Μάθημα 4

ΔΙΑΚΡΙΤΗ ΠΡΟΣΕΓΓΙΣΗ

4.1 Μέθοδος ελάχιστων τετραγώνων

4.1.1 Εισαγωγή

Στο Μάθημα 3 εξετάστηκε το πρόβλημα της εύρεσης του πολυωνύμου παρεμβολής, δηλαδή του πολυωνύμου που συνέπιπτε ή διαφορετικά διέρχονταν 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 **best fitting**) ένα σύνολο τιμών (data) της μορφής (Σχ. 4.1 - 1):

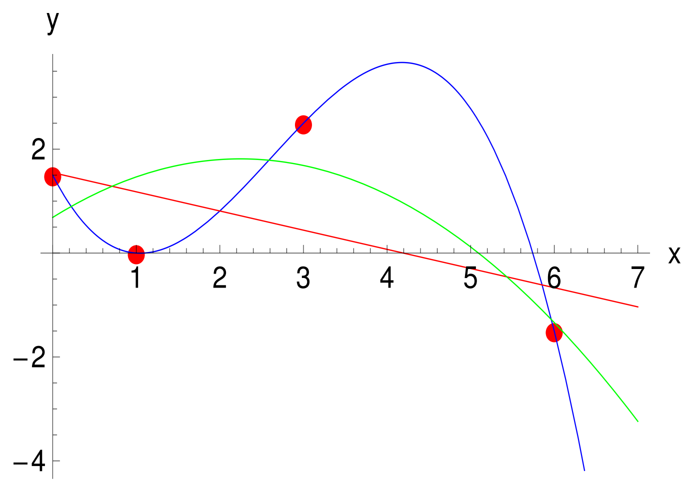
$$S = \{(x_i, y_i) \quad \text{με} \quad i = 1,\, 2,\, \ldots,\, n\}\,. \tag{4.1 - 1}$$

Έστω ότι $y_i = f\,(x_i)$; $i = 1,\, 2,\, \ldots,\, n$ 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 της προσέγγισης στην περίπτωση αυτή είναι το μικρότερο δυνατό.

4.1.2 Περίπτωση I πολυώνυμο 1ου βαθμού

Έστω ότι το σύνολο των σημείων S στην $(4.1 - 1)$ προσεγγίζεται από ένα πολυώνυμο 1ου βαθμού της μορφής

$$P_1(x) = P(x) = ax + b, \tag{4.1 - 2}$$



Σχήμα 4.1 - 1: Δεδομένα: $S = \{(0,1.5),(1,0),(3,2.5),(6,-1.5)\}$. Προσέγγιση με: παρεμβολή (μπλε καμπύλη) και με διακριτή προσέγγιση: 1ου βαθμού - ευθεία (κόκκινη) και 2ου βαθμού - παραβολή (πράσινη) καμπύλη

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

$$\tilde{E} = \tilde{e}_1 + \ldots + \tilde{e}_n = |y_1 - (ax_1 + b)| + \ldots + |y_n - (ax_n + b)|\,. \qquad (4.1 - 3)$$

Προφανώς $\tilde{E} = \tilde{E}(a,b)$, δηλαδή το ολικό σφάλμα είναι μια συνάρτηση των a, b. Άρα το πρόβλημα ανάγεται στον υπολογισμό των a και b, έτσι ώστε το σφάλμα $\tilde{E}$ στην $(4.1 - 3)$ να είναι ελάχιστο. Τότε, όπως είναι ήδη γνωστό στον αναγνώστη από τη μελέτη συναρτήσεων πολλών μεταβλητών, οι **αναγκαίες** συνθήκες για να συμβαίνει αυτό είναι:

$$\frac{\partial \tilde{E}}{\partial a} = 0 \quad \text{και} \quad \frac{\partial \tilde{E}}{\partial b} = 0. \qquad (4.1 - 4)$$

Εύκολα όμως διαπιστώνεται ότι η $(4.1-4)$ λόγω και του απολύτου δεν παραγωγίζεται[1], οπότε το πρόβλημα στη μορφή αυτή δε λύνεται.

Στη **διακριτή μέθοδο των ελάχιστων τετραγώνων** (discrete least squares method), σε αντίθεση με την $(4.1-3)$, προσδιορίζονται οι σταθερές a και b, έτσι ώστε το ολικό τετραγωνικό σφάλμα E, δηλαδή το

$$E = e_1^2 + \ldots + e_n^2 = [y_1 - (ax_1 + b)]^2 + \ldots + [y_n - (ax_n + b)]^2 \qquad (4.1 - 5)$$

να είναι **ελάχιστο**. Τότε από την $(4.1-5)$ προκύπτει το σύστημα

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

που τελικά μετά τις πράξεις γράφεται ως εξής:

$$\begin{aligned} a\sum_{i=1}^{n}x_i^2 + b\sum_{i=1}^{n}x_i &= \sum_{i=1}^{n}x_i\,y_i \\ a\sum_{i=1}^{n}x_i + b\overbrace{\sum_{i=1}^{n}x_i^0}^{n} &= \sum_{i=1}^{n}y_i. \end{aligned} \qquad (4.1 - 6)$$

Το γραμμικό σύστημα $(4.1-6)$ λέγεται τότε και **σύστημα κανονικών εξισώσεων** (normal equations) και από τη λύση του προκύπτει ότι:

$$a = \frac{n\left(\sum_{i=1}^{n}x_i\,y_i\right) - \left(\sum_{i=1}^{n}x_i\right)\left(\sum_{i=1}^{n}y_i\right)}{n\left(\sum_{i=1}^{n}x_i^2\right) - \left(\sum_{i=1}^{n}x_i\right)^2}, \qquad (4.1 - 7)$$

[1] Έστω για ευκολία ότι απαλείφονται τα απόλυτα. Τότε

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

άτοπο.

Πίνακας 4.1 - 1: Παράδειγμα 4.1 - 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

$$b = \frac{\left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i\right) - \left(\sum_{i=1}^{n} x_i\, y_i\right)\left(\sum_{i=1}^{n} x_i\right)}{n\left(\sum_{i=1}^{n} x_i^2\right) - \left(\sum_{i=1}^{n} x_i\right)^2}. \qquad (4.1 - 8)$$

Παράδειγμα 4.1 - 1

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

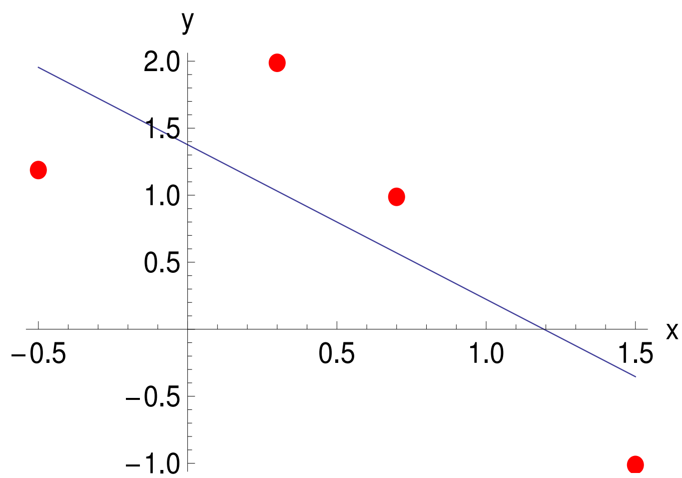
x_i	-0.5	0.3	0.7	1.5
y_i	1.2	2.0	1.0	-1.0

Λύση. Για την εφαρμογή των τύπων $(4.1-7)$ και $(4.1-8)$ απαιτείται η δημιουργία του Πίνακα 4.1 - 1.

Τότε έχουμε

$$\begin{aligned} a &= \frac{4\cdot(-0.8) - 2\cdot(3.2)}{4\cdot(3.08) - 2^2} \approx -1.1539 \quad \text{και} \\ b &= \frac{(3.08)\cdot(3.2) - (-0.8)\cdot 2}{4\cdot(3.08) - 2^2} \approx 1.3769, \end{aligned}$$

δηλαδή $P(x) = -1.1539\, x + 1.3769$ (Σχ. 4.1 - 2). ■



Σχήμα 4.1 - 2: Παράδειγμα 4.1 - 1. Η εξίσωση της ευθείας είναι $y = -1.1539\,x + 1.3769$

Παράδειγμα 4.1 - 2

Όμοια το πολυώνυμο 1ου βαθμού που προσεγγίζει τα δεδομένα (x_i, y_i) του Πίνακα 4.1 - 3.

Λύση. Σύμφωνα με τους τύπους $(4.1-8)$ και $(4.1-8)$ προκύπτει ότι

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

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

Άρα $P(x) = 1.1044\,x + 1.1667$ (Σχ. 4.1 - 3). ■

4.1.3 Περίπτωση ΙΙ πολυώνυμο m-βαθμού

Στην περίπτωση αυτή ζητείται η προσέγγιση του συνόλου S στην $(4.1-1)$ με ένα πολυώνυμο m-βαθμού της μορφής

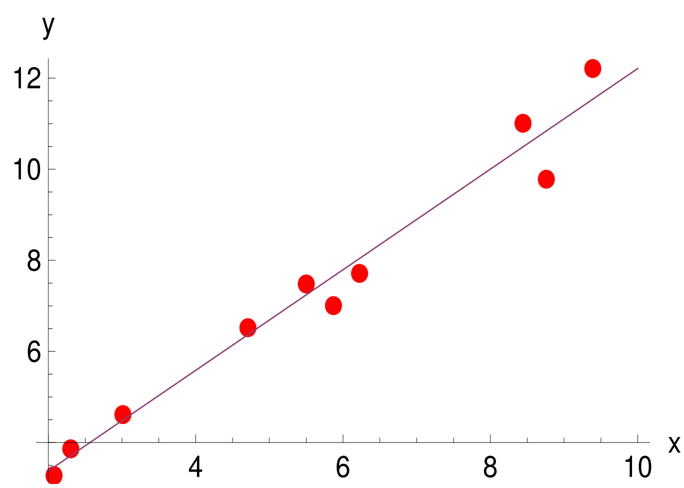
$$P_m(x) = a_0 + a_1 x + \ldots + a_m x^m,$$

όταν

$$\mathbf{m < n - 1}. \qquad (4.1 - 9)$$

Πίνακας 4.1 - 2: Παράδειγμα 4.1 - 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	485.9487	380.5423



Σχήμα 4.1 - 3: Παράδειγμα 4.1 - 2. Η εξίσωση της ευθείας είναι $y = 1.1044\,x + 1.1667$

Τότε, όμοια με την Περίπτωση I, εκλέγονται οι σταθερές a_0, a_1, …, a_m, έτσι ώστε το σφάλμα[2]

$$E = e_1^2 + \ldots + e_n^2 = [y_1 - P_m\,(x_1)]^2 + \ldots + [y_n - P_m\,(x_n)]^2$$

να είναι ελάχιστο.

Όπως και στην περίπτωση του πολυωνύμου 1ου βαθμού οι ανάλογες αναγκαίες συνθήκες των $(4.1 - 4)$ είναι οι εξής:

$$\frac{\partial E}{\partial a_j} = 0 \quad \text{για κάθε} \quad j = 0,\, 1,\, \ldots,\, m. \tag{4.1 - 10}$$

Από την $(4.1 - 10)$ τελικά[3] έχουμε το παρακάτω γραμμικό σύστημα **κανονικών εξισώσεων** με $m+1$ εξισώσεις και $m+1$ αγνώστους τους συντελεστές a_0, a_1, …, a_m του πολυωνύμου

$$\begin{aligned}
a_0 \sum_{i=1}^{n} x_i^0 + a_1 \sum_{i=1}^{n} x_i^1 + \ldots + a_m \sum_{i=1}^{n} x_i^m &= \sum_{i=1}^{n} y_i\, x_i^0 \\
a_0 \sum_{i=1}^{n} x_i^1 + a_1 \sum_{i=1}^{n} x_i^2 + \ldots + a_m \sum_{i=1}^{n} x_i^{m+1} &= \sum_{i=1}^{n} y_i\, x_i^1 \\
&\;\vdots \qquad\qquad (4.1 - 11) \\
a_0 \sum_{i=1}^{n} x_i^m + a_1 \sum_{i=1}^{n} x_i^{m+1} + \ldots + a_m \sum_{i=1}^{n} x_i^{2m} &= \sum_{i=1}^{n} y_i\, x_i^m\,.
\end{aligned}$$

[2] Βλέπε βιβλιογραφία και Α. Μπράτσος [2] Κεφ. 7.

[3] Το σύστημα $(4.1 - 10)$ γράφεται $\frac{\partial E}{\partial a_j} = -2\sum_{i=1}^{n} y_i x_i^j + 2\sum_{k=0}^{m} a_k \sum_{i=1}^{n} x_i^{j+k} = 0$, οπότε $\sum_{k=0}^{n}\sum_{i=1}^{m} x_i^{j+k} = \sum_{i=1}^{m} y_i x_i^j$ για κάθε $j = 0,\, 1,\, \ldots,\, n$. Η απόδειξη να παραλειφθεί σε πρώτη ανάγνωση.

Παρατήρησεις 4.1 - 1

i) Ο πίνακας των συντελεστών των αγνώστων

$$\begin{bmatrix} \sum_{i=1}^{n} x_i^0 & \sum_{i=1}^{n} x_i^1 & \dots & \sum_{i=1}^{n} x_i^m \\ \sum_{i=1}^{n} x_i^1 & \sum_{i=1}^{n} x_i^2 & \dots & \sum_{i=1}^{n} x_i^{m+1} \\ \vdots & \vdots & & \vdots \\ \sum_{i=1}^{n} x_i^m & \sum_{i=1}^{n} x_i^{m+1} & \dots & \sum_{i=1}^{n} x_i^{2m} \end{bmatrix} \tag{4.1 - 12}$$

είναι **συμμετρικός**. Η ιδιότητα αυτή είναι σημαντική και περιορίζει τον αριθμό των πράξεων[4], που απαιτούνται για τη λύση του συστήματος $(4.1-11)$.

ii) Αποδεικνύεται ότι το σύστημα $(4.1-11)$ έχει ακριβώς μία λύση, όταν τα σημεία x_i; $i = 1, 2, \dots, n$ είναι **διαφορετικά** μεταξύ τους.

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 $(4.1-10)$ σε μη γραμμικό σύστημα. Επομένως η μέθοδος με την απαίτηση αυτή δεν είναι εφαρμόσιμη.

Παράδειγμα 4.1 - 3

Να προσδιοριστεί με τη διακριτή μέθοδο των ελάχιστων τετραγώνων το πολυώνυμο 2ου βαθμού που προσεγγίζει τα δεδομένα του Παραδείγματος 4.1 - 1.

Λύση. Επειδή ο αριθμός των σημείων είναι $n = 4$, σύμφωνα με τη συνθήκη $(4.1-9)$ ο μεγαλύτερος δυνατός βαθμός m του πολυωνύμου θα είναι $m < 4-1$, δηλαδή $m = 2$. Έστω

$$P_2(x) = a_0 + a_1 x + a_2 x^2$$

[4] Βλέπε βιβλιογραφία μέθοδος Cholesky και Α. Μπράτσος [2] Κεφ. 8.

Πίνακας 4.1 - 3: Παράδειγμα 4.1 - 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

το ζητούμενο πολυώνυμο. Τότε το σύστημα $(4.1-11)$ γράφεται

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

οπότε σύμφωνα με τον Πίνακα 4.1 - 3 έχουμε

$$\begin{array}{rcrcrcr} 4a_0 & + & 2.0a_1 & + & 3.08a_2 & = & 3.2 \\ 2.0a_0 & + & 3.08a_1 & + & 3.62a_2 & = & -0.8 \\ 3.08a_0 & + & 3.62a_1 & + & 5.3732a_2 & = & -1.28. \end{array}$$

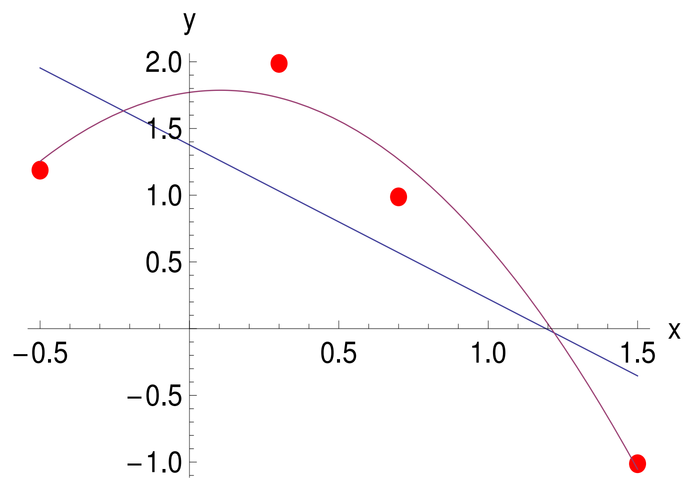
Από τη λύση[5] του συστήματος προκύπτει ότι το πολυώνυμο είναι (Σχ. 4.1 - 4)

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

Τότε το σφάλμα της προσέγγισης είναι $E = \sum_{i=1}^{5}\left[y_i - P_2\left(x_i\right)\right]^2 \approx 2.76E-04$, που είναι και το ελάχιστο που προκύπτει από προσέγγιση με πολυώνυμο 2ου βαθμού για τα παραπάνω δεδομένα.

■

[5] Η λύση του συστήματος δεν είναι στην εξεταστέα ύλη.



Σχήμα 4.1 - 4: Παράδειγμα 4.1 - 2. Η καμπύλη ορίζεται από το πολυώνυμο $P_2(x) = -1.4583\,x^2 + 0.3045\,x + 1.7707$, ενώ η ευθεία (Παράδειγμα 4.1 - 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 ελάχιστων 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

6

E-mail: bratsos@teiath.gr URL: http://users.teiath.gr/bratsos/

Βιβλιογραφία

[1] Ακρίβης, Γ., Δουγαλής, Β. (1995), Εισαγωγή στην Αριθμητική Ανάλυση, Πανεπιστημιακές Εκδόσεις Κρήτης, Αθήνα, ISBN 978–960–524–022–6.

[2] Μπράτσος, Α. (2011), Εφαρμοσμένα Μαθηματικά, Εκδόσεις Α. Σταμούλη, Αθήνα, ISBN 978–960–351–874–7.

[3] Μπράτσος, Α. (2002), Ανώτερα Μαθηματικά, Εκδόσεις Α. Σταμούλη, Αθήνα, ISBN 960–351–453–5/978–960–351–453–4.

[4] Στεφανάκος, Χ., Προγραμματισμός Η/Υ με MATLAB, Γκούρδας Εκδοτική, ISBN 978–960–387–856–8.

[5] Burden, Richard L. and Faires, J. Douglas (2000), Numerical Analysis (7th ed.), Brooks/Cole, ISBN 978–0–534–38216–2.

[6] Conte, S. D., Carl de Boor (1981), Elementary Numerical Analysis: An Algorithmic Approach (3rd ed.), McGraw-Hill Book Company, ISBN 978–0–07–012447–9.

[7] Don, E., Schaum's Outlines – Mathematica (2006), Εκδόσεις Κλειδάριθμος, ISBN 978–960–461–000–6.

[8] Kendell A. Atkinson (1989), An Introduction to Numerical Analysis (2nd ed.), John Wiley & Sons, ISBN 0–471–50023–2.

[9] Leader, Jeffery J. (2004), Numerical Analysis and Scientific Computation, Addison Wesley, ISBN 978–0–201–73499–7.

[10] Schatzman, M. (2002), Numerical Analysis: A Mathematical Introduction, Clarendon Press, Oxford, ISBN 978–0–19–850279–1.

[11] Stoer, Josef; Bulirsch, Roland (2002), Introduction to Numerical Analysis (3rd ed.), Springer, ISBN 978–0–387–95452–3.

[12] Sli, E. and 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://eqworld.ipmnet.ru/index.htm
- http://mathworld.wolfram.com/
- http://eom.springer.de/