Η ΤΕΧΝΗ ΤΟΥ ΔΙΑΒΑΣΜΑΤΟΣ ΜΕΤΑΞΥ ΤΩΝ ΑΡΙΘΜΩΝ

Μιχάλης Τζούμας

4ο ΓΥΜΝΑΣΙΟ ΑΓΡΙΝΙΟΥ

Περιεχόμενα

1. Εισαγωγή
2. Η Παρεμβολή
   - Πολυώνυμο του Lagrange
   - Το πολυώνυμο των Newton - Gregory
3. Προσέγγιση
   - Τα ελάχιστα τετράγωνα
   - min-max ευθεία και άλλες συναρτήσεις
4. Επίλογος

Συναρτήσεις

Τύπος $\longrightarrow$ Πίνακας Τιμών (Γράφημα)

Τύπος $\xleftarrow{?}$ Πίνακας Τιμών (Γράφημα)

Το πρόβλημα

Στην πράξη τα πράγματα εμφανίζονται «ανάποδα», δηλ. πρώτα έχουμε τον πίνακα τιμών και στο τέλος τον τύπο της συνάρτησης, που πιθανόν να περιγράφει το φαινόμενο.

Η εύρεση ενός τέτοιου τύπου, που επαληθεύει τις ήδη υπάρχουσες τιμές και μας επιτρέπει να παρεμβάλλουμε κι άλλες, λέγεται **παρεμβολή**,

ενώ η εύρεση ενός τύπου, που προσεγγίζει τις ήδη υπάρχουσες τιμές και μας επιτρέπει να προσεγγίσουμε κι άλλες, λέγεται **προσέγγιση**.

Η λύση

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.

Ένας απλός τρόπος

Ένας απλός τρόπος, που ίσως θα μπορούσε να διδάσκεται και στο Λύκειο, είναι τα πολυώνυμα.

Θεώρημα του Weierstrass:

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

Υποθέσεις

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

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

όπου $f_i = f(x_i)$ και f η άγνωστη συνάρτηση. Επίσης, θεωρούμε το n-οστού βαθμού πολυώνυμο,

$$P_n(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

η γραφική παράσταση του οποίου διέρχεται από τα σημεία αυτά.

Το Σύστημα

Προφανώς, αφού τα σημεία (x_i, f_i) ανήκουν στο πολυώνυμο, ορίζονται οι επόμενες σχέσεις

$$\begin{aligned}
f_0 &= a_0 + a_1 x_0 + a_2 x_0^2 + \cdots + a_n x_0^n \\
f_1 &= a_0 + a_1 x_1 + a_2 x_1^2 + \cdots + a_n x_1^n \\
&\vdots \\
f_n &= a_0 + a_1 x_n + a_2 x_n^2 + \cdots + a_n x_n^n
\end{aligned}$$

Η λύση του συστήματος μας δίνει τους συντελεστές του πολυώνυμου.

Η Ορίζουσα

Η ορίζουσα των συντελεστών των αγνώστων είναι η γνωστή ορίζουσα του Vandermonde:

$$D_{n+1} = \begin{vmatrix} 1 & x_0 & \cdots & x_0^n \\ 1 & x_1 & \cdots & x_1^n \\ & \vdots & & \\ 1 & x_n & \cdots & x_n^n \end{vmatrix} = \prod_{\substack{i=1 \\ j=1 \\ i>j}}^{n} (x_i - x_j)$$

Αυτή είναι, προφανώς, διαφορετική από το μηδέν.

Το Θεώρημα

Έτσι, από $n+1$ σημεία με διαφορετικές μεταξύ τους τετμημένες περνά ένα μοναδικό πολυώνυμο $P_n(x)$ βαθμού (το πολύ) n.

Το Θεώρημα

Έτσι, από $n+1$ σημεία με διαφορετικές μεταξύ τους τετμημένες περνά ένα μοναδικό πολυώνυμο $P_n(x)$ βαθμού (το πολύ) n.

Υπάρχουν άπειρα πολυώνυμα βαθμού μεγαλύτερου του n, π.χ. 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 είναι δυνατόν να προκύψουν μόνο στην περίπτωση, που από τη λύση του συστήματος βρεθεί ο συντελεστής του μεγιστοβάθμιου όρου να είναι μηδέν.

Η ιδέα

Ο Lagrange είχε την εξής καταπληκτική ιδέα: Το $P_n(x)$ θα προκύπτει από το άθροισμα $n+1$ πολυώνυμων, βαθμού n, της μορφής $L_i(x)f_i$ και απαίτησε, για να περνά το πολυώνυμο από τα σημεία μας, δυο πράγματα:

1) $L_i(x_j) = 0, \; i,j = 0(1)n$ με $i \neq j$
2) $L_i(x_i) = 1, \; i = 0(1)n$

Το Πολυώνυμο - Το Σφάλμα

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

Το Πολυώνυμο - Το Σφάλμα

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

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

Το Πολυώνυμο - Το Σφάλμα

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

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

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

Το Πολυώνυμο - Το Σφάλμα

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

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

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

$$f(x) - P_n(x) = E(x) = A(x) \prod_{i=0}^{n} (x - x_i)$$

Οι Δυσκολίες

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.

Υποθέσεις

Πολλές φορές τα σημεία 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 $x_{i+1} - x_i = h, \ \forall i \in I$. Γενικά ισχύει:

$$x_i = x_0 + ih, \ \ i \in I$$

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

$$x = x_0 + \theta h, \ \ \theta \in (0, i)$$

Τα Πολυώνυμα

Το πολυώνυμο 1ου βαθμού που διέρχεται από τα σημεία (x_0, f_0) και (x_1, f_1) είναι

$$P_1(x) = f_0 + \frac{x - x_0}{x_1 - x_0}(f_1 - f_0) = \cdots = f_0 + \theta \Delta f_0$$

Τα Πολυώνυμα

Το πολυώνυμο 1ου βαθμού που διέρχεται από τα σημεία (x_0, f_0) και (x_1, f_1) είναι

$$P_1(x) = f_0 + \frac{x - x_0}{x_1 - x_0}(f_1 - f_0) = \cdots = f_0 + \theta \Delta f_0$$

Το δε σφάλμα είναι

$$f(x) - P_1(x) = E(x) = A(x - x_0)(x - x_1)$$

Τα Πολυώνυμα

Το πολυώνυμο 1ου βαθμού που διέρχεται από τα σημεία (x_0, f_0) και (x_1, f_1) είναι

$$P_1(x) = f_0 + \frac{x - x_0}{x_1 - x_0}(f_1 - f_0) = \cdots = f_0 + \theta \Delta f_0$$

Το δε σφάλμα είναι

$$f(x) - P_1(x) = E(x) = A(x - x_0)(x - x_1)$$

οπότε η συνάρτηση f έχει τη μορφή

$$f(x) = P_1(x) + A(x - x_0)(x - x_1)$$

Τα Πολυώνυμα

Απαιτώντας το σημείο (x_2, f_2) να δίνει σφάλμα μηδέν προσδιορίζουμε το A

$$A = \frac{f(x_2) - P_1(x_2)}{(x_2 - x_0)(x_2 - x_1)}$$

Τα Πολυώνυμα

Απαιτώντας το σημείο (x_2, f_2) να δίνει σφάλμα μηδέν προσδιορίζουμε το A

$$A = \frac{f(x_2) - P_1(x_2)}{(x_2 - x_0)(x_2 - x_1)}$$

οπότε

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

Τα Πολυώνυμα

Απαιτώντας το σημείο (x_2, f_2) να δίνει σφάλμα μηδέν προσδιορίζουμε το A

$$A = \frac{f(x_2) - P_1(x_2)}{(x_2 - x_0)(x_2 - x_1)}$$

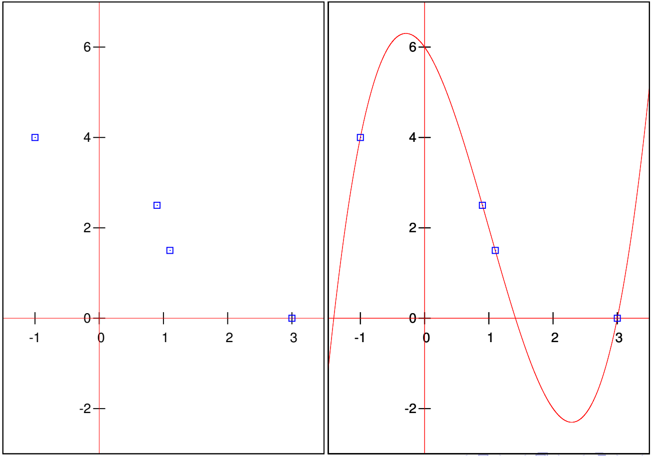
οπότε

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

και γενικά

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

Προβλήματα!!!



Τα Σφάλματα

Το πρόβλημα της προσέγγισης είναι περισσότερο πολύπλοκο και μάλλον δεν μπορεί να διδάσκεται στο Λύκειο. Ας θεωρήσουμε πάλι τα $n+1$ γνωστά σημεία, τα $(x_i, f_i), i = 0(1)n$. 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 $\varepsilon = (\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_n)^T$ το διάνυσμα σφάλματος, τότε ένα καλό κριτήριο για την επιλογή της συνάρτησης y είναι η ελαχιστοποίηση της νόρμας $\| \varepsilon \|$.

Η Ευθεία

Αν επιλέξουμε ως νόρμα την Ευκλείδεια νόρμα $\| \varepsilon \|_2$ και ως συνάρτηση το πρωτοβάθμιο πολυώνυμο $y = ax + b$ 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

$$\left\{\begin{array}{l} \frac{\partial E}{\partial a} = 0 \\ \frac{\partial E}{\partial a} = 0 \end{array}\right. \Rightarrow \cdots \Rightarrow \left\{\begin{array}{c} a\sum_{i=0}^{n} x_i^2 + b\sum_{i=0}^{n} x_i = \sum_{i=0}^{n} f_i x_i \\ a\sum_{i=0}^{n} x_i + (n+1)b = \sum_{i=0}^{n} f_i \end{array}\right.$$

Οι συνθήκες

Πρώτης τάξης

$$\left\{\begin{array}{l} \frac{\partial E}{\partial a} = 0 \\ \frac{\partial E}{\partial a} = 0 \end{array}\right. \Rightarrow \cdots \Rightarrow \left\{\begin{array}{c} a\sum_{i=0}^{n} x_i^2 + b\sum_{i=0}^{n} x_i = \sum_{i=0}^{n} f_i x_i \\ a\sum_{i=0}^{n} x_i + (n+1)b = \sum_{i=0}^{n} f_i \end{array}\right.$$

και δεύτερης τάξης

$$\left\{\begin{array}{c} \frac{\partial^2 E}{\partial a^2} > 0 \\ \frac{\partial^2 E}{\partial a^2}\frac{\partial^2 E}{\partial b^2} - \left(\frac{\partial^2 E}{\partial a \partial b}\right)^2 > 0 \end{array}\right. \Leftrightarrow \cdots \Leftrightarrow \left\{\begin{array}{c} 2\sum_{i=0}^{n} x_i^2 > 0 \\ n\sum_{i=0}^{n} x_i^2 - \left(\sum_{i=0}^{n} x_i\right)^2 > 0 \end{array}\right.$$

Η Παραβολή και άλλα Πολυώνυμα

Η διατήρηση της Ευκλείδειας νόρμας και η αλλαγή του πολυώνυμου σε πολυώνυμο βαθμού $1 < m < n$, αποδεικνύεται μάλλον πολύπλοκα ότι μας δίνει μοναδικά βέλτιστα πολυώνυμα $P_m(x)$. Όμως, τα συστήματα που πρέπει να επιλυθούν για να βρούμε τους συντελεστές των πολυώνυμων αυτών δεν είναι «ευσταθή» για $m > 7$, με αποτέλεσμα να δημιουργούνται προβλήματα στη λύση τους.

Η min-max ευθεία

Η αλλαγή της νόρμας σε $\| \varepsilon \|_\infty$, με μια διαδικασία μάλλον πολύπλοκη, μας δίνει τη min-max ευθεία.

Οι συναρτήσεις ce^{ax} και cx^a

Πολλές φορές η γραφική παράσταση των τιμών οδηγεί σε εκθετικές συναρτήσεις της μορφής $y = ce^{ax}$ ή $y = cx^a$. Η απ' ευθείας χρήση της μεθόδου των ελαχίστων τετραγώνων δημιουργεί προβλήματα. Ωστόσο, αν λογαριθμίσουμε τις σχέσεις αυτές παίρνουμε

$$\ln y = ax + \ln c \quad \text{ή} \quad \ln y = a \ln x + \ln c$$

Οι συναρτήσεις ce^{ax} και cx^a

Πολλές φορές η γραφική παράσταση των τιμών οδηγεί σε εκθετικές συναρτήσεις της μορφής $y = ce^{ax}$ ή $y = cx^a$. Η απ' ευθείας χρήση της μεθόδου των ελαχίστων τετραγώνων δημιουργεί προβλήματα. Ωστόσο, αν λογαριθμίσουμε τις σχέσεις αυτές παίρνουμε

$$\ln y = ax + \ln c \quad \text{ή} \quad \ln y = a \ln x + \ln c$$

ή ισοδύναμα

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

Οι άλλες συναρτήσεις

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

$$\{1, x, x^2, \ldots, x^n\}$$

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

$$\{1, \cos x, \cos 2x, \ldots, \cos nx\} \text{ ή } \{1, e^{\lambda_1 x}, e^{\lambda_2 x}, \ldots, e^{\lambda_n x}\}$$

χωρίς αλλαγές στη θεωρία του.

Επίλογος

ΣΥΜΠΕΡΑΣΜΑ

- ★ Η Παρεμβολή μάλλον μπορεί να διδάσκεται στο Λύκειο.

Επίλογος

ΣΥΜΠΕΡΑΣΜΑ

- ★ Η Παρεμβολή μάλλον μπορεί να διδάσκεται στο Λύκειο.
- ★ Η Προσέγγιση ΔΕΝ ΜΠΟΡΕΙ να διδάσκεται.

Επίλογος

ΣΥΜΠΕΡΑΣΜΑ

- ★ Η Παρεμβολή μάλλον μπορεί να διδάσκεται στο Λύκειο.
- ★ Η Προσέγγιση ΔΕΝ ΜΠΟΡΕΙ να διδάσκεται.

- ΟΜΩΣ ΟΛΟΙ ΓΝΩΡΙΖΟΥΜΕ ΟΤΙ
- ★ Η Παρεμβολή ΔΕ διδάσκεται!
- ★ Η Προσέγγιση διδάσκεται!

Επίλογος

ΣΥΜΠΕΡΑΣΜΑ

- ★ Η Παρεμβολή μάλλον μπορεί να διδάσκεται στο Λύκειο.
- ★ Η Προσέγγιση ΔΕΝ ΜΠΟΡΕΙ να διδάσκεται.

- ΟΜΩΣ ΟΛΟΙ ΓΝΩΡΙΖΟΥΜΕ ΟΤΙ
- ★ Η Παρεμβολή ΔΕ διδάσκεται!
- ★ Η Προσέγγιση διδάσκεται!

Ευχαριστώ !!!