Η ΤΕΧΝΗ ΤΟΥ ΔΙΑΒΑΣΜΑΤΟΣ ΜΕΤΑΞΥ ΤΩΝ ΑΡΙΘΜΩΝ (ΠΑΡΕΜΒΟΛΗ ΚΑΙ ΠΡΟΣΕΓΓΙΣΗ)

ΜΙΧΑΛΗΣ ΤΖΟΥΜΑΣ
ΔΕΣΠΟΤΑΤΟΥ 2
30100 ΑΓΡΙΝΙΟ

1. ΠΕΡΙΛΗΨΗ

Η έννοια της συνάρτησης είναι στενά συνυφασμένη με τον πίνακα τιμών και τη γραφική παράσταση. Μαθητές και φοιτητές είναι εξοικειωμένοι με τις έννοιες αυτές και μάλιστα εύκολα αναγνωρίζουν ή μπορούν να δημιουργούν τα παραπάνω από τον τύπο της συνάρτησης. Στην πράξη όμως, ίσως τις περισσότερες φορές, τα πράγματα εμφανίζονται «ανάποδα», δηλ. πρώτα έχουμε τον πίνακα τιμών (συνήθως πειραματικά αποτελέσματα ή δεδομένα άλλων μετρήσεων), στη συνέχεια ένα γράφημα αυτού και στο τέλος τον τύπο της συνάρτησης, που πιθανόν να περιγράφει το φαινόμενο, που μας έδωσε τα πειραματικά αποτελέσματα ή τα αποτελέσματα των μετρήσεών μας. Κάτι τέτοιο είναι εξαιρετικά χρήσιμο, αφού ένας τύπος μας επιτρέπει να βλέπουμε μεταξύ των τιμών ή να προβλέπουμε πέρα από αυτές. Η εύρεση ενός τέτοιου τύπου, που μας επιτρέπει να παρεμβάλουμε τιμές στις ήδη υπάρχουσες, γίνεται με **παρεμβολή**, ενώ η εύρεση ενός τύπου, που μας επιτρέπει να προσεγγίσουμε και άλλες τιμές από τις ήδη υπάρχουσες με **προσέγγιση**.

2. ΕΙΣΑΓΩΓΗ

Το πρόβλημα της δημιουργίας ενός τύπου από έναν πίνακα τιμών έχει άπειρες λύσεις, δηλαδή μπορούμε να βρούμε άπειρους τύπους συναρτήσεων, που να έχουν ως πίνακα τιμών τον πίνακα των αποτελεσμάτων μας. Το γεγονός αυτό κάνει το πρόβλημα εξαιρετικά δύσκολο και είναι ένας από τους λόγους που έχουν αναπτυχθεί πολλοί τρόποι για τον προσδιορισμό του τύπου αυτού. Θα μπορούσε, ίσως, κάποιος να αναφέρει τεχνικές, ανάλογα με τις συνθήκες που ο ερευνητής απαιτεί ο εν λόγω τύπος να πληροί, όπως π.χ. τα πολυώνυμα Hermite ή την παρεμβολή με Splines. Όμως, ο απλούστερος τρόπος, ένας τρόπος που ίσως θα μπορούσε να διδάσκεται και στο

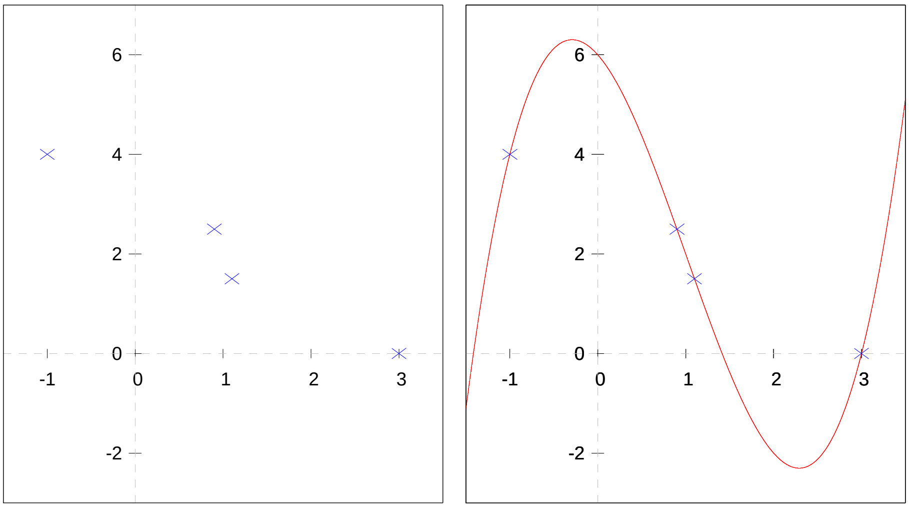
Λύκειο, είναι τα πολυώνυμα. Το Θεώρημα του Weierstrass ([1]), που παρατίθεται στη συνέχεια, επιτρέπει κάτι τέτοιο.

Θεώρημα 1. **Αν μια συνάρτηση *f* είναι συνεχής σε ένα διάστημα** $[a,b]$**, τότε για κάθε** ε **θετικό** $(\forall\varepsilon > 0)$ **υπάρχει ένα πολυώνυμο** $P(x)$**, έτσι ώστε**

$$|P(x) - f(x)| < \varepsilon,\ \forall x \in [a,b]. \tag{2.1}$$

Τον τρόπο επιλογής του πολυώνυμου αυτού, δηλ. του πολυώνυμου παρεμβολής, θα μελετήσουμε στην επόμενη παράγραφο.

Η δημιουργία ενός τύπου, από έναν πίνακα τιμών, δε σημαίνει ότι λύνει και το πρόβλημά μας. Για παράδειγμα στα σημεία Α(-1,4), Β(0.9,2.499), Γ(1.1,1.501) και Δ(3,0) αντιστοιχεί το πολυώνυμο $P(x) = x^3 - 3x^2 - 2x + 6$ (η γραφική παράσταση των σημείων και του πολυωνύμου φαίνεται στο Σχήμα 1) .



Σχήμα 1. Τα τέσσερα σημεία και το πολυώνυμο παρεμβολής

Από τη θέση των σημείων, στη γραφική παράσταση αυτών, φαίνεται ότι ο τύπος που προέρχεται από το πολυώνυμο παρεμβολής δύσκολα περιγράφει το πείραμα, που μας έδωσε τα εν λόγω σημεία. Θα μπορούσε κάποιος να ισχυρισθεί ότι τα σημεία αυτά βρίσκονται πάνω σε μια ευθεία και μικρές αποκλίσεις από αυτήν ερμηνεύονται από πιθανά σφάλματα. Τον τρόπο επιλογής της ευθείας (πολυώνυμο πρώτου βαθμού) ή καλύτερα της συνάρτησης προσέγγισης των σημείων αυτών θα μελετήσουμε στην τελευταία παράγραφο του άρθρου.

3. ΠΑΡΕΜΒΟΛΗ

3.1. Το πολυώνυμο του **Lagrange**

Θεωρούμε τα $n+1$ διαφορετικά μεταξύ τους σημεία, π.χ. τα

$$(x_i, f_i),\ i = 0(1)n,\ \text{με}\ x_i \neq x_j\ \text{για}\ i \neq j, \tag{3.1}$$

όπου $f_i = f(x_i)$ και f η άγνωστη συνάρτηση, που λύνει το πρόβλημά μας και πολύ θα θέλαμε να γνωρίζουμε. Επίσης, θεωρούμε το n-οστού βαθμού πολυώνυμο $P_n(x) = a_0 + a_1 x + \cdots + a_n x^n$, η γραφική παράσταση του οποίου διέρχεται από τα σημεία αυτά. Προφανώς, αφού τα σημεία της σχέσης (3.1) ανήκουν στο πολυώνυμο, ορίζονται οι επόμενες σχέσεις, δημιουργώντας το σύστημα

$$\begin{aligned}
f_0 &= a_0 + a_1 x_0 + \cdots + a_{n-1} x_0^{n-1} + a_n x_0^n \\
f_1 &= a_0 + a_1 x_1 + \cdots + a_{n-1} x_1^{n-1} + a_n x_1^n \\
&\ \ \vdots \\
f_n &= a_0 + a_1 x_n + \cdots + a_{n-1} x_n^{n-1} + a_n x_n^n
\end{aligned} \tag{3.2}$$

Η λύση του συστήματος (3.2) μας δίνει τους συντελεστές του πολυώνυμου $P_n(x)$. Η ορίζουσα των συντελεστών των αγνώστων είναι η γνωστή ορίζουσα του Vandermonde:

$$D_n = \begin{vmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{vmatrix} \tag{3.3}$$

***Λήμμα* 1. Η ορίζουσα του Vandermonde έχει τη μορφή**

$$D_n = \prod_{\substack{i=0,\, j=0 \\ i>j}}^{n} (x_i - x_j) \tag{3.4}$$

Πράγματι για $n = 2$, η (3.4) ισχύει. Υποθέτουμε ότι ισχύει για τυχαίο k και θα δείξουμε την αλήθεια της πρότασης για $k+1$. Θεωρούμε το πολυώνυμο

$$D_{k+1}(x) = \begin{vmatrix} 1 & x & x^2 & \dots & x^{k+1} \\ 1 & x_0 & x_0^2 & \dots & x_0^{k+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_k & x_k^2 & \dots & x_k^{k+1} \end{vmatrix},$$

του οποίου ρίζες είναι οι αριθμοί x_i, $i = 0(1)k$, ενώ ο συντελεστής του μεγιστοβάθμιου όρου x^{k+1} είναι $(-1)^{k+3} D_k$. Αυτό, τότε, γράφεται ως εξής:

$$D_{k+1}(x) = (-1)^{k+3} D_k \prod_{i=0}^{k} (x - x_i)$$

Η τιμή του πολυώνυμου αυτού για $x = x_{k+1}$, μετά από $k+1$ αλλαγές της πρώτης γραμμής, μας δίνει τη μορφή (3.4).

Η (3.4) είναι, προφανώς, διαφορετική από το μηδέν. Έτσι αποδείξαμε το επόμενο Θεώρημα.

Θεώρημα 2. **Από $n+1$ σημεία με διαφορετικές μεταξύ τους τετμημένες περνά ένα μοναδικό πολυώνυμο $P_n(x)$ βαθμού (το πολύ) n.**

Προφανώς, πολυώνυμα βαθμού μεγαλύτερου του n υπάρχουν άπειρα, αφού κάθε πολυώνυμο της μορφής $Q(x) = P_n(x) + R_m(x)\prod_{i=0}^{n}(x - x_i)$ διέρχεται από τα σημεία μας, ενώ πολυώνυμα βαθμού μικρότερου του n είναι δυνατόν να προκύψουν μόνο στην περίπτωση, που από τη λύση του συστήματος βρεθεί ο συντελεστής του μεγιστοβάθμιου όρου να είναι μηδέν.

Ο Lagrange, για τον προσδιορισμό αυτού του πολυώνυμου, είχε την εξής καταπληκτική ιδέα: Το $P_n(x)$ θα προκύπτει από το άθροισμα $n+1$ πολυωνύμων της μορφής $L_i(x) f_i$, $i = 0(1)n$, δηλαδή

$$P_n(x) = \sum_{i=0}^{n} L_i(x) f_i \tag{3.5}$$

Επιπλέον, για να περνά το πολυώνυμο (3.5) από τα σημεία μας, απαίτησε δυο πράγματα:

1) $L_i(x_j) = 0, \forall i, j = 0(1)n$ με $i \neq j$ και

2) $L_i(x_i) = 1$.

Η πρώτη, από τις παραπάνω απαιτήσεις, μας δείχνει ότι η μορφή των επιμέρους πολυωνύμων είναι

$$L_i(x) = A_i \prod_{\substack{j=0 \\ j\neq i}}^{n} (x - x_j),$$

ενώ η δεύτερη μας προσδιορίζει τα A_i, δηλαδή

$$A_i = \frac{1}{\prod_{\substack{j=0\\ j\neq i}}^{n} (x_i - x_j)}$$

Έτσι, το πολυώνυμο παρεμβολής του Lagrange, τελικά, έχει τη μορφή:

$$P_n(x) = \sum_{i=0}^{n} \frac{\prod_{\substack{j=0\\ j\neq i}}^{n} (x - x_j)}{\prod_{\substack{j=0\\ j\neq i}}^{n} (x_i - x_j)} f_i \tag{3.6}$$

Όταν οι τιμές $f_i = f(x_i)$ είναι τιμές της συνάρτησης f, το σφάλμα που γίνεται, όταν τις τιμές $f(x)$ τις προσεγγίζουμε από τις τιμές του πολυώνυμου $P_n(x)$, πρέπει να έχει τη μορφή

$$\varepsilon = A(x) \prod_{i-0}^{n} (x - x_i), \tag{3.7}$$

αφού αυτό μηδενίζεται για $x = x_i$, $i = 0(1)n$. Ο συντελεστής $A(x)$ προσδιορίζεται στα περισσότερα βιβλία της Αριθμητικής Ανάλυσης ([2]) και δεν θα αναφερθούμε σ' αυτόν.

Το πολυώνυμο παρεμβολής του Lagrange είχε μέχρι τελευταία, ως μειονέκτημα, τη δυσκολία που παρουσίαζε στον προγραμματισμό του σε ηλεκτρονικό υπολογιστή (Η/Υ). Σήμερα, υπάρχουν υπολογιστικά περιβάλλοντα, όπως εκείνα του Maple και του Mathematica αλλά και γλώσσες προγραμματισμού, όπως η C++, που διαχειρίζονται λίστες και τα πράγματα έχουν αλλάξει.

3.2. Το πολυώνυμο των Newton – Gregory

Πολλές φορές τα σημεία στα οποία γίνεται η παρεμβολή ισαπέχουν, δηλ. δυο διαδοχικά σημεία έχουν σταθερή διαφορά. Αν με h παραστήσουμε αυτή τη διαφορά και με I το υποσύνολο εκείνο των ακεραίων, που εκλέξαμε ως δείκτες στα σημεία μας, θα ισχύει $x_{i+1} - x_i = h,\ \forall i, i+1 \in I$. Αν με I_0 παραστήσουμε το προηγούμενο σύνολο που περιέχει το μηδέν, τότε θα ισχύει

$$x_i = x_0 + ih,\ \ \forall i \in I_0,$$

ενώ για οποιοδήποτε $x \in (x_0, x_1)$ έχουμε

$$x = x_0 + \theta h,\ \text{με}\ \theta \in (0,1). \tag{3.8}$$

Θεωρούμε αρχικά δυο σημεία, τα (x_0, f_0) και (x_1, f_1), στα οποία γίνεται η παρεμβολή. Προφανώς το πολυώνυμο $1^{ου}$ βαθμού που διέρχεται από αυτά είναι το

$$\begin{aligned} P_1(x) &= f_0 + \frac{x - x_0}{x_1 - x_0}(f_1 - f_0) \\ &= f_0 + \frac{\theta h}{h}\Delta f_0 \\ &= f_0 + \theta \Delta f_0, \end{aligned}$$

όπου Δf_0 είναι η διαφορά πρώτης τάξης στο f_0, ενώ, λαμβάνοντας υπόψη το σφάλμα για κάθε άλλο σημείο εκτός της ευθείας, από τη σχέση (3.7) θα ισχύει

$$f(x) = f_0 + \theta \Delta f_0 + A(x - x_0)(x - x_1). \tag{3.9}$$

Θεωρώντας τώρα τρία σημεία, τα (x_0, f_0), (x_1, f_1) και (x_2, f_2), στα οποία γίνεται η παρεμβολή και απαιτώντας το τρίτο να επαληθεύει την (3.9) μπορούμε να προσδιορίσουμε το A. Θέτοντας $x = x_2$ στην (3.8) παίρνουμε $\theta = 2$, ενώ από την (3.9) έχουμε $f_2 = f_0 + 2\Delta f_0 + 2Ah^2$. Έτσι, μετά από λίγες πράξεις παίρνουμε

$$P_2(x) = f_0 + \theta \Delta f_0 + \frac{\theta(\theta - 1)}{2}\Delta^2 f_0,$$

όπου $\Delta^2 f_0$ είναι η διαφορά δεύτερης τάξης στο f_0. Επαγωγικά μπορούμε να γενικεύσουμε την παραπάνω διαδικασία και να έχουμε

$$P_n(x) = f_0 + \theta \Delta f_0 + \cdots + \frac{\theta(\theta - 1)\cdots(\theta - n + 1)}{n!}\Delta^n f\ .$$

Εφαρμογή 1. Να βρεθεί το άθροισμα $S_2 = 1^2 + 2^2 + \cdots + n^2$.

Λύση: Το παραπάνω άθροισμα είναι η τιμή, για $x = n$, της συνάρτησης $f(x) = 1^2 + 2^2 + \cdots + ([x] - 1)^2 + x^2$, με $x > 1$. Ο πίνακας διαφορών αυτής έχει την επόμενη μορφή, στην οποία φαίνεται ότι η συνάρτηση f είναι ένα πολυώνυμο τρίτου βαθμού, αφού οι διαφορές τρίτης τάξης, για ισαπέχουσες

x	$f(x)$	Δf	$\Delta^2 f$	$\Delta^3 f$
1	1			
		4		
2	5		5	
		9		2
3	14		7	
		16		2
4	30		9	
		25		
5	55			

τιμές της μεταβλητής, είναι σταθερές.

Χρησιμοποιώντας τον τύπο παρεμβολής των Newton – Gregory, με $x_0 = 1$ και $x = n$, οπότε $\theta = n - 1$, προκύπτει:

$$f(n) = S_3 = 1 + 4\theta + \frac{5}{2}\theta(\theta-1) + \frac{2}{6}\theta(\theta-1)(\theta-2)$$
$$= \frac{1}{6}n(n+1)(2n+1).$$

4. ΠΡΟΣΕΓΓΙΣΗ

Το πρόβλημα της προσέγγισης είναι περισσότερο πολύπλοκο και μάλλον δεν μπορεί να διδάσκεται στο Λύκειο. Ας θεωρήσουμε πάλι τα $n+1$ γνωστά σημεία, τα (x_i, f_i), με $i = 0(1)n$. Επιπλέον, ας θεωρήσουμε τη συνάρτηση y που προσεγγίζει αυτά. Για κάθε τιμή x_i, από την προηγούμενη συνάρτηση προκύπτει μια τιμή y_i, οπότε το σφάλμα που γίνεται κάθε φορά είναι

$$\varepsilon_i = f_i - y_i, \ i = 0(1)n .$$

Αν $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)^T$ το διάνυσμα σφάλματος, τότε ένα καλό κριτήριο για την επιλογή της συνάρτησης y είναι η ελαχιστοποίηση της νόρμας $\|\varepsilon\|$.

4.1. Η ευθεία των ελαχίστων τετραγώνων

Αν επιλέξουμε ως νόρμα την Ευκλείδεια νόρμα $\|\varepsilon\|_2$ και ως συνάρτηση το πρωτοβάθμιο πολυώνυμο

$$y = ax + b , \tag{4.1}$$

τότε έχουμε, ως συνάρτηση προσέγγισης, την αποκαλούμενη «ευθεία των ελαχίστων τετραγώνων». Στην περίπτωση αυτή ουσιαστικά έχουμε να ελαχιστοποιήσουμε τη συνάρτηση:

$$E(a,b) = \|\varepsilon\|_2^2 = \sum_{i=0}^{n} (f_i - ax_i - b)^2$$

Οι συνθήκες πρώτης τάξης, που μας δίνουν τα κρίσιμα σημεία ([1],[3]), είναι το σύστημα

$$\begin{cases} \dfrac{\partial E}{\partial a} = 0 \\ \dfrac{\partial E}{\partial b} = 0 \end{cases} \Leftrightarrow \cdots \Leftrightarrow \begin{cases} a\sum_{i=0}^{n} x_i^2 + b\sum_{i=0}^{n} x_i = \sum_{i=0}^{n} f_i x_i \\ a\sum_{i=0}^{n} x_i + (n+1)b = \sum_{i=0}^{n} f_i \end{cases} \tag{4.2}$$

Η λύση του συστήματος (4.2) επαληθεύει τις συνθήκες δεύτερης τάξης ([1], [3]), δηλ. ισχύει

$$\begin{cases} \dfrac{\partial^2 E}{\partial a^2} = 2\sum_{i=0}^{n} x_i^2 \quad \text{και} \\ \dfrac{\partial^2 E}{\partial a^2}\dfrac{\partial^2 E}{\partial b^2} - \left(\dfrac{\partial^2 E}{\partial a \partial b}\right)^2 = \cdots = n\sum_{i=0}^{n} x_i^2 - \left(\sum_{i=0}^{n} x_i\right)^2 > 0 \end{cases} \tag{4.3}$$

Το τελευταίο προκύπτει άμεσα από την ανισότητα του Schwarz ([4])

$$\left|x^T y\right| \le \|x\| \cdot \|y\|$$

με $x = (x_1, x_2, \ldots, x_n)^T$ και $y = (1,1,\ldots,1)^T$. Έτσι η «ευθεία των ελαχίστων τετραγώνων» προκύπτει από τη λύση του συστήματος (4.2).

Η αλλαγή της νόρμας σε $\|\varepsilon\|_\infty$, με μια διαδικασία μάλλον πολύπλοκη, μας δίνει τη min-max ευθεία. Η διατήρηση της Ευκλείδειας νόρμας και η αλλαγή του πολυώνυμου σε πολυώνυμο βαθμού $1 < m < n$, αποδεικνύεται ότι μας δίνει μοναδικά βέλτιστα πολυώνυμα βαθμού m. Όμως, τα συστήματα που πρέπει να επιλυθούν για να βρούμε τους συντελεστές των πολυωνύμων αυτών δεν είναι «ευσταθή» για $m > 7$, με αποτέλεσμα να δημιουργούνται προβλήματα στη λύση τους ([5], [6]).

4.2. Προσεγγίσεις με άλλες συναρτήσεις

Πολλές φορές η γραφική παράσταση των τιμών οδηγεί σε εκθετικές συναρτήσεις της μορφής

$$y = ce^{ax} \qquad \text{ή} \qquad y = cx^a . \tag{4.4}$$

Η απευθείας χρήση της μεθόδου των ελαχίστων τετραγώνων δημιουργεί προβλήματα. Ωστόσο, αν λογαριθμίσουμε τις σχέσεις (4.4) παίρνουμε

$$\ln(y) = ax + \ln(b) \qquad \text{ή} \qquad \ln(y) = a\ln(x) + \ln(c) . \tag{4.5}$$

Αν θέσουμε $\hat{y} = \ln(y)$, $\hat{x} = \ln(x)$ και $b = \ln(c)$ στη σχέση (4.5), προκύπτουν

$$\hat{y} = ax + b \quad \text{ή} \quad \hat{y} = a\hat{x} + b. \tag{4.6}$$

Οι σχέσεις (4.6) μοιάζουν με εκείνες της (4.1). Μπορούμε, τώρα, να ακολουθήσουμε τη διαδικασία της προηγούμενης παραγράφου, με την παρατήρηση ότι οι $\hat{y} = \ln(y)$ και $\hat{x} = \ln(x)$ προσεγγίζουν τις τιμές $\ln(f_i)$ και $\ln(x_i)$ αντίστοιχα.

Η μέθοδος των ελαχίστων τετραγώνων χρησιμοποιήθηκε μέχρι εδώ με στοιχεία από το σύνολο των συναρτήσεων

$$P = \{1, x, x^2, \ldots, x^m\}.$$

Όμως, η μέθοδος δουλεύει εξίσου καλά και με στοιχεία από οποιοδήποτε σύνολο με γραμμικά ανεξάρτητες συναρτήσεις. Έτσι, κάποιος θα μπορούσε να δουλέψει π.χ. με τα σύνολα

$$T = \{1, \cos(x), \cos(2x), \ldots, \cos(mx)\} \text{ ή } S = \{1, e^{\lambda_1 x}, e^{\lambda_2 x}, \ldots, e^{\lambda_m x}\},$$

χωρίς αλλαγές στη θεωρία του.

ΒΙΒΛΙΟΓΡΑΦΙΑ

[1] **L. Brand:** *Μαθηματική Ανάλυση*, Ελληνική Μαθηματική Εταιρία, Αθήνα, 1984.

[2] **Γ. Ακρίβης, Β. Δουγαλής:** *Εισαγωγή στην Αριθμητική Ανάλυση*, Παν/κές εκδόσεις Κρήτης, 1997.

[3] **G. Tomas, R. Finney:** *Απειροστικός Λογισμός, τόμος I και II*, Παν/κές Εκδόσεις Κρήτης, 1993.

[4] **G. Strang:** *Γραμμική Άλγεβρα*, Παν/κές εκδόσεις Κρήτης, Ηράκλειο, 1996.

[5] **Α. Χατζηδήμος:** *Αριθμητική Ανάλυση I και II*, Δεύτερη έκδοση, Ιωάννινα, 1981

[6] **Α. Γιέγιος:** *Αριθμητικές μέθοδοι, τόμος I*, Ιωάννινα, 1986.