} : εκτιμήσεις των πινάκων A, B .

- Το σφάλμα αναγνώρισης είναι $e = x - \hat{x}$ και παραγωγίζοντάς το ως προς τον χρόνο παίρνουμε:

$$\dot{e} = Ax + Bu - \hat{A}x - \hat{B}u - Ce \xrightarrow{\substack{\tilde{A}=\hat{A}-A \\ \tilde{B}=\hat{B}-B}} \dot{e} = -Ce - \tilde{A}x - \tilde{B}u \quad (1-3)$$

- Για την ανάλυσή μας ως συνάρτηση Lyapunov παίρνουμε την:

$$V = \frac{1}{2}e^\top e + \frac{1}{2}\text{tr}\{\tilde{A}^\top \tilde{A}\} + \frac{1}{2}\text{tr}\{\tilde{B}^\top \tilde{B}\} \quad (1-4)$$

όπου με $\text{tr}\{\cdot\}$ συμβολίζουμε το ίχνος (trace) ενός πίνακα. Αν παραγωγίσουμε την V ως προς τον χρόνο κατά μήκος της λύσης της (1-3) βρίσκουμε:

$$\dot{V} = e^\top \dot{e} + \text{tr}\{\tilde{A}^\top \dot{\tilde{A}}\} + \text{tr}\{\tilde{B}^\top \dot{\tilde{B}}\} \stackrel{(1-3)}{=} -e^\top Ce - e^\top \tilde{A}x - e^\top \tilde{B}u + \text{tr}\{\tilde{A}^\top \dot{\tilde{A}}\} + \text{tr}\{\tilde{B}^\top \dot{\tilde{B}}\}. \quad (1-5)$$

Επίσης γνωρίζουμε από τις ιδιότητες του trace ότι:

$$e^\top \tilde{A}x = \text{tr}\{\tilde{A}xe^\top\} \quad \& \quad \text{tr}\{\tilde{A}^\top \dot{\tilde{A}}\} = \text{tr}\{\tilde{A}\dot{\tilde{A}}^\top\} \quad (\text{ομοίως για } B)$$

και έτσι η (1-5) γράφεται:

$$\dot{V} = -e^\top Ce + \text{tr}\{\tilde{A}\dot{\tilde{A}}^\top + \tilde{B}\dot{\tilde{B}}^\top - \tilde{A}xe^\top - \tilde{B}ue^\top\}. \quad (1-6)$$

- Για να απαλειφθούν οι όροι των οποίων το πρόσημο δεν γνωρίζουμε στην παράγωγο της Lyapunov, επιλέγουμε:

$$\dot{\tilde{A}}^\top = xe^\top \quad \& \quad \dot{\tilde{B}}^\top = ue^\top \quad \Rightarrow \quad \dot{\tilde{A}} = ex^\top \quad \& \quad \dot{\tilde{B}} = eu^\top, \quad (1-7)$$

και τότε η (1-6) γίνεται:

$$\dot{V} = -e^\top Ce \leq 0, \quad \text{όπου ο πίνακας } C \text{ πρέπει να είναι θετικά (ημι)ορισμένος.} \quad (1-8)$$

Άρα ο αλγόριθμος με προβολή που θα αναπτύξουμε στην συνέχεια θα ακολουθεί την λύση (1-7) που παρέχει η παραπάνω μέθοδος (με εξαίρεση τις περιπτώσεις που φτάνουμε στο σύνορο, όπως θα εξηγήσουμε). Με το παραπάνω βλέπουμε επίσης πως με κατάλληλη επιλογή του πίνακα C η μέθοδος (χωρίς προβολή) είναι ευσταθής. (Για την σύγκληση των παραμετρικών σφαλμάτων στο μηδέν πρέπει επίσης η είσοδος να ικανοποιεί μια ΣΕΔ).

[1.1.2] Αλγόριθμος Lyapunov με Προβολή

Για την ανάπτυξη του αλγορίθμου με προβολή ορίζουμε:

$$\theta = [a_{11} \ a_{12} \ a_{21} \ a_{22} \ b_1 \ b_2]^\top \in \mathbb{R}^6, \quad (1-9)$$

συνεπώς λαμβάνοντας υπόψιν την (1-7) θα έχουμε (για την περίπτωση χωρίς περιορισμούς):

$$\dot{\theta} = [\dot{a}_{11} \ \dot{a}_{12} \ \dot{a}_{21} \ \dot{a}_{22} \ \dot{b}_1 \ \dot{b}_2]^\top \stackrel{(1-7)}{=} [e_1 x_1 \ e_1 x_2 \ e_2 x_1 \ e_2 x_2 \ e_1 u \ e_2 u]^\top, \quad (1-10)$$

η οποία σχέση προέκυψε αντικαθιστώντας στην (1-7) τα $x = [x_1 \ x_2]^\top$ και $e = [e_1 \ e_2]^\top$.

Ορίζουμε τώρα το κυρτό σύνολο Θ όλων των $\hat{\theta}$ που να ικανοποιούν τους περιορισμούς:

$$\Theta = \{\hat{\theta} \in \mathbb{R}^6 : g_1(\hat{\theta}) \leq 0, g_2(\hat{\theta}) \leq 0, g_3(\hat{\theta}) \leq 0\}, \quad (1-11)$$

όπου τα $g_i(\hat{\theta})$, για $i = 1, 2, 3$, ορίζονται παρακάτω έτσι, ώστε οι ανισότητες στη (1-11) να εκφράζουν τους αρχικούς μας περιορισμούς:

$$g_1(\hat{\theta}) = -\hat{a}_{11} - 3, \quad g_2(\hat{\theta}) = \hat{a}_{11} + 1, \quad g_3(\hat{\theta}) = -\hat{b}_2 + 1.$$

Σχόλιο

Ο διανυσματικός ορισμός του $g(\hat{\theta})$ μπορεί να διατυπωθεί ως:

$$g(\hat{\theta}) = [g_1 \ g_2 \ g_3]^\top \in \mathbb{R}^3.$$

Τότε στην έκφραση των περιορισμών στην (1-11) θα είχαμε την διανυσματική ανισότητα: $g(\hat{\theta}) < 0_{3 \times 1}$. Αυτή η μορφή είναι μαθηματικά ισοδύναμη με την ανάλυση καθενός από τους επιμέρους όρους/συνιστώσες $g_i < 0$, για $i = 1, 2, 3$.

Για την ανάπτυξη του αλγορίθμου με προβολή, εκκινούμε από ένα αρχικό σημείο $\hat{\theta}(0) = \hat{\theta}_0$ το οποίο βρίσκεται στο εσωτερικό του συνόλου Θ , και ακολουθούμε τη δυναμική που υπαγορεύεται από τη Σχέση (1-10) (που δεν λαμβάνει υπόψιν τους περιορισμούς). Όταν η τροχιά φτάσει στο σύνορο του Θ —δηλαδή όταν για κάποιο $i \in 1, 2, 3$ ισχύει $g_i(\hat{\theta}) = 0$ —εξετάζεται αν το $\hat{\theta}$ τείνει να κινηθεί εκτός του επιτρεπτού συνόλου Θ , και έχουμε περιπτώσεις:

- Εάν αυτό δεν συμβαίνει, συνεχίζουμε να εφαρμόζουμε την ίδια δυναμική.
- Σε αντίθετη περίπτωση, **προστίθεται κατάλληλος όρος**, ο οποίος εξαναγκάζει την κίνηση του $\hat{\theta}$ προς το εσωτερικό του Θ ή κατά μήκος του συνόρου του.

Σύμφωνα με τα παραπάνω προκύπτει ο αλγόριθμος με προβολή ως:

$$\dot{\hat{\theta}} = \begin{cases} \Gamma \begin{bmatrix} e_1 x_1 \\ e_1 x_2 \\ e_2 x_1 \\ e_2 x_2 \\ e_1 u \\ e_2 u \end{bmatrix}, & \text{αν } \hat{\theta} \in \Theta_{in} \text{ ή αν } \hat{\theta} \in \Theta_b \text{ \& } \dot{\hat{\theta}}_{(1-10)}^\top \cdot \nabla g \leq 0_{3 \times 1} \\ \Gamma \begin{bmatrix} e_1 x_1 \\ e_1 x_2 \\ e_2 x_1 \\ e_2 x_2 \\ e_1 u \\ e_2 u \end{bmatrix} - \Gamma \frac{\nabla g \nabla g^\top}{\nabla g^\top \Gamma \nabla g} \Gamma \begin{bmatrix} e_1 x_1 \\ e_1 x_2 \\ e_2 x_1 \\ e_2 x_2 \\ e_1 u \\ e_2 u \end{bmatrix}, & \text{διαφορετικά} \end{cases} \quad (1-12)$$

όπου Θ_{in} είναι το εσωτερικό και Θ_b το σύνορο του Θ . Στην (1-12) πρέπει $\hat{\theta}(0) \in \Theta$.

Για την (1-12) ισχύουν τα εξής:

- Ο πίνακας $\Gamma = \begin{bmatrix} \gamma_1 & 0 & \cdots & 0 \\ 0 & \gamma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \gamma_6 \end{bmatrix} \in \mathbb{R}^{6 \times 6}$ περιλαμβάνει όλες τις μεταβλητές $\gamma_i > 0$ (learning rate) με τις οποίες θα πολλαπλασιαστεί καθεμία συνιστώσα του διανύσματος $\dot{\hat{\theta}}$.
- Όταν το g εκφράζει διανυσματικά όλους τους περιορισμούς ισχύει:

$$\nabla g = [\nabla g_1 \quad \nabla g_2 \quad \nabla g_3] \in \mathbb{R}^{6 \times 3},$$

όπου:

$$\nabla g_1 = [-1 \ 0 \ 0 \ 0 \ 0 \ 0]^\top, \quad \nabla g_2 = [1 \ 0 \ 0 \ 0 \ 0 \ 0]^\top, \quad \nabla g_3 = [0 \ 0 \ 0 \ 0 \ 0 \ -1]^\top \quad (1-13)$$

- Ο λόγος $\frac{\nabla g \nabla g^\top}{\nabla g^\top \Gamma \nabla g} \in \mathbb{R}^{6 \times 6}$ εκφράζει τον μηχανισμό με τον οποίο προβάλλεται το διάνυσμα $\hat{\theta}$ εντός του επιτρεπτού χώρου Θ , όταν διαπιστώνεται ότι τείνει να κινηθεί εκτός αυτού—δηλαδή, όταν ο πρώτος κλάδος είτε οδηγεί σε τιμή εκτός του Θ , είτε δίνει σημείο στο σύνορό του, αλλά με παράγωγο $\dot{\hat{\theta}}$ που έχει φορά προς το εξωτερικό του συνόλου. Για τον δε υπολογισμό του λόγου αυτού η μαθηματική ισοδυναμία ώστε να οδηγεί σε προβολή στο Θ είναι: $\nabla g (\nabla g^\top \Gamma \nabla g)^{-1} \nabla g^\top$.

Παρατήρηση

Για οποιοδήποτε διαγώνιο πίνακα Γ κατόπιν αντικατάστασης και εύρεσης του $\nabla g^\top \Gamma \nabla g$, παρατηρούμε ότι αυτός είναι ιδιάζον (μη αντιστρέψιμος/με ορίζουσα μηδέν). Αυτό συμβαίνει επειδή οι περιορισμοί που εκφράζουν οι συναρτήσεις g_1 και g_2 αναφέρονται στην ίδια παράμετρο a_{11} και δεν γίνεται και οι δύο ταυτόχρονα να είναι ≥ 0 . Άρα η έκφραση του λόγου ως αντίστροφος πίνακας, όπως δείξαμε προηγουμένως, δεν μπορεί να εφαρμοστεί στην περίπτωση μας.

Για την επίλυση του προβλήματος, στην περίπτωση όπου πρέπει να πάρουμε τον δεύτερο κλάδο της (1-12), θα προβάλλουμε την λύση λαμβάνοντας μόνο τους ενεργούς περιορισμούς. Ένας τρόπος να το κάνουμε αυτό είναι ο παρακάτω:

Τροποποιούμε τον αλγόριθμο που παρουσιάζει η (1-12)—και συγκεκριμένα τον δεύτερο κλάδο που κάνει την προβολή—ώστε να χρησιμοποιούμε άθροισμα επί των περιορισμών. Επομένως, στον τροποποιημένο αλγόριθμο, ο όρος προβολής ανανεώνεται ως:

$$\dot{\hat{\theta}} = \Gamma \begin{bmatrix} e_1 x_1 \\ e_1 x_2 \\ e_2 x_1 \\ e_2 x_2 \\ e_1 u \\ e_2 u \end{bmatrix} - \sum_{i \in \mathcal{A}} \Gamma \frac{\nabla g_i \nabla g_i^\top}{\nabla g_i^\top \Gamma \nabla g_i} \Gamma \begin{bmatrix} e_1 x_1 \\ e_1 x_2 \\ e_2 x_1 \\ e_2 x_2 \\ e_1 u \\ e_2 u \end{bmatrix}, \quad (1-14)$$

ή για απλότητα:

$$\dot{\hat{\theta}}_{new} = \Gamma \dot{\hat{\theta}}_{current} - \sum_{i \in \mathcal{A}} \Gamma \frac{\nabla g_i \nabla g_i^\top}{\nabla g_i^\top \Gamma \nabla g_i} \Gamma \dot{\hat{\theta}}_{current}, \quad (1-15)$$

όπου \mathcal{A} είναι το σύνολο των ενεργών περιορισμών—δηλαδή, οι δείκτες i για τους οποίους είτε $g_i(\hat{\theta}) > 0$, είτε $g_i(\hat{\theta}) = 0$ & $\dot{\hat{\theta}}_{(1-10)} \cdot \nabla g_i > 0$.

Με $\dot{\hat{\theta}}_{(1-10)}$ συμβολίζουμε την χωρίς περιορισμούς επόμενη ενημέρωση του διανύσματος $\dot{\hat{\theta}}_{current}$ (ακολουθώντας την μέθοδο Lyapunov για την M δομή).

Το άθροισμα επί των ενεργών περιορισμών εμφανίζεται επειδή, όταν πολλαπλοί περιορισμοί είναι ενεργοί (παρατηρώντας τους περιορισμούς μπορούμε να έχουμε \max δύο ενεργούς περιορισμούς: g_1 με g_3 ή g_2 με g_3), η ενημέρωση πρέπει να προβληθεί ως προς όλους τους ενεργούς περιορισμούς.

Ο όρος $\Gamma \frac{\nabla g_i \nabla g_i^\top}{\nabla g_i^\top \Gamma \nabla g_i} \Gamma \dot{\hat{\theta}}$ προβάλλει την ενημέρωση $\dot{\hat{\theta}}$ στην κάθετη κατεύθυνση της i -οστής επιφάνειας περιορισμού, $g_i(\hat{\theta}) = 0$. Η αφαίρεση του αποτελέσματος της προβολής αφαιρεί τη συνιστώσα του $\dot{\hat{\theta}}$ που θα παραβίαζε το $g_i(\hat{\theta}) \leq 0$. Όταν πολλαπλοί περιορισμοί είναι ενεργοί, αθροίζουμε αυτές τις προβολές για να αφαιρέσουμε τις συνιστώσες κατά μήκος όλων των σχετικών κάθετων κατευθύνσεων, διασφαλίζοντας ότι η ενημέρωση του $\dot{\hat{\theta}}$ παραμένει στο σύνορο Θ_b .

Ο λόγος για τον οποίο το άθροισμα στην (1-15) είναι μαθηματικά σωστό είναι επειδή σε περίπτωση που έχουμε πολλαπλούς ενεργούς περιορισμούς, τότε όπως εξηγήσαμε αυτό σημαίνει ότι $\mathcal{A} = \{1, 3\}$ ή $\{2, 3\}$. Άρα, επειδή από τις (1-13) έχουμε $\nabla g_1^\top \cdot \nabla g_3 = 0$ και $\nabla g_2^\top \cdot \nabla g_3 = 0$, επομένως οι προβολές ως προς αυτά τα ζεύγη είναι μεταξύ τους ορθογώνιες και έπεται πως μπορούν να υπολογιστούν και να αφαιρεθούν ανεξάρτητα οι συνιστώσες που δίνουν. Εναλλακτικά, θα έπρεπε να σχηματίζουμε κάθε φορά ένα διάνυσμα g' με όλους τους ενεργούς περιορισμούς.

Έχοντας πλέον ολοκληρώσει την θεωρητική ανάλυση του προβλήματος μπορούμε να συνεχίσουμε με την συνοπτική ανάλυση του τρόπου προσομοίωσης/επίλυσης του προβλήματος στο MATLAB.

[1.1.3] Προσομοίωση στο MATLAB

Συνοψίζοντας την ανάλυση της προηγούμενης ενότητας, για την εύρεση των προς εκτίμηση παραμέτρων, έχουμε καταλήξει στο εξής σύστημα προς επίλυση:

- Η πρώτη διαφορική εξίσωση που θέλουμε να λύσουμε είναι αυτή του πραγματικού συστήματος, καθώς για κάθε χρονική στιγμή πρέπει να γνωρίζουμε την $x(t)$ (η μέθοδος είναι πραγματικού χρόνου επομένως λύνουμε με *ode45* αυτήν την διαφορική, ταυτόχρονα με όλες τις υπόλοιπες):

$$\dot{x}(t) = A_{\text{true}}x(t) + B_{\text{true}}u(t), \quad (1-16)$$

όπου $x(t) \in \mathbb{R}^2$ είναι το διάνυσμα κατάστασης και $u(t) \in \mathbb{R}$ είναι ένα γνωστό σήμα εισόδου της επιλογής μας (θα αναφερθούμε στην επιλογή του παρακάτω).

- Η δεύτερη διαφορική αναφέρεται στο σύστημα αναγνώρισης. Οι πίνακες $A_{\text{true}} = A \in \mathbb{R}^{2 \times 2}$ και $B_{\text{true}} = B \in \mathbb{R}^{2 \times 1}$ για το πάνω σύστημα θεωρούνται γνωστοί, όμως θα πρέπει να εκτιμηθούν από το σύστημα αναγνώρισης το οποίο ορίστηκε στην προηγούμενη ανάλυση με την παρακάτω μεικτή δομή:

$$\dot{\hat{x}}(t) = A_{\text{hat}}x(t) + B_{\text{hat}}u(t) + C(x(t) - \hat{x}(t)), \quad (1-17)$$

όπου:

$$\hat{A} = A_{\text{hat}} = \begin{bmatrix} \hat{\theta}_1 & \hat{\theta}_2 \\ \hat{\theta}_3 & \hat{\theta}_4 \end{bmatrix}, \quad \hat{B} = B_{\text{hat}} = \begin{bmatrix} \hat{\theta}_5 \\ \hat{\theta}_6 \end{bmatrix}, \quad e(t) = x(t) - \hat{x}(t), \quad C = \theta_m \cdot I_2,$$

με το θ_m να πρέπει να είναι θετικός αριθμός της επιλογής μας, ώστε να ικανοποιείται η (1-8).

- Τέλος το σύστημά μας πρέπει να περιλαμβάνει τον νόμο ενημέρωσης παραμέτρων (με περιορισμούς). Ο αντίστοιχος νόμος χωρίς περιορισμούς δίνεται από την (1-10), και κλιμακώνεται από έναν θετικά ορισμένο διαγώνιο πίνακα $\Gamma = \text{diag}(\gamma_1, \dots, \gamma_6)$. Για την δε επιβολή περιορισμών στις παραμέτρους, εφαρμόζεται η μέθοδος προβολής της (1-15).

Άρα, το πλήρες σύστημα που θα δώσουμε στην *ode45* για επίλυση είναι το εξής:

$$\frac{d}{dt} \begin{bmatrix} x(t) \\ \hat{x}(t) \\ \theta(t) \end{bmatrix} = \begin{bmatrix} A_{\text{true}}x(t) + B_{\text{true}}u(t) \\ A_{\text{hat}}x(t) + B_{\text{hat}}u(t) + C(x(t) - \hat{x}(t)) \\ \text{Proj}(\hat{\theta}_{\text{current}}, \Gamma \dot{\hat{\theta}}_{\text{current}}) \end{bmatrix}, \quad (1-18)$$

με κατάλληλες αρχικές τιμές για τα x_0 , \hat{x}_0 και $\hat{\theta}_0$.

Σχόλιο

Η συνάρτηση προβολής $\text{Proj}(\hat{\theta}_{\text{current}}, \Gamma \dot{\hat{\theta}}_{\text{current}})$ παίρνει ως όρισμα και το $\hat{\theta}_{\text{current}}$ για να ελέγξει το εσωτερικό γεγόμενο αυτού με τις παραγώγους των συναρτήσεων περιορισμών, (1-13)—μιας και με βάση το πρόσημο του αποτελέσματος βρίσκει αν και ποιοι εκ των περιορισμών είναι ενεργοί ώστε να κάνει την αντίστοιχη προβολή.

Στο αρχείο *src/constrainedSystemDynamics.m* έχει εκφραστεί το παραπάνω σύστημα, το οποίο λύνεται (με *ode45*) στο εκτελέσιμο script: [mainPart1a.m](#) που αντιστοιχεί στο Ερώτημα α του Θέματος 1.

[1.1.4] Μελέτη Ευστάθειας Συστήματος Εκτίμησης & Επιλογή Εισόδου $u(t)$

Στην μαθηματική ανάλυση της μεθόδου Lyapunov για μεικτή δομή συστήματος αναγνώρισης δείξαμε ότι το σύστημα εκτίμησης είναι ευσταθές αν ο πίνακας C της (1-6) είναι θετικά ημιορισμένος.

Παρόλα αυτά η παραπάνω ανάλυση αφορούσε την περίπτωση που έχουμε ενημέρωση των παραμέτρων θ , χωρίς περιορισμούς, μέσω της (1-10). Για την περίπτωση που η δυναμική του θ δίνεται από την (1-15), γεννιέται το ερώτημα αν αυτό θα επηρεάσει την ευστάθεια του συστήματος εκτίμησης. Παρόλα αυτά, με την προβολή του θ σύνορο—φράσσοντας δηλαδή το επιτρεπτό πεδίο στο οποίο μπορούν να κινηθούν τα $\hat{\theta}$ —όχι απλώς δεν "χαλάει" η συνθήκη ευστάθειας της (1-6), αλλά στην πράξη γίνεται ακόμα πιο αρνητική.

Πράγματι, αντικαθιστώντας την (1-14) στην (1-6), παρατηρούμε ότι μέσα στο trace θα μείνουν μόνο τα στοιχεία, που αναφέρονται στην προβολή. Με αυτά, το αποτέλεσμα του trace δίνει μία αρνητική ποσότητα η οποία δεν επηρεάζει την συνθήκη $\dot{V} < 0$. Προφανώς η επιλογή κατάλληλου πίνακα C είναι σε κάθε περίπτωση σημαντική για την ευστάθεια.

Ένας τρόπος που μπορούμε να επαληθεύσουμε και στην πράξη την ευστάθεια του συστήματος εκτίμησης είναι να δοκιμάσουμε να τρέξουμε την προσομοίωση με διάφορες εισόδους $u(t)$ και να δούμε ότι όντως στο σύστημά μας υπάρχει ευστάθεια. Η επιλογή κατάλληλης εισόδου είναι επίσης σημαντική μιας και από αυτήν (όπως αναφέραμε και στο τέλος της μαθηματική ανάλυση της μεθόδου Lyapunov) θα εξαρτηθεί αν τα παραμετρικά σφάλματα θα συγχλίνουν στο μηδέν με την πάροδο του χρόνου.

Ικανή και αναγκαία συνθήκη για να συγχλίνει η εκτίμηση των παραμέτρων $\hat{\theta}$ στην πραγματική τους τιμή θ είναι η ικανοποίηση μιας ΣΕΔ. Εφόσον μας δίνεται η δυνατότητα η $u(t)$ να είναι δικής μας επιλογής, επιλέγουμε την:

$$u(t) = 2.5 \sin(t) + 1.5 \cos(1.5t) + 0.8 \sin(3t). \quad (1-19)$$

Στην πράξη, για το ΓΧΑ σύστημα 2ης τάξης που μελετάμε, μας ενδιαφέρει να εκτιμήσουμε και τις 6 ανεξάρτητες παραμέτρους του. Γι' αυτόν τον λόγο, επιλέγουμε να χρησιμοποιήσουμε συνολικά τρεις συχνότητες. Με αυτόν τον τρόπο εξασφαλίζουμε ότι το σήμα εισόδου έχει αρκετές συχνότητες ώστε να έχουμε την δυνατότητα να εκτιμήσουμε με αρκετή ακρίβεια καθεμία από τις 6 παραμέτρους του (με σφάλμα εκτίμησης κάθε παραμέτρου να μπορεί ιδανικά να τείνει στο μηδέν).

Στην συνέχεια παραθέτουμε τις ζητούμενες γραφικές παραστάσεις και εξάγουμε σχετικά συμπεράσματα.

[1.1.5] Αποτελέσματα MATLAB και Συμπεράσματα

Στο Σχήμα 1 παρουσιάζονται οι γραφικές παραστάσεις των δύο καταστάσεων $x(t) = [x_1(t) \ x_2(t)]^T$. Όπως είναι φανερό και οπτικά, αλλά και όπως δείχνουν οι αντίστοιχες γραφικές παραστάσεις των σφαλμάτων στο Σχήμα 2, το σφάλμα της μοντελοποίησης τείνει στο μηδέν με την πάροδο του χρόνου—όπως ακριβώς περιμέναμε να συμβαίνει και σύμφωνα με την θεωρητική ανάλυση.

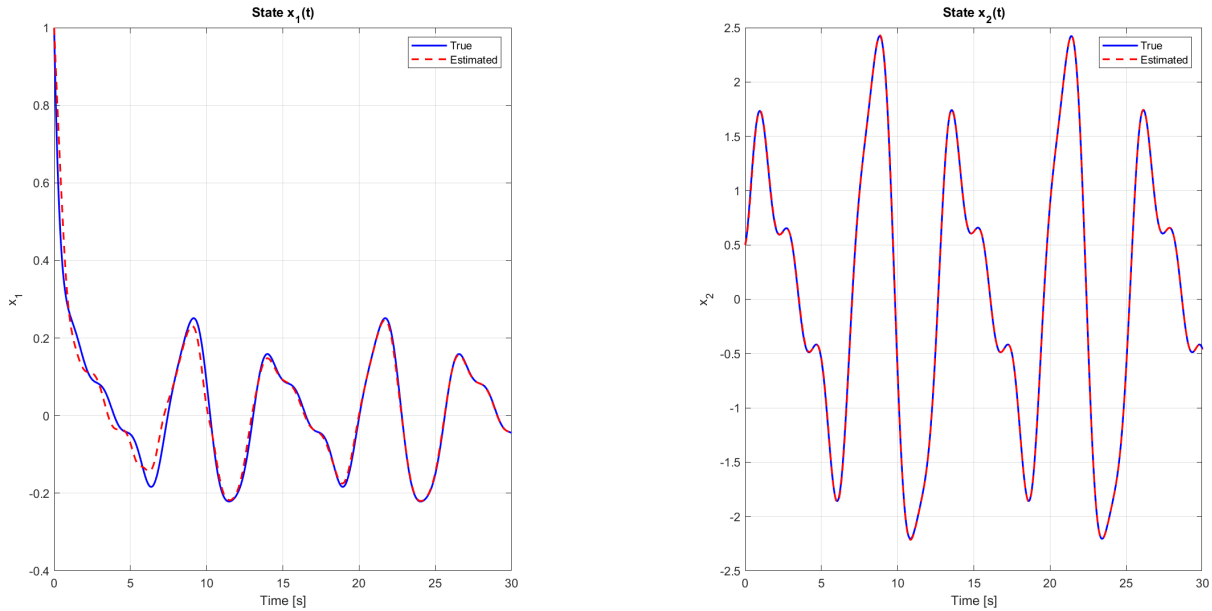
Στο αρχείο *mainTopic1a.m* μπορούμε να αλλάξουμε τον συνολικό χρόνο προσομοίωσης μέσω της μεταβλητής `T_final`, όπως επίσης και τις τιμές των θ_m και Γ μέσω των `c_coeff` και `gamma` αντίστοιχα. Μπορούμε επίσης να αλλάξουμε τις αρχικές τιμές/σημείο εκκίνησης, προσέχοντας όμως η αρχική τιμή θ_0 να ανήκει στο Θ .

Στο Σχήμα 3 βλέπουμε τις γραφικές παραστάσεις των $[A]_{ij}$ ως προς τον χρόνο. Όπως είναι φανερό, με το πέρασμα του χρόνου, καθεμία από αυτές τείνει να συγχλίνει προς την πραγματική της τιμή. Αξίζει να επισημάνουμε ότι στην περίπτωση της γραφικής παράστασης του a_{11} , αρχικά η λύση πάει να βγει εκτός επιτρεπτών τιμών αλλά παρόλα αυτά προβάλλεται στο σύνορο Θ_b , όπως ακριβώς αναλύσαμε στην θεωρία ότι θα συνέβαινε.

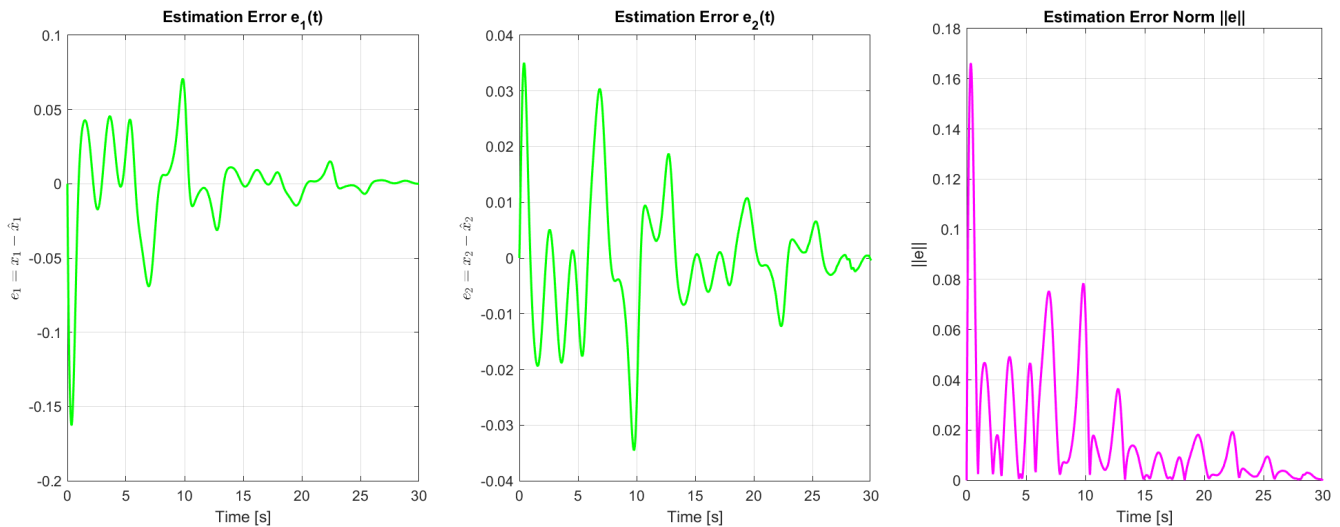
Τέλος το Σχήμα 4 περιέχει τις προβλέψεις που αφορούν τις παραμέτρους $[B]_i$ με τον χρόνο. Και σε αυτήν την περίπτωση, όπως και με τις προηγούμενες, βλέπουμε ότι όσο περνάει ο χρόνος οι τιμές

των παραμέτρων προσαρμόζονται κατάλληλα και τείνουν σε αυτές του πραγματικού συστήματος. Αυτό περιμέναμε να το δούμε, μιας και η μέθοδος που έχουμε εφαρμόσει είναι πραγματικού χρόνου.

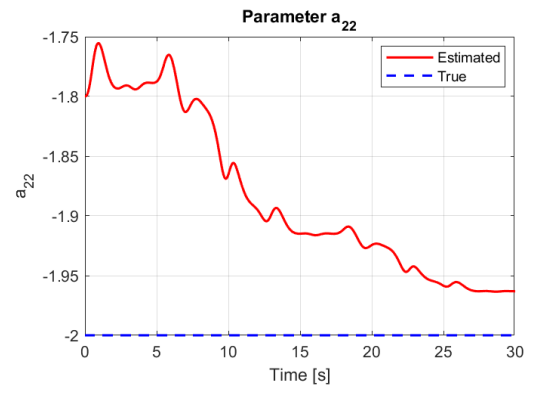
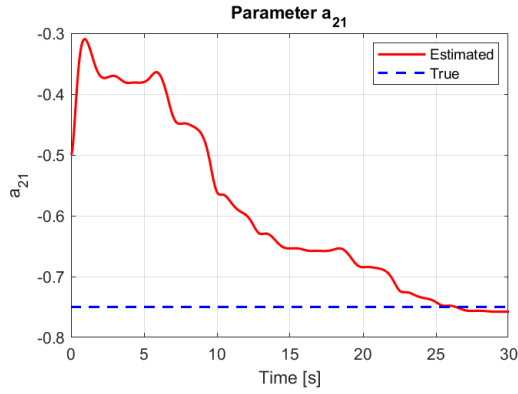
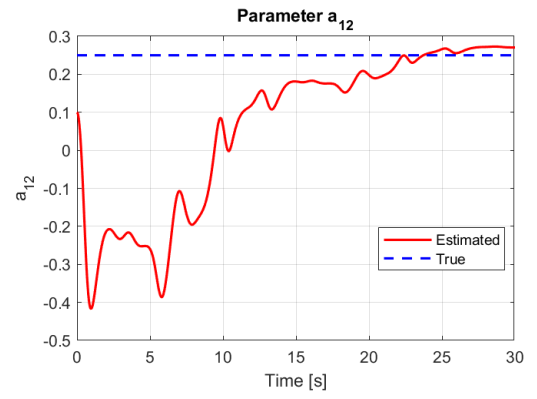
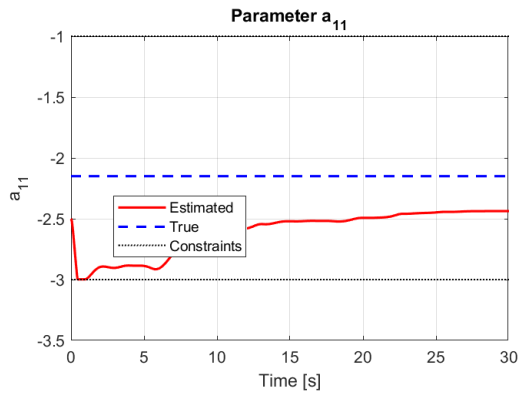
Τέλος παραθέτουμε τα logs που τυπώνουμε, τα οποία αναφέρονται στο σχετικό σφάλμα της τελικής εκτίμησης σε σχέση με την πραγματική τιμή για την κάθε παράμετρο. Με χρήση διαφορετικής εισόδου ή με την αλλαγή των Γ και θ_m , μπορούμε να πετύχουμε διαφορετικά tradeoff μεταξύ των προβλέψεων των παραμέτρων.



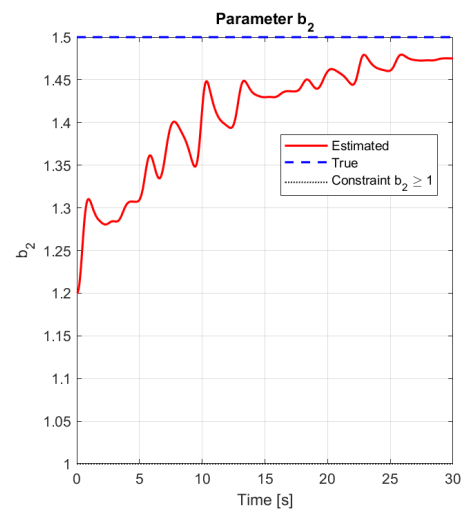
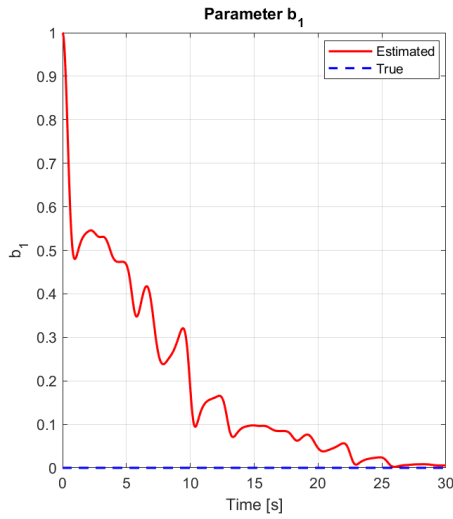
Σχήμα 1



Σχήμα 2



Σχήμα 3



Final Parameter Estimates:
a11: -2.4370 (True: -2.1500 - Error: -13.35%)
a12: 0.2706 (True: 0.2500 - Error: 8.24%)
a21: -0.7577 (True: -0.7500 - Error: -1.03%)
a22: -1.9631 (True: -2.0000 - Error: -1.85%)
b1: 0.0055 (True: 0.0000 - Error: 0.55%)
b2: 1.4751 (True: 1.5000 - Error: 1.66%)

Σχήμα 4

[1.2] Ερώτημα Β

Στην ανάλυση του προβλήματός μας ως τώρα υποθέσαμε την **πλήρη γνώση της δομής του συστήματος** που μελετάμε και με βάση αυτήν σχεδιάσαμε έναν τρόπο εκτίμησης των άγνωστων παραμέτρων του (λαμβάνοντας υπόψιν τους δοθέντες περιορισμούς). Σε πρακτικές εφαρμογές όμως αυτό είναι ανεφάρμοστο, μιας και πάντα υπάρχουν άγνωστες μη γραμμικότητες που δεν μοντελοποιούνται.

Γι' αυτόν τον λόγο, σε αυτό το ερώτημα θα εισάγουμε στο σύστημα **σφάλμα πόλωσης** $\omega \in \mathbb{R}^2$ που να ικανοποιεί $\|\omega(t)\| \leq \bar{\omega}, \forall t \geq 0$, για κάποια άγνωστη σταθερά $\bar{\omega} > 0$.

[1.2.1] Γραμμικά Παραμετροποιημένη Μορφή

Πριν το κάνουμε αυτό, θα φέρουμε το σύστημα (1-1) σε γραμμικά παραμετροποιημένη μορφή, ακολουθώντας τα παρακάτω βήματα:

- Αρχικά, ορίζουμε ένα διάνυσμα θ^* (όπως κάναμε και στο Ερώτημα Α) στο οποίο συγκεντρώνουμε όλες τις προς εκτίμηση παραμέτρους της (1-1):

$$\theta^* = [a_{11} \ a_{12} \ a_{21} \ a_{22} \ b_1 \ b_2]^\top \in \mathbb{R}^6. \quad (1-20)$$

Επίσης, ορίζουμε πίνακα $\Delta \in \mathbb{R}^{2 \times 6}$ που περιλαμβάνει όλα τα σήματα εισόδου-εξόδου, ως εξής:

$$\Delta = \begin{bmatrix} x_1 & x_2 & 0 & 0 & u & 0 \\ 0 & 0 & x_1 & x_2 & 0 & u \end{bmatrix}. \quad (1-21)$$

Σύμφωνα με τα παραπάνω το αρχικό μας σύστημα μπορεί να γραφεί στην μορφή:

$$\dot{x} = \Delta \cdot \theta^* \quad (1-22)$$

Η παραπάνω μορφή εκφράζει την παράγωγο του διανύσματος $x(t) = [x_1(t) \ x_2(t)]^\top$ ως το γινόμενο του διανύσματος των παραμέτρων θ^* με τον πίνακα όλων των (γνωστών) σημάτων, Δ .

- Το σύστημα έχει δύο εξόδους, τις x_1 και x_2 , συνεπώς ορίζουμε $y(t) = x(t)$. Για να φέρουμε το σύστημα (1-22) στην μορφή $y = \Phi \theta^*$ —με το Φ εδώ να είναι πίνακας στο $\mathbb{R}^{2 \times 6}$ —πρέπει να απαλλαχθούμε από την παράγωγο $\dot{x}(t)$. Για να πετύχουμε το ζητούμενο, φιλτράρουμε και τα δύο μέλη της (1-22) (κάθε συνιστώσα ξεχωριστά) με κατάλληλης τάξης ευσταθές φίλτρο, $\frac{1}{\Lambda(s)}$. Για απλότητα και για τις δύο συνιστώσες, μπορούμε να επιλέξουμε ένα φίλτρο πρώτης τάξης, μορφής:

$$\frac{1}{\Lambda(s)}, \quad \text{με } \Lambda(s) = s + \lambda, \ \lambda \in \mathbb{R}, \quad (1-23)$$

με την ρίζα του πολωνύμου αυτού να κυμαίνεται στο αριστερό ανοιχτό ημιεπίπεδο, δηλαδή $\lambda > 0$, για να έχουμε ευστάθεια. Υπενθυμίζουμε ότι και το αρχικό μας σύστημα είναι ευσταθές, μιας και—όπως αναφέρθηκε και στην εισαγωγική ανάλυση—οι ιδιοτιμές του πίνακα A , της (1-1), βρίσκονται στο ανοιχτό αριστερό ημιεπίπεδο.

- Από την (1-22), παίρνουμε (περνώντας στο πεδίο του Laplace):

$$\left. \begin{aligned} sX_1 &= a_{11}X_1 + a_{12}X_2 + b_1U \\ sX_2 &= a_{21}X_1 + a_{22}X_2 + b_2U \end{aligned} \right\} \Rightarrow \left. \begin{aligned} (s + \lambda_1)X_1 &= (a_{11} + \lambda_1)X_1 + a_{12}X_2 + b_1U \\ (s + \lambda_2)X_2 &= a_{21}X_1 + (a_{22} + \lambda_2)X_2 + b_2U \end{aligned} \right\} \quad (1-24)$$

$$\Rightarrow \left. \begin{aligned} X_1 &= \frac{a_{11} + \lambda_1}{s + \lambda_1}X_1 + \frac{a_{12}}{s + \lambda_1}X_2 + \frac{b_1}{s + \lambda_1}U \\ X_2 &= \frac{a_{21}}{s + \lambda_2}X_1 + \frac{a_{22} + \lambda_2}{s + \lambda_2}X_2 + \frac{b_2}{s + \lambda_2}U \end{aligned} \right\}$$

και όπως βλέπουμε τώρα, καταφέραμε να διώξουμε την παράγωγο και έχουμε απλώς την x .

Σχόλιο

Όπως παρατηρούμε, στην (1-24) δεν χρειάζεται για κάθε συνιστώσα να έχουμε το ίδιο ακριβώς φίλτρο. Παρόλα αυτά, για απλότητα (και συμμετρία μεταξύ των συνιστωσών) στην ανάλυση στο MATLAB θα έχουμε $\lambda_1 = \lambda_2 = \lambda$.

Από την (1-24) έχουμε ισοδύναμα:

$$(1-24) \Rightarrow \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} \frac{X_1}{s+\lambda_1} & \frac{X_2}{s+\lambda_1} & 0 & 0 & \frac{U}{s+\lambda_1} & 0 \\ 0 & 0 & \frac{X_1}{s+\lambda_2} & \frac{X_2}{s+\lambda_2} & 0 & \frac{U}{s+\lambda_2} \end{bmatrix} \begin{bmatrix} a_{11} + \lambda_1 \\ a_{12} \\ a_{21} \\ a_{22} + \lambda_2 \\ b_1 \\ b_2 \end{bmatrix} \quad (1-25)$$

$$\stackrel{ILT}{\Rightarrow} y = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1^{[\lambda_1]} & x_2^{[\lambda_1]} & 0 & 0 & u^{[\lambda_1]} & 0 \\ 0 & 0 & x_1^{[\lambda_2]} & x_2^{[\lambda_2]} & 0 & u^{[\lambda_2]} \end{bmatrix} \begin{bmatrix} a_{11} + \lambda_1 \\ a_{12} \\ a_{21} \\ a_{22} + \lambda_2 \\ b_1 \\ b_2 \end{bmatrix}$$

και επιστρέφοντας στο πεδίο του χρόνου βλέπουμε ότι έχουμε φέρει το σύστημα (1-1) σε μορφή $y = \Phi \theta_\lambda^*$ —όπου όλα τα στοιχεία του πίνακα $\Phi \in \mathbb{R}^{2 \times 6}$ είναι γνωστά (μπορούν να υπολογιστούν - αποτελούν τις εξόδους των φίλτρων).

Παρατηρήσεις

1. Στον τύπο (1-25), ο συμβολισμός $x_i^{[\lambda_j]}(t) \in \mathbb{R}$, $i, j \in \{1, 2\}$, χρησιμοποιείται όταν περνάμε με τον αντίστροφο μετασχηματισμό Laplace (ILT) στο πεδίο του χρόνου και αναφέρεται στο σήμα που προκύπτει όταν περάσουμε το αρχικό σήμα $x_i(t)$ από το φίλτρο πρώτης τάξης (της μορφής της (1-23), για $\lambda = \lambda_j$).
2. Ο τύπος της (1-25) μπορεί να γραφεί ξεχωριστά για καθεμία από τις συνιστώσες της $y(t) = [y_1(t) \ y_2(t)]^\top$ οι οποίες είναι σε γραμμικά παραμετροποιημένη μορφή, μιας και:

$$y_1 = x_1 = \begin{bmatrix} x_1^{[\lambda_1]} & x_2^{[\lambda_1]} & 0 & 0 & u^{[\lambda_1]} & 0 \end{bmatrix} \begin{bmatrix} a_{11} + \lambda_1 \\ a_{12} \\ a_{21} \\ a_{22} + \lambda_2 \\ b_1 \\ b_2 \end{bmatrix} = \phi_1^\top \cdot \theta_\lambda^* = \theta_\lambda^{*\top} \cdot \phi_1,$$

και αντίστοιχα για την $y_2(t)$.

Συνοψίζοντας, φέραμε το σύστημα (1-1) στην (γραμμικά παραμετροποιημένη) μορφή $y = \Phi \theta_\lambda^*$ (1-25). Σε αυτήν, όπως εξηγήσαμε όλα τα μεγέθη είναι γνωστά και μετρήσιμα/υπολογίσιμα εκτός από το διάνυσμα θ_λ —οπότε μπορούμε να προχωρήσουμε την ανάλυσή μας.

[1.2.2] Εισαγωγή Σφάλματος Πόλωσης στο Σύστημα

Για να καταλήξουμε στην (1-25) θεωρήσαμε γνωστή, μέσω της (1-1), την δομή του συστήματος που μελετάμε. Όπως αναφέραμε όμως, σε πρακτικές εφαρμογές πάντα υπάρχουν άγνωστες μη γραμμικότητες. Έτσι, ακόμα και αν γνωρίζαμε το προς εκτίμηση θ^* , η έξοδος του συστήματος στην πράξη δεν θα περιγράφονταν από την (1-25) αλλά από την:

$$y = \Phi\theta_\lambda^* + \omega, \quad (1-26)$$

στην οποία, ο όρος ω μπορεί να θεωρηθεί ως το σφάλμα προσέγγισης της εξόδου y του πραγματικού συστήματος από το $\Phi\theta_\lambda^*$ και στην ανάλυση που κάνουμε για αυτό ισχύει ότι: $\|\omega(t)\| \leq \bar{\omega}, \forall t \geq 0$, για κάποια άγνωστη σταθερά $\bar{\omega} > 0$.

Για να λύσουμε το πρόβλημα μπορούμε να χρησιμοποιήσουμε την διακοπτική σ -τροποποίηση η οποία, σύμφωνα με την θεωρία, εξασφαλίζει τόσο ότι δεν θα έχουμε παραμετρική ολίσθηση όσο και σύγκλιση του σφάλματος εξόδου στο 0 (θεωρητικά). Εφόσον έχουμε ήδη φέρει το σύστημα στην μορφή (1-25), στην συνέχεια θα σχεδιάσουμε/εφαρμόσουμε την μέθοδο κλίσης με συνεχή διακοπτική σ -τροποποίηση για την εύρεση των παραμέτρων θ_λ^* .

Άρα, σύμφωνα με την θεωρία, το σύστημα που χρειάζεται να λύσουμε για την εύρεση των παραμέτρων $\hat{\theta}_\lambda^*$, αν δεν είχαμε περιορισμούς για τις επιτρεπτές τιμές αυτών, είναι το εξής:

$$\begin{aligned} y &= \Phi\theta_\lambda^* + \omega, \quad |\theta_\lambda^*| \leq M \\ \hat{y} &= \Phi\hat{\theta}_\lambda^* \\ \dot{\hat{\theta}}_\lambda^* &= -\sigma_\delta \Gamma \hat{\theta}_\lambda^* - \Gamma \nabla K(\hat{\theta}_\lambda^*) \\ \sigma_\delta &= \begin{cases} 0, & \text{αν } |\hat{\theta}_\lambda^*| < M \\ \sigma \left(\frac{|\hat{\theta}_\lambda^*|}{M} - 1 \right), & \text{αν } M \leq |\hat{\theta}_\lambda^*| \leq 2M, \\ \sigma, & \text{αν } |\hat{\theta}_\lambda^*| > 2M \end{cases} \end{aligned} \quad (1-27)$$

όπου:

- Το $K(\hat{\theta}_\lambda^*)$ είναι η συνάρτηση κόστους που θέλουμε να ελαχιστοποιήσουμε για $\hat{\theta}_\lambda^* \in \Theta$. Ως συνάρτηση κόστους επιλέγουμε την $K(\hat{\theta}_\lambda^*) = \frac{\|e\|^2}{2} = \frac{\|y - \Phi\hat{\theta}_\lambda^*\|^2}{2}$. Παραγωγίζοντάς την παίρνουμε $\nabla K(\hat{\theta}_\lambda^*) = -\Phi^\top e$.
- Το $M > 0$ είναι σταθερά την οποία πρέπει να θέσουμε σε κατάλληλη τιμή—μέσω trial and error—για να ικανοποιεί την: $|\theta_\lambda^*| \leq M$ (άρα πρέπει να οριστεί ίση με έναν αρκούντως μεγάλο θετικό αριθμό).
- Το $\sigma > 0$ είναι ο συντελεστής της σ -τροποποίησης. Θα μπορούσαμε να ορίζουμε διαφορετικό σ για κάθε μία εκ των παραμέτρων προς εκτίμηση, $\hat{\theta}_i$, όμως για απλότητα διατηρούμε ίδια τιμή σε όλους.

Για να αξιοποιεί ο αλγόριθμος πραγματικού χρόνου που σχεδιάζουμε την πληροφορία ότι $\hat{\theta}_\lambda^* \in {}^\lambda\Theta$, στην ενημέρωση του $\hat{\theta}_\lambda^*$ που παραθέσαμε στο παραπάνω σύστημα της (1-27) θα εφαρμόσουμε προβολή στο σύνολο ${}^\lambda\Theta_b$ μέσω της σχέσης (1-15) (που αναπτύξαμε για τις συνθήκες περιορισμού του προβλήματος). Έτσι, μπορούμε να εγγυηθούμε ότι το $\hat{\theta}_\lambda^*$ δεν θα βγει ποτέ εκτός του συνόλου ${}^\lambda\Theta$.

Συμβολισμός: Το σύνολο ${}^\lambda\Theta$ αποτελεί το σύνολο όλων $\hat{\theta}_\lambda^*$ που ικανοποιούν τους περιορισμούς για τα a_{11} και b_2 . Υπάρχει ένα προς ένα αντιστοιχία των στοιχείων του με αυτών του συνόλου Θ .

Παρατήρηση

Ο αλγόριθμος που θέλουμε να σχεδιάσουμε είναι πραγματικού χρόνου. Επομένως πρέπει στο σύστημα εξισώσεων της (1-27) να συμπεριλάβουμε και τον μηχανισμό με τον οποίο βρίσκουμε τις μη μηδενικές συνιστώσες του πίνακα Φ , για κάθε δεδομένη χρονική στιγμή. Από την (1-25) βλέπουμε ότι για τα μη μηδενικά στοιχεία του πίνακα Φ (και για $\lambda_1 = \lambda_2 = \lambda$) ισχύει:

$$\Phi_{11} = \frac{X_1}{s + \lambda} \xrightarrow{ILT} \dot{\Phi}_{11}(t) = x_1(t) - \lambda\Phi_{11}(t), \quad \Phi_{11}(t) = \Phi_{23}(t) \text{ (για κοινό φίλτρο } \lambda),$$

$$\Phi_{12} = \frac{X_2}{s + \lambda} \xrightarrow{ILT} \dot{\Phi}_{12}(t) = x_2(t) - \lambda\Phi_{12}(t), \quad \Phi_{12}(t) = \Phi_{24}(t) \text{ (-//-),}$$

$$\Phi_{15} = \frac{U}{s + \lambda} \xrightarrow{ILT} \dot{\Phi}_{15}(t) = u(t) - \lambda\Phi_{15}(t), \quad \Phi_{15}(t) = \Phi_{26}(t) \text{ (-//-).}$$

Επομένως, για να έχουμε κάθε χρονική στιγμή την τρέχουσα τιμή του πίνακα $\Phi(t)$, στην επίλυση του συστήματος των διαφορικών με *ode45* συμπεριλαμβάνουμε και τις παραπάνω τρεις.

[1.2.3] Μοντελοποίηση ω στο MATLAB

Στην προηγούμενη ενότητα αναφέραμε τον τρόπο με τον οποίο θα εφαρμόσουμε την διακοπτική σ-τροποποίηση (συνεχής εκδοχή) για την εύρεση των παραμέτρων $\hat{\theta}^*$. Με αυτήν εξασφαλίζουμε πως δεν θα έχουμε παραμετρική ολίσθηση και ότι το σφάλμα εξόδου, $e = x - \hat{x}$, θα μπορεί να τείνει στο μηδέν. Το μόνο που μένει να μοντελοποιήσουμε ώστε να μπορούμε να περάσουμε το σύστημα της (1-27)¹ στο MATLAB είναι το ω .

Το $\omega(t) \in \mathbb{R}^2$ πολώνει την (1-25) και έτσι μας οδηγήσει στην (1-26). Γενικά, το $\omega(t)$ μπορεί να μοντελοποιηθεί ως μικρή διαταραχή χαμηλής συχνότητας. Ένας απλός τρόπος μοντελοποίησης θα ήταν να ορίσουμε και τις δύο συνιστώσες ίσες με κάποιο ημίτονο συχνότητας ω_f , δηλαδή:

$$\omega(t) = \begin{bmatrix} 0.8\bar{\omega} \sin(\omega_f t) \\ 0.6\bar{\omega} \sin(\omega_f t) \end{bmatrix}.$$

Παρόλο που το παραπάνω αποτελεί έναν απλό τρόπο μοντελοποίησης του ω , θα ισχύει $\|\omega\| = \bar{\omega}, \forall t$.

Για να αποφύγουμε περιπτώσεις σταθερού πλάτους επιλέγουμε για κάθε συνιστώσα να δώσουμε μία διαφορετική μικρή συχνότητα και μια διαφορά φάσης μεταξύ τους. Έτσι, στο αρχείο *src/computeBiasError.m* μοντελοποιήσαμε την παρακάτω συνάρτηση για το σφάλμα πόλωσης:

$$\omega(t) = \begin{bmatrix} 0.7\bar{\omega} \sin(\omega_{f1} t) \\ 0.6\bar{\omega} \cos(\omega_{f2} t + \pi/4) \end{bmatrix}, \quad (1-28)$$

κάνοντας κατάλληλο έλεγχο έτσι ώστε, αν έχουμε κάποιο πλάτος μεγαλύτερο του $\bar{\omega}$ να το κάνουμε scale.

Τα ω_{f1}, ω_{f2} ορίζονται στο script *mainTopic1b.m*, στην μεταβλητή **omega_freq**, και είναι ίσα με 0.7 και 1.2 αντίστοιχα, για όλες τις περιπτώσεις $\bar{\omega}$. Στο αρχείο *normOmega.m* υπολογίζουμε και κάνουμε plot την νόρμα του ω για να δούμε το πως αυτή κυμαίνεται και αν ικανοποιεί τους περιορισμούς.

¹Όπως το αναλύσαμε στην προηγούμενη ενότητα—μαζί με τα σχόλια και τις παρατηρήσεις που το συνοδεύουν.

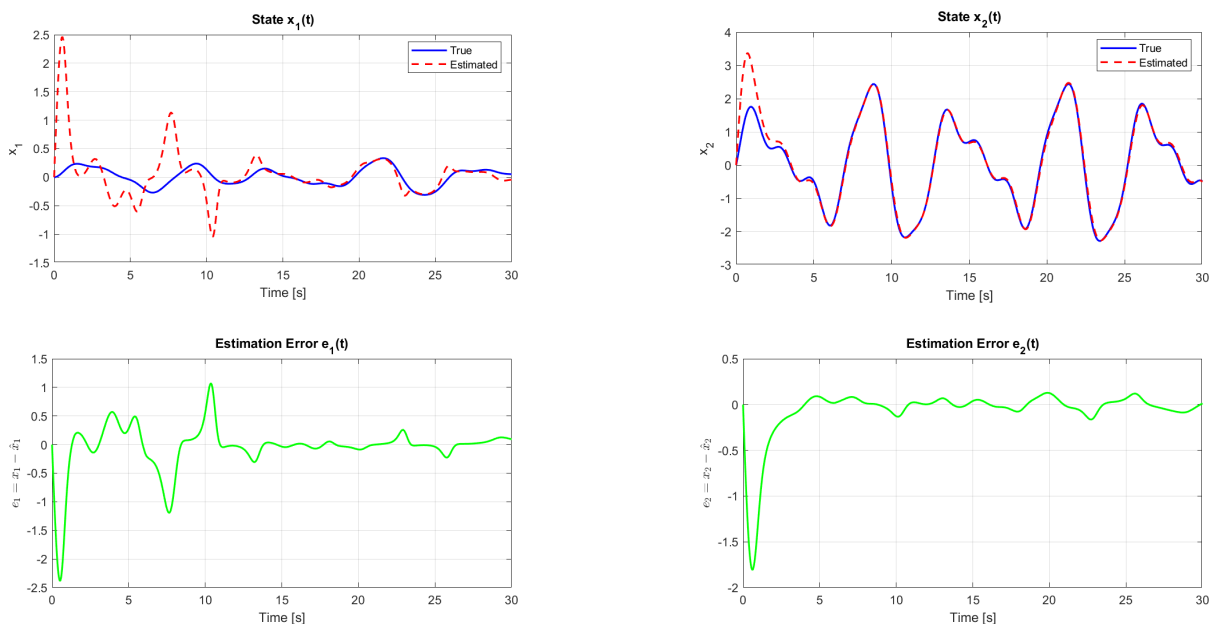
Σημειώσεις

1. Από το αρχείο *normOmega.m* είναι φανερό ότι με τον τύπο που δώσαμε για την $\omega(t)$, το $\bar{\omega}$ αποτελεί το κατώτερο άνω φράγμα για την συνάρτηση της νόρμας $\|\omega(t)\|$ —ακόμη και αν δεν κάνουμε έλεγχο για πλάτος μεγαλύτερο του $\bar{\omega}$ και scale. Επομένως, ο έλεγχος αυτός είναι καθαρά προαιρετικός για την επιλεγμένη μοντελοποίηση που παραθέσαμε.
2. Θα μπορούσαμε να μοντελοποιήσουμε το σφάλμα πόλωσης $\omega(t)$, ως μια στοχαστική διαδικασία. Σε αυτήν την περίπτωση θα μπορούσαμε να είχαμε για παράδειγμα το πλάτος $\|\omega(t)\|$ να είναι λευκός θόρυβος, με έλεγχο για την τήρηση του περιορισμού $\|\omega(t)\| \leq \bar{\omega}$. Ωστόσο, επιλέξαμε να διατηρήσουμε τη συνάρτηση πλήρως ντετερμινιστική, ώστε να αποφύγουμε την ανάγκη εκτέλεσης στατιστικού ελέγχου επί των αποτελεσμάτων και να διασφαλίσουμε την αναπαραγωγικότητα των ίδιων γραφικών παραστάσεων και τιμών σε κάθε εκτέλεση του κώδικα.

[1.2.4] Αποτελέσματα MATLAB και Συμπεράσματα

Εκτελώντας το script του αρχείου *mainTopic1b.m* για $\text{omega_bar} = 0.3$, προκύπτουν τα παρακάτω αποτελέσματα. Για την είσοδο $u(t)$ χρησιμοποιήθηκε η ίδια με αυτήν που αναφέρθηκε στο Ερώτημα 1Α.

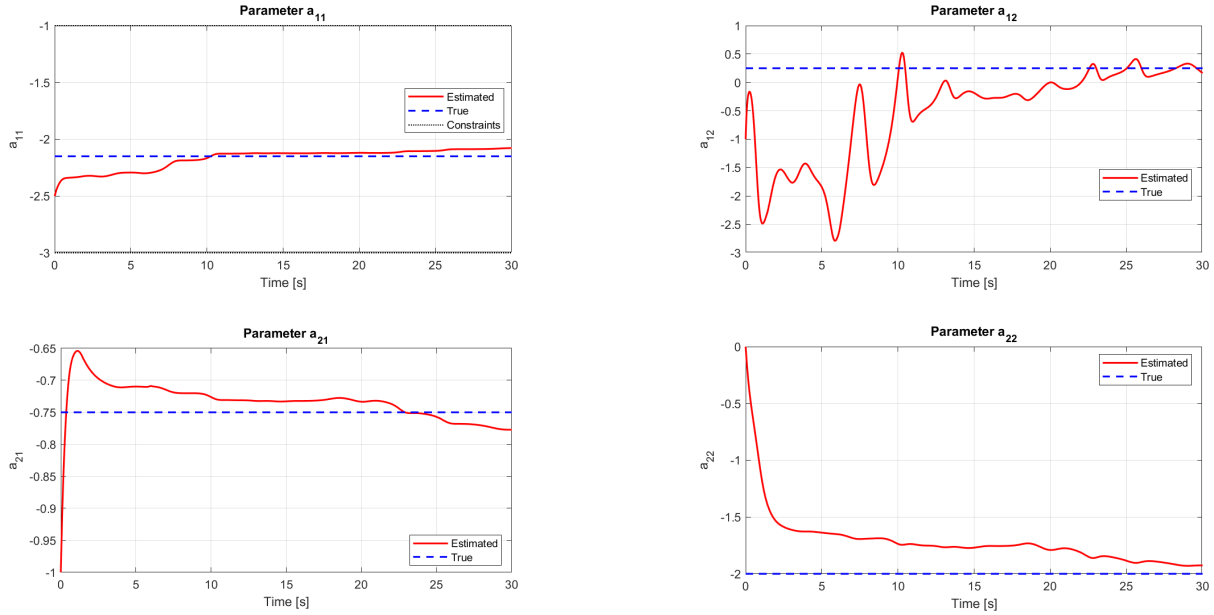
Από τις γραφικές παραστάσεις του Σχήματος 5, παρατηρούμε ότι το σφάλμα του συστήματος εκτίμησης, αν και αρχικά παρουσιάζει μεγαλύτερη απόκλιση από την πραγματική συμπεριφορά του συστήματος, μειώνεται σταδιακά με την πάροδο του χρόνου και τελικά συγχλίνει κοντά στο μηδέν. Αξιοσημείωτο είναι επίσης ότι η εκτίμηση του x_1 είναι εμφανώς λιγότερο ακριβής σε σύγκριση με εκείνη του x_2 , η οποία δείχνει να προσαρμόζεται καλύτερα στην πραγματική τιμή με την πάροδο του χρόνου. Αυτό οφείλεται στις επιλογές των παραμέτρων που έχουμε κάνει (λ , Γ , σ , κλπ...) αλλά και στην επιλογή της εισόδου μας.



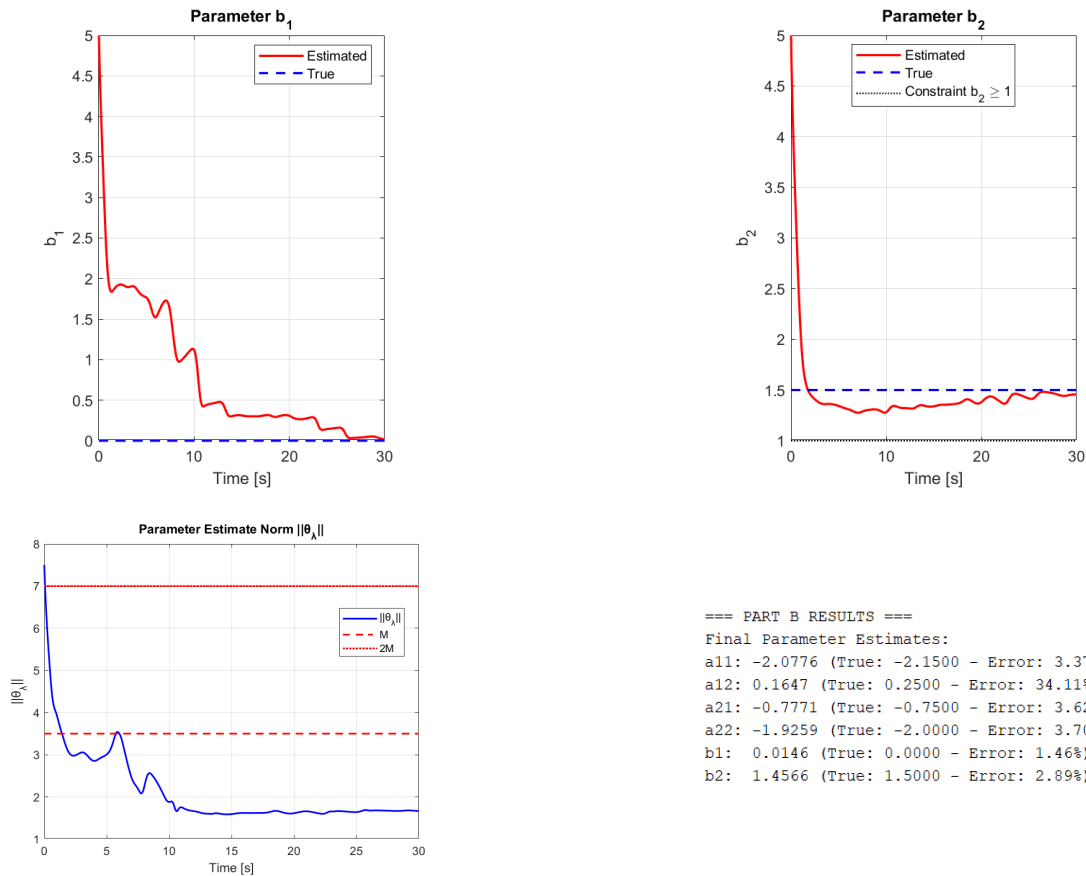
Σχήμα 5

Όσον αφορά τις γραφικές παραστάσεις των Σχημάτων 6 και 7, παρατηρείται ότι με τον σχε-

διασμένο αλγόριθμο αποφεύγεται η παραμετρική ολίσθηση και ότι οι εκτιμήσεις των παραμέτρων τείνουν να συγκλίνουν στις πραγματικές τους τιμές.



Σχήμα 6



Σχήμα 7

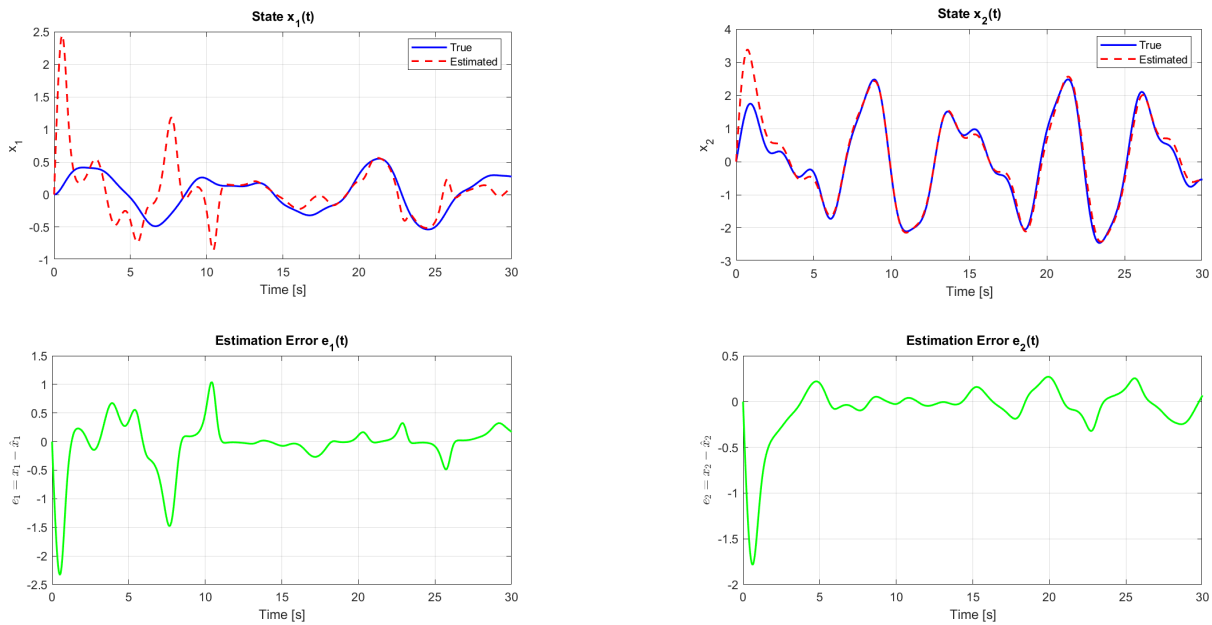
Τέλος, στο Σχήμα 7, στην κάτω σειρά, απεικονίζεται αριστερά το μέτρο του διανύσματος $\hat{\theta}_\lambda^*$. Από αυτή την πληροφορία μπορούμε να κατανοήσουμε ποιος κλάδος της διακοπτικής σ-τροποποίησης (συνεχής εκδοχή) ενεργοποιείται για τον υπολογισμό του νέου $\hat{\theta}_\lambda^*$. Είναι σαφές ότι για μεγάλες τιμές του μέτρου $\|\hat{\theta}_\lambda^*\|$ παρατηρούνται και υψηλοί ρυθμοί μεταβολής του, καθώς η τιμή της $\bar{\sigma}$ είναι μη μηδενική όταν $\|\hat{\theta}_\lambda^*\| > M$ και φτάνει τη μέγιστη τιμή της για $\|\hat{\theta}_\lambda^*\| > 2M$.

Στην κάτω δεξιά περιοχή του Σχήματος 7 παρουσιάζονται τα σχετικά σφάλματα των τελικών εκτιμήσεων, τα οποία είναι ιδιαίτερα μικρά και συγκρίσιμα ως τάξη μεγέθους με αυτά του Ερωτήματος 1Α.

Εκτελώντας εκ νέου το script του αρχείου *mainTopic1b.m*, αυτή τη φορά τροποποιώντας μόνο την τιμή της παραμέτρου **omega_bar** σε 1, προκύπτουν τα εξής νέα αποτελέσματα.

Από τις γραφικές παραστάσεις του Σχήματος 8, παρατηρείται ότι το σφάλμα του συστήματος εκτίμησης δεν μειώνεται πλέον τόσο αποτελεσματικά όσο στην προηγούμενη περίπτωση με τη μικρότερη τιμή του $\bar{\omega}$. Το σύστημα αναγνώρισης φαίνεται να απαιτεί περισσότερο χρόνο για να προσαρμοστεί "κοντά" στις πραγματικές τιμές, χωρίς ωστόσο να επιτυγχάνει το ίδιο επίπεδο ακρίβειας. Παρότι το σφάλμα μειώνεται σταδιακά, δεν φαίνεται να σταθεροποιείται στο μηδέν².

Επιπλέον, η εκτίμηση του x_1 παραμένει λιγότερο ακριβής σε σχέση με εκείνη του x_2 , η οποία, παρά την αύξηση του πλάτους του σφάλματος πόλωσης, καταφέρνει τελικά να προσεγγίσει καλύτερα την πραγματική συμπεριφορά. Το γεγονός αυτό υποδεικνύει ότι ο αλγόριθμος παρουσιάζει ορισμένο βαθμό ανοχής σε μεταβολές της τιμής του $\bar{\omega}$.



Σχήμα 8

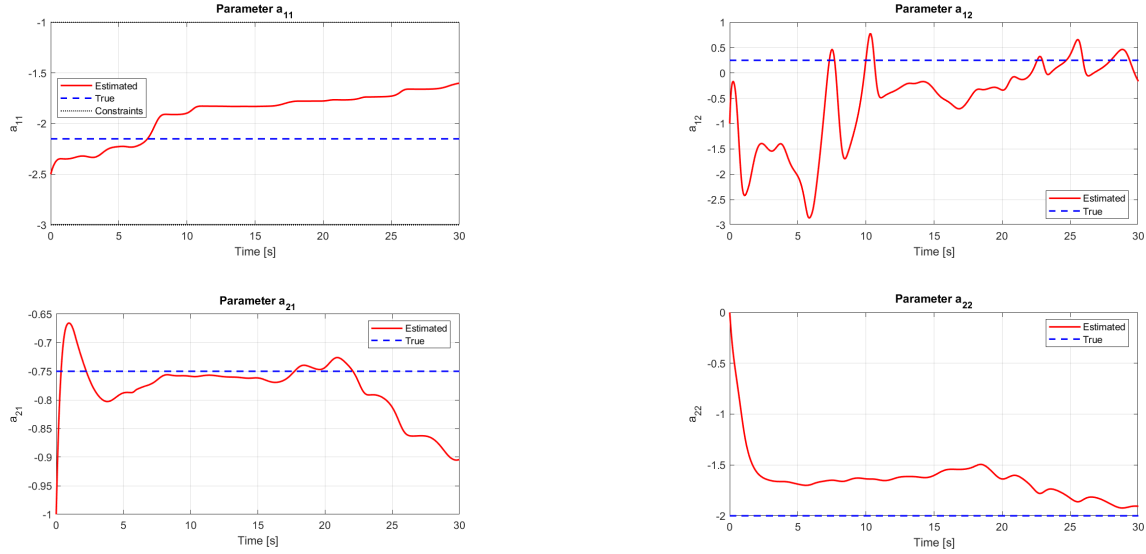
Αναφορικά με τις παραστάσεις των Σχημάτων 9 και 10, παρατηρείται ότι ο αλγόριθμος συνεχίζει να αποτρέπει την παραμετρική ολίσθηση, με τις εκτιμήσεις των παραμέτρων να περιορίζονται σε περιοχές κοντά στις πραγματικές τιμές. Παρ' όλα αυτά, τα αποτελέσματα είναι χειρότερης ακρίβειας σε σύγκριση με την προηγούμενη εκτέλεση.

Για τη βελτίωση της απόδοσης του αλγορίθμου υπό υψηλότερες τιμές του $\bar{\omega}$, θα μπορούσε να εξεταστεί η τροποποίηση συγκεκριμένων παραμέτρων, όπως για παράδειγμα οι τιμές του διαγώνιου

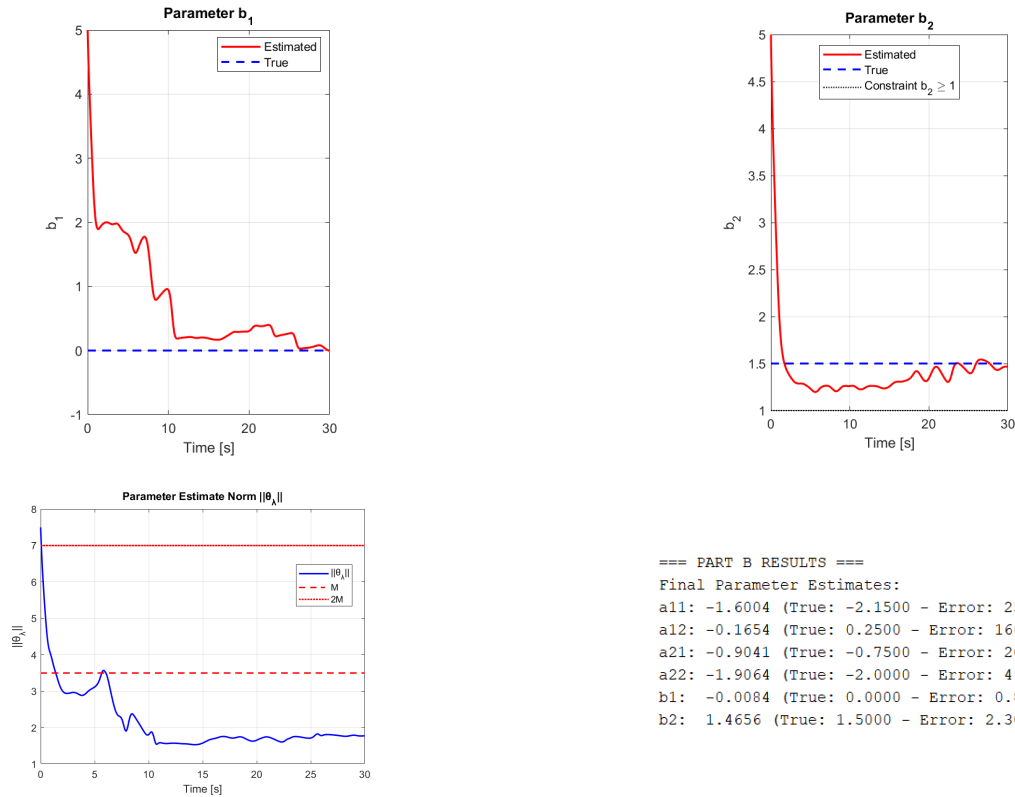
²Παρόλα αυτά, όπως προκύπτει από τη σύγκριση των γραφικών παραστάσεων για τις διαφορετικές τιμές του $\bar{\omega}$, το σφάλμα φράσσεται από τιμή που εξαρτάται από το $\bar{\omega}$.

πίνακα Γ (learning rate). Μέσω fine tuning, είναι πιθανό να επιτυγχάνονταν εκτιμήσεις παραμέτρων πιο κοντά στις επιθυμητές, ακόμη κι αν η προσέγγιση του πραγματικού μη γραμμικού συστήματος παρέμενε περιορισμένη λόγω του αυξημένου σφάλματος πόλωσης $\bar{\omega}$.

Η απόφαση να μην πραγματοποιηθεί τέτοια ρύθμιση έγινε, προκειμένου να μελετηθεί απομονωμένα η επίδραση της μεταβολής του $\bar{\omega}$, χωρίς την ταυτόχρονη αλλαγή άλλων παραμέτρων του συστήματος.



Σχήμα 9



```

=== PART B RESULTS ===
Final Parameter Estimates:
a11: -1.6004 (True: -2.1500 - Error: 25.56%)
a12: -0.1654 (True: 0.2500 - Error: 166.16%)
a21: -0.9041 (True: -0.7500 - Error: 20.55%)
a22: -1.9064 (True: -2.0000 - Error: 4.68%)
b1: -0.0084 (True: 0.0000 - Error: 0.84%)
b2: 1.4656 (True: 1.5000 - Error: 2.30%)

```

Σχήμα 10

Θέμα 2

Στόχος του θέματος αυτού είναι η μελέτη της διαδικασίας επιλογής και αξιολόγησης μοντέλου για την προσέγγιση αγνώστου μη-γραμμικού δυναμικού συστήματος της μορφής:

$$\dot{x}(t) = f(x(t), u(t), \theta), \quad f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}, \quad (2-1)$$

με $u(t) \in \mathbb{R}$ και $x(t) \in \mathbb{R}$ να είναι μετρήσιμα και να αποτελούν την είσοδο και την έξοδο του συστήματος, αντίστοιχα, ενώ $\theta = [\theta_1, \theta_2]^\top \in \mathbb{R}^2$ είναι ένα σταθερό διάνυσμα παραμέτρων.

Για το πραγματικό μη γραμμικό σύστημα που περιγράφεται από την (2-1) ισχύει:

$$f(x, u, \theta) = -x^3 + \theta_1 \tanh(x) + \theta_2 \frac{1}{1+x^2} + u, \quad (2-2)$$

όπου $\theta_1, \theta_2 \in [0.5, 2]$ και $u(t)$ είναι δικής μας επιλογής.

Σε αυτό το θέμα το δοθέν σύστημα αποτελεί ένα "Μαύρο Κουτί", μιας και στην πράξη δεν έχουμε καμία πληροφορία για την δομή του. Αν S είναι το πραγματικό σύστημα, τότε αυτό παίρνει ως είσοδο μια οποιαδήποτε συνάρτηση της επιλογής μας, $u(t)$, και δίνει στην έξοδό του την (επίσης μετρήσιμη) συνάρτηση $x(t) = S[u(t)]$.

[2.1] Ανάλυση Βημάτων

Ένα πρώτο βήμα για την κατανόηση του συστήματός μας (μιας και δεν έχουμε καμία πληροφορία για αυτό) είναι να ελέγξουμε αν αυτό είναι γραμμικό και χρονοαμετάβλητο. Για να το κάνουμε αυτό μπορούμε να δοκιμάσουμε να δώσουμε στο σύστημα διάφορα ζεύγη εισόδων $\{u_1(t), u_2(t)\}$ και να ελέγχουμε αν για αυτά ισχύει η σχέση $S[au_1(t) + bu_2(t)] = aS[u_1(t)] + bS[u_2(t)]$, $a, b \in \mathbb{R}$. Αντίστοιχα για τον έλεγχο του αμετάβλητου ως προς το χρόνο μπορούμε να κάνουμε τον έλεγχο για το αν ισχύει η συνεπαγωγή: $S[u(t)] = x(t) \Rightarrow S[u(t - t_0)v(t - t_0)] = x(t - t_0)v(t - t_0)$, όπου με $v(t)$ ορίζουμε την βηματική συνάρτηση.

Τα παραπάνω βέβαια αποτελούν απλώς μία πρώτη επαφή με το σύστημά μας, και μπορούν μόνο να χρησιμοποιηθούν, όχι τόσο για να αποδείξουν, αλλά για να καταρρίψουν την υπόθεση για γραμμικό ή χρονοαμετάβλητο σύστημα.

Στην συνέχεια της ανάλυσης θα παρουσιάσουμε την διαδικασία επιλογής δομής που ακολουθήσαμε για την προσέγγιση του αγνώστου συστήματος της (2-1).

[2.1.1] Διαδικασία Επιλογής Δομής

Ο στόχος εδώ είναι να προσεγγίσουμε την άγνωστη συνάρτηση $f(x, u, \theta)$ χρησιμοποιώντας μια δομή μοντέλου που συντίθεται από *βασικές συναρτήσεις*. Για την διαδικασία επιλογής δομής μοντέλου—δηλαδή την επιλογή ενός συνόλου βασικών συναρτήσεων **και** τον προσδιορισμό της πολυπλοκότητάς τους, όπως για παράδειγμα το πλήθος των όρων του—ακολουθούμε τα παρακάτω βήματα:

1. Επιλογή Βασικών Συναρτήσεων:

Επιλέγουμε μια "οικογένεια" βασικών συναρτήσεων για να αναπαραστήσουμε τη $f(x, u, \theta)$. Συνήθεις επιλογές περιλαμβάνουν πολυώνυμα $\{1, x, x^2, x^3, \dots, u, u^2, \dots\}$, γκαουσιανές βασικές συναρτήσεις (RBFs) $\phi_i(x) = \exp\left(-\frac{(x-c_i)^2}{2\sigma_i^2}\right)$, ρητές συναρτήσεις όπως η ίδια η $\frac{1}{1+x^2}$, κ.α.

2. Επιλογή Υποψήφιων Μοντέλων:

Ορίζουμε πολλαπλές δομές μοντέλων με αυξανόμενη πολυπλοκότητα. Οι επιλογές που κάνουμε παρατίθενται παρακάτω:

- Μοντέλο 1 (Απλό Πολυώνυμο):

$$\hat{f} = \hat{\theta}_0 + \hat{\theta}_1 x + \hat{\theta}_2 x^2 + \hat{\theta}_3 x^3 + \hat{\theta}_4 u.$$

- Μοντέλο 2 (Πολυώνυμο + Tanh):

$$\hat{f} = \hat{\theta}_0 + \hat{\theta}_1 x + \hat{\theta}_2 x^3 + \hat{\theta}_3 \tanh(x) + \hat{\theta}_4 u.$$

- Μοντέλο 3 (Γκαουσιανές RBFs):

$$\hat{f} = \sum_{i=0}^k \hat{\theta}_i \exp\left(-\frac{(x - c_i)^2}{2\sigma^2}\right) + \hat{\theta}_{k+1} u, \text{ με κέντρα } c_i \text{ και πλάτος } \sigma.$$

Σχόλιο

Σε κανένα εκ των μοντέλων που αναφέρουμε δεν έχουμε βάλει όρο που να περιλαμβάνει την είσοδο σε κάποια δύναμη. Επίσης, δεν έχουμε δώσει κάποια μη-γραμμική σχέση που να "μπλέκει" την είσοδο με την έξοδο σε έναν όρο. Ο λόγος που το κάνουμε αυτό είναι επειδή ξεκινάμε ελέγχοντας πρώτα αν απλά μοντέλα μπορούν να προσεγγίσουν το πραγματικό σύστημα. Αν το αποτέλεσμα δεν είναι ικανοποιητικό, τότε μπορούμε να περάσουμε σε πιο περίπλοκα μοντέλα.

4. Επιλογές Εισόδου $u(t)$:

Η είσοδος $u(t)$ θα πρέπει να επιλεγεί κατάλληλα ώστε να διεγείρει το σύστημα έτσι, ώστε να σχηματίζεται η δυναμική του. Καλές επιλογές που θα χρησιμοποιηθούν στην προσομοίωση είναι αθροίσματα ημιτονοειδών, μιας και με αυτά καλύπτεται ένα εύρος συχνοτήτων.

5. Προσομοίωση του Συστήματος:

Τέλος, χρησιμοποιούμε το αληθινό σύστημα $\dot{x} = -x^3 + \theta_1 \tanh(x) + \theta_2 \frac{1}{1+x^2} + u$ με επιλεγμένα $\theta_1, \theta_2 \in [0.5, 2]$ (επιλέγουμε για την συνέχεια της ανάλυσής μας $\theta_1 = 1.5, \theta_2 = 2.0$) για να παράγουμε δεδομένα εξόδου για την προσαρμογή του μοντέλου.

[2.1.2] Εκτίμηση Παραμέτρων σε Πραγματικό Χρόνο

Για την εκτίμηση παραμέτρων σε πραγματικό χρόνο, το μοντέλο πρέπει να προσαρμόζει τις παραμέτρους του, $\hat{\theta}$, καθώς νέα δεδομένα $(x(t), u(t))$ γίνονται διαθέσιμα. Δεδομένου ότι το σύστημα είναι μη γραμμικό, μπορούμε να χρησιμοποιήσουμε μεθόδους βασισμένες σε κλίση όπως την αναδρομική μορφή της μεθόδου ελαχίστων τετραγώνων (RLS) ή την μέθοδο κλίσης, όπως την παραθέσαμε και στην ανάλυση του προηγούμενου θέματος.

Από την θεωρία έχουμε ότι τα RLS και η μέθοδος κλίσης μπορούν να εφαρμοστούν σε ένα μοντέλο γραμμικό ως προς τις παραμέτρους:

$$\dot{x}(t) \approx \hat{f}(x, u, \hat{\theta}) = \phi(x, u)^\top \hat{\theta} = \hat{\theta}^\top \phi(x, u),$$

όπου $\phi(x, u) = [\phi_0(x), \dots, \phi_n(x), u]^\top$ είναι το διάνυσμα των οπισθοδρομητών και $\hat{\theta} = [\hat{\theta}_0, \dots, \hat{\theta}_{n+1}]^\top$.

Προφανώς, ακολουθώντας την διαδικασία που αναλύσαμε στο Ερώτημα Β του θέματος 1, μπορούμε με χρήση κατάλληλου φίλτρου πρώτης τάξης να διώξουμε την παράγωγο $\dot{x}(t)$ και να φέρουμε το σύστημα σε μορφή $y = \theta^\top \phi$.

Η μέθοδος κλίσης για το πρόβλημά μας δίνει την ανανέωση του θ σύμφωνα με την:

$$\dot{\hat{\theta}} = \gamma(y - \hat{\theta}^\top \phi)\phi = \gamma e\phi,$$

όπου $e = y - \hat{\theta}^\top \phi$.

Λύνοντας το σύστημα των διαφορικών που προκύπτουν (με ανάλογο τρόπο με αυτόν που αναλύσαμε ήδη στο Θέμα 1) θα είμαστε σε θέση να εκτιμήσουμε το διάνυσμα $\hat{\theta}$.

[2.1.3] Αξιολόγηση και Σύγκριση Μοντέλων

Για να συγκρίνουμε τις υποψήφιες δομές μοντέλων, χρειάζεται να αξιολογήσουμε την απόδοσή τους χρησιμοποιώντας ποσοτικές μετρικές και ποιοτική ανάλυση (οπτικοποίηση προβλέψεων).

Μερικές μετρικές αξιολόγησης που μπορούμε να λάβουμε υπόψιν είναι:

1. Μέσο Τετραγωνικό Σφάλμα (MSE)

$$\text{MSE} = \frac{1}{T} \int_0^T (x(t) - \hat{x}(t))^2 dt,$$

προσεγγιζόμενο διακριτά σε χρονικά βήματα. Χαμηλότερο MSE δείχνει καλύτερη προσαρμογή.

2. Κριτήριο Πληροφορίας Akaike (AIC)

$$\text{AIC} = 2k + N \ln(\text{MSE}),$$

όπου k είναι ο αριθμός των παραμέτρων, και N είναι ο αριθμός των σημείων δεδομένων. Το AIC εξισορροπεί την προσαρμογή του μοντέλου (MSE) και την πολυπλοκότητα (αριθμός παραμέτρων).

3. Εγκάρσια Αξιολόγηση (Cross Validation - CV)

Όλες οι δομές υποψήφιων μοντέλων είναι χρονοαμετάβλητα συστήματα³.

- Εκτελούμε την προσομοίωση του πραγματικού συστήματος για το χρονικό διάστημα $[0, T]$ και καταγράφουμε τα αντίστοιχα δεδομένα εισόδου-εξόδου $(x(t), y(t))$.
- Διαχωρίζουμε τα δεδομένα αυτά σε τέσσερα ισομεγέθη χρονικά διαστήματα (υποσύνολα):
 - (α') $[0, T/4]$
 - (β') $[T/4, T/2]$
 - (γ') $[T/2, 3T/4]$
 - (δ') $[3T/4, T]$
- Εκπαιδεύουμε το μοντέλο χρησιμοποιώντας τα τρία από τα τέσσερα παραπάνω υποσύνολα ως σύνολο εκπαίδευσης (training set), ενώ το εναπομείναν υποσύνολο χρησιμοποιείται ως σύνολο ελέγχου (test set), στο οποίο εφαρμόζονται οι αντίστοιχες τεχνικές αξιολόγησης. Η διαδικασία αυτή επαναλαμβάνεται συνολικά τέσσερις φορές, ώστε κάθε υποσύνολο να χρησιμοποιηθεί μία φορά ως test set.

[2.2] Αποτελέσματα

Ακολουθώντας τα βήματα που περιγράφηκαν προηγουμένως, υλοποιούμε και εκτελούμε τον κώδικα του αρχείου *mainTopic2.m*. Σε αυτόν ορίζουμε τη συνάρτηση εισόδου καθώς και τις συναρτήσεις βάσης, όπως απεικονίζονται στο Σχήμα 11.

Κατά την εκτέλεση του κώδικα, τυπώνονται logs (Σχήμα 12) που περιέχουν τα αποτελέσματα όλων των υπολογιζόμενων μετρικών για κάθε μοντέλο που εφαρμόστηκε στην προσπάθεια προσαρμογής (fit) των δεδομένων μέσω cross-validation.

³Και το σύστημα που έχουμε στην διάθεσή μας είναι χρονοαμετάβλητο—έχουμε ήδη εξηγήσει στην εισαγωγή του ερωτήματος Α, έναν απλό τρόπο να ελέγξουμε αν η υπόθεση αυτή απορρίπτεται, και εφόσον δεν απορρίπτεται μπορούμε να επιλέξουμε τις πρώτες υποψήφιες δομές μοντέλου να είναι ως προς τον χρόνο αμετάβλητες.

```

%% 2. Design Input Signal u(t)
% Rich input signal - sum of sinusoids
input_signal = @(t) sin(0.5*t) + 0.8*sin(1.2*t) + 0.6*sin(2.3*t) + 0.4*sin(3.7*t);

% Model 1: Simple Polynomial
model1_basis = @(x, u) [1; x; x^2; x^3; u]; % 5 parameters

% Model 2: Polynomial + Tanh
model2_basis = @(x, u) [1; x; x^3; tanh(x); u]; % 5 parameters

% Model 3: Gaussian RBFs
rbf_centers = [-2, -1, 0, 1, 2]; % RBF centers
rbf_sigma = 1.0; % RBF width
model3_basis = @(x, u) [exp(-(x - rbf_centers(1))^2/(2*rbf_sigma^2));
    exp(-(x - rbf_centers(2))^2/(2*rbf_sigma^2));
    exp(-(x - rbf_centers(3))^2/(2*rbf_sigma^2));
    exp(-(x - rbf_centers(4))^2/(2*rbf_sigma^2));
    exp(-(x - rbf_centers(5))^2/(2*rbf_sigma^2));
    u]; % 6 parameters

```

Σχήμα 11

Στη συνέχεια, με βάση τις μετρικές του cross-validation, επιλέγουμε το μοντέλο που παρουσιάζει το ελάχιστο σφάλμα μοντελοποίησης. Για το επιλεγμένο μοντέλο, προχωρούμε σε νέα εκπαίδευση χρησιμοποιώντας ολόκληρο το σύνολο δεδομένων (χωρίς διαχωρισμό για cross-validation) και παραθέτουμε τις τελικές τιμές των παραμέτρων θ , καθώς και τις αντίστοιχες τιμές των δεικτών root mean square error (RMSE) και mean absolute error (MAE). Οι τελικές τιμές των θ αποτελούν την απαιτούμενη πληροφορία για την πλήρη αναπαράσταση του μοντέλου (δεδομένων των συναρτήσεων βάσης του).

Σύμφωνα με τα αποτελέσματα του cross-validation, το μοντέλο με συναρτήσεις βάσης Γκαουσιανής (RBFs) επιτυγχάνει τη βέλτιστη προσαρμογή, ενώ το πολυωνυμικό μοντέλο εμφανίζει τη χαμηλότερη επίδοση.

```

==== MODEL EVALUATION RESULTS ====
Model          | CV MSE      | CV Std      | Final MSE    | AIC
-----|-----|-----|-----|-----|
Polynomial      | 0.067883    | 0.026199    | 0.026161     | -3637.12
Polynomial + Tanh | 0.056510    | 0.021945    | 0.025988     | -3643.78
Gaussian RBFs   | 0.015359    | 0.001584    | 0.024125     | -3716.24

Best model based on CV MSE: Gaussian RBFs

==== BEST MODEL PARAMETER ANALYSIS ====
Best Model: Gaussian RBFs
Final Parameters:
theta_1 = 0.0119
theta_2 = 0.1018
theta_3 = 0.4014
theta_4 = 0.7101
theta_5 = 0.5234
theta_6 = 0.3433

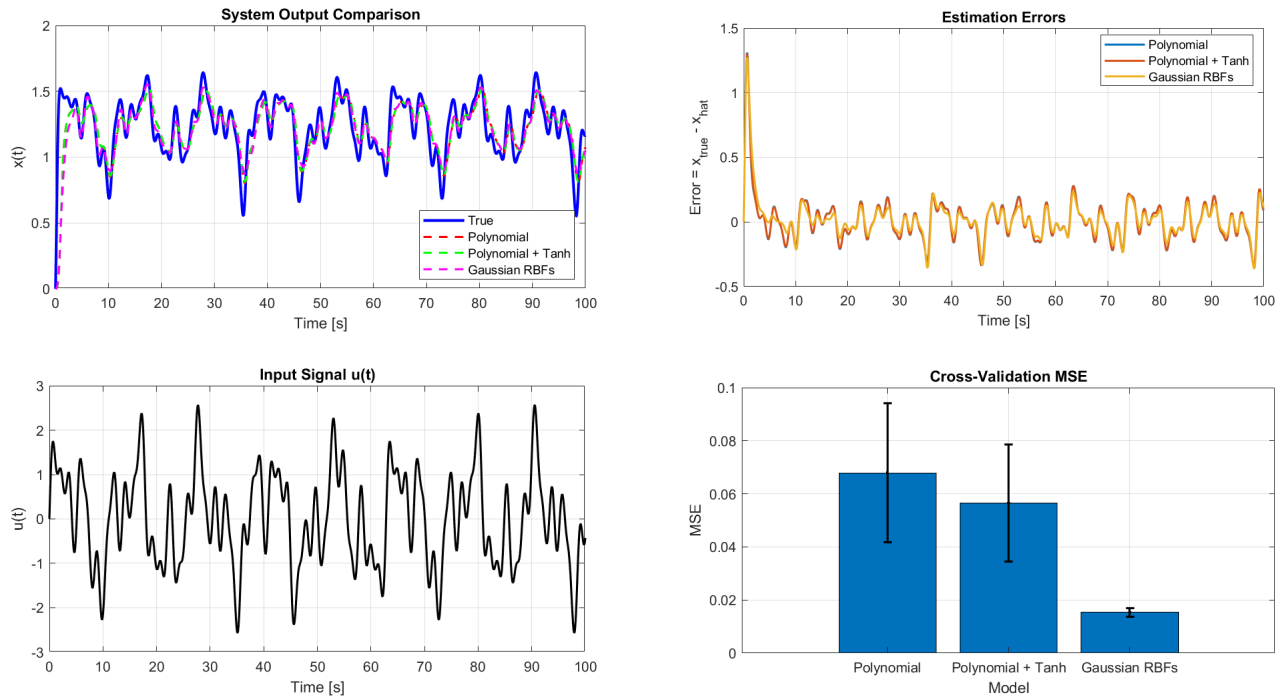
Final Error Statistics for Best Model:
RMSE: 0.155322
MAE: 0.089686

```

Σχήμα 12

Στο Σχήμα 13, παρουσιάζεται στην επάνω σειρά η απόκριση $x(t)$ του πραγματικού συστήματος και των προσεγγιστικών μοντέλων στην ίδια είσοδο, καθώς και οι αντίστοιχες γραφικές παραστάσεις

των σφαλμάτων $e(t) = x(t) - \hat{x}(t)$. Στην κάτω σειρά, απεικονίζονται η είσοδος του συστήματος (από την οποία επιλέχθηκαν χρονικά διαστήματα για το training στο CV), καθώς και το μέσο τετραγωνικό σφάλμα (mean square error) από το cross-validation συνοδευόμενο από τις αντίστοιχες τυπικές αποκλίσεις.



Σχήμα 13

Τα Σχήματα 14 περιλαμβάνουν τις γραφικές παραστάσεις των τιμών του MSE, τόσο για την περίπτωση του cross-validation όσο και για την περίπτωση απευθείας εκπαίδευσης χωρίς διαχωρισμό των δεδομένων, καθώς και του κριτηρίου Akaike (AIC). Και οι δύο μετρικές συγκλίνουν στο συμπέρασμα ότι το Gaussian μοντέλο υπερτερεί των υπολοίπων.

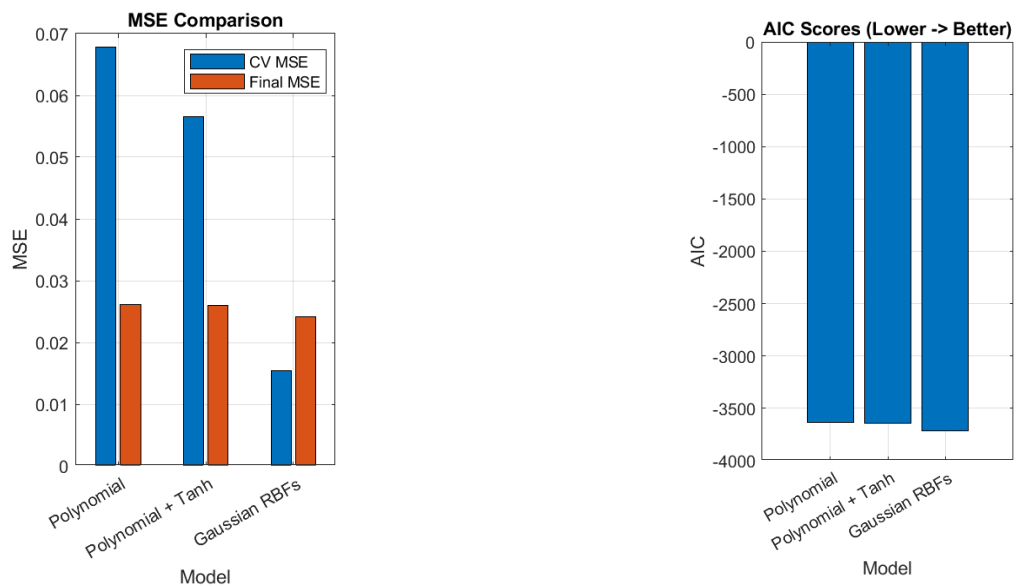
Τέλος, για λόγους πληρότητας και για πιο clean απεικόνιση των αποτελεσμάτων του τελικού/βέλτιστου⁴ μοντέλου, στο Σχήμα 15 παρατίθεται η χρονική εξέλιξη των $x(t)$ του πραγματικού συστήματος και των $\hat{x}(t)$ του επιλεγμένου μοντέλου, συνοδευόμενη από τη γραφική παράσταση του σφάλματος εκτίμησης.

Αφού διαπιστώσαμε ότι το σύστημά μας μπορεί να προσεγγιστεί ικανοποιητικά με τη χρήση συναρτήσεων βάσης τύπου Gaussian (RBFs), προχωρήσαμε στη δοκιμή διαφόρων τιμών για τις παραμέτρους γ και λ ⁵. Παράλληλα, εξετάσαμε την επίδραση της προσθαφαίρεσης συναρτήσεων βάσης, με στόχο να αξιολογήσουμε σε ποιο βαθμό επηρεάζεται η ακρίβεια του μοντέλου (κοιτώντας τα logs των μετρικών που τυπώνουμε) και κατά πόσο είναι εφικτή η βελτίωση της προσέγγισης χωρίς σημαντική αύξηση της πολυπλοκότητας του μοντέλου.

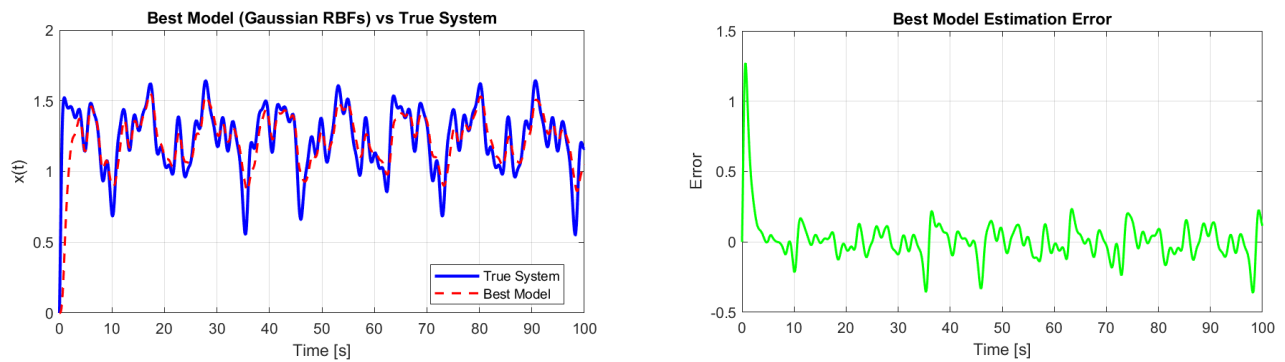
Μέσω trial and error—και μη έχοντας κάποιες προδιαγραφές για το επιθυμητό επίπεδο ακρίβειας προσεγγίσεις του πραγματικού συστήματος (μιας και αυτό εξαρτάται από την εφαρμογή στο κάθε πεδίο)—καταλήξαμε ότι το μοντέλο με τα RBFs της 11 που παρουσιάζεται στο Σχήμα 15 και έχει

⁴Βέλτιστο σε σχέση με τα υπόλοιπα που επιλέχθηκαν.

⁵Σε αυτό το σημείο εξετάσαμε και την επίδραση της αλλαγής της συνάρτησης εισόδου στο πόσο καλά μπορεί να κάνει fit το μοντέλο. Τα Gaussian RBFs μπορούν σε συνεχή σύνολα να κάνουν αρκούντως καλές προσεγγίσεις ώστε η τάξη μεγέθους του σφάλματος να μην αλλάζει δραματικά (εξού και το μικρό standard deviation στο σχήμα 13).



Σχήμα 14



Σχήμα 15

θ που φαίνονται στο 12, αποτελεί μια απλή αλλά επαρκώς ακριβή προσέγγιση της συμπεριφοράς του πραγματικού συστήματος.