

## Μάθημα 5

# ΠΡΟΣΕΓΓΙΣΗ ΕΛΑΧΙΣΤΩΝ ΤΕΤΡΑΓΩΝΩΝ

### 5.1 Διακριτή προσέγγιση

#### 5.1.1 Εισαγωγή

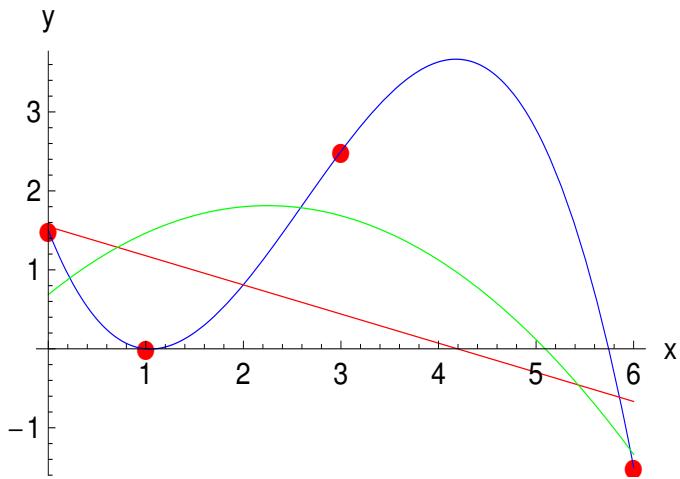
Στο Μάθημα Πολυωνυμική παρεμβολή εξετάστηκε το πρόβλημα της εύρεσης του πολυωνύμου παρεμβολής, δηλαδή του πολυωνύμου που συνέπιπτε ή διαφορετικά διερχόταν από ορισμένα σημεία μιας συνάρτησης. Το πρόβλημα που θα εξεταστεί στο μάθημα αυτό είναι ο προσδιορισμός ενός πολυωνύμου που προσεγγίζει με τον καλύτερο δυνατό ή διαφορετικά **άριστο τρόπο** (best approximation ή **curve fitting**) ένα σύνολο τιμών (data) της μορφής ( $\Sigma \chi$ . 5.1.1 - 1):<sup>1</sup>

$$S = \{(x_i, y_i) \quad \text{με} \quad i = 1, 2, \dots, n\}. \quad (5.1.1 - 1)$$

Έστω ότι  $y_i = f(x_i); i = 1, 2, \dots, n$  όπου  $f$  μια άγνωστη γενικά συνάρτηση, δηλαδή μια συνάρτηση, που ενώ είναι γνωστό ότι υπάρχει, είναι άγνωστος ο τύπος της. Η έννοια της άριστης προσέγγισης σημαίνει τότε ότι το σφάλμα

---

<sup>1</sup>Βλέπε βιβλιογραφία και: [https://en.wikipedia.org/wiki/Curve\\_fitting](https://en.wikipedia.org/wiki/Curve_fitting)



**Σχήμα 5.1.1 - 1:** Δεδομένα:  $S = \{(0, 1.5), (1, 0), (3, 2.5), (6, -1.5)\}$ . Προσέγγιση με: παρεμβολή (υπλε καμπύλη) και με διακριτή προσέγγιση: 1ου βαθμού - ευθεία (χόκκινη) και 2ου βαθμού - παραβολή (πράσινη) καμπύλη

της προσέγγισης στην περίπτωση αυτή είναι το μικρότερο δυνατό.

### 5.1.2 Πολυώνυμο 1ου βαθμού

Έστω ότι το σύνολο των σημείων  $S$  στην (5.1.1 - 1) προσεγγίζεται από ένα πολυώνυμο 1ου βαθμού της μορφής

$$P_1(x) = P(x) = ax + b, \quad (5.1.2 - 1)$$

δηλαδή η προσέγγιση των δεδομένων γίνεται με μια ευθεία. Αν θεωρηθεί το τυχόν σημείο  $(x_i, y_i) \in S$ , τότε η τιμή  $y_i$  προσεγγίζεται από την

$$\tilde{y}_i = P(x_i) = ax_i + b,$$

οπότε το αντίστοιχο απόλυτο σφάλμα της προσέγγισης στην περίπτωση αυτή θα είναι

$$e_i = |y_i - \tilde{y}_i| = |y_i - (ax_i + b)|.$$

Επομένως για το ολικό σφάλμα, έστω  $\tilde{E}$ , θα έχουμε

$$\begin{aligned}\tilde{E} &= \tilde{e}_1 + \dots + \tilde{e}_n \\ &= |y_1 - (ax_1 + b)| + \dots + |y_n - (ax_n + b)|.\end{aligned}\quad (5.1.2 - 2)$$

Προφανώς  $\tilde{E} = \tilde{E}(a, b)$ , δηλαδή το ολικό σφάλμα είναι μια συνάρτηση των σταθερών  $a, b$ . Άρα το πρόβλημα ανάγεται στον υπολογισμό των  $a$  και  $b$ , έτσι ώστε το σφάλμα  $\tilde{E}$  στην (5.1.2 - 2) να είναι ελάχιστο. Τότε όμως, όπως είναι ήδη γνωστό στον αναγνώστη από τη μελέτη συναρτήσεων πολλών μεταβλητών, οι **αναγκαίες** συνθήκες για να συμβαίνει αυτό είναι:<sup>2</sup>

$$\frac{\partial \tilde{E}}{\partial a} = 0 \quad \text{και} \quad \frac{\partial \tilde{E}}{\partial b} = 0.\quad (5.1.2 - 3)$$

Εύκολα όμως διαπιστώνεται ότι η (5.1.2 - 3) λόγω και του απολύτου δεν παραγωγίζεται, οπότε το πρόβλημα στη μορφή αυτή δεν λύνεται.<sup>3</sup>

Στη **διακριτή μέθοδο των ελαχίστων τετραγώνων** (discrete least squares method), σε αντίθεση με την (5.1.2 - 2), προσδιορίζονται οι σταθερές  $a$  και  $b$ , έτσι ώστε το ολικό τετραγωνικό σφάλμα

$$\begin{aligned}E &= e_1^2 + \dots + e_n^2 \\ &= [y_1 - (ax_1 + b)]^2 + \dots + [y_n - (ax_n + b)]^2\end{aligned}\quad (5.1.2 - 4)$$

να είναι **ελάχιστο**.

Τότε από την (5.1.2-4), αν εφαρμοστούν οι συνθήκες (5.1.2-3), προκύπτει το σύστημα

$$\begin{aligned}\frac{\partial E}{\partial a} &= -2 \sum_{i=1}^n (y_i - ax_i - b) x_i = 0, \quad \text{και} \\ \frac{\partial E}{\partial b} &= -2 \sum_{i=1}^n (y_i - ax_i - b) = 0\end{aligned}$$

<sup>2</sup>Βλέπε βιβλιογραφία και Α. Μπράτσος [1] Κεφ. 7.

<sup>3</sup>Έστω για ευκολία ότι απαλείφονται τα απόλυτα. Τότε

$$\frac{\partial \tilde{E}}{\partial a} = -x_1 - \dots - x_n = 0 \quad \text{και} \quad \frac{\partial \tilde{E}}{\partial b} = -1 - \dots - 1 = -n = 0$$

άτοπο.

που τελικά μετά τις πράξεις γράφεται ως εξής:

$$\begin{aligned} a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i &= \sum_{i=1}^n x_i y_i \\ a \sum_{i=1}^n x_i + b \sum_{i=1}^n \overbrace{x_i^0}^n &= \sum_{i=1}^n y_i. \end{aligned} \quad (5.1.2 - 5)$$

Το σύστημα (5.1.2-5), που λέγεται και **σύστημα κανονικών εξισώσεων** (normal equations), είναι γραμμικό ως προς τους αγνώστους  $a$  και  $b$ . Από τη λύση του προκύπτει τότε ότι οι ζητούμενοι συντελεστές του πολυωνύμου  $P_1(x)$  στην περίπτωση αυτή είναι:

$$\begin{aligned} a &= \frac{n \left( \sum_{i=1}^n x_i y_i \right) - \left( \sum_{i=1}^n x_i \right) \left( \sum_{i=1}^n y_i \right)}{n \left( \sum_{i=1}^n x_i^2 \right) - \left( \sum_{i=1}^n x_i \right)^2}, \\ b &= \frac{\left( \sum_{i=1}^n x_i^2 \right) \left( \sum_{i=1}^n y_i \right) - \left( \sum_{i=1}^n x_i y_i \right) \left( \sum_{i=1}^n x_i \right)}{n \left( \sum_{i=1}^n x_i^2 \right) - \left( \sum_{i=1}^n x_i \right)^2}. \end{aligned} \quad (5.1.2 - 6)$$

### Παράδειγμα 5.1.2 - 1

Να προσδιοριστεί με τη μέθοδο των ελαχίστων τετραγώνων το πολυώνυμο 1ου βαθμού που προσεγγίζει τα δεδομένα:

$x_i$	-0.5	0.3	0.7	1.5
$y_i$	1.2	2.0	1.0	-1.0

**Λύση.** Για την εφαρμογή των τύπων (5.1.2 - 6) απαιτείται η δημιουργία του Πίνακα 5.1.2 - 1.

**Πίνακας 5.1.2 - 1: Παράδειγμα 5.1.2 - 1**

$i$	$x_i$	$y_i$	$x_i y_i$	$x_i^2$
1	-0.5	1.2	-0.6	0.25
2	0.3	2.0	0.6	0.09
3	0.7	1.0	0.7	0.49
4	1.5	-1.0	-1.5	2.25
	2.0	3.2	-0.8	3.08

Τότε σύμφωνα με τα αποτελέσματα του Πίνακα 5.1.2 - 1 έχουμε

$$a = \frac{4 \cdot (-0.8) - 2 \cdot (3.2)}{4 \cdot (3.08) - 2^2} \approx -1.1539, \quad \text{και}$$

$$b = \frac{(3.08) \cdot (3.2) - (-0.8) \cdot 2}{4 \cdot (3.08) - 2^2} \approx 1.3769,$$

δηλαδή το ζητούμενο πολυώνυμο του 1ου βαθμού - ευθεία γραμμή - (Σχ. 5.1.2 - 1) είναι

$$P(x) = -1.1539x + 1.3769.$$

■

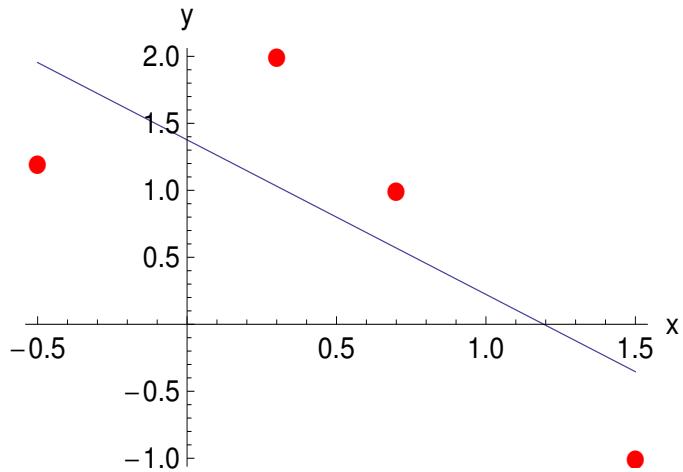
**Παράδειγμα 5.1.2 - 2**

Όμοια το πολυώνυμο 1ου βαθμού που προσεγγίζει τα δεδομένα  $(x_i, y_i)$  του Πίνακα 5.1.3 - 1.

**Λύση.** Σύμφωνα με τους τύπους (5.1.2-6) και τα αποτελέσματα του Πίνακα 5.1.3 - 1 προκύπτει ότι

$$a = \frac{10 \cdot 485.9487 - 56.2933 \cdot 73.8373}{10 \cdot 380.5423 - 56.2933^2} \approx 1.1044 \quad \text{και}$$

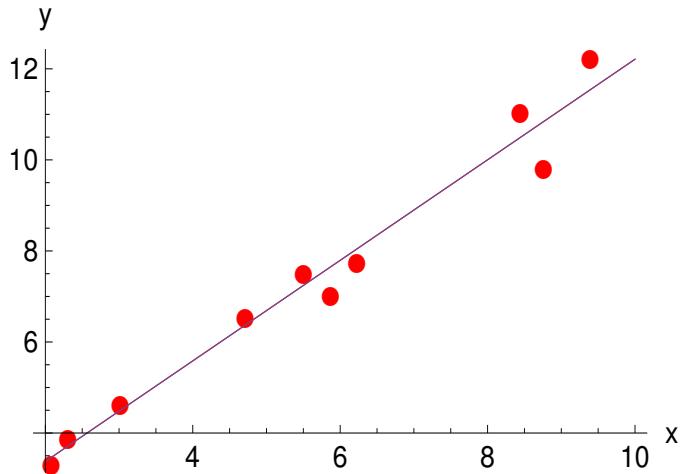
$$b = \frac{380.5423 \cdot 73.8373 - 56.2933 \cdot 485.9487}{10 \cdot 380.5423 - 56.2933^2} \approx 1.1667.$$



**Σχήμα 5.1.2 - 1:** Παράδειγμα 5.1.2 - 1. Η εξίσωση της ευθείας είναι  $y = -1.1539x + 1.3769$

**Πίνακας 5.1.2 - 2: Παράδειγμα 5.1.2 - 2**

$i$	$x_i$	$y_i$	$x_i y_i$	$x_i^2$
1	2.0774	3.3123	6.8810	4.3156
2	2.3049	3.8982	8.9850	5.3126
3	3.0125	4.6500	14.0081	9.0752
4	4.7092	6.5576	30.8810	22.1766
5	5.5016	7.5173	41.3572	30.2676
6	5.8704	7.0415	41.3364	34.4616
7	6.2248	7.7497	48.2403	38.7481
8	8.4431	11.0451	93.2549	71.2859
9	8.7594	9.8179	85.9989	76.7271
10	9.3900	12.2477	115.0059	88.1721
		56.2933	73.8373	380.5423



**Σχήμα 5.1.2 - 2:** Παράδειγμα 5.1.2 - 2. Η εξίσωση της ευθείας είναι  $y = 1.1044x + 1.1667$

Άρα όμοια το ζητούμενο πολυώνυμο του 1ου βαθμού - ευθεία γραμμή - (Σχ. 5.1.2 - 2) είναι

$$P(x) = 1.1044x + 1.1667.$$

■

### 5.1.3 Πολυώνυμο $m$ -βαθμού

Στην περίπτωση αυτή ζητείται η προσέγγιση του συνόλου  $S$  στην (5.1.1 - 1), όπου

$$S = \{(x_i, y_i) \quad \text{με} \quad i = 1, 2, \dots, n\},$$

με ένα πολυώνυμο  $m$ -βαθμού της μορφής

$$P_m(x) = a_0 + a_1x + \dots + a_mx^m.$$

Αποδεικνύεται στα Μαθηματικά ότι στην περίπτωση αυτή για τον βαθμό  $m$  του πολυωνύμου πρέπει να ισχύει ότι

$$\mathbf{m} < \mathbf{n} - \mathbf{1}. \tag{5.1.3 - 1}$$

Ανάλογα με την περίπτωση του πολυωνύμου 1ου βαθμού της Παραγράφου 5.1.2 και στην περίπτωση αυτή οι  $m + 1$  σταθερές  $a_0, a_1, \dots, a_m$  εκλέγονται, έτσι ώστε το ολικό τετραγωνικό σφάλμα

$$\begin{aligned} E &= e_1^2 + \dots + e_n^2 \\ &= [y_1 - P_m(x_1)]^2 + \dots + [y_n - P_m(x_n)]^2 \end{aligned}$$

να είναι **ελάχιστο**.

Το παραπάνω σφάλμα  $E$ , όταν οι όροι του δεύτερου μέλους υψηλού στο τετράγωνο, γράφεται

$$\begin{aligned} E &= \sum_{i=1}^n y_i^2 - 2 \sum_{i=1}^n P_m(x_i) y_i + \sum_{i=1}^n [P_m(x_i)]^2 \\ &= \sum_{i=1}^n y_i^2 - 2 \sum_{i=1}^n \left( \sum_{j=0}^m a_j x_i^j \right) y_i + \sum_{i=1}^n \left( \sum_{j=0}^m a_j x_i^j \right)^2 \\ &= \sum_{i=1}^n y_i^2 - 2 \sum_{i=1}^n \left( \sum_{j=0}^m a_j x_i^j \right) y_i + \sum_{i=0}^m \sum_{k=0}^m a_j a_k \left( \sum_{i=1}^n x_i^{j+k} \right), \end{aligned}$$

οπότε οι  $m + 1$  για την περίπτωση αυτή αναγκαίες συνθήκες (5.1.2 – 3) γράφονται ως εξής:

$$\frac{\partial E}{\partial a_j} = 0 \quad \text{για κάθε } j = 0, 1, \dots, m. \quad (5.1.3 - 2)$$

Από τις συνθήκες (5.1.3 – 2) προκύπτει τότε ότι

$$\frac{\partial E}{\partial a_j} = -2 \sum_{i=1}^n y_i x_i^j + 2 \sum_{k=0}^m a_k \sum_{i=1}^n x_i^{j+k} = 0,$$

δηλαδή

$$\sum_{k=0}^n \sum_{i=1}^m x_i^{j+k} = \sum_{i=1}^m y_i x_i^j \quad \text{για κάθε } j = 0, 1, \dots, n,$$

από την οποία τελικά μετά τις πράξεις προκύπτει στην περίπτωση αυτή το παρακάτω γραμμικό σύστημα **κανονικών εξισώσεων**:

$$\begin{aligned} a_0 \sum_{i=1}^n x_i^0 + a_1 \sum_{i=1}^n x_i^1 + \dots + a_m \sum_{i=1}^n x_i^m &= \sum_{i=1}^n y_i x_i^0 \\ a_0 \sum_{i=1}^n x_i^1 + a_1 \sum_{i=1}^n x_i^2 + \dots + a_m \sum_{i=1}^n x_i^{m+1} &= \sum_{i=1}^n y_i x_i^1 \\ &\vdots & (5.1.3 - 3) \\ a_0 \sum_{i=1}^n x_i^m + a_1 \sum_{i=1}^n x_i^{m+1} + \dots + a_m \sum_{i=1}^n x_i^{2m} &= \sum_{i=1}^n y_i x_i^m. \end{aligned}$$

Το σύστημα αυτό έχει  $m+1$  εξισώσεις και  $m+1$  αγνώστους τους συντελεστές  $a_0, a_1, \dots, a_m$  του πολυωνύμου  $P_m(x) = a_0 + a_1 x + \dots + a_m x^m$ .

### Παρατηρήσεις 5.1.3 - 1

i) Ο πίνακας των συντελεστών των αγνώστων

$$\left[ \begin{array}{cccc} \sum_{i=1}^n x_i^0 & \sum_{i=1}^n x_i^1 & \dots & \sum_{i=1}^n x_i^m \\ \sum_{i=1}^n x_i^1 & \sum_{i=1}^n x_i^2 & \dots & \sum_{i=1}^n x_i^{m+1} \\ \vdots & \vdots & & \vdots \\ \sum_{i=1}^n x_i^m & \sum_{i=1}^n x_i^{m+1} & \dots & \sum_{i=1}^n x_i^{2m} \end{array} \right] \quad (5.1.3 - 4)$$

είναι **συμμετρικός**. Η ιδιότητα αυτή είναι σημαντική και περιορίζει τον αριθμό των πράξεων<sup>4</sup> που απαιτούνται για τη λύση του συστήματος (5.1.3 - 3).

ii) Αποδεικνύεται ότι το σύστημα (5.1.3 - 3) έχει ακριβώς μία λύση, όταν τα σημεία  $x_i; i = 1, 2, \dots, n$  είναι **διαφορετικά** μεταξύ τους.

<sup>4</sup>Βλέπε Μάθημα Γραμμική Άλγεβρα - Μέθοδος Cholesky και A. Μπράτσος [1] Κεφ. 8.

- iii) Κρίνεται σκόπιμο στο σημείο αυτό να διευκρινιστεί ότι μια γενίκευση του παραπάνω προβλήματος, δηλαδή η προσέγγιση των σημείων  $S$  με ένα πολυώνυμο  $P_m(x)$  με την απαίτηση το άθροισμα των σφαλμάτων  $E = \sum_{i=1}^n e_i^k$  με  $k \geq 3$ , να είναι ελάχιστο, καταλήγει μετά και την εφαρμογή της συνθήκης (5.1.3-2) σε μη γραμμικό σύστημα. Επομένως, η μέθοδος με την απαίτηση αυτή δεν είναι εφαρμόσιμη.

### Παράδειγμα 5.1.3 - 1

Να προσδιοριστεί με τη διακριτή μέθοδο των ελαχίστων τετραγώνων το πολυώνυμο 2ου βαθμού που προσεγγίζει τα δεδομένα του Παραδείγματος 5.1.2 - 1.

**Λύση.** Επειδή ο αριθμός των σημείων είναι  $n = 4$ , σύμφωνα με τη συνθήκη (5.1.3 - 1) ο μεγαλύτερος δυνατός βαθμός  $m$  του πολυωνύμου θα είναι

$$m < 4 - 1, \quad \text{δηλαδή} \quad m = 2.$$

Έστω

$$P_2(x) = a_0 + a_1x + a_2x^2$$

το ζητούμενο πολυώνυμο. Τότε το σύστημα (5.1.3 - 3) γράφεται

$$\begin{aligned} a_0 \sum_{i=1}^4 x_i^0 + a_1 \sum_{i=1}^4 x_i^1 + a_2 \sum_{i=1}^4 x_i^2 &= \sum_{i=1}^4 y_i x_i^0 \\ a_0 \sum_{i=1}^4 x_i^1 + a_1 \sum_{i=1}^4 x_i^2 + a_2 \sum_{i=1}^4 x_i^3 &= \sum_{i=1}^4 y_i x_i^1 \\ a_0 \sum_{i=1}^4 x_i^2 + a_1 \sum_{i=1}^4 x_i^3 + a_2 \sum_{i=1}^4 x_i^4 &= \sum_{i=1}^4 y_i x_i^2, \end{aligned}$$

οπότε σύμφωνα με τα αποτελέσματα του Πίνακα 5.1.3 - 1 έχουμε το σύστημα

$$4a_0 + 2.0a_1 + 3.08a_2 = 3.2$$

$$2.0a_0 + 3.08a_1 + 3.62a_2 = -0.8$$

$$3.08a_0 + 3.62a_1 + 5.3732a_2 = -1.28.$$

**Πίνακας** 5.1.3 - 1: Παράδειγμα 5.1.2 - 2

$x_i$	$y_i$	$x_i y_i$	$x_i^2$	$x_i^3$	$x_i^4$	$x_i^2 y_i$
-0.5	1.2	-0.6	0.25	-0.125	0.0625	0.30
0.3	2.0	0.6	0.09	0.027	0.0081	0.18
0.7	1.0	0.7	0.49	0.343	0.2401	0.49
1.5	-1.0	-1.5	2.25	3.375	5.0625	-2.25
2.0	3.2	-0.8	3.08	3.62	5.3732	-1.28

Από τη λύση του συστήματος προκύπτει ότι το ζητούμενο πολυώνυμο είναι  
(Σχ. 5.1.3 - 1)

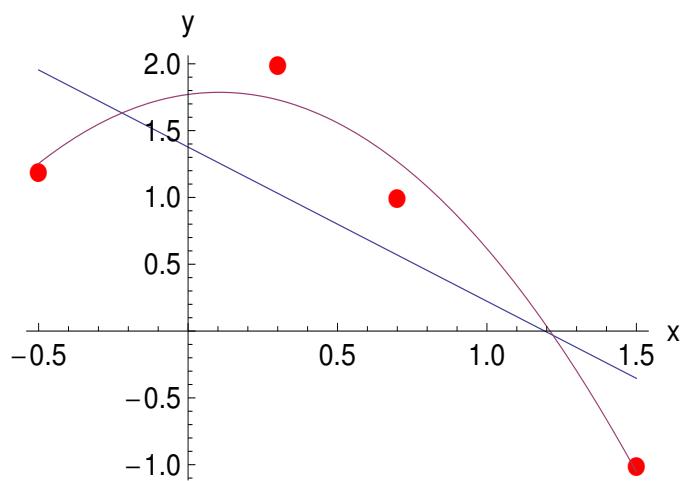
$$P_2(x) = -1.4583 x^2 + 0.3045 x + 1.7707.$$

Τότε το σφάλμα της προσέγγισης είναι

$$E = \sum_{i=1}^5 [y_i - P_2(x_i)]^2 \approx 2.76E-04,$$

που είναι και το ελάχιστο που προκύπτει από προσέγγιση με πολυώνυμο 2ου βαθμού για τα παραπάνω δεδομένα.

■



**Σχήμα 5.1.3 - 1:** Παράδειγμα 5.1.2 - 2. Η καμπύλη ορίζεται από το πολυώνυμο  $P_2(x) = -1.4583 x^2 + 0.3045 x + 1.7707$ , ενώ η ευθεία (Παράδειγμα 5.1.2 - 1) έχει εξίσωση  $y = -1.1539 x + 1.3769$

## Ασκήσεις

1. Έστω τα δεδομένα:

$x_i$	0.500	0.150	0.250	0.400	0.550	0.700
$y_i$	1.235	1.750	2.020	-1.550	-2.345	0.435

Με τη διακριτή μέθοδο των ελαχίστων τετραγώνων να υπολογιστούν:

- i) το πολυώνυμο 1ου, αντίστοιχα 2ου βαθμού που τα προσεγγίζει και να γίνει η γραφική τους παράσταση,
- ii) το αντίστοιχο σε κάθε περίπτωση σφάλμα της προσέγγισης.

2. Όμοια τα δεδομένα

$x_i$	0.3	0.5	0.7	1.4	1.8	2.2	3.5
$y_i$	0.0647	0.0985	0.2490	1.0395	1.5393	3.5941	4.0549

## Απαντήσεις

1. (i)  $P_1(x) = 2.086\,564 - 4.303\,681x$ , (ii)  $P_2(x) = 5.321\,234 - 23.546\,340x + 23.037\,480x^2$ .
2. (i)  $P_1(x) = -0.564\,6827 + 1.403\,152x$ , (ii)  $P_2(x) = -0.712\,560 + 1.654\,234x - 0.068\,171x^2$ .



# Βιβλιογραφία

- [1] Αχρίβης, Γ. & Δουγαλής, Β. (1995). *Εισαγωγή στην Αριθμητική Ανάλυση*. Πανεπιστημιακές Εκδόσεις Κρήτης. ISBN 978-960-524-022-6.
- [2] Μπράτσος, Α. (2011). *Εφαρμοσμένα Μαθηματικά*. Εκδόσεις Α. Σταμούλη. ISBN 978-960-351-874-7.
- [3] Μπράτσος, Α. (2002). *Ανώτερα Μαθηματικά*. Εκδόσεις Α. Σταμούλη. ISBN 960-351-453-5/978-960-351-453-4.
- [4] Στεφανάκος, Χ. (2009). *Προγραμματισμός H/Y με MATLAB*. Γκιούρδας Εκδοτική. ISBN 978-960-387-856-8.
- [5] Atkinson, K. E. (1989). *An Introduction to Numerical Analysis*. John Wiley & Sons (2nd ed.). ISBN 0-471-50023-2.
- [6] Burden, R. L. & Faires, D. J. (2010). *Numerical Analysis*. Brooks/Cole (7th ed.). ISBN 978-0-534-38216-2.
- [7] Conte, S. D. & de Boor, C. (1980). *Elementary Numerical Analysis: An Algorithmic Approach*. McGraw-Hill Inc (3rd ed.). ISBN 978-0-07-012447-9.
- [8] Don, E. (2006). *Schaum's Outlines – Mathematica*. Εκδόσεις Κλειδάριθμος. ISBN 978-960-209-961-2.
- [9] Leader, L. J. (2004). *Numerical Analysis and Scientific Computation*. Addison-Wesley. ISBN 978-0-201-73499-7.

- [10] Schatzman, M. (2002). *Numerical Analysis: A Mathematical Introduction*. Oxford: Clarendon Press. ISBN 978-0-19-850279-1.
- [11] Stoer, J. & Bulirsch, R. (2002). *Introduction to Numerical Analysis*. Springer (3rd ed.). ISBN 978-0-387-95452-3.
- [12] Sli, E. & Mayers, D. (2003). *An Introduction to Numerical Analysis*. Cambridge University Press. ISBN 978-0-521-00794-8.

### Μαθηματικές βάσεις δεδομένων

- [http://en.wikipedia.org/wiki/Main\\_Page](http://en.wikipedia.org/wiki/Main_Page)
- <http://eqworld.ipmnet.ru/index.htm>
- <http://mathworld.wolfram.com/>
- <http://eom.springer.de/>