ΥΠΟΥΡΓΕΙΟ ΕΘΝΙΚΗΣ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ ΕΠΕΑΕΚ
ΕΥΡΩΠΑΪΚΗ ΕΝΩΣΗ
ΣΥΓΧΡΗΜΑΤΟΔΟΤΗΣΗ
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ
ΕΥΡΩΠΑΪΚΟ ΤΑΜΕΙΟ ΠΕΡΙΦΕΡΕΙΑΚΗΣ ΑΝΑΠΤΥΞΗΣ

ΠΑΙΔΕΙΑ ΜΠΡΟΣΤΑ
2ο Επιχειρησιακό Πρόγραμμα Εκπαίδευσης και Αρχικής Επαγγελματικής Κατάρτισης

Πρόγραμμα Αναμόρφωσης Προπτυχιακών Σπουδών

Τμήματος Μαθηματικών

Πανεπιστήμιο Πατρών

Μάθημα: Αριθμητική Ανάλυση II

ΑΑΙΙ: 8η Εβδομαδα - 11 Απριλιου 2005

Πολυώνυμα Ελάχιστων Τετραγώνων (Least Squares Polynomials)

1. Πολυώνυμα ελάχιστων τετραγώνων

Κατ᾽ αρχή παρατηρούμε ότι στον προσδιορισμό του προσεγγιστικού πολυωνύμου ο βαθμός του είναι κατά ένα μικρότερος του πλήθους των σημείων παρεμβολής. Επί πλέον τα σημεία παρεμβολής είναι συνήθως πολυάριθμα, γι᾽ αυτό και **ο βαθμός του προσεγγιστικού πολυωνύμου είναι μεγάλος**. Έξ άλλου, πολλές φορές, οι τιμές $\varphi(x_k)$, $k = 0, 1, 2, \ldots, \nu$, βρίσκονται από παρατηρήσεις όπου η εισδοχή σφαλμάτων είναι αναπόφευκτη. Αυτό πάλι συνεπάγεται την **γνωστή ιδιάζουσα ευαισθησία των πολυωνύμων υψηλού βαθμού σε σφάλματα των συντελεστών τους**. Στις αρχές του 19ου αιώνα, **ο C.F. Gauss και ο A.M. Legendre επινόησαν ένα τρόπο προσδιορισμού καταλλήλου πολυωνύμου μικρότερου βαθμού** (ο βαθμός είναι σχεδόν αυθαίρετος και συνάρτηση κυρίως των αναγκών του προβλήματος), **με συντελεστές αυξημένης ακριβείας**.

Πιο συγκεκριμένα, ας υποθέσουμε ότι σ' ένα πείραμα έχουμε τα ζεύγη των παρατηρήσεων: $\{x_k, \varphi(x_k)\}$, $k = 0, 1, 2, \ldots, \nu$, τότε, είναι δυνατόν να προσδιοριστεί ένα πολυώνυμο $E(x)$ βαθμού $\mu(< \nu)$:

$$E(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_\mu x^\mu, \tag{1}$$

με τέτοιο τρόπο, ώστε ενώ δεν είναι πλέον δυνατό να πληρούνται οι σχέσεις ταύτισης: $E(x_k) = \varphi(x_k)$, $k = 0, 1, 2, \ldots, \nu$, αντ' αυτών, που εγκαταλείπονται, απαιτείται η ελαχιστοποίηση κάποιας από τις εκφράσεις:

$$\sum_{\mathbf{k=0}}^{\nu} |\mathbf{E}(\mathbf{x_k}) - \varphi(\mathbf{x_k})| \left(\equiv \|\bar{\sigma}\|_{\mathbf{1}}\right), \sum_{\mathbf{k=0}}^{\nu} \{\mathbf{E}(\mathbf{x_k}) - \varphi(\mathbf{x_k})\}^{\mathbf{2}} \left(\equiv \|\bar{\sigma}\|_{\mathbf{2}}^{\mathbf{2}}\right),$$

$$\max_{0 \le k \le \nu} |E(x_k) - \varphi(x_k)| \left(\equiv \|\bar{\sigma}\|_\infty\right),$$

όπου το διάνυσμα $\bar{\sigma}$ παρέχει τις αποκλίσεις - σφάλματα στα σημεία x_k, $k = 0, 1, 2, \ldots, \nu$:

$$\sigma \equiv \{E(x_0) - \varphi(x_0), E(x_1) - \varphi(x_1), \ldots, E(x_\nu) - \varphi(x_\nu)\}.$$

Η πιο διαδεδομένη και πλατιά χρησιμοποιούμενη εκλογή από τις τρεις, με πολυάριθμες εφαρμογές σε περιπτώσεις φυσικών και τεχνολογικών προβλημάτων, **είναι** η μεσαία περίπτωση, όπου ελαχιστοποιείται **η l_2 στάθμη του διανύσματος των αποκλίσεων** $\bar{\sigma}$. Η τεχνική αυτή καλείται και **Μ**έθοδος των **Ε**λάχιστων **Τ**ετραγώνων - Μ.Ε.Τ. (Least Squares Method).

2. Ο Μηχανσιμός της Μεθόδου των Ελαχίστων Τετραγώνων

Για εξέταση του μηχανισμού της Μ.Ε.Τ. ας μορφώσουμε με τη βοήθεια τη ς (1) την έκφραση:

$$\|\bar{\sigma}\|_2^2 = \sum_{k=0}^{\nu} \left\{a_0 + a_1 x_k + a_2 x_k^2 + \cdots + a_\mu x_k^\mu - \varphi(x_k)\right\}^2 . \tag{2}$$

Προφανώς, το β΄ μέλος της (2) αποτελεί μία συνάρτηση των παραμέτρων: $a_0, a_1, \ldots, a_\mu$· δηλαδή ισχύει:

$$\|\bar{\sigma}\|_2 = \Phi\left(a_0, a_1, a_2, \cdots, a_\mu\right). \tag{3}$$

Ως γνωστό όμως από την Ανάλυση, οι αναγκαίες συνθήκες για την ελαχιστοποίηση της (3) είναι οι εξής:

$$\frac{\partial \Phi}{\partial a_k} = 0, \quad k = 0, 1, 2, \ldots, \mu. \tag{4}$$

Οι εξισώσεις (4) είναι πλήθους $\mu + 1$ (και οι άγνωστοι $a_0, a_1, a_2, \ldots, a_\mu$ είναι ισάριθμοι) **είναι γραμμικές**, είναι δε δυνατό να αποδειχθεί ότι **η ορίζουσα των συντελεστών των αγνώστων είναι μη μηδενική**·

άρα υπάρχει πάντοτε μια λύση. Αναλυτικότερα, οι (4) δίνουν:

$$
(5) \qquad \begin{cases}
\dfrac{\partial \Phi}{\partial a_0} = 2\displaystyle\sum_{k=0}^{\nu} \Big(a_0 + a_1 x_k + a_2 x_k^2 + \cdots + a_\mu x_k^\mu - \varphi(x_k)\Big) = 0, \\
\dfrac{\partial \Phi}{\partial a_1} = 2\displaystyle\sum_{k=0}^{\nu} \Big(a_0 + a_1 x_k + a_2 x_k^2 + \cdots + a_\mu x_k^\mu - \varphi(x_k)\Big) = 0, \\
\dfrac{\partial \Phi}{\partial a_2} = 2\displaystyle\sum_{k=0}^{\nu} \Big(a_0 + a_1 x_k + a_2 x_k^2 + \cdots + a_\mu x_k^\mu - \varphi(x_k)\Big) = 0, \\
\qquad \vdots \\
\dfrac{\partial \Phi}{\partial a_\mu} = 2\displaystyle\sum_{k=0}^{\nu} \Big(a_0 + a_1 x_k + a_2 x_k^2 + \cdots + a_\mu x_k^\mu - \varphi(x_k)\Big) = 0.
\end{cases}
$$

Οι εξισώσεις (5), εάν χρησιμοποιήσουμε τον κομψό συμβολισμό του

Gauss: $\left\{x^k\right\} = \sum\limits_{\lambda=0}^{\nu} x_\lambda^k$, $k = 0, 1, 2, \ldots, 2\mu$, γράφονται ως εξής:

$$
(6) \qquad \begin{cases}
a_0\left\{x^0\right\} + a_1\left\{x^1\right\} + \cdots + a_\mu\left\{x^\mu\right\} = \{\varphi\}, \\
a_0\left\{x^1\right\} + a_1\left\{x^2\right\} + \cdots + a_\mu\left\{x^{\mu+1}\right\} = \{x\varphi\}, \\
a_0\left\{x^2\right\} + a_1\left\{x^3\right\} + \cdots + a_\mu\left\{x^{\mu+2}\right\} = \left\{x^2\varphi\right\}, \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
a_0\left\{x^\mu\right\} + a_1\left\{x^{\mu+1}\right\} + \cdots + a_\mu\left\{x^{2\mu}\right\} = \left\{x^\mu\varphi\right\}.
\end{cases}
$$

Οι εξισώσεις (6) καλούνται **«κανονικές» (normal) εξισώσεις της Μ.Ε.Τ.** Στο παραπάνω σύστημα (6) είναι ενδιαφέρουσα η συμμετρική ιδιότητα· άρα αντί των ν^2 το πλήθος των συντελεστών $\left\{x^\lambda\right\}$, απαιτούνται κατ᾽ αρχή, μόνο $\nu(\nu+1)/2$ προσδιοριστέοι συντελεστές. Τέλος, εάν συμβολίσουμε με X **τον πίνακα των παρατηρήσεων** και ψ το διάνυσμα **των τιμών** της συνάρτησης, δηλαδή:

$$X = \begin{bmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^\mu \\ 1 & x_1 & x_1^2 & \cdots & x_1^\mu \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_\nu & x_\nu^2 & \cdots & x_\nu^\mu \end{bmatrix}, \quad \bar{\psi} = \begin{bmatrix} \varphi(x_0) \\ \varphi(x_1) \\ \vdots \\ \varphi(x_\nu) \end{bmatrix},$$

τότε το σύστημα (6) γράφεται:

$$\boldsymbol{X^T X \bar{a} = X^T \bar{\psi}.} \tag{7}$$

Η (7) αποτελεί την **μορφή των κανονικών εξισώσεων** που συνήθως αξιοποιούμε για την εύρεση των παραμέτρων $\bar{a}$.

3. Εφαρμογές

(α). Ας υποθέσουμε τις παρακάτω παρατηρήσεις από κάποιο φαινόμενο: $(1,10),(2,2),(3,2),(4,5),(5,4)$, και ας εφαρμόσουμε την (7) για να προσδιορίσουμε το πρωτοβάθμιο προσεγγιστικό πολυώνυμο της Μ.Ε.Τ.:

$$E(x) = a_0 + a_1 x.$$

Πράγματι έχουμε

$$X^\top = \left[\begin{array}{ccccc} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \end{array}\right], \quad \bar{\psi} = (0, 2, 2, 5, 4),$$

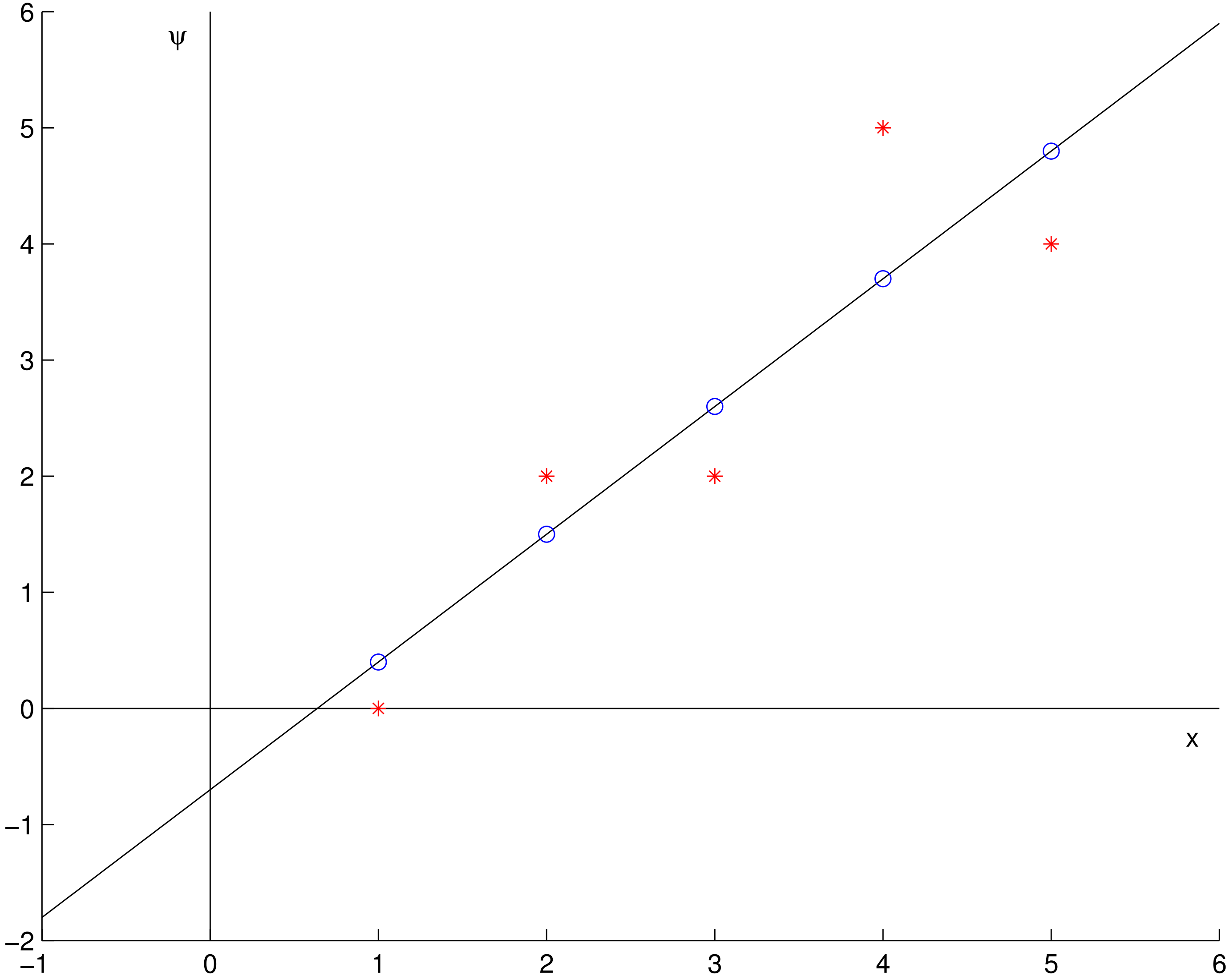
οπότε η (7) μας δίδει το σύστημα:

$$\begin{array}{rcrcl} 5a_0 & + & 15a_1 & = & 13 \\ 15a_0 & + & 55a_1 & = & 50, \end{array}$$

με λύση: $a_0 = -0.7$, $a_1 = 1.1$ και $E(x) = -0.7 + 1.1x$.

Τέλος, στο παρακάτω σχήμα περιγράφεται η εικόνα της προσέγγισης· οι αστερίσκοι δίδουν τις τιμές που παρατηρήθηκαν ενώ οι τελείες δίδουν τις τιμές του προσεγγιστικού πολυωνύμου. Η **όλη τεχνική αποδίδει**

- όταν δεν υπάρχουν μεγάλες αποκλίσεις στα δεδομένα, πράγμα που σημαίνει ότι χαρακτηρίζονται από - **μιά ορισμένη τάση**.


ψ
x
6
5
4
3
2
1
0
−1
−2
−1
0
1
2
3
4
5
6

(β). Στα ίδια δεδομένα αν θέλαμε να είχαμε αντί του πρωτοβαθμίου πολυωνύμου της προηγούμενης περίπτωσης, την μορφή του σταθερού πολυωνύμου (μηδενικού βαθμού) θα είχαμε εκ των (6), την εξίσωση:

$$a_0 \left\{x^0\right\} = \{\varphi\},$$

ή

$$a_0(\nu + 1) = \sum_{k=0}^{\nu} \varphi_k,$$

και τελικά:

$$E_0(x) = \frac{1}{\nu+1} \sum_{k=0}^{\nu} \varphi_k \qquad \text{(τον μέσο όρο των τιμών).} \tag{8}$$

Παρατηρήσεις:

(i) Η (8) παρέχει μια ενδιαφέρουσα εκδοχή του πολυωνύμου των ελαχί-στων τετραγώνων, όταν οι παρατηρούμενες τιμές είναι ***διαφορετικές***

μετρήσεις του αυτού μεγέθους, *οπότε από την (8) έχουμε το πρακτικό αποτέλεσμα ότι* ***ο μέσος όρος των μετρήσεων*** *(πράγμα που συνήθως γίνεται στην πράξη) αποτελεί την τιμή με* ***το ελάχιστο «τετραγωνικό» σφάλμα***.

(ii) Η εξίσωση (7) είναι δυνατόν ***να εξαχθεί διανυσματικά μ΄ έναν κομψό τρόπο***, *τον ακόλουθο: Προφανώς το διάνυσμα*

$$\bar{a}^\top = (a_0, a_1, a_2, \ldots, a_\mu),$$

των συντελεστών του πολυωνύμου (1), είναι εκείνο που ελαχιστοποιεί την έκφραση:

$$\min_{\bar{a}} \|X\bar{a} - \bar{\psi}\|_2 \tag{9}$$

που με χρήση του ορισμού διαδοχικά γράφεται:

$$\begin{aligned}\|X\bar{a}-\bar{\psi}\|_2^2 &= \left(X\bar{a}-\bar{\psi}\right)^\top \left(X\bar{a}-\bar{\psi}\right) = \\ &= \bar{a}X^\top X\bar{a} - \bar{\psi}X\bar{a} - \bar{a}X^\top\bar{\psi} + \bar{\psi}^\top\bar{\psi} = \\ &= \left(X^\top X\bar{a} - X^\top\bar{\psi}\right)^\top \left(\bar{X}^\top X\right)^{-1} \left(X^\top X\bar{a} - X^\top\bar{\psi}\right) - \\ &\quad - \bar{\psi}^\top X \left(X^\top X\right)^{-1} X^\top\bar{\psi} + \bar{\psi}^\top\bar{\psi}.\end{aligned}$$

Στην τριμελή ανάλυση του τετραγώνου της στάθμης, το δεύτερο και τρίτο μέλος είναι σταθερές ποσότητες, ενώ το πρώτο μέλος είναιμεταβλητό και εξαρτώμενο μόνο από το διάνυσμα $\bar{a}$, άρα η (8) θα λαμβάνει την ελάχιστη τιμή όταν το τρίτο μέλος γίνει μηδέν, πράγμα που υλοποιείται όταν το διάνυσμα: $\boldsymbol{X}^\top\boldsymbol{X}\bar{\boldsymbol{a}} - \boldsymbol{X}^\top\bar{\psi}$ γίνει το μηδενικό, αφού ο πίνακας $\boldsymbol{X}^\top\boldsymbol{X}$ εύκολα αποδεικνύεται ότι είναι ***θετικά ορισμένος****. Αλλά ο μηδενισμός του παραπάνω διανύσματος αποτελεί ακριβώς την εξίσωση των ελαχίστων τετραγώνων.*

(iii) Η φιλοσοφία της μεθόδου, εφαρμόζεται σε διάφορες περιπτώσεις,

π.χ. ***τριγωνομετρικών πολυωνύμων*** *της μορφής:*

$$\Pi_1(x) = \sum \left(a_k \eta\mu kx + \beta_k \sigma\upsilon\nu kx\right), \tag{10}$$

ή, ακόμη και ***εκθετικών πολυωνύμων της μορφής****:*

$$\Pi_2(x) = \sum_{k=0}^{\nu} a_k e^{-kx}, \tag{11}$$

αρκεί ο πίνακας των παρατηρήσεων να ικανοποιεί τις ανάγκες της μεθόδου, δηλαδή πλήθος παρατηρήσεων υπέρτερο του πλήθους των εμπεριεχομένων παραμέτρων στα πολυώνυμα (10) και (11).

Εφαρμογές:

(I) Στα παρακάτω δεδομένα:

x_k	0	$\frac{\pi}{4}$	$\frac{\pi}{2}$
φ_k	0	1	1

είναι δυνατόν να προσδιορισθεί μια τριγωνομετρική συνάρτηση:

$$\Pi(x) = a_0 \eta\mu x + \beta_0 \sigma\upsilon\nu x, \tag{12}$$

με την Μ.Ε.Τ. (τρία σημεία συνάρτησης, 2 παράμετροι).

*Προς τούτο δημιουργούμε τον **πίνακα παρατηρήσεων*** X *και το διάνυσμα* $\bar{\psi}$*, δηλαδή:*

$$\begin{array}{lcccc} \textit{σημείο } (0,0) & \rightarrow & a_0\eta\mu 0^\circ + \beta_0\sigma\upsilon\nu 0^\circ & \rightsquigarrow & 0 \\ \textit{σημείο } (\frac{\pi}{4},1) & \rightarrow & a_0\eta\mu\frac{\pi}{4} + \beta_0\sigma\upsilon\nu\frac{\pi}{4} & \rightsquigarrow & 1 \\ \textit{σημείο } (\frac{\pi}{2},1) & \rightarrow & a_0\eta\mu\frac{\pi}{2} + \beta_0\sigma\upsilon\nu\frac{\pi}{2} & \rightsquigarrow & 1. \end{array}$$

Άρα, θα έχουμε:

$$X = \begin{bmatrix} 0 & 1 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ 1 & 0 \end{bmatrix}, \quad \bar{\psi} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix},$$

και την εξίσωση της Μ.Ε.Τ.:

$$X^\top X\bar{a} = X^\top \bar{\psi},$$

δηλαδή

$$\begin{bmatrix} \frac{3}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{3}{2} \end{bmatrix} \begin{bmatrix} a_0 \\ \beta_0 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2} + 1 \\ \frac{\sqrt{2}}{2} \end{bmatrix},$$

από την οποία εύκολα έχουμε: $a_0 = \frac{\sqrt{2}+3}{4}$ *και* $\beta_0 = \frac{\sqrt{2}-1}{4}$. *Κατά συνέπεια η συνάρτηση* (12) *της Μ.Ε.Τ. θα είναι:*

$$\Pi(x) = \frac{\sqrt{2}+3}{4}\eta\mu x + \frac{\sqrt{2}-1}{4}\sigma\upsilon\nu x.$$

(II) Στα επόμενα δεδομένα

x_k	0	3	5
φ_k	4	9	13

είναι δυνατόν να προσαρμοστεί μια συνάρτηση M.E.T. της μορφής:

$$\Pi_1(x) = a_0 e^{-x} + a_1 e^{x}. \tag{13}$$

Ο πίνακας παρατηρήσεων στην προκειμένη περίπτωση είναι:

$$\begin{array}{lcrcr} \textit{σημείο } (0,4) & \rightarrow & a_0 e^{-0} + a_1 e^{0} & = & 4 \\ \textit{σημείο } (3,9) & \rightarrow & a_0 e^{-3} + a_1 e^{3} & = & 9 \\ \textit{σημείο } (5,13) & \rightarrow & a_0 e^{-5} + a_1 e^{5} & = & 13. \end{array}$$

Άρα θα έχουμε:

$$X = \begin{bmatrix} 1 & 1 \\ 0.05 & 20.085 \\ 0.007 & 148.413 \end{bmatrix}, \text{ και } \bar{\psi} = \begin{bmatrix} 4 \\ 9 \\ 13 \end{bmatrix},$$

Οπότε η εξίσωση της M.E.T. θα είναι:

$$\begin{bmatrix} 1.002 & 3.04 \\ 3.04 & 22430.93 \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ \beta_0 \end{bmatrix} = \begin{bmatrix} 4.541 \\ 2114.134 \end{bmatrix},$$

απ΄ όπου έχουμε την λύση: $a_0 = 4.251$, $a_1 = 0.093$. *Άρα η ζητούμενη συνάρτηση* (13) *της Μ.Ε.Τ. θα είναι:*

$$\Pi_1(x) = 4.251e^{-x} + 0.093e^{x}.$$

(iv) Η βασική εξίσωση της Μ.Ε.Τ. (7), *παρ΄ όλο που ο πίνακας των συντελεστών της* $(X^\top X)$ *είναι θετικά ορισμένος, εν τούτοις ο συντελεστής αστάθειάς της μπορεί να είναι μεγάλος, οπότε η λύση του συστήματος είναι ενδεχόμενο να είναι ασταθής. Η δυσκολία αυτή μπορεί να υπερκερασθεί εάν* ***αντί της ακολουθίας των μονονήμων*** *που χρησιμοποιήσαμε στην δομή της* $E(x)$*:*

$$x^0, x^1, x^2, x^3, \ldots, x^\nu, \ldots$$

μπορούμε να χρησιμοποιήσουμε όρους ***μιας ακολουθίας ορθογωνίων πολυωνύμων****:*

$$T_0, T_1, T_2, \ldots, T_\nu,$$

οπότε το τελικό αποτέλεσμα θα είναι η αντίστοιχη εξίσωση της Μ.Ε.Τ. να έχει πίνακα των συντελεστών ***διαγώνιο*** *και* ***κατάσταση πίνακα βελτιωμένη σημαντικά***.