ΘΕΩΡΙΑ ΠΡΟΣΕΓΓΙΣΗΣ

ΔΗΜΗΤΡΙΟΣ ΝΟΥΤΣΟΣ
ΚΑΘΗΓΗΤΗΣ
ΤΜΗΜΑΤΟΣ ΜΑΘΗΜΑΤΚΩΝ

Ιωάννινα 2014

Περιεχόμενα

Κεφάλαιο 1

ΕΙΣΑΓΩΓΗ

Η Θεωρία Προσέγγισης είναι κλάδος της επιστήμης της Μαθηματικής Ανάλυσης αν τη δούμε υπό το πρίσμα της θεωρίας που χρειάζεται για την επίτευξη της επιθυμητής προσέγγισης. Θεωρείται όμως και κλάδος της Αριθμητικής Ανάλυσης αν τη δούμε υπό το πρίσμα του τελικού προϊόντος που είναι μια Αριθμητική μέθοδος και τελικά ένας αλγόριθμος για την εύρεση της προσέγγισης. Όπως θα δούμε στη συνέχεια τα μαθηματικά εργαλεία της Θεωρίας Προσέγγισης προέρχονται από τον Απειροστικό Λογισμό, τη Γραμμική Άλγεβρα, την Τοπολογία και τη Συναρτησιακή Ανάλυση.

Το γενικότερο πρόβλημα της Θεωρίας Προσέγγισης συνίσταται στην εύρεση λύσης στο παρακάτω πρόβλημα:

Πρόβλημα 1 *Έστω V ένας σταθμητός γραμμικός χώρος εφοδιασμένος με τη νόρμα $\|\cdot\|$ και W ένας υποχώρος του V. Δοθέντος $v \in V$, Να βρεθεί $w^* \in W$ τέτοιο ώστε*

$$\|v - w^*\| \leq \|v - w\| \quad \forall w \in W. \tag{1.1}$$

Θα λέμε τότε ότι w^* είναι η βέλτιστη προσέγγιση του v στον W.

Με άλλα λόγια θέλουμε να βρούμε το πλησιέστερο στοιχείο στο v που ανήκει στον W. Αυτό το κάνουμε επειδή έχουμε περισσότερα μαθηματικά εργαλεία για να μελετήσουμε ένα πρόβλημα στον υποχώρο W από εκείνα που έχουμε στο χώρο V.

Θα ασχοληθούμε εδώ κυρίως με την προσέγγιση συναρτήσεων στον υποχώρο P_n των πολυωνύμων n βαθμού. Ως χώρο αναφοράς V θα έχουμε κυρίως το χώρο των συνεχών συναρτήσεων στο κλειστό διάστημα $[a, b]$ που θα το συμβολίζουμε με $C[a, b]$ ή το χώρο των συναρτήσεων που ορίζονται σε ένα σύνολο διακριτών σημείων $\{x_1, x_2, \cdots x_m\}$ που θα το συμβολίζουμε με X_m.

Στο εισαγωγικό αυτό κεφάλαιο θα δώσουμε κάποια πρώτα στοιχεία για τις νόρμες που θα χρησιμοποιήσουμε και στη συνέχεια μια γενική θεωρία ύπαρξης και μονοσήμαντου της βέλτιστης προσέγγισης.

1.1 Νόρμες

Ο γενικός ορισμός της νόρμας είναι ο ακόλουθος: Έστω V ένας γραμμικός χώρος. Νόρμα στο V θα είναι μια συνάρτηση από το V στους μη αρνητικούς πραγματικούς αριθμούς $\mathbb{R}^{+,0}$ ($\|\cdot\| : V \to \mathbb{R}^{+,0}$) που πληροί τις παρακάτω τρεις ιδιότητες:

(i) $\|v\| \geq 0$ και $\|v\| = 0$ αν και μόνο άν (ανν) $v = 0$.
(ii) $\|\lambda v\| = |\lambda| \|v\|$ για κάθε σταθερό $\lambda \in \mathbb{R}$ (ή $\lambda \in \mathbb{C}$).
(iii) $\|v + u\| \leq \|v\| + \|u\|$ (τριγωνική ιδιότητα).

Μια νόρμα ορίζει μια απόσταση στο V τον οποίο καθιστά μετρικό γραμμικό χώρο. Αν v και u είναι στοιχεία του V, τότε $\|v - u\|$ είναι η απόσταση μεταξύ αυτών των στοιχείων.

Οι κυριότερες νόρμες που χρησιμοποιούνται είναι οι νόρμες 1, 2 και ∞ ή l_1, l_2 και l_∞ αντίστοιχα. Αν ως χώρο V έχουμε το χώρο συνεχών συναρτήσεων $C[a,b]$ τότε οι αντίστοιχες νόρμες ορίζονται ως:

$$l_1 : \quad \|v\|_1 = \int_a^b |v(x)| dx.$$

$$l_2 : \quad \|v\|_2 = \left(\int_a^b |v(x)|^2 dx \right)^{\frac{1}{2}}.$$

$$l_\infty : \quad \|v\|_\infty = \max_{a \leq x \leq b} |v(x)|.$$

Αυτές αποτελούν ειδικές περιπτώσεις της l_p νόρμας που ορίζεται ως:

$$l_p : \quad \|v\|_p = \left(\int_a^b |v(x)|^p dx \right)^{\frac{1}{p}}, \quad p \geq 1.$$

Ας σημειωθεί ότι ο p μπορεί να είναι θετικός πραγματικός (≥ 1) και όχι απαραίτητα ακέραιος. Αποδεικνύεται δε, σχετικά εύκολα, ότι $\lim_{p \to \infty} \|v\|_p = \|v\|_\infty$.

Αντίστοιχα, αν ως χώρο V έχουμε το χώρο X_m, των συναρτήσεων σε διακριτό σύνολο, αυτές είναι γνωστές από τη Γραμμική Άλγεβρα ως διανυσματικές νόρμες και ορίζονται ως εξής:

$$l_1 : \quad \|v\|_1 = \sum_{i=1}^{m} |v_i|.$$

$$l_2 : \quad \|v\|_2 = \left(\sum_{i=1}^{m} |v_i|^2 \right)^{\frac{1}{2}}.$$

$$l_\infty : \quad \|v\|_\infty = \max_{i \in \{1,2,\cdots,m\}} |v_i|.$$

Αυτές αποτελούν ειδικές περιπτώσεις της l_p νόρμας που ορίζεται ως:

$$l_p : \quad \|v\|_p = \left(\sum_{i=1}^{n} |v_i|^p \right)^{\frac{1}{p}}, \quad p \geq 1.$$

Αποδεικνύεται και εδώ, σχετικά εύκολα, ότι $\lim_{p\to\infty}\|v\|_p = \|v\|_\infty$.

Είναι εύκολο να αποδειχτεί η ισχύς των τριών ιδιοτήτων για τις νόρμες l_1 και l_∞ ενώ για την l_2 ή γενικότερα για την l_p είναι δύσκολο να αποδειχτεί η τριγωνική ιδιότητα. Δίνουμε εδώ, χωρίς απόδειξη τις ανισότητες, γνωστές ως ανισότητες Minkowski, που αποδεικνύουν την ιδιότητα αυτή.

$$\left(\int_a^b |v(x)+u(x)|^p dx\right)^{\frac{1}{p}} \leq \left(\int_a^b |v(x)|^p dx\right)^{\frac{1}{p}} + \left(\int_a^b |u(x)|^p dx\right)^{\frac{1}{p}}$$

και

$$\left(\sum_{i=1}^n |v_i+u_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^n |v_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^n |u_i|^p\right)^{\frac{1}{p}}.$$

Δυο άλλες χρήσιμες ανισότητες που αναφέρονται στην l_p νόρμα είναι οι ανισότητες Hölder:

$$\int_a^b |v(x)u(x)|dx \leq \left(\int_a^b |v(x)|^p dx\right)^{\frac{1}{p}} \cdot \left(\int_a^b |u(x)|^q dx\right)^{\frac{1}{q}}, \quad p>1, \quad \frac{1}{p}+\frac{1}{q}=1$$

και

$$\sum_{i=1}^n |v_i u_i| \leq \left(\sum_{i=1}^n |v_i|^p\right)^{\frac{1}{p}} \cdot \left(\sum_{i=1}^n |u_i|^q\right)^{\frac{1}{q}}, \quad p>1, \quad \frac{1}{p}+\frac{1}{q}=1.$$

Ειδικές περιπτώσεις αυτών για $p=q=2$ είναι οι γνωστες ως ανισότητες Cauchy-Schwarz:

$$\int_a^b |v(x)u(x)|dx \leq \left(\int_a^b |v(x)|^2 dx\right)^{\frac{1}{2}} \cdot \left(\int_a^b |u(x)|^2 dx\right)^{\frac{1}{2}}$$

και

$$\sum_{i=1}^n |v_i u_i| \leq \left(\sum_{i=1}^n |v_i|^2\right)^{\frac{1}{2}} \cdot \left(\sum_{i=1}^n |u_i|^2\right)^{\frac{1}{2}}.$$

Δίνουμε εδώ μια απόδειξη για την πρώτη των ανισοτήτων Cauchy-Schwarz, η δεύτερη αποδεικνύεται με τον ίδιο ακριβώς τρόπο.

Απόδειξη: Θεωρούμε το ολοκλήρωμα

$$\int_a^b (|v(x)|+\theta|u(x)|)^2 dx = \int_a^b |v(x)|^2 dx + 2\theta \int_a^b |v(x)u(x)|dx + \theta^2 \int_a^b |u(x)|^2 dx,$$

όπου θ είναι μια πραγματική παράμετρος. Προφανώς το ολοκλήρωμα αυτό παίρνει μη αρνητικές τιμές αφού είναι ολοκλήρωμα μη αρνητικής συνάρτησης. Παρατηρούμε επίσης ότι το δεύτερο μέλος της παραπάνω ισότητας είναι τριώνυμο ως προς θ. Αφού είναι τριώνυμο που παίρνει μη αρνητικές τιμές, θα έχει διακρίνουσα μη θετική. Θα έχουμε επομένως

$$\Delta = 4\left(\int_a^b |v(x)u(x)|dx\right)^2 - 4\int_a^b |v(x)|^2 dx \cdot \int_a^b |u(x)|^2 dx \leq 0,$$

από την οποία προκύπτει εύκολα η ανισότητα Cauchy-Schwarz.

1.2 Ύπαρξη και μονοσήμαντο της βέλτιστης προσέγγισης

Θα δώσουμε εδώ μερικά γενικά θεωρήματα ύπαρξης και μοναδικότητας της βέλτιστης προσέγγισης, ξεκινώντας από το θεώρημα ύπαρξης.

Θεώρημα 1 *(Ύπαρξη βέλτιστης προσέγγισης) Αν V είναι ένας σταθμητός γραμμικός χώρος εφοδιασμένος με τη νόρμα $\|\cdot\|$ και W ένας υποχώρος του V πεπερασμένης διάστασης, τότε δοθέντος $v \in V$, $\exists w^* \in W$ τέτοιο ώστε*

$$\|v - w^*\| \leq \|v - w\|$$

για όλα τα $w \in W$.

Απόδειξη: Επειδή $0 \in W$, αν αυτό ήταν η βέλτιστη προσέγγιση, η ελάχιστη νόρμα θα ήταν $\|v - 0\| = \|v\| = M$. Αν w είναι τέτοιο ώστε $\|v - w\| > \|v\|$, δε θα μπορούσε να είναι βέλτιστη προσέγγιση επειδή το 0 θα έδινε καλύτερη προσέγγιση. Επομένως τη βέλτιστη προσέγγιση θα πρέπει να την αναζητούμε στο σύνολο των w για τα οποία

$$\|v - w\| \leq \|v\| = M.$$

Τότε

$$\|w\| = \|-w\| = \|v - w - v\| \leq \|v - w\| + \|v\| \leq 2M.$$

Έστω ότι ο W είναι ένας $k-$διάστατος χώρος και ότι $w_1, w_2, \cdots w_k$ είναι μια βάση αυτού. Τότε το στοιχείο $w \in W$ γράφεται ως $w = \lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_k w_k$. Ορίζουμε τη συνάρτηση

$$f(\lambda_1, \lambda_2, \cdots, \lambda_k) = \|v - w\| = \|v - (\lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_k w_k)\|$$

και θα αποδείξουμε ότι αυτή παίρνει ελάχιστη τιμή όταν το σημείο $(\lambda_1, \lambda_2, \cdots, \lambda_k)$ κινείται στον $k-$διάστατο χώρο, με τον περιορισμό

$$\|w\| = \|\lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_k w_k\| \leq 2M. \tag{1.2}$$

Είναι φανερό ότι η συνάρτηση

$$g(\lambda_1, \lambda_2, \cdots, \lambda_k) = \|w\| = \|\lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_k w_k\|$$

είναι συνεχής ως προς $(\lambda_1, \lambda_2, \cdots, \lambda_k)$ και επομένως θα παίρνει ελάχιστη τιμή στο συμπαγές σύνολο

$$|\lambda_1| + |\lambda_2| + \cdots + |\lambda_k| = 1, \tag{1.3}$$

έστω m αυτό το ελάχιστο. Θα έχουμε $m > 0$ γιατί σε αντίθετη περίπτωση δε θα πληρούται ο περιορισμός (1.3).

Επιστρέφουμε ξανά στη γενική περίπτωση όπου ισχύει ο περιορισμός (1.2) και υποθέτουμε καταρχήν ότι $\sum_{i=1}^k |\lambda_i| \neq 0$. Τότε

$$g\left(\frac{\lambda_1}{\sum_{i=1}^k |\lambda_i|}, \frac{\lambda_2}{\sum_{i=1}^k |\lambda_i|}, \cdots, \frac{\lambda_k}{\sum_{i=1}^k |\lambda_i|}\right) \geq m > 0$$

ή ισοδύναμα, αν λάβουμε υπόψη την ιδιότητα (ii) των νορμών,

$$g(\lambda_1, \lambda_2, \cdots, \lambda_k) \geq m \sum_{i=1}^k |\lambda_i|.$$

Παρατηρούμε ότι αυτή ισχύει και για $\sum_{i=1}^k |\lambda_i| = 0$, επομένως θα ισχύει για όλα τα $(\lambda_1, \lambda_2, \cdots, \lambda_k) \in I\!R^k$.
Η παραπάνω σχέση μαζί με την (1.2) δίνουν

$$2M \geq g(\lambda_1, \lambda_2, \cdots, \lambda_k) \geq m \sum_{i=1}^k |\lambda_i|$$

από την οποία προκύπτει ότι

$$\sum_{i=1}^k |\lambda_i| \leq \frac{2M}{m}.$$

Αυτή με τη σειρά της δίνει

$$|\lambda_i| \leq \frac{2M}{m}, \quad i = 1, 2, \cdots, k.$$

Η τελευταία ορίζει έναν υπερκύβο που είναι συμπαγές σύνολο και κατά συνέπεια η f που είναι συνεχής, παίρνει ελάχιστη τιμή στο σύνολο αυτό, που αντιστοιχεί στη βέλτιστη προσέγγιση. •

Δίνουμε στη συνέχεια κάποια παραδείγματα:

Παράδειγμα 1 Έστω $V = C[a, b]$, ο γραμμικός χώρος όλων των συνεχών συναρτήσεων στο διάστημα $[a, b]$ με νόρμα την ομοιόμορφη ή Chebyshev νόρμα:

$$\|f\|_\infty = \max_{a \leq x \leq b} |f(x)|.$$

Έστω επίσης W ο υποχώρος $n + 1$ διάστασης P_n, όλων των πολυωνύμων το πολύ n βαθμού ορισμένων στο διάστημα $[a, b]$. Από το Θεώρημα 1 προκύπτει ότι υπάρχει η βέλτιστη προσέγγιση $p^* \in P_n$ τέτοια ώστε

$$\|f - p^*\|_\infty = \max_{a \leq x \leq b} |f(x) - p^*(x)| \leq \|f - p\|_\infty = \max_{a \leq x \leq b} |f(x) - p(x)| \quad \forall p \in P_n$$

ή αλλιώς

$$\min_{p \in P_n} \max_{a \leq x \leq b} |f(x) - p(x)| = \max_{a \leq x \leq b} |f(x) - p^*(x)|.$$

Από την τελευταία σχέση η προσέγγιση αυτή έλαβε και το όνομα min-max προσέγγιση.

Παράδειγμα 2 Έστω το σύνολο των διακριτών σημείων $\{x_1, x_2, \cdots, x_m\}$ με $x_1 < x_2 < \cdots < x_m$ και $V = E_m$, ο γραμμικός χώρος όλων των συναρτήσεων $f = (f_1, f_2, \cdots, f_m)$ ορισμένων στα σημεία αυτά. Η ομοιόμορφη νόρμα στον E_m ορίζεται ως

$$\|f\|_\infty = \max_{i=1,2,\cdots,m} |f_i|.$$

Έστω επίσης W ο υποχώρος $n+1$ διάστασης P_n, όλων των πολυωνύμων το πολύ n βαθμού ορισμένων στα σημεία αυτά. Από το Θεώρημα 1 προκύπτει ότι υπάρχει η βέλτιστη προσέγγιση $p^* \in P_n$ τέτοια ώστε

$$\|f - p^*\|_\infty = \max_{i=1,2,\cdots,m} |f_i - p^*(x_i)| \le \|f - p\|_\infty = \max_{i=1,2,\cdots,m} |f_i - p(x_i)| \quad \forall p \in P_n$$

ή αλλιώς

$$\min_{p \in P_n} \max_{i=1,2,\cdots,m} |f_i - p(x_i)| = \max_{i=1,2,\cdots,m} |f_i - p^*(x_i)|.$$

Αν στα δυο παραπάνω παραδείγματα αντί της ομοιόμορφης νόρμας ορίσουμε την l_2 νόρμα έχουμε την προσέγγιση ελάχιστων τετραγώνων. Εξασφαλίζεται και εδώ η ύπαρξη της βέλτιστης προσέγγισης από το Θεώρημα 1. Επίσης εξασφαλίζεται η ύπαρξη της βέλτιστης προσέγγισης και στις περιπτώσεις όπου ως νόρμα λάβουμε οποιαδήποτε νόρμα, αρκεί ο υποχώρος W να είναι πεπερασμένης διάστασης.

Παράδειγμα 3 Έστω $V = C_{2\pi}[0, 2\pi]$, ο γραμμικός χώρος όλων των συνεχών περιοδικών συναρτήσεων, με περίοδο 2π στο διάστημα $[0, 2\pi]$ με νόρμα την l_2 (ή την ευκλείδεια νόρμα):

$$\|f\|_2 = \left(\int_0^{2\pi} |f(x)|^2 dx \right)^{\frac{1}{2}}.$$

Έστω επίσης W ο υποχώρος όλων των τριγωνομετρικών πολυωνύμων S_k, το πολύ k βαθμού της μορφής

$$\alpha_0 + \sum_{i=1}^{k} \alpha_i \cos(ix) + \sum_{i=1}^{k} \beta_i \sin(ix).$$

Αποδεικνύεται εύκολα ότι οι περιοδικές συναρτήσεις $\cos(ix), \quad \sin(ix), \quad i = 1, 2, \cdots, k$ καθώς και η σταθερά συνάρτηση 1, είναι γραμμικά ανεξάρτητες μεταξύ των και αποτελούν μια βάση για τον S_k. Αυτό σημαίνει ότι ο S_k είναι διάστασης $2k+1$ και επομένως από το Θεώρημα 1 προκύπτει ότι υπάρχει η βέλτιστη προσέγγιση $s^* \in S_k$ τέτοια ώστε

$$\|f - s^*\|_2 \le \|f - s\|_2 \quad \forall s \in S_k.$$

Το επόμενο παράδειγμα δείχνει ότι η απαίτηση ο υποχώρος W να είναι πεπερασμένης διάστασης, είναι υποχρεωτική.

Παράδειγμα 4 Έστω $V = C[0, \frac{1}{2}]$, εφοδιασμένος με την ομοιόμορφη νόρμα. Έστω επίσης W ο υποχώρος όλων των πολυωνύμων οποιουδήποτε βαθμού ορισμένων στο διάστημα $[0, \frac{1}{2}]$ (Ο W είναι άπειρης διάστασης γιατί δεν υπάρχει άνω φράγμα στο βαθμό).

Θα αποδείξουμε ότι η συνάρτηση $f(x) = \frac{1}{1-x}$ δεν έχει βέλτιστη προσέγγιση στο $[0, \frac{1}{2}]$. Γνωρίζουμε ότι για δοθέν $\epsilon > 0$, υπάρχει ακέραιος N τέτοιος ώστε

$$|f(x) - (1 + x + x^2 + \cdots + x^N)| < \epsilon, \quad 0 \le x \le \frac{1}{2}.$$

Έτσι, αν υπήρχε βέλτιστη προσέγγιση $p^* \in W$ θα ίσχυε

$$\|f - p^*\|_\infty = 0 \Leftrightarrow p^* = \frac{1}{1-x}$$

που δεν ανήκει στον W αφού δεν είναι πολυώνυμο, επομένως δεν υπάρχει βέλτιστη προσέγγιση.

Το παραπάνω θεώρημα μας δίνει τις συνθήκες ύπαρξης βέλτιστης προσέγγισης. Πριν προβούμε στη διατύπωση και απόδειξη των συνθηκών μοναδικότητας θα δώσουμε ένα χαρακτηρισμό του συνόλου των βέλτιστων προσεγγίσεων, αφού πρώτα θυμηθούμε τον ορισμό του κυρτού συνόλου.

Ορισμός 1 *Ένα σύνολο S λέγεται κυρτό σε ένα γραμμικό χώρο αν για $s_1, s_2 \in S$ συνεπάγεται ότι*

$$\lambda_1 s_1 + \lambda_2 s_2 \in S$$

με $\lambda_1 \ge 0, \quad \lambda_2 \ge 0$ και $\lambda_1 + \lambda_2 = 1$.

Θεώρημα 2 *Αν $v \in V$ και W είναι ένας υποχώρος του V, τότε το σύνολο W^* όλων των βέλτιστων προσεγγίσεων του v στον W είναι κυρτό σύνολο.*

Απόδειξη: Αν το W^* είναι το κενό σύνολο τότε είναι από μόνο του κυρτό. Έστω w_1^* και w_2^* είναι δυο βέλτιστες προσεγγίσεις του v στον W, τότε

$$\|v - w_1^*\| = \|v - w_2^*\| = \rho.$$

Θεωρούμε $\lambda_1 \ge 0, \quad \lambda_2 \ge 0$, $\lambda_1 + \lambda_2 = 1$. Τότε

$$\|v-(\lambda_1 w_1^*+\lambda_2 w_2^*)\| = \|\lambda_1(v-w_1^*)+\lambda_2(v-w_2^*)\| \le \lambda_1\|v-w_1^*\|+\lambda_2\|v-w_2^*\| = (\lambda_1+\lambda_2)\rho = \rho.$$

Επομένως $\lambda_1 w_1^* + \lambda_2 w_2^* \in W^*$ και έτσι το W^* είναι κυρτό σύνολο. •

Ορισμός 2 *Ένας σταθμητός γραμμικός χώρος V λέμε ότι έχει αυστηρά κυρτή νόρμα, όταν το σύνολο $B = \{v : \|v\| \le 1\}$ (μοναδιαία μπάλα) είναι αυστηρά κυρτό.*

Υπενθυμίζουμε ότι για να είναι αυστηρά κυρτό ένα σύνολο απαιτείται η ισχύς της πρότασης: Αν $v_1 \ne v_2$ με $\|v_1\| = \|v_2\| = 1$, τότε $\|\lambda_1 v_1 + \lambda_2 v_2\| < 1$, όταν $\lambda_1 > 0, \quad \lambda_2 > 0$ και $\lambda_1 + \lambda_2 = 1$.

Θεώρημα 3 *Αν V είναι ένας γραμμικός χώρος με μια αυστηρά κυρτή νόρμα και W είναι ένας υποχώρος του V, τότε το στοιχείο $v \in V$ έχει το πολύ μια βέλτιστη προσέγγιση στον W.*

Απόδειξη: Υποθέτουμε ότι w_1^* και w_2^* είναι δυο διαφορετικές βέλτιστες προσεγγίσεις του v στον W. Από το Θεώρημα 2 θα έχουμε ότι και $\frac{w_1^*+w_2^*}{2}$ θα είναι επίσης βέλτιστη προσέγγιση του v στον W. Υποθέτουμε ότι

$$\|v - w_1^*\| = \|v - w_2^*\| = \rho$$

και θέτουμε με $v_1 = \frac{v-w_1^*}{\rho}$ και $v_2 = \frac{v-w_2^*}{\rho}$. Τότε προφανώς $v_1 \neq v_2$, $\|v_1\| = \|v_2\| = 1$ και επειδή η νόρμα είναι αυστηρά κυρτή,

$$\left\|\frac{1}{2}v_1 + \frac{1}{2}v_2\right\| = \left\|\frac{1}{2\rho}(v - w_1^*) + \frac{1}{2\rho}(v - w_2^*)\right\| = \frac{1}{\rho}\left\|v - \frac{w_1^* + w_2^*}{2}\right\| < 1$$

ή $\left\|v - \frac{w_1^*+w_2^*}{2}\right\| < \rho$ το οποίο είναι άτοπο. •

Ο συνδυασμός του Θεωρήματος 1 με το Θεώρημα 3 δίνει ως αποτέλεσμα το παρακάτω πόρισμα.

Πόρισμα 1 *Αν V είναι ένας γραμμικός χώρος με μια αυστηρά κυρτή νόρμα και W είναι ένας υποχώρος του V πεπερασμένης διάστασης, τότε κάθε στοιχείο $v \in V$ έχει ακριβώς μια βέλτιστη προσέγγιση στον W.*

Αποδεικνύεται ότι οι l_p νόρμες, για $p > 1$, είναι όλες αυστηρά κυρτές νόρμες όταν $V = C[a, b]$ ή στη διακριτή περίπτωση όπου $V = E_m$. Επομένως για όλες αυτές τις νόρμες εξασφαλίζεται η μοναδικότητα της βέλτιστης πολυωνυμικής προσέγγισης. Ειδικότερα για την l_2 ή ευκλείδεια νόρμα εξασφαλίζεται η μοναδικότητα της βέλτιστης πολυωνυμικής προσέγγισης "Ελαχίστων Τετραγώνων". Αντίθετα, για τις νόρμες l_1 και l_∞ έχουμε ότι δεν είναι αυστηρά κυρτές νόρμες. Επομένως δεν εξασφαλίζεται από το γενικό αυτό Θεώρημα η μοναδικότητα της βέλτιστης προσέγγισης "Πρώτης Δύναμης" ούτε η μοναδικότητα της βέλτιστης "Ομοιόμορφης" προσέγγισης. Παρόλα αυτά, ειδική μελέτη κάθε μιας από τις παραπάνω περιπτώσεις, αποδεικνύει τη μοναδικότητα της βέλτιστης προσέγγισης και στις περιπτώσεις αυτές. Αυτό θα το δούμε στο επόμενο κεφάλαιο για την περίπτωση της ομοιόμορφης προσέγγισης. Δε θα αποδείξουμε εδώ ότι είναι αυστηρά κυρτές οι l_p νόρμες για $p > 1$ αλλά θα δώσουμε παραδείγματα από τα οποία θα φαίνεται ότι οι l_1 και l_∞ δεν είναι αυστηρά κυρτές νόρμες.

Παράδειγμα 5 Έστω $V = C[0, 1]$, εφοδιασμένος με την l_1 νόρμα. Για τις συναρτήσεις $v_1(x) = 2x$ και $v_2(x) = 2(1 - x)$, έχουμε ότι $v_1 \neq v_2$ με $\|v_1\| = \int_0^1 2x dx = \|v_2\| = \int_0^1 2(1 - x)dx = 1$. Αν θεωρήσουμε $\lambda_1 = \lambda_2 = \frac{1}{2}$ παίρνουμε ότι $\|v\| = \|\frac{v_1+v_2}{2}\| = \|\frac{2x+2(1-x)}{2}\| = \|1\| = 1$. Αυτό σημαίνει ότι η l_1 νόρμα δεν είναι αυστηρά κυρτή νόρμα.

Παράδειγμα 6 Έστω $V = C[0,1]$, εφοδιασμένος με την l_∞ νόρμα. Για τις συναρτήσεις $v_1(x) = x$ και $v_2(x) = 2x-1$, έχουμε ότι $v_1 \neq v_2$ με $\|v_1\| = \max_{x\in[0,1]}\{x\} = \|v_2\| = \max_{x\in[0,1]}\{|2x-1|\} = 1$. Αν θεωρήσουμε $\lambda_1 = \lambda_2 = \frac{1}{2}$ παίρνουμε ότι $\|v\| = \|\frac{v_1+v_2}{2}\| = \|\frac{x+2x-1)}{2}\| = \max_{x\in[0,1]} |\frac{3x-1}{2}| = 1$. Αυτό σημαίνει ότι η l_∞ νόρμα δεν είναι αυστηρά κυρτή νόρμα.

Παράδειγμα 7 Έστω $V = E_m$, $m > 1$, εφοδιασμένος με την l_∞ νόρμα. Για τις συναρτήσεις $v_1 = (1, 1, \cdots, 1)$ και $v_2 = (1, 1-\frac{1}{m}, 1-\frac{2}{m}, \cdots, \frac{1}{m})$, έχουμε ότι $v_1 \neq v_2$ με $\|v_1\| = \|v_2\| = 1$. Αν θεωρήσουμε $\lambda_1 = \lambda_2 = \frac{1}{2}$ παίρνουμε ότι $\|v\| = \|\frac{v_1+v_2}{2}\| = \max(1, 1-\frac{1}{2m}, 1-\frac{2}{2m}, \cdots, \frac{1}{2}+\frac{1}{2m}) = 1$. Αυτό σημαίνει ότι η l_∞ νόρμα δεν είναι αυστηρά κυρτή νόρμα.

Ολοκληρώσαμε εδώ τη γενική θεωρία ύπαρξης και μοναδικότητας της βέλτιστης προσέγγισης.

Ασκήσεις

1. Να αποδειχτεί ότι η έκφραση $\int_a^b |xf(x)|dx, \quad f \in C[a,b]$ ορίζει μια νόρμα στο χώρο $C[a,b]$

2. Δίνεται το σύνολο των διακριτών σημείων $\{x_1, x_2, \cdots, x_m\}$ και E_m, ο γραμμικός χώρος όλων των συναρτήσεων $f = (f_1, f_2, \cdots, f_m)$ ορισμένων στα σημεία αυτά. Να αποδειχτεί ότι η έκφραση $\sum_{i=1}^m i|f_i|, \quad f \in E_m$ ορίζει μια νόρμα στο χώρο E_m, ενώ η έκφραση $\sum_{i=1}^m (i-1)|f_i|, \quad f \in E_m$ δεν ορίζει νόρμα στο χώρο E_m.

3. Να αποδειχτούν οι ανισότητες
$$\|f\|_2^2 \leq \|f\|_1 \cdot \|f\|_\infty \quad \text{και} \quad \|f\|_1 \leq \sqrt{2} \cdot \|f\|_2, \quad \forall f \in C[-1,1].$$

4. Δίνεται το σύνολο των διακριτών σημείων $\{x_1, x_2, \cdots, x_m\}$ και E_m, ο γραμμικός χώρος όλων των συναρτήσεων $f = (f_1, f_2, \cdots, f_m)$ ορισμένων στα σημεία αυτά. Να αποδειχτούν οι ανισότητες
$$\|f\|_2^2 \leq \|f\|_1 \cdot \|f\|_\infty \quad \text{και} \quad \|f\|_1 \leq \sqrt{m} \cdot \|f\|_2, \quad \forall f \in E_m.$$

5. Να βρεθεί το σύνολο όλων των βέλτιστων προσεγγίσεων του σημείου $(1 \;\; 1 \;\; 1)^T \in I\!R^3$ στο χώρο $I\!R^2 = (x \;\; y \;\; 0)^T, x, y \in I\!R$, θεωρώντας ως νόρμα την $\|\cdot\|_\infty$.

Κεφάλαιο 2

ΟΜΟΙΟΜΟΡΦΗ ΠΡΟΣΕΓΓΙΣΗ

Έστω ότι P_n είναι το σύνολο των πολυωνύμων το πολύ n βαθμού. Το Θεώρημα 1 μας δίνει ότι για δοθείσα συνάρτηση $f \in C[a,b]$ υπάρχει πολυώνυμο $p_n^* \in P_n$ τέτοιο ώστε

$$\|f - p_n^*\| \leq \|f - p\| \quad \forall p \in P_n,$$

όπου $\|\cdot\|$ είναι η ομοιόμορφη νόρμα (l_∞ ή νόρμα άπειρο). Θα συμβολίζουμε με $E_n(f;[a,b]) = \|f - p_n^*\|$ το σφάλμα της ομοιόμορφης βέλτιστης προσέγγισης. Όταν θα είναι συγκεκριμένο το διάστημα $[a,b]$ και δε θα υπάρχει λόγος σύγχυσης, για συντομία θα το συμβολίζουμε και με $E_n(f)$.

Για να έχει έννοια η πολυωνυμική προσέγγιση θα πρέπει, όσο αυξάνει ο βαθμός του πολυωνύμου, τόσο να μικραίνει το σφάλμα. Θα αποδείξουμε εδώ ότι $E_n(f) \to 0$ για $n \to \infty$, που είναι το γνωστό Θεώρημα της πολυωνυμικής προσέγγισης του Weierstrass.

Θεώρημα 4 *(Weierstrass) Δοθέντων $f \in C[a,b]$ και $\epsilon > 0$, υπάρχει πολυώνυμο p τέτοιο ώστε*

$$\|f - p\| < \epsilon.$$

Απόδειξη: Θα κατασκευάσουμε ένα τέτοιο πολυώνυμο, χρησιμοποιώντας τα πολυώνυμα Bernstein. Μετασχηματίζουμε πρώτα το διάστημα $[a,b]$ στο $[0,1]$. Ο μετασχηματισμός είναι γραμμικός και $1-1$. Θέτοντας $x = (b-a)t+a$ έχουμε $f(x) = f((b-a)t+a) = g(t)$, όπου η $g(t)$ ορίζεται στο $[0,1]$ και είναι συνεχής. Μπορούμε επομένως να αποδείξουμε το θεώρημα για συναρτήσεις ορισμένες στο $[0,1]$.

Δοθείσας συνάρτησης $h(t)$ φραγμένης στο $[0,1]$ ορίζουμε το πολυώνυμο Bernstein m βαθμού $B_m(h;t)$ ως εξής:

$$B_m(h;t) = \sum_{k=0}^{m} h\left(\frac{k}{m}\right) \binom{m}{k} t^k (1-t)^{m-k}. \tag{2.1}$$

Είναι φανερό ότι
i) $B_m(h;t) \in P_m$
ii) $B_m(\alpha h;t) = \alpha B_m(h;t)$

iii) $B_m(h_1 + h_2; t) = B_m(h_1; t) + B_m(h_2; t)$
iv) Αν $h_1(t) \leq h_2(t)$ τότε $B_m(h_1; t) \leq B_m(h_2; t)$
Από τις (ii) και (iii) προκύπτει ότι

$$B_m(\alpha h_1 + \beta h_2; t) = \alpha B_m(h_1; t) + \beta B_m(h_2; t)$$

που σημαίνει ότι ο τελεστής B_m είναι ένας γραμμικός τελεστής. Η ιδιότητα (iv) προκύπτει εύκολα αφού, για δεδομένο $t \in [0, 1]$ εκείνο που αλλάζει στον τύπο (2.1) είναι οι παράγοντες $h_1(t)$ και $h_2(t)$, $h_1(t) \leq h_2(t)$, ενώ ο σταθερός παράγοντας $\begin{pmatrix} m \\ k \end{pmatrix} t^k(1-t)^{m-k}$ είναι μη αρνητικός. Η ιδιότητα αυτή καθιστά τον τελεστή B_m ένα θετικό τελεστή αφού για $h = h_2 - h_1 \geq 0$ έχουμε ότι

$$B_m(h; t) \geq 0.$$

Έχουμε να αποδείξουμε ότι για $\epsilon > 0$ και για $h \in C[0, 1]$, υπάρχει ακέραιος m_0 τέτοιος ώστε

$$\|h - B_{m_0}(h)\| < \epsilon.$$

Θα μελετήσουμε τώρα τη συμπεριφορά του τελεστή B_m, όταν εφαρμοστεί στις συναρτήσεις μονώνυμα $1, t$ και t^2.
Για $h(t) = 1$,

$$B_m(1; t) = \sum_{k=0}^{m} \begin{pmatrix} m \\ k \end{pmatrix} t^k(1-t)^{m-k} = (t + 1 - t)^m = 1.$$

Για $h(t) = t$, έχουμε

$$\begin{aligned} B_m(t; t) &= \textstyle\sum_{k=0}^{m} \frac{k}{m} \begin{pmatrix} m \\ k \end{pmatrix} t^k(1-t)^{m-k} = \sum_{k=1}^{m} \begin{pmatrix} m-1 \\ k-1 \end{pmatrix} t^k(1-t)^{m-k} \\ &= t \textstyle\sum_{j=0}^{m-1} \begin{pmatrix} m-1 \\ j \end{pmatrix} t^j(1-t)^{m-1-j} = t(t+1-t)^{m-1} = t. \end{aligned}$$

Τέλος, για $h(t) = t^2$, έχουμε

$$B_m(t^2; t) = B_m\left(t\left(t - \frac{1}{m}\right); t\right) + \frac{1}{m} B_m(t; t) = B_m\left(t\left(t - \frac{1}{m}\right); t\right) + \frac{1}{m} t.$$

Υπολογίζουμε τώρα το $B_m\left(t\left(t - \frac{1}{m}\right); t\right)$.

$$B_m\left(t\left(t - \frac{1}{m}\right); t\right) = \sum_{k=0}^{m} \frac{k}{m} \frac{k-1}{m} \begin{pmatrix} m \\ k \end{pmatrix} t^k(1-t)^{m-k}$$

αλλά

$$\begin{aligned} \frac{k}{m} \frac{k-1}{m} \begin{pmatrix} m \\ k \end{pmatrix} &= \frac{k(k-1)m!}{m^2 k!(m-k)!} = \frac{(m-1)!}{m(k-2)!(m-k)!} \\ &= \frac{(m-1)}{m} \frac{(m-2)!}{(k-2)!(m-k)!} = \left(1 - \frac{1}{m}\right) \begin{pmatrix} m-2 \\ k-2 \end{pmatrix} \end{aligned}$$

οπότε

$$B_m\left(t\left(t-\frac{1}{m}\right);t\right)=\left(1-\frac{1}{m}\right)t^2\sum_{j=0}^{m-2}\begin{pmatrix} m-2 \\ j \end{pmatrix}t^j(1-t)^{m-2-j}=\left(1-\frac{1}{m}\right)t^2.$$

Τελικά έχουμε ότι

$$B_m(t^2;t)=\left(1-\frac{1}{m}\right)t^2+\frac{1}{m}t=t^2+\frac{1}{m}t(1-t).$$

Αυτές οι τρεις συναρτήσεις ελέγχου είναι αρκετές για να αποδείξουμε το ζητούμενο για κάθε συνάρτηση h.
Υποθέτουμε ότι $\|h\|=M$ και ότι $s,t\in[0,1]$, τότε προκύπτει εύκολα ότι

$$-2M\leq h(t)-h(s)\leq 2M. \tag{2.2}$$

Από την άλλη πλευρά, επειδή η h είναι ομοιόμορφα συνεχής στο $[0,1]$ έχουμε ότι, για κάθε $\epsilon_1>0$ υπάρχει $\delta(\epsilon_1)>0$ τέτοιο ώστε

$$-\epsilon_1\leq h(t)-h(s)\leq \epsilon_1,\ \text{όταν}\ |t-s|<\delta. \tag{2.3}$$

Από τις (2.2) και (2.3) προκύπτει ότι ισχύει η

$$-\epsilon_1-\frac{2M}{\delta^2}(t-s)^2\leq h(t)-h(s)\leq \epsilon_1+\frac{2M}{\delta^2}(t-s)^2 \tag{2.4}$$

για όλα τα $s,t\in[0,1]$. Πραγματικά αν $|t-s|<\delta$ τότε η (2.4) ισχύει αφού τα φράγματά της είναι εκείνα της (2.3) διευρυμένα κατά τον όρο $\frac{2M}{\delta^2}(t-s)^2$, ενώ αν $|t-s|\geq\delta$ τότε η (2.4) ισχύει αφού τα φράγματα της είναι εκείνα της (2.2) διευρυμένα κατά παράγοντα $\frac{(t-s)^2}{\delta^2}\geq 1$ και κατά τον όρο ϵ_1.
Στη συνέχεια, χρησιμοποιώντας τη γραμμικότητα και τη θετικότητα του τελεστή B_m, τον εφαρμόζουμε στις ανισότητες (2.4) θεωρώντας τις συναρτήσεις ως προς τη μεταβλητή t ενώ το s το παίρνουμε σταθερό. Έτσι έχουμε

$$B_m\left(-\epsilon_1-\frac{2M}{\delta^2}(t-s)^2;t\right)\leq B_m\left(h(t)-h(s);t\right)\leq B_m\left(\epsilon_1+\frac{2M}{\delta^2}(t-s)^2;t\right)$$

ή ισοδύναμα

$$-\epsilon_1-\frac{2M}{\delta^2}B_m((t-s)^2;t)\leq B_m(h(t);t)-h(s)\leq \epsilon_1+\frac{2M}{\delta^2}B_m((t-s)^2;t). \tag{2.5}$$

Αλλά

$$\begin{aligned}B_m((t-s)^2;t)&=B_m(t^2;t)-2sB_m(t;t)+s^2B_m(1;t)\\&=t^2+\frac{1}{m}t(1-t)-2st+s^2=(t-s)^2+\frac{1}{m}t(1-t).\end{aligned}$$

Λαμβάνοντας τώρα το όριο της παραπάνω σχέσης για t τείνοντος στο s έχουμε

$$\lim_{t \to s} B_m((t-s)^2; t) = \frac{1}{m} s(1-s) \leq \frac{1}{4m},$$

αφού $s \in [0,1]$. Θεωρώντας επίσης τα όρια σε όλα τα μέλη της (2.5) για t τείνοντος στο s, παίρνουμε

$$-\epsilon_1 - \frac{M}{2\delta^2 m} \leq B_m(h; s) - h(s) \leq \epsilon_1 + \frac{M}{2\delta^2 m}$$

ή

$$|B_m(h; s) - h(s)| \leq \epsilon_1 + \frac{M}{2\delta^2 m}. \tag{2.6}$$

Τώρα αν επιλέξουμε $\epsilon_1 = \frac{\epsilon}{2}$, τότε για κάθε $m_0 > \frac{M}{\delta^2 \epsilon}$ έχουμε το ζητούμενο

$$|B_m(h; s) - h(s)| < \epsilon, \qquad 0 \leq s \leq 1. \tag{2.7}$$

•

Αποδείξαμε εδώ το Θεώρημα του Weierstrass θεωρώντας τα πολυώνυμα Bernstein. Θα μπορούσε να γίνει η απόδειξη χρησιμοποιώντας κάποια άλλη ακολουθία πολυωνύμων αρκεί να πληρούνται κάποιες προϋποθέσεις. Ο Korovkin έδωσε μια γενικότερη απόδειξη της πολυωνυμικής προσέγγισης. Δίνουμε εδώ το Θεώρημα Korovkin χωρίς απόδειξη αφού αυτή βασίζεται στην ίδια τεχνική.

Θεώρημα 5 *(Korovkin) Έστω $C[a,b]$ το σύνολο των συνεχών συναρτήσεων στο διάστημα $[a,b]$. Θεωρούμε το σύνολο των συναρτήσεων $T = \{1, x, x^2\}$, που καλείται σύνολο Korovkin. Έστω $\{\Phi_n\}$ μια ακολουθία γραμμικών και θετικών τελεστών στο χώρο $C[a,b]$. Αν για κάθε $g \in T$, η $\Phi_n(g)$ συγκλίνει ομοιόμορφα στην g, τότε η $\{\Phi_n\}$ αποτελεί μια διαδικασία προσέγγισης, δηλαδή*

$$\Phi_n(f) \;\; \textit{συγκλίνει ομοιόμορφα στην} \;\; f \;\; \forall f \in C[a,b].$$

Το Θεώρημα Korovkin γενικεύεται και για συναρτήσεις πολλών μεταβλητών. Το δίνουμε εδώ στην πιο γενική του μορφή.

Θεώρημα 6 *(Korovkin) Έστω K ένα συμπαγές σύνολο του $\mathbb{R}^d$ και έστω $C(K)$ ο χώρος Banach των συνεχών συναρτήσεων (πραγματικών ή μιγαδικών) στο K, εφοδιασμένος με τη νόρμα άπειρο. Θεωρούμε το σύνολο των συναρτήσεων $T = \{1, x_i, \|x\|_2^2, \;\; i = 1, \ldots, d\}$, που καλείται σύνολο Korovkin. Έστω $\{\Phi_n\}$ μια ακολουθία γραμμικών και θετικών τελεστών στο χώρο $C(K)$. Αν για κάθε $g \in T$, η $\Phi_n(g)$ συγκλίνει ομοιόμορφα στην g, τότε η $\{\Phi_n\}$ αποτελεί μια διαδικασία προσέγγισης, δηλαδή*

$$\Phi_n(f) \;\; \textit{συγκλίνει ομοιόμορφα στην} \;\; f \;\; \forall f \in C(K).$$

Παρατηρούμε εδώ ότι όταν το συμπαγές σύνολο είναι διάστασης d, αρκούν $d+2$ συναρτήσεις στο σύνολο Korovkin T, εκ των οποίων η μια είναι δεύτερου βαθμού, που είναι απαραίτητη για να μας δώσει μια ανισότητα αντίστοιχη της (2.2) στην περίπτωση των τελεστών Bernstein. Ακόμη πρέπει να τονίσουμε ότι το Θεώρημα Korovkin μπορεί να επεκταθεί και για την περίπτωση προσέγγισης με άλλου είδους συναρτήσεις πέραν των πολυωνυμικών, αρκεί να επιλεγεί κατάλληλα το σύνολο ελέγχου T. Ένα τέτοιο παράδειγμα αποτελεί η προσέγγιση των περιοδικών συναρτήσεων με τριγωνομετρικά πολυώνυμα.

Τα παραπάνω θεωρήματα δίνουν απάντηση στο πρόβλημα της ύπαρξης της πολυωνυμικής προσέγγισης. Δε μιλήσαμε καθόλου για το πόσο καλή είναι αυτή η προσέγγιση, δηλαδή για το σφάλμα αποκοπής κατά την πολυωνυμική προσέγγιση με συγκεκριμένο το βαθμό του πολυωνύμου. Δε θα επιμείνουμε στο θέμα αυτό γιατί χρειάζεται αρκετή θεωρία, απλώς θα αναφέρουμε ότι η σύγκλιση της πολυωνυμικής προσέγγισης εξαρτάται από τη συνάρτηση f και πιο συγκεκριμένα από το είδος της συνέχειας της f. Ορίζεται εδώ το *Μέτρο Συνέχειας* $\omega(\delta), \quad \delta > 0$ μιας συνάρτησης f ως

$$\omega(\delta) = \sup_{x_1,x_2\in[a,b],|x_1-x_2|\leq\delta} |f(x_1) - f(x_2)|. \tag{2.8}$$

Όσο ταχύτερα συγκλίνει η συνάρτηση ω στο μηδέν του δ τείνοντος στο μηδέν, τόσο μικρότερο είναι το σφάλμα της πολυωνυμικής προσέγγισης.

Για τη συνάρτηση $f(x) = x$, για παράδειγμα, έχουμε

$$\omega(\delta) = \sup_{x_1,x_2\in[a,b],|x_1-x_2|\leq\delta} |x_1 - x_2| = \delta,$$

ενώ για τη συνάρτηση $f(x) = x^2$,

$$\omega(\delta) = \sup_{x_1,x_2\in[a,b],|x_1-x_2|\leq\delta} |x_1^2 - x_2^2| = \sup_{x_1,x_2\in[a,b],|x_1-x_2|\leq\delta} |x_1 + x_2||x_1 - x_2| \leq 2b\delta.$$

Πρέπει να τονίσουμε επίσης ότι τα πολυώνυμα Bernstein καθώς και τα πολυώνυμα που προκύπτουν από ακολουθίες τελεστών Korovkin, αποτελούν απλώς μια πολυωνυμική προσέγγιση και όχι τη βέλτιστη ομοιόμορφη πολυωνυμική προσέγγιση. Στο εξής θα ασχοληθούμε με τη βέλτιστη ομοιόμορφη πολυωνυμική προσέγγιση. Θα αποδείξουμε κάποιες ιδιότητες που τη χαρακτηρίζουν, ώστε να μπορέσουμε να καταλήξουμε σε αλγόριθμο εύρεσης αυτής.

2.1 Χαρακτηρισμός της βέλτιστης ομοιόμορφης προσέγγισης

Έστω p_n^* η βέλτιστη ομοιόμορφη προσέγγιση της $f \in C[a,b]$, στο χώρο P_n των πολυωνύμων το πολύ n βαθμού, τότε η συνάρτηση σφάλμα δίνεται ως

$$e(x) = f(x) - p_n^*(x),$$

με $\|e(x)\| = E_n(f;[a,b])$. Στο Θεώρημα που ακολουθεί θα δώσουμε μια πρώτη ιδιότητα της βέλτιστης προσέγγισης.

Θεώρημα 7 *Υπάρχουν τουλάχιστο δυο διαφορετικά σημεία $x_1, x_2 \in [a,b]$ τέτοια ώστε*

$$|e(x_1)| = |e(x_2)| = E_n(f;[a,b])$$

και $e(x_1) = -e(x_2)$.

Απόδειξη: Η καμπύλη $y = e(x)$ θα βρίσκεται μεταξύ των ευθειών $y = \pm E_n(f)$ και θα εφάπτεται τουλάχιστον σε μια από αυτές. Χωρίς βλάβη της γενικότητας υποθέτουμε ότι εφάπτεται μόνο στην $y = E_n(f)$ και ότι

$$\min_{a \le x \le b} e(x) = m > -E_n(f).$$

Θεωρούμε $c = \frac{E_n(f)+m}{2} > 0$. Τότε για το πολυώνυμο $q_n = p_n^* + c \in P_n$ ισχύει ότι $f(x) - q_n(x) = e(x) - c$ και επομένως

$$-(E_n(f) - c) = m - c \le e(x) - c \le E_n(f) - c$$

που σημαίνει ότι

$$\|f - q_n\| \le E_n(f) - c.$$

Αυτό όμως είναι άτοπο αφού υποθέσαμε ότι p_n^* είναι η βέλτιστη προσέγγιση.

Πόρισμα 2 *Η βέλτιστη προσεγγιστική σταθερά (πολυώνυμο μηδενικού βαθμού) είναι*

$$p_0^* = \frac{1}{2}(\max_{a \le x \le b} f(x) + \min_{a \le x \le b} f(x))$$

με

$$E_0(f) = \frac{1}{2}(\max_{a \le x \le b} f(x) - \min_{a \le x \le b} f(x)).$$

Θα δώσουμε τώρα έναν ορισμό για να μπορέσουμε στη συνέχεια να διατυπώσουμε το βασικό θεώρημα που χαρακτηρίζει τη βέλτιστη προσέγγιση.

Ορισμός 3 *Ένα σύνολο σημείων $x_0, x_1, \ldots, x_k$ με $a \le x_0 < x_1 < \cdots < x_k \le b$ λέγεται εναλλασσόμενο σύνολο σημείων για τη συνάρτηση σφάλμα $e = f - p_n^*$, αν*

$$|e(x_j)| = |f(x_j) - p_n^*(x_j)| = \|f - p_n^*\|, \quad j = 0, 1, \ldots, k$$

και

$$f(x_j) - p_n^*(x_j) = -(f(x_{j+1}) - p_n^*(x_{j+1})), \quad j = 0, 1, \ldots, k-1.$$

Θεώρημα 8 *Έστω $f \in C[a,b]$. Το πολυώνυμο p_n^* είναι η βέλτιστη προσέγγιση στο διάστημα $[a,b]$ της f, στο σύνολο πολυωνύμων P_n αν και μόνο αν (ανν) υπάρχει εναλλασσόμενο σύνολο σημείων για τη συνάρτηση σφάλμα $e = f - p_n^*$, αποτελούμενο από $n+2$ σημεία.*

Απόδειξη: ($\Leftarrow$) Έστω ότι υπάρχει εναλλασσόμενο σύνολο σημείων $x_0, x_1, \dots, x_{n+1}$ για την $e = f - p_n^*$. Έστω επίσης ότι υπάρχει $q_n \in P_n$ τέτοιο ώστε

$$\|f - q_n\| < \|f - p_n^*\|.$$

Τότε

$$|f(x_j) - q_n(x_j)| < \|f - p_n^*\| = |f(x_j) - p_n^*(x_j)|, \quad j = 0, 1, \dots, n+1$$

και από τον ορισμό του εναλλασσομένου συνόλου προκύπτει ότι: Αν

$$f(x_j) - p_n^*(x_j) = \|f - p_n^*\| > 0 \Rightarrow f(x_j) - p_n^*(x_j) > f(x_j) - q_n(x_j)$$

ή

$$f(x_j) - p_n^*(x_j) - (f(x_j) - q_n(x_j)) = q_n(x_j) - p_n^*(x_j) > 0,$$

ενώ αν

$$f(x_j) - p_n^*(x_j) = -\|f - p_n^*\| < 0 \Rightarrow f(x_j) - p_n^*(x_j) < f(x_j) - q_n(x_j)$$

ή

$$f(x_j) - p_n^*(x_j) - (f(x_j) - q_n(x_j)) = q_n(x_j) - p_n^*(x_j) < 0.$$

Αυτό σημαίνει ότι το πολυώνυμο $q_n - p_n^*$ εναλλάσσει συνεχώς πρόσημο καθώς το j μεταβάλλεται από το 0 στο $n+1$. Δηλαδή υπάρχει ρίζα σε κάθε διάστημα (x_j, x_{j+1}), $j = 0, 1, \dots, n$. Έτσι το πολυώνυμο $q_n - p_n^* \in P_n$ έχει $n+1$ ρίζες, το οποίο είναι άτοπο.

($\Rightarrow$) Έστω ότι p_n^* είναι η βέλτιστη προσέγγιση της f, όπου η f δεν είναι πολυώνυμο n βαθμού. Έστω επίσης ότι το μέγιστο εναλλασσόμενο σύνολο σημείων αποτελείται από $m+1$ σημεία $a \le x_0 < x_1 < \cdots < x_m \le b$. Από το Θεώρημα 7 προκύπτει ότι $m \ge 1$. Θα αποδείξουμε ότι $m \ge n+1$.

Έστω ότι $m \le n$ και ότι

$$\|f - p_n^*\| = \rho.$$

Θεωρούμε τη διαμέριση $t_0, t_1, t_2, \dots, t_s \in [a,b]$ τέτοια ώστε $a = t_0 < t_1 < t_2 < \cdots < t_s = b$ και η συνάρτηση $e = f - p_n^*$ να πληροί την

$$|e(\xi) - e(\eta)| < \frac{1}{2}\rho \quad \forall \xi, \eta \in [t_j, t_{j+1}], \quad j = 0, 1, \dots, s-1.$$

Κάθε υποδιάστημα που περιέχει ακραίο θετικό σημείο t για τη συνάρτηση e ($e(t) = \rho$), το καλούμε (+) υποδιάστημα ενώ κάθε υποδιάστημα που περιέχει ακραίο αρνητικό σημείο t για τη συνάρτηση e ($e(t) = -\rho$), το καλούμε (–) υποδιάστημα. Γράφουμε από αριστερά προς τα δεξιά μόνο τα $(\pm)$ υποδιαστήματα $I_1, I_2, \ldots, I_N$ και έστω ότι είναι $(+)$ το υποδιάστημα I_1. Χωρίζουμε στη συνέχεια το σύνολο $I_1, I_2, \ldots, I_N$ σε υποσύνολα που θα περιέχουν τα διαδοχικά (+) ή (–) υποδιαστήματα. Έστω ότι αυτά είναι:

$$\begin{array}{rcl} \{I_1, I_2, \ldots, I_{k_1}\} & (+) & \text{υποδιαστήματα} \\ \{I_{k_1+1}, I_{k_1+2}, \ldots, I_{k_2}\} & (-) & \text{υποδιαστήματα} \\ \vdots & \vdots & \vdots \\ \{I_{k_m+1}, I_{k_m+2}, \ldots, I_{k_{m+1}}\} & (-)^m & \text{υποδιαστήματα} \end{array}$$

Κάθε υποσύνολο περιέχει τουλάχιστον ένα υποδιάστημα, ενώ $2 \leq m+1 \leq n+1$ από τις υποθέσεις που κάναμε. Είναι φανερό ότι τα υποδιαστήματα I_{k_j} και I_{k_j+1} είναι ξένα μεταξύ των αφού υπάρχει ολόκληρο διάστημα που τα χωρίζει. Αυτό συμβαίνει διότι $e(x) > \frac{1}{2}\rho, x \in I_{k_j}$ και $e(x) < -\frac{1}{2}\rho, x \in I_{k_j+1}$ αν I_{k_j} και I_{k_j+1} είναι (+) και (–) υποδιαστήματα, αντίστοιχα ή αντίστροφα αν αυτά είναι (–) και (+) υποδιαστήματα, αντίστοιχα. Θεωρούμε τώρα m σημεία $z_1, z_2, \ldots, z_m$ τέτοια ώστε

$$z_j > x \quad \forall x \in I_{k_j} \quad \text{και} \quad z_j < x \quad \forall x \in I_{k_j+1}.$$

Ορίζουμε στη συνέχεια το πολυώνυμο $q \in P_m$:

$$q(x) = (z_1 - x)(z_2 - x)\cdots(z_m - x).$$

Παρατηρούμε ότι το $q(x)$ έχει το ίδιο πρόσημο με τη συνάρτηση $e(x)$ σε όλα τα υποδιαστήματα $I_1, I_2, \ldots, I_N$. Πραγματικά, αν το x ανήκει στο πρώτο σύνολο υποδιαστημάτων $\{I_1, I_2, \ldots, I_{k_1}\}$, που είναι (+) υποδιαστήματα, το $q(x)$ είναι θετικό αφού $z_i > x, \quad i = 1, 2, \ldots, m$, ενώ κάθε φορά που το x μεταβαίνει σε επόμενο υποσύνολο το πολυώνυμο q αλλάζει πρόσημο αφού κάθε φορά και ένα από τα z_i γίνεται μικρότερο του x. Ονομάζουμε με R την ένωση των υποδιαστημάτων που δεν είναι $(\pm)$ υποδιαστήματα. Τότε,

$$\max_{x \in R} |e(x)| = \rho' < \rho.$$

Υποθέτουμε ότι $\|q\| = M$ και επιλέγουμε ένα $\lambda > 0$ τόσο μικρό ώστε

$$\lambda M < \min\{\rho - \rho', \frac{\rho}{2}\}.$$

Θα αποδείξουμε τώρα ότι το πολυώνυμο $p(x) = \lambda q(x) + p_n^*(x) \in P_n$ είναι καλύτερη προσέγγιση από το $p_n^*(x)$.
Έστω ότι $x \in R$, τότε

$$|f(x) - p(x)| = |f(x) - \lambda q(x) - p_n^*(x)| = |e(x) - \lambda q(x)| \leq |e(x)| + \lambda |q(x)| \leq \rho' + \lambda M < \rho.$$

Αν το x ανήκει σε ένα από τα $I_j, \quad j = 1, 2, \ldots, N$, $e(x)$ και $\lambda q(x)$ έχουν το ίδιο πρόσημο και ισχύουν οι ανισότητες $|e(x)| > \frac{\rho}{2} > \lambda|q(x)|$. Τότε,

$$|f(x) - p(x)| = |f(x) - \lambda q(x) - p_n^*(x)| = |e(x) - \lambda q(x)| = |e(x)| - \lambda |q(x)| \leq \rho - \lambda |q(x)| < \rho.$$

Καταλήξαμε έτσι σε άτοπο οπότε ισχύει το ζητούμενο και ολοκληρώθηκε η απόδειξη του Θεωρήματος. •

Το θεώρημα αυτό δε μας λέει ότι το εναλλασσόμενο σύνολο σημείων είναι μοναδικό, που σημαίνει ότι μπορούν να υπάρχουν περισσότερα από $n+2$ εναλλασσόμενα σημεία. Αυτό φαίνεται καθαρά στο επόμενο παράδειγμα.

Παράδειγμα 8 Να βρεθεί η βέλτιστη προσέγγιση της συνάρτησης $\sin(4x)$ στο διάστημα $[-\pi,\pi]$, στον P_6.

Παρατηρούμε ότι η συνάρτηση $\sin(4x)$ έχει ένα εναλλασσόμενο σύνολο σημείων στο διάστημα $[-\pi,\pi]$ αποτελούμενο από 8 σημεία, τα $\{-\pi+\frac{\pi}{8}, -\pi+\frac{3\pi}{8}, -\pi+\frac{5\pi}{8}, -\pi+\frac{7\pi}{8}, -\pi+\frac{9\pi}{8}, -\pi+\frac{11\pi}{8}, -\pi+\frac{13\pi}{8}, -\pi+\frac{15\pi}{8}\}$. Αν θεωρήσουμε την ίδια συνάρτηση ως τη συνάρτηση σφάλμα $e(x)$, τότε το βέλτιστο πολυώνυμο μέχρι και έκτου βαθμού είναι το μηδενικό πολυώνυμο. Έχουμε λοιπόν ως συμπέρασμα ότι

$$p_0^* = p_1^* = p_2^* = p_3^* = p_4^* = p_5^* = p_6^* = 0.$$

Από το Θεώρημα 1 της Εισαγωγής έχουμε εξασφαλίσει την ύπαρξη της βέλτιστης ομοιόμορφης προσέγγισης, το Θεώρημα 3 όμως δε μας εξασφαλίζει τη μοναδικότητα αυτής, αφού η νόρμα άπειρο δεν είναι αυστηρά κυρτή νόρμα. Θα αποδείξουμε εδώ τη μοναδικότητα της βέλτιστης ομοιόμορφης προσέγγισης, χρησιμοποιώντας τη χαρακτηριστική ιδιότητα που αποδείξαμε στο Θεώρημα 8.

Θεώρημα 9 *(Μοναδικότητα) Η βέλτιστη ομοιόμορφη προσέγγιση p_n^* της $f \in C[a,b]$ στον P_n είναι μοναδική.*

$$\|f-p\| > \|f-p_n^*\| \quad \forall p \in P_n,\ p \neq p_n^*$$

Απόδειξη: Υποθέτουμε ότι υπάρχει πολυώνυμο $p \in P_n$ τέτοιο ώστε

$$\|f-p\| = \|f-p_n^*\| = E_n(f).$$

Τότε, σύμφωνα με το Θεώρημα 2, και ο κυρτός συνδυασμός $q = \frac{p+p_n^*}{2}$ θα είναι βέλτιστη προσέγγιση. Έστω ότι $\{x_0, x_1, \ldots, x_{n+1}\}$ συνιστά ένα εναλλασσόμενο σύνολο σημείων για την $f-q$ τέτοιο ώστε

$$f(x_j)-q(x_j) = \frac{f(x_j)-p(x_j)}{2} + \frac{f(x_j)-p_n^*(x_j)}{2} = (-1)^{l+j}E_n(f),\ j=0,1,\ldots,n+1$$

όπου το l παίρνει τις τιμές 0 ή 1 αν το σφάλμα στο σημείο x_0 είναι θετικό ή αρνητικό, αντίστοιχα. Επειδή όμως

$$\left|\frac{f(x_j)-p(x_j)}{2}\right| \le \frac{E_n(f)}{2}, \quad \left|\frac{f(x_j)-p_n^*(x_j)}{2}\right| \le \frac{E_n(f)}{2}$$

συμπεραίνουμε ότι

$$f(x_j) - p(x_j) = f(x_j) - p_n^*(x_j) = (-1)^{l+j} E_n(f),\ j = 0, 1, \dots, n+1$$

ή

$$p(x_j) = p_n^*(x_j),\ j = 0, 1, \dots, n+1$$

που αποδεικνύει ότι $p \equiv p_n^*$. •

Θα δώσουμε εδώ κάποια απλά παραδείγματα εύρεσης της βέλτιστης προσέγγισης.

Παράδειγμα 9 Να βρεθεί η βέλτιστη προσέγγιση της συνάρτησης x^2 στο διάστημα $[-1, 1]$, στον P_1.

Αναζητούμε ένα πολυώνυμο το πολύ πρώτου βαθμού. Υποθέτουμε ότι αυτό θα είναι της μορφής $\alpha x + \beta$, οπότε η συνάρτηση σφάλμα θα είναι $e(x) = x^2 - \alpha x - \beta$. Αναζητούμε λοιπόν τα 3 ακρότατα αυτής της συνάρτησης στο διάστημα $[-1, 1]$. Επειδή αυτή είναι συνεχής και παραγωγίσιμη στο $(-1, 1)$, βρίσκουμε τα εσωτερικά ακρότατα μηδενίζοντας την παράγωγο. Αυτή μας δίνει το σημείο $\frac{\alpha}{2}$. Προφανώς τα άλλα δυο ακρότατα θα είναι αναγκαστικά τα άκρα του διαστήματος -1 και 1. Θα έχουμε λοιπόν, σύμφωνα με το νόμο των εναλλασσόμενων σημείων, ότι

$$e(-1) = -e(\frac{\alpha}{2}) = e(1) \Leftrightarrow 1 + \alpha - \beta = -(\frac{\alpha^2}{4} - \frac{\alpha^2}{2} - \beta) = 1 - \alpha - \beta.$$

Λύνοντας το σύστημα αυτό βρίσκουμε ότι $\alpha = 0$ και $\beta = \frac{1}{2}$. Έχουμε επομένως ότι

$$p_1^*(x) = \frac{1}{2} = p_0^*(x).$$

Παράδειγμα 10 Να βρεθεί η βέλτιστη προσέγγιση της συνάρτησης x^3, $x \in [-1, 1]$, στον P_1.

Εργαζόμαστε με τον ίδιο τρόπο όπως και στο προηγούμενο παράδειγμα. Υποθέτουμε ότι $p_1^*(x) = \alpha x + \beta$, οπότε η συνάρτηση σφάλμα θα είναι $e(x) = x^3 - \alpha x - \beta$. Παραγωγίζοντας και μηδενίζοντας την παράγωγο, βρίσκουμε δυο εσωτερικά ακρότατα τα $\pm\sqrt{\frac{\alpha}{3}}$. Προφανώς το τρίτο ακρότατο θα είναι αναγκαστικά ένα από τα δυο άκρα του διαστήματος, το -1 ή το 1. Θα έχουμε λοιπόν, σύμφωνα με το νόμο των εναλλασσόμενων σημείων, ότι

$$e(-1) = -e\left(-\sqrt{\frac{\alpha}{3}}\right) = e\left(\sqrt{\frac{\alpha}{3}}\right) \Leftrightarrow -1 + \alpha - \beta = -\left(\frac{2\alpha}{3}\sqrt{\frac{\alpha}{3}} - \beta\right) = -\frac{2\alpha}{3}\sqrt{\frac{\alpha}{3}} - \beta.$$

ή

$$e\left(-\sqrt{\frac{\alpha}{3}}\right) = -e\left(\sqrt{\frac{\alpha}{3}}\right) = e(1) \Leftrightarrow \frac{2\alpha}{3}\sqrt{\frac{\alpha}{3}} - \beta = -\left(-\frac{2\alpha}{3}\sqrt{\frac{\alpha}{3}} - \beta\right) = 1 - \alpha - \beta.$$

Λύνοντας τα δυο παραπάνω συστήματα, βρίσκουμε την κοινή λύση $\alpha = \frac{3}{4}$ και $\beta = 0$. Έχουμε επομένως ότι

$$p_1^*(x) = \frac{3}{4}x.$$

Επειδή βρέθηκαν τελικά τέσσερα τα σημεία στο εναλλασσόμενο σύνολο, το πολυώνυμο που βρήκαμε θα είναι βέλτιστη προσέγγιση και στον P_2, θα έχουμε δηλαδή ότι $p_2^*(x) = \frac{3}{4}x$.

Παράδειγμα 11 Να βρεθεί η βέλτιστη προσέγγιση της συνάρτησης $|x|$, $x \in [-1, 1]$, στον P_2.

Υποθέτουμε ότι $p_2^*(x) = \alpha x^2 + \beta x + \gamma$, οπότε η συνάρτηση σφάλμα θα είναι $e(x) = |x| - \alpha x^2 - \beta x - \gamma$. Εδώ έχουμε το πρόβλημα ότι δεν ορίζεται η παράγωγος στο 0. Στην περίπτωση αυτή παραγωγίζουμε στα υποδιαστήματα $(-1, 0)$ και $(0, 1)$, έτσι έχουμε

$$e(x) = -x - \alpha x^2 - \beta x - \gamma = -\alpha x^2 - (\beta + 1)x - \gamma,$$

$$e'(x) = -2\alpha x - (\beta + 1) = 0 \Leftrightarrow x = -\frac{\beta + 1}{2\alpha}, \quad x \in (-1, 0)$$

και

$$e(x) = x - \alpha x^2 - \beta x - \gamma = -\alpha x^2 - (\beta - 1)x - \gamma,$$

$$e'(x) = -2\alpha x - (\beta - 1) = 0 \Leftrightarrow x = -\frac{\beta - 1}{2\alpha}, \quad x \in (0, 1).$$

Έχουμε βρεί λοιπόν δυο ακρότατα, ένα για κάθε υποδιάστημα. Τα υπόλοιπα δυο που μας χρειάζονται θα ληφθούν από τα άκρα -1 και 1 ή από το 0 που δεν ορίζεται η παράγωγος. Έχουμε επομένως να επιλέξουμε μέσα από τρεις επιλογές τετράδων, τις $\{-1, -\frac{\beta+1}{2\alpha}, 0, -\frac{\beta-1}{2\alpha\}}$ ή $\{-\frac{\beta+1}{2\alpha}, 0, -\frac{\beta-1}{2\alpha}, 1\}$ ή $\{-1, -\frac{\beta+1}{2\alpha}, -\frac{\beta-1}{2\alpha}, 1\}$. Δοκιμάζοντας την πρώτη επιλογή παίρνουμε

$$e(-1) = -e(-\frac{\beta + 1}{2\alpha}) = e(0) = -e(-\frac{\beta - 1}{2\alpha})$$

$$\Leftrightarrow -\alpha + \beta + 1 - \gamma = -\frac{(\beta + 1)^2}{4\alpha} + \gamma = -\gamma = -\frac{(\beta - 1)^2}{4\alpha} + \gamma$$

Λύνοντας το σύστημα βρίσκουμε $\alpha = 1$, $\beta = 0$ και $\gamma = \frac{1}{8}$. Το εναλλασσόμενο σύνολο θα είναι $\{-1, -\frac{1}{2}, 0, \frac{1}{2}\}$. Έχουμε επομένως ότι

$$p_2^*(x) = x^2 + \frac{1}{8}.$$

Υπολογίζουμε στη συνέχεια τα σφάλματα στο εναλλασσόμενο σύνολο: $e(-1) = -\frac{1}{8}$, $e(-\frac{1}{2}) = \frac{1}{8}$, $e(0) = -\frac{1}{8}$, $e(\frac{1}{2}) = \frac{1}{8}$. Αν υπολογίσουμε το σφάλμα της λύσης που βρήκαμε και στο σημείο 1, θα βρούμε $e(1) = -\frac{1}{8}$ που σημαίνει ότι υπάρχει ένα εναλλασσόμενο σύνολο αποτελούμενο από πέντε σημεία, δηλαδή η λύση που βρήκαμε θα είναι βέλτιστη προσέγγιση και στον P_3, θα έχουμε δηλαδή ότι $p_3^*(x) = x^2 + \frac{1}{8}$. Λόγω της μοναδικότητας, αφού βρήκαμε τη λύση, είναι ανώφελο να εξετάσουμε τις άλλες δυο επιλογές. Προφανώς η δεύτερη επιλογή θα μας δώσει την ίδια λύση ενώ η τρίτη δε θα μας δώσει λύση, αφού αν εξαιρέσουμε το ενδιάμεσο σημείο 0 δε μπορεί τα υπόλοιπα σημεία να συνεχίζουν να αποτελούν εναλλασσόμενο σύνολο.

Παρατηρούμε στα τρία παραπάνω παραδείγματα ότι εκεί όπου η συνάρτηση είναι άρτια, έδωσε βέλτιστο πολυώνυμο ένα άρτιο πολυώνυμο, ενώ εκεί όπου είναι περιττή, έδωσε περιττό. Αυτό συμβαίνει επειδή το διάστημα $[-1, 1]$ των τριών παραδειγμάτων, είναι συμμετρικό ως προς το 0. Μπορεί να αποδειχτεί γενικά ότι οι άρτιες συναρτήσεις ορισμένες σε συμμετρικό ως προς το 0 διάστημα δίνουν άρτιο βέλτιστο πολυώνυμο, ενώ οι περιττές, περιττό.

Η διαδικασία που ακολουθήθηκε στα παραδείγματα, δεν είναι εύκολο να ακολουθηθεί γενικά, για το λόγο ότι η συνάρτηση f μπορεί να είναι πολύπλοκη ή το n είναι σχετικά μεγάλο και είναι δύσκολο (ή αδύνατο) να προσδιορισθούν οι ρίζες της παραγώγου. Παρακάτω θα δούμε ότι υπάρχει αλγόριθμος προσδιορισμού του βέλτιστου πολυωνύμου, που βασίζεται στην προσέγγιση συναρτήσεων ορισμένων σε πεπερασμένο σύνολο σημείων. Πρώτα όμως θα μελετήσουμε το πρόβλημα της προσέγγισης πολυωνύμων με πολυώνυμα μικρότερου βαθμού.

2.2 Βέλτιστη ομοιόμορφη προσέγγιση πολυωνύμων – Πολυώνυμα Chebyshev

Ορισμός 4 *Το πολυώνυμο Chebyshev βαθμού n ορίζεται ως*

$$T_n(x) = \cos(n\theta), \qquad x = \cos(\theta). \tag{2.9}$$

Το πολυώνυμο Chebyshev μπορεί να γραφεί και ως

$$T_n(x) = \cos(n \arccos x).$$

Από την ιδιότητα των συνημιτόνων:

$$\cos((n+1)\theta) + \cos((n-1)\theta) = 2\cos(\theta)\cos(n\theta),$$

προκύπτει η αναδρομική σχέση των πολυωνύμων Chebyshev:

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x). \tag{2.10}$$

Από τη σχέση αυτή μπορούμε να υπολογίσουμε το $T_n(x)$ για οποιοδήποτε n, ξεκινώντας με $T_0(x) = 1$ και $T_1(x) = x$ που είναι προφανή από τη (2.9). Δίνουμε εδώ τους τρεις επόμενους όρους:

$$T_2(x) = 2xT_1(x) - T_0(x) = 2x^2 - 1$$

$$T_3(x) = 2xT_2(x) - T_1(x) = 2x(2x^2 - 1) - x = 4x^3 - 3x$$

$$T_4(x) = 2xT_3(x) - T_2(x) = 2x(4x^3 - 3x) - (2x^2 - 1) = 8x^4 - 8x^2 + 1.$$

Αποδεικνύεται πολύ εύκολα με επαγωγή ότι ο συντελεστής του μεγιστοβαθμίου όρου του $T_n(x)$ είναι 2^{n-1}, καθώς και ότι το $T_n(x)$ περιέχει μόνο τις άρτιες δυνάμεις του x αν ο n είναι άρτιος ή μόνο τις περιττές αν ο n είναι περιττός.

Τα πολυώνυμα Chebyshev έχουν πολύ καλές ιδιότητες. Μια από αυτές είναι η ορθογωνιότητά τους ως προς τη συνάρτηση βάρους $\frac{1}{\sqrt{1-x^2}}$, στο διάστημα $[-1,1]$. Η ιδιότητα αυτή δίνει ότι

$$\int_{-1}^{1}(1-x^2)^{-\frac{1}{2}}T_n(x)T_m(x)dx = 0, \quad m \neq n \tag{2.11}$$

και

$$\int_{-1}^{1}(1-x^2)^{-\frac{1}{2}}[T_n(x)]^2dx = \begin{cases} \frac{\pi}{2}, & n \geq 1 \\ \pi, & n = 0 \end{cases} \tag{2.12}$$

Την ιδιότητα αυτή θα τη δούμε να εφαρμόζεται στην προσέγγιση ελάχιστων τετραγώνων στο επόμενο Κεφάλαιο. Εδώ θα χρησιμοποιήσουμε μια άλλη πολύ σημαντική ιδιότητα τη λεγόμενη minmax ιδιότητα. Αυτή διατυπώνεται ως εξής

$$\min_{p \in P_n^c} \max_{x \in [-1,1]} |p(x)| = \max_{x \in [-1,1]} \left| \frac{c}{2^{n-1}} T_n(x) \right| = \frac{|c|}{2^{n-1}}, \tag{2.13}$$

όπου P_n^c είναι το σύνολο των πολυωνύμων n βαθμού με συντελεστή μεγιστοβαθμίου όρου c. Αυτό αποδεικνύεται εύκολα αφού, από τον ορισμό του (2.9), το πολυώνυμο $T_n(x)$ έχει $n+1$ ακρότατα στο διάστημα $[-1,1]$, εκεί όπου το $\cos(n\theta)$ παίρνει τιμές ± 1. Αυτά είναι τα σημεία

$$x_j = \cos\left(\frac{j\pi}{n}\right), \quad j = 0, 1, 2, \ldots, n.$$

Εύκολα φαίνεται ότι $T_n(x_j) = (-1)^j$, επομένως τα x_j αποτελούν ένα εναλλασσόμενο σύνολο από $n+1$ σημεία στο $[-1,1]$. Αν τώρα θεωρήσουμε τη βέλτιστη προσέγγιση $p_n^* \in P_n$ του $p \in P_{n+1}^c$ στον P_n, θα έχουμε ότι το σφάλμα

$$e(x) = p(x) - p_n^*(x) \in P_{n+1}$$

έχει ένα εναλλασσόμενο σύνολο αποτελούμενο από $n+2$ σημεία και τον ίδιο συντελεστή μεγιστοβαθμίου όρου με εκείνον του p, δηλαδή τον c. Επειδή η βέλτιστη προσέγγιση είναι μοναδική, το σφάλμα $e(x)$ δεν μπορεί να είναι άλλο παρά το $\frac{c}{2^n}T_{n+1}(x)$, αφού 2^n είναι ο συντελεστής μεγιστοβαθμίου όρου του $T_{n+1}(x)$. Αυτό το πολυώνυμο, επειδή είναι σφάλμα βέλτιστης προσέγγισης, θα πληροί και τη minmax ιδιότητα (2.13), σύμφωνα με τον ορισμό της βέλτιστης προσέγγισης.

Από την παραπάνω ανάλυση προκύπτει συγχρόνως και ο αλγόριθμος της εύρεσης της βέλτιστης προσέγγισης ενός πολυωνύμου $n+1$ βαθμού, στον P_n, αρκεί να αφαιρέσουμε το $\frac{c}{2^n}T_{n+1}(x)$. Θα δώσουμε εδώ κάποια παραδείγματα.

Παράδειγμα 12 Να βρεθεί η βέλτιστη προσέγγιση της συνάρτησης x^{n+1} στον P_n, στο διάστημα $[-1,1]$.

Σύμφωνα με την παραπάνω ανάλυση, αυτή θα είναι

$$p_n^*(x) = x^{n+1} - \frac{1}{2^n}T_{n+1}(x).$$

Για $n = 2$ αυτή γίνεται

$$p_2^*(x) = x^3 - \frac{1}{2^2}T_3(x) = x^3 - \frac{1}{4}(4x^3 - 3x) = \frac{3}{4}x.$$

Παρατηρούμε εδώ ότι αυτή είναι βέλτιστη προσέγγιση και στον P_1. Γενικά η βέλτιστη προσέγγιση του x^{n+1} θα είναι η ίδια και στον P_{n-1} αφού το $T_{n+1}(x)$ έχει μόνο άρτιες ή περιττές δυνάμεις του x.
Προφανώς εδώ βρήκαμε την ίδια ακριβώς λύση με εκείνη του Παραδείγματος 10, αφού λύσαμε το ίδιο πρόβλημα με διαφορετικό όμως τρόπο.

Για $n = 3$ έχουμε

$$p_3^*(x) = x^4 - \frac{1}{2^3}T_4(x) = x^4 - \frac{1}{8}(8x^4 - 8x^2 + 1) = x^2 - \frac{1}{8}.$$

Επειδή η παραπάνω ανάλυση ισχύει μόνο για το διάστημα $[-1, 1]$, για να αντιμετωπίσουμε προβλήματα που ορίζονται στο γενικότερο διάστημα $[a, b]$, χρησιμοποιούμε και εδώ το γραμμικό μετασχηματισμό

$$x = \frac{b-a}{2}t + \frac{b+a}{2}$$

που απεικονίζει το διάστημα $[-1, 1]$ στο $[a, b]$. Βρίσκουμε έτσι τη συνάρτηση

$$g(t) = f\left(\frac{b-a}{2}t + \frac{b+a}{2}\right), \quad t \in [-1, 1].$$

Στη συνέχεια, προσδιορίζουμε τη βέλτιστη προσέγγιση της g με τον παραπάνω αλγόριθμο και επιστρέφουμε στο διάστημα $[a, b]$, με τον αντίστροφο μετασχηματισμό

$$t = \frac{2}{b-a}x - \frac{b+a}{b-a}.$$

Παράδειγμα 13 Να βρεθεί η βέλτιστη προσέγγιση της συνάρτησης x^3 στον P_2, στο διάστημα $[-2, 2]$.

Ο γραμμικός μετασχηματισμός που μας πάει στο $[-1, 1]$ είναι $x = 2t$, οπότε

$$g(t) = f(2t) = (2t)^3 = 8t^3.$$

Σύμφωνα με την παραπάνω ανάλυση, η βέλτιστη προσέγγιση θα είναι

$$q_2^*(t) = 8t^3 - \frac{8}{2^2}T_3(x) = 8t^3 - 2(4t^3 - 3t) = 6t.$$

Επανερχόμαστε στο διάστημα $[-2, 2]$ με τον αντίστροφο μετασχηματισμό $t = \frac{x}{2}$, οπότε

$$p_2^*(x) = q_2^*(\frac{x}{2}) = 6\frac{x}{2} = 3x.$$

2.3 Ομοιόμορφη προσέγγιση σε πεπερασμένο σύνολο σημείων

Έστω ότι X_m είναι ένα σύνολο διακριτών σημείων

$$x_1 < x_2 < \cdots < x_m$$

στο διάστημα $[a, b]$ και $f(x_i), \;\; i = 1, 2, \ldots, m$ είναι οι τιμές της συνάρτησης f που ορίζεται στο X_m.
Το πρόβλημα με το οποίο θα ασχοληθούμε εδώ είναι η εύρεση της βέλτιστης ομοιόμορφης προσέγγισης της f στο σύνολο πολυωνύμων P_n που ορίζεται στο X_m.
Αν $m \leq n + 1$, τότε υπάρχει $p \in P_n$ τέτοιο ώστε

$$p(x_i) = f(x_i), \quad i = 1, 2, \ldots, m$$

που είναι το πολυώνυμο παρεμβολής της f, για οποιαδήποτε f ορισμένη στο X_m. Τότε το πολυώνυμο p είναι και η βέλτιστη προσέγγιση της f στον P_n.
Για να μελετήσουμε το γενικότερο πρόβλημα της βέλτιστης ομοιόμορφης προσέγγισης υποθέτουμε ότι $m \geq n + 2$.

Η βέλτιστη ομοιόμορφη προσέγγιση της f, ορισμένης στο X_m, στον P_n, χαρακτηρίζεται ακριβώς όπως και στη συνεχή περίπτωση. Ο χαρακτηρισμός αυτός δίνεται από το ακόλουθο θεώρημα.

Θεώρημα 10 *Έστω f μια συνάρτηση ορισμένη στο X_m. Το πολυώνυμο p_n^* είναι η βέλτιστη προσέγγιση της f στο σύνολο πολυωνύμων P_n ανν υπάρχει εναλλασσόμενο σύνολο σημείων για τη συνάρτηση σφάλμα $e = f - p_n^*$, που ορίζεται στο X_m, αποτελούμενο από $n + 2$ σημεία του X_m.*

Απόδειξη: Η απόδειξη είναι βήμα προς βήμα ανάλογη με εκείνη του Θεωρήματος 8 που αναφέρεται σε συνεχείς συναρτήσεις. Θα τη δώσουμε όμως για να φανούν κάποιες λεπτές διαφορές.

($\Leftarrow$) Έστω ότι υπάρχει εναλλασσόμενο σύνολο σημείων $\{y_1, y_2, \ldots, y_{n+2}\} \subset X_m$ για την $e = f - p_n^*$. Έστω επίσης ότι το p_n^* δεν είναι βέλτιστη προσέγγιση, άρα υπάρχει $q_n \in P_n$ τέτοιο ώστε

$$\|f - q_n\| < \|f - p_n^*\|.$$

Τότε

$$|f(y_j) - q_n(y_j)| < \|f - p_n^*\| = |f(y_j) - p_n^*(y_j)|, \quad j = 1, 2, \ldots, n + 2$$

και από τον ορισμό του εναλλασσομένου συνόλου προκύπτει ότι: Αν

$$f(y_j) - p_n^*(y_j) = \|f - p_n^*\| > 0 \Rightarrow f(y_j) - p_n^*(y_j) > f(y_j) - q_n(y_j)$$

ή

$$f(y_j) - p_n^*(y_j) - (f(y_j) - q_n(y_j)) = q_n(y_j) - p_n^*(y_j) > 0,$$

ενώ αν

$$f(y_j) - p_n^*(y_j) = -\|f - p_n^*\| < 0 \Rightarrow f(y_j) - p_n^*(y_j) < f(y_j) - q_n(y_j)$$

ή

$$f(y_j) - p_n^*(y_j) - (f(y_j) - q_n(y_j)) = q_n(y_j) - p_n^*(y_j) < 0.$$

Αυτό σημαίνει ότι το πολυώνυμο $q_n - p_n^*$ εναλλάσσει συνεχώς πρόσημο καθώς το j μεταβάλλεται από το 1 στο $n+2$. Δηλαδή, αν επεκτείνουμε τα πολυώνυμα q_n και p_n^* σε ολόκληρο το διάστημα $[a, b]$, υπάρχει ρίζα σε κάθε διάστημα $(y_j, y_{j+1}),\ \ j = 1, 2, \ldots, n+1$. Έτσι το πολυώνυμο $q_n - p_n^* \in P_n$ έχει $n+1$ ρίζες, το οποίο είναι άτοπο.

($\Rightarrow$) Έστω ότι p_n^* είναι η βέλτιστη προσέγγιση της f, όπου η f δεν είναι πολυώνυμο n βαθμού. Έστω επίσης ότι το μέγιστο εναλλασσόμενο σύνολο σημείων αποτελείται από $m+1$ σημεία $y_1 < y_2 < \cdots < y_{m+1}$. Είναι φανερό ότι και στη διακριτή περίπτωση ισχύει το Θεώρημα 7, οπότε $m \geq 1$. Θα αποδείξουμε ότι $m \geq n+1$.

Έστω ότι $m \leq n$ και ότι

$$\|f - p_n^*\| = \rho.$$

Θεωρούμε ότι τα N σημεία $u_1 < u_2 < \cdots < u_N$ του X_m, με $N \geq m+1$, δίνουν το μέγιστο σφάλμα $\|e(u_j)\| = \rho, j = 1, 2, \ldots, N$. Ονομάζουμε (+) σημεία εκείνα για τα οποία η συνάρτηση e παίρνει θετική τιμή ρ και (−) σημεία εκείνα για τα οποία η συνάρτηση e παίρνει αρνητική τιμή $-\rho$. Χωρίζουμε στη συνέχεια το σύνολο $\{u_1, u_2, \ldots, u_N\}$ σε υποσύνολα που θα περιέχουν τα διαδοχικά (+) ή (−) σημεία. Έστω ότι αυτά είναι:

$$\begin{array}{rcl}
\{u_1, u_2, \ldots, u_{k_1}\} & (+) & \text{υποσύνολο} \\
\{u_{k_1+1}, u_{k_1+2}, \ldots, u_{k_2}\} & (-) & \text{υποσύνολο} \\
\vdots & \vdots & \vdots \\
\{u_{k_m+1}, u_{k_m+2}, \ldots, u_{k_{m+1}}\} & (-)^m & \text{υποσύνολο}
\end{array}$$

Κάθε υποσύνολο περιέχει τουλάχιστον ένα σημείο, ενώ $2 \leq m+1 \leq n+1$ από τις υποθέσεις που κάναμε. Θεωρούμε τώρα m σημεία $z_1, z_2, \ldots, z_m$ τέτοια ώστε

$$z_j > u_{k_j} \ \text{ και } \ z_j < u_{k_j+1}.$$

Ορίζουμε στη συνέχεια το πολυώνυμο $q \in P_m$:

$$q(x) = (z_1 - x)(z_2 - x) \cdots (z_m - x).$$

Παρατηρούμε ότι το $q(x)$ έχει το ίδιο πρόσημο με τη συνάρτηση $e(x)$ σε όλα τα σημεία $u_1, u_2, \ldots, u_N$. Πραγματικά, αν το x ανήκει στο πρώτο υποσύνολο $\{u_1, u_2, \ldots, u_{k_1}\}$, που είναι (+) υποσύνολο, το $q(x)$ είναι θετικό αφού $z_i > x,\ \ i = 1, 2, \ldots, m$, ενώ κάθε φορά που το x μεταβαίνει σε επόμενο υποσύνολο το πολυώνυμο q αλλάζει πρόσημο αφού κάθε φορά και ένα από τα z_i γίνεται μικρότερο του x. Ονομάζουμε με U_N το σύνολο των $(\pm)$ σημείων και με R το σύνολο των σημείων που δεν είναι $(\pm)$ σημεία $(R = X_m \setminus U_N)$. Τότε,

$$\max_{x\in R} |e(x)| = \rho' < \rho.$$

Υποθέτουμε ότι $\|q\| = M$ και επιλέγουμε ένα $\lambda > 0$ τόσο μικρό ώστε

$$\lambda M < \rho - \rho'.$$

Θα αποδείξουμε τώρα ότι το πολυώνυμο $p(x) = \lambda q(x) + p_n^*(x) \in P_n$ είναι καλύτερη προσέγγιση από το $p_n^*(x)$.
Έστω ότι $x \in R$, τότε

$$|f(x)-p(x)| = |f(x)-\lambda q(x)-p_n^*(x)| = |e(x)-\lambda q(x)| \leq |e(x)|+\lambda|q(x)| \leq \rho'+\lambda M < \rho.$$

Αν το x ανήκει στο U_N, τότε $e(x)$ και $\lambda q(x)$ έχουν το ίδιο πρόσημο με $|e(x)| > \lambda|q(x)|$. Τότε,

$$|f(x)-p(x)| = |f(x)-\lambda q(x)-p_n^*(x)| = |e(x)-\lambda q(x)| = |e(x)|-\lambda|q(x)| \leq \rho-\lambda|q(x)| < \rho.$$

Καταλήξαμε έτσι σε άτοπο οπότε ισχύει το ζητούμενο και ολοκληρώθηκε η απόδειξη του Θεωρήματος. •

Και στην περίπτωση αυτή ισχύει η παρατήρηση ότι το εναλλασσόμενο σύνολο σημείων δεν είναι μοναδικό, που σημαίνει ότι μπορούν να υπάρχουν περισσότερα από $n+2$ εναλλασσόμενα σημεία.

Απομένει να δώσουμε το θεώρημα της μοναδικότητας αφού η ύπαρξη εξασφαλίζεται από το Θεώρημα 1 της Εισαγωγής. Θα το δώσουμε απλώς χωρίς απόδειξη, επειδή η απόδειξη είναι ίδια βήμα προς βήμα με εκείνη του Θεωρήματος 9.

Θεώρημα 11 *(Μοναδικότητα) Η βέλτιστη ομοιόμορφη προσέγγιση p_n^* της f, ορισμένης στο X_m, στον P_n είναι μοναδική.*

$$\|f-p\| > \|f-p_n^*\| \quad \forall p \in P_n,\ p \neq p_n^*.$$

Θα αποδείξουμε τώρα ότι η βέλτιστη ομοιόμορφη προσέγγιση της f που ορίζεται στο διάστημα $[a,b]$ ή στο σύνολο σημείων X_m, στον P_n μπορεί να αναχθεί τελικά σε αντίστοιχο πρόβλημα πάνω σε κάποιο σύνολο σημείων X_{n+2}.

Θεώρημα 12 *Αν p_n^* είναι η βέλτιστη ομοιόμορφη προσέγγιση της $f \in C[a,b]$, στον P_n, τότε υπάρχει $X_{n+2}^* \subset [a,b]$ τέτοιο ώστε:*

$$E_n(f;[a,b]) = E_n(f;X_{n+2}^*) = \max_{x\in X_{n+2}^*} |f(x)-p_n^*(x)| < \max_{x\in X_{n+2}^*} |f(x)-p(x)| \qquad (2.14)$$

για κάθε $p \in P_n, \quad p \neq p_n^*$. *Επιπλέον, για κάθε* $X_{n+2} \in [a,b]$, *ισχύει ότι*

$$\begin{aligned} E_n(f;X_{n+2}) &= \min_{p\in P_n} \max_{x\in X_{n+2}} |f(x)-p(x)| \\ &= \max_{x\in X_{n+2}} |f(x)-p_n^*(X_{n+2};x)| \leq E_n(f;[a,b]) \\ &= E_n(f;X_{n+2}^*), \end{aligned} \tag{2.15}$$

όπου η ισότητα είναι δυνατή μόνο αν $p_n^*(X_{n+2}) = p_n^*$.

Απόδειξη: Η πρώτη σχέση (2.14) είναι προφανής αν θεωρήσουμε το X_{n+2}^* ως το εναλλασσόμενο σύνολο σημείων. Για την απόδειξη της δεύτερης (2.15), θεωρούμε ότι $p_n^*(X_{n+2}) \neq p_n^*$. Τότε από τη μοναδικότητα έχουμε ότι

$$E_n(f;X_{n+2}) = \|f - p_n^*(X_{n+2})\| < \|f - p_n^*\| = E_n(f;[a,b]) = E_n(f;X_{n+2}^*),$$

αφού το p_n^* είναι διαφορετικό από το βέλτιστο πολυώνυμο $p_n^*(X_{n+2})$ στο σύνολο X_{n+2}. Αν $p_n^*(X_{n+2}) = p_n^*$, τότε προφανώς έχουμε ισότητα, που σημαίνει ότι και το X_{n+2} αποτελεί εναλλασσόμενο σύνολο σημείων. •

Αν αντικαταστήσουμε με X_m το $[a,b]$ έχουμε το αντίστοιχο θεώρημα, η απόδειξη του οποίου είναι ακριβώς ίδια.

Θεώρημα 13 *Αν* p_n^* *είναι η βέλτιστη ομοιόμορφη προσέγγιση της* f, *ορισμένης στο* X_m, *στον* P_n, *τότε υπάρχει* $X_{n+2}^* \subseteq X_m$ *τέτοιο ώστε:*

$$E_n(f;X_m) = E_n(f;X_{n+2}^*) = \max_{x\in X_{n+2}^*} |f(x) - p_n^*(X_m;x)| < \max_{x\in X_{n+2}^*} |f(x)-p(x)| \tag{2.16}$$

για κάθε $p \in P_n, \quad p \neq p_n^*$. *Επιπλέον, για κάθε* $X_{n+2} \subseteq X_m$, *ισχύει ότι*

$$E_n(f;X_{n+2}) \leq E_n(f;X_m) = E_n(f;X_{n+2}^*), \tag{2.17}$$

όπου η ισότητα είναι δυνατή μόνο αν $p_n^*(X_{n+2}) = p_n^*$.

Τα παραπάνω θεωρήματα μας δίνουν τη δυνατότητα να αναζητούμε για τη βέλτιστη ομοιόμορφη προσέγγιση, σε σύνολα σημείων αποτελούμενα από $n+2$ σημεία. Αν έχουμε το πρόβλημα στο διάστημα $[a,b]$, τότε υπάρχουν άπειρα τέτοια σύνολα. Αν όμως εργαζόμαστε στο X_m, ο συνολικός αριθμός των συνόλων αυτών είναι $\binom{m}{n+2}$. Τότε το πρόβλημα ανάγεται στην εύρεση του

$$\max_{i=1,2,\ldots,\binom{m}{n+2}} E_n(f;X_{m,i}) = E_n(f;X_{m,i^*}), \tag{2.18}$$

όπου με $X_{m,i}$ έχουμε συμβολίσει το i στη σειρά σύνολο αποτελούμενο από $n+2$ σημεία. Τότε,

$$p_n^*(X_m) = p_n^*(X_{m,i^*}).$$

Αυτό όμως απαιτεί το να μπορούμε να βρούμε το σφάλμα $E_n(f;X_{n+2})$ και το πολυώνυμο $p_n^*(X_{n+2})$, δοθέντος του συνόλου X_{n+2} αποτελούμενο από $n+2$ σημεία.

2.3.1 Εύρεση ομοιόμορφης προσέγγισης στο σύνολο X_{n+2}

Θεωρούμε το πολυώνυμο παρεμβολής $p \in P_{n+1}$ της f στα σημεία $x_i,\ i = 1, 2, \ldots, n+2$ του συνόλου X_{n+2}

$$p(x_i) = f(x_i), \qquad i = 1, 2, \ldots, n+2.$$

Το p είναι μοναδικό και δίνεται από τον τύπο του Lagrange

$$p(x) = \sum_{i=1}^{n+2} \prod_{j=1, j\neq i}^{n+2} \frac{x - x_j}{x_i - x_j} f(x_i).$$

Είναι εύκολο να δούμε ότι αν θέσουμε

$$\omega(x) = \prod_{i=1}^{n+2} (x - x_i),$$

το πολυώνυμο παρεμβολής γράφεται

$$p(x) = \sum_{i=1}^{n+2} \frac{\omega(x)}{x - x_i} \, \frac{f(x_i)}{\omega'(x_i)}.$$

Από τη σχέση αυτή φαίνεται καθαρά ότι ο συντελεστής του μεγιστοβαθμίου όρου του p είναι ο $\sum_{i=1}^{n+2} \frac{f(x_i)}{\omega'(x_i)}$. Αν αυτός είναι μηδέν, τότε θα έχουμε ότι $p \in P_n$, που θα αποτελεί και τη βέλτιστη ομοιόμορφη προσέγγιση, δηλαδή $p^*(X_{n+2}) = p$ με σφάλμα $E_n(f; X_{n+2}) = 0$.

Υποθέτουμε εδώ ότι $p \notin P_n$. Θεωρούμε στη συνέχεια το πολυώνυμο $l_{n+1} \in P_{n+1}$ τέτοιο ώστε

$$l_{n+1}(x_i) = (-1)^i, \qquad i = 1, 2, \ldots, n+2.$$

Ως πολυώνυμο παρεμβολής στο σύνολο σημείων X_{n+2}, αυτό θα είναι μοναδικό και θα δίνεται από τον τύπο παρεμβολής του Lagrange ως εξής:

$$l_{n+1}(x) = \sum_{i=1}^{n+2} \frac{\omega(x)}{x - x_i} \, \frac{(-1)^i}{\omega'(x_i)}.$$

Από τη σχέση

$$\omega'(x_i) = \prod_{j=1, j\neq i}^{n+2} (x_i - x_j) = \prod_{j=1}^{i-1} (x_i - x_j) \prod_{j=i+1}^{n+2} (x_j - x_i) \ (-1)^{n+2-i},$$

προκύπτει ότι

$$\operatorname{sign}(\omega'(x_i)) = \operatorname{sign}((-1)^{n-i}) \ \text{ ή } \ \operatorname{sign}\left(\frac{(-1)^i}{\omega'(x_i)}\right) = \operatorname{sign}((-1)^n).$$

Αυτό σημαίνει ότι οι όροι του αθροίσματος $\sum_{i=1}^{n+2} \frac{(-1)^i}{\omega'(x_i)}$, που αποτελεί το συντελεστή του μεγιστοβαθμίου όρου του πολυωνύμου l_{n+1}, δεν αλλάζουν πρόσημο και επομένως

$$\sum_{i=1}^{n+2} \frac{(-1)^i}{\omega'(x_i)} \neq 0.$$

Θεωρούμε τώρα το πολυώνυμο

$$q = p - \lambda l_{n+1},$$

όπου p, l_{n+1} τα προαναφερθέντα πολυώνυμα παρεμβολής και

$$\lambda = \frac{\sum_{i=1}^{n+2} \frac{f(x_i)}{\omega'(x_i)}}{\sum_{i=1}^{n+2} \frac{(-1)^i}{\omega'(x_i)}}.$$

Τότε όμως $q \in P_n$ επειδή p και λl_{n+1} έχουν τον ίδιο μεγιστοβάθμιο όρο. Επίσης έχουμε ότι

$$e(x_i) = f(x_i) - (p(x_i) - \lambda l_{n+1}(x_i)) = \lambda l_{n+1}(x_i) = \lambda(-1)^i, \qquad i = 1, 2, \ldots, n+2,$$

που σημαίνει ότι τα σημεία $x_1, x_2, \ldots, x_{n+2}$ αποτελούν ένα εναλλασσόμενο σύνολο σημείων για την $e(x)$, επομένως το πολυώνυμο q είναι η βέλτιστη ομοιόμορφη προσέγγιση της f στο X_{n+2}, $q = p^*(X_{n+2})$. Επίσης προκύπτει ότι

$$E_n(f; X_{n+2}) = |\lambda|.$$

Θεωρώντας τώρα ότι έχουμε το σύνολο $X_m,\ m > n+2$, παίρνουμε όλα τα υποσύνολα αποτελούμενα από $n+2$ σημεία και υπολογίζουμε όλα τα λ. Τότε το $p^*(X_m) = p^*(X^*_{n+2})$ αντιστοιχεί στο σύνολο σημείων X^*_{n+2} που αντιστοιχεί στο μέγιστο $|\lambda|$, σύμφωνα με το Θεώρημα 13.

Θα δούμε καλύτερα τη διαδικασία υπολογισμού του βέλτιστου πολυωνύμου, δια μέσου του επόμενου παραδείγματος:

Παράδειγμα 14 Δίνεται το σύνολο σημείων $X_5 = \{-1, -\frac{1}{2}, 0, \frac{1}{2}, 1\}$ και θέλουμε να βρούμε τη βέλτιστη προσέγγιση δευτέρου βαθμού για τη συνάρτηση $f(x) = |x|$.

Ο πίνακας τιμών της συνάρτησης είναι

x_i	-1	$-\frac{1}{2}$	0	$\frac{1}{2}$	1
f_i	1	$\frac{1}{2}$	0	$\frac{1}{2}$	1

Πρέπει να βρούμε όλα τα δυνατά υποσύνολα σημείων αποτελούμενα από $n+2=4$ σημεία και να ακολουθήσουμε τη διαδικασία που περιγράψαμε.

1) $X_{5,1} = \{-1, -\frac{1}{2}, 0, \frac{1}{2}\}$:

$$\begin{array}{rclclclcl}
\omega'(x_1) &=& \omega'(-1) &=& (-1+\frac{1}{2})(-1-0)(-1-\frac{1}{2}) &=& (-\frac{1}{2})(-1)(-\frac{3}{2}) &=& -\frac{3}{4} \\
\omega'(x_2) &=& \omega'(-\frac{1}{2}) &=& (-\frac{1}{2}+1)(-\frac{1}{2}-0)(-\frac{1}{2}-\frac{1}{2}) &=& \frac{1}{2}(-\frac{1}{2})(-1) &=& \frac{1}{4} \\
\omega'(x_3) &=& \omega'(0) &=& (0+1)(0+\frac{1}{2})(0-\frac{1}{2}) &=& 1\frac{1}{2}(-\frac{1}{2}) &=& -\frac{1}{4} \\
\omega'(x_4) &=& \omega'(\frac{1}{2}) &=& (\frac{1}{2}+1)(\frac{1}{2}+\frac{1}{2})(\frac{1}{2}-0) &=& \frac{3}{2}1\frac{1}{2} &=& \frac{3}{4}
\end{array}$$

$$\lambda^{(1)} = \frac{\sum_{i=1}^{4} \frac{f(x_i)}{\omega'(x_i)}}{\sum_{i=1}^{4} \frac{(-1)^i}{\omega'(x_i)}} = \frac{\frac{1}{-\frac{3}{4}} + \frac{\frac{1}{2}}{\frac{1}{4}} + \frac{0}{-\frac{1}{4}} + \frac{\frac{1}{2}}{\frac{3}{4}}}{\frac{-1}{-\frac{3}{4}} + \frac{1}{\frac{1}{4}} + \frac{-1}{-\frac{1}{4}} + \frac{1}{\frac{3}{4}}} = \frac{\frac{4}{3}}{\frac{32}{3}} = \frac{1}{8}.$$

2) $X_{5,2} = \{-1, -\frac{1}{2}, 0, 1\}$:

$$\begin{array}{lll|lll} \omega'(-1) = & (-1+\frac{1}{2})(-1)(-1-1) & = -1 & \omega'(0) = & 1\frac{1}{2}(-1) & = -\frac{1}{2} \\ \omega'(-\frac{1}{2}) = & (-\frac{1}{2}+1)(-\frac{1}{2})(-\frac{1}{2}-1) & = \frac{3}{8} & \omega'(1) = & (1+1)(1+\frac{1}{2})(1) & = 3 \end{array}$$

$$\lambda^{(2)} = \frac{\sum_{i=1}^{4} \frac{f(x_i)}{\omega'(x_i)}}{\sum_{i=1}^{4} \frac{(-1)^i}{\omega'(x_i)}} = \frac{\frac{1}{-1} + \frac{\frac{1}{2}}{\frac{3}{8}} + \frac{0}{-\frac{1}{2}} + \frac{1}{3}}{\frac{-1}{-1} + \frac{1}{\frac{3}{8}} + \frac{-1}{-\frac{1}{2}} + \frac{1}{\frac{1}{3}}} = \frac{\frac{2}{3}}{6} = \frac{1}{9}.$$

3) $X_{5,3} = \{-1, -\frac{1}{2}, \frac{1}{2}, 1\}$:

$$\begin{array}{lll|lll} \omega'(-1) = & (-1+\frac{1}{2})(-1-\frac{1}{2})(-1-1) & = -\frac{3}{2} & \omega'(\frac{1}{2}) = & (\frac{1}{2}+1)(\frac{1}{2}+\frac{1}{2})(\frac{1}{2}-1) & = -\frac{3}{4} \\ \omega'(-\frac{1}{2}) = & (-\frac{1}{2}+1)(-\frac{1}{2}-\frac{1}{2})(-\frac{1}{2}-1) & = \frac{3}{4} & \omega'(1) = & (1+1)(1+\frac{1}{2})(1-\frac{1}{2}) & = \frac{3}{2} \end{array}$$

$$\lambda^{(3)} = \frac{\sum_{i=1}^{4} \frac{f(x_i)}{\omega'(x_i)}}{\sum_{i=1}^{4} \frac{(-1)^i}{\omega'(x_i)}} = \frac{\frac{1}{-\frac{3}{2}} + \frac{\frac{1}{2}}{\frac{3}{4}} + \frac{\frac{1}{2}}{-\frac{3}{4}} + \frac{1}{\frac{3}{2}}}{\frac{-1}{-\frac{3}{2}} + \frac{1}{\frac{3}{4}} + \frac{-1}{-\frac{3}{4}} + \frac{1}{\frac{3}{2}}} = \frac{0}{4} = 0.$$

4) $X_{5,4} = \{-1, 0, \frac{1}{2}, 1\}$:

$$\begin{array}{lll|lll} \omega'(-1) = & (-1)(-1-\frac{1}{2})(-1-1) & = -3 & \omega'(\frac{1}{2}) = & (\frac{1}{2}+1)\frac{1}{2}(\frac{1}{2}-1) & = -\frac{3}{8} \\ \omega'(0) = & 1(-\frac{1}{2})(-1) & = \frac{1}{2} & \omega'(1) = & (1+1)1(1-\frac{1}{2}) & = 1 \end{array}$$

$$\lambda^{(4)} = \frac{\sum_{i=1}^{4} \frac{f(x_i)}{\omega'(x_i)}}{\sum_{i=1}^{4} \frac{(-1)^i}{\omega'(x_i)}} = \frac{\frac{1}{-3} + \frac{0}{\frac{1}{2}} + \frac{\frac{1}{2}}{-\frac{3}{8}} + \frac{1}{1}}{\frac{-1}{-3} + \frac{1}{\frac{1}{2}} + \frac{-1}{-\frac{3}{8}} + \frac{1}{1}} = \frac{-\frac{2}{3}}{6} = -\frac{1}{9}.$$

5) $X_{5,5} = \{-\frac{1}{2}, 0, \frac{1}{2}, 1\}$:

$$\begin{array}{lll|lll} \omega'(-\frac{1}{2}) = & -\frac{1}{2}(-\frac{1}{2}-\frac{1}{2})(-\frac{1}{2}-1) & = -\frac{3}{4} & \omega'(\frac{1}{2}) = & (\frac{1}{2}+\frac{1}{2})\frac{1}{2}(\frac{1}{2}-1) & = -\frac{1}{4} \\ \omega'(0) = & \frac{1}{2}(-\frac{1}{2})(-1) & = \frac{1}{4} & \omega'(1) = & (1+\frac{1}{2})1(1-\frac{1}{2}) & = \frac{3}{4} \end{array}$$

$$\lambda^{(5)} = \frac{\sum_{i=1}^{4} \frac{f(x_i)}{\omega'(x_i)}}{\sum_{i=1}^{4} \frac{(-1)^i}{\omega'(x_i)}} = \frac{\frac{\frac{1}{2}}{-\frac{3}{4}} + \frac{0}{\frac{1}{4}} + \frac{\frac{1}{2}}{-\frac{1}{4}} + \frac{1}{\frac{3}{4}}}{\frac{-1}{-\frac{3}{4}} + \frac{1}{\frac{1}{4}} + \frac{-1}{-\frac{1}{4}} + \frac{1}{\frac{3}{4}}} = \frac{-\frac{4}{3}}{\frac{32}{3}} = -\frac{1}{8}.$$

Παρατηρούμε ότι $|\lambda_{\max}| = |\lambda^{(1)}| = |\lambda^{(5)}| = \frac{1}{8}$, επομένως η βέλτιστη προσέγγιση δίνεται από δυο σύνολα σημείων το $X_{5,1} = \{-1, -\frac{1}{2}, 0, \frac{1}{2}\}$ και το $X_{5,5} = \{-\frac{1}{2}, 0, \frac{1}{2}, 1\}$. Προφανώς, λόγω της μοναδικότητας, και τα δυο αυτά σύνολα θα δώσουν το ίδιο βέλτιστο πολυώνυμο. Για να το επιβεβαιώσουμε θα το υπολογίσουμε, σύμφωνα με τη θεωρία που αναπτύχθηκε, χρησιμοποιώντας και τα δυο σύνολα.

$$\begin{aligned} p^*(X_5;x) &= p^*(X_{5,1};x) = p_3(x) - \lambda^{(1)} l_3(x) = \textstyle\sum_{i=1}^{4} \frac{\omega(x)}{x-x_i} \frac{f(x_i)}{\omega'(x_i)} - \lambda^{(1)} \sum_{i=1}^{4} \frac{\omega(x)}{x-x_i} \frac{(-1)^i}{\omega'(x_i)} \\ &= (x+\tfrac{1}{2})x(x-\tfrac{1}{2})\frac{1}{-\frac{3}{4}} + (x+1)x(x-\tfrac{1}{2})\frac{\frac{1}{2}}{\frac{1}{4}} + (x+1)(x+\tfrac{1}{2})x\frac{\frac{1}{2}}{\frac{3}{4}} \\ &- \tfrac{1}{8}\Big((x+\tfrac{1}{2})x(x-\tfrac{1}{2})\frac{-1}{-\frac{3}{4}} + (x+1)x(x-\tfrac{1}{2})\frac{1}{\frac{1}{4}} + (x+1)(x+\tfrac{1}{2})(x-\tfrac{1}{2})\frac{-1}{-\frac{1}{4}} \\ &+ (x+1)(x+\tfrac{1}{2})x\frac{1}{\frac{3}{4}}\Big) = x^2 + \tfrac{1}{8}. \end{aligned}$$

$$\begin{aligned}
p^*(X_5;x) &= p^*(X_{5,5};x) = p_3(x) - \lambda^{(5)} l_3(x) = \textstyle\sum_{i=1}^4 \frac{\omega(x)}{x-x_i}\,\frac{f(x_i)}{\omega'(x_i)} - \lambda^{(1)} \sum_{i=1}^4 \frac{\omega(x)}{x-x_i}\,\frac{(-1)^i}{\omega'(x_i)} \\
&= x(x-\tfrac{1}{2})(x-1)\frac{\frac{1}{2}}{-\frac{3}{4}} + (x+\tfrac{1}{2})x(x-1)\frac{\frac{1}{2}}{-\frac{1}{4}} + (x+\tfrac{1}{2})x(x-\tfrac{1}{2})\frac{1}{\frac{3}{4}} \\
&+ \tfrac{1}{8}\Big(x(x-\tfrac{1}{2})(x-1)\frac{-1}{-\frac{3}{4}} + (x+\tfrac{1}{2})(x-\tfrac{1}{2})(x-1)\frac{1}{\frac{1}{4}} + (x+\tfrac{1}{2})x(x-1)\frac{-1}{-\frac{1}{4}} \\
&+ (x+\tfrac{1}{2})x(x-\tfrac{1}{2})\frac{1}{\frac{3}{4}} \Big) = x^2 + \tfrac{1}{8}.
\end{aligned}$$

Ας υπολογίσουμε τώρα το σφάλμα $e(x) = |x| - (x^2 + \frac{1}{8})$ στο σύνολο σημείων X_5:

$$\begin{gathered}
e(-1) = |-1| - ((-1)^2 + \frac{1}{8}) = -\frac{1}{8} \\
e\left(-\frac{1}{2}\right) = \left|-\frac{1}{2}\right| - \left(\left(-\frac{1}{2}\right)^2 + \frac{1}{8}\right) = \frac{1}{8} \\
e(0) = |0| - (0^2 + \frac{1}{8}) = -\frac{1}{8} \\
e\left(\frac{1}{2}\right) = \left|\frac{1}{2}\right| - \left(\left(\frac{1}{2}\right)^2 + \frac{1}{8}\right) = \frac{1}{8} \\
e(1) = |1| - (1^2 + \frac{1}{8}) = -\frac{1}{8}.
\end{gathered}$$

Παρατηρούμε εδώ ότι το σύνολο X_5 αποτελεί εναλλασσόμενο σύνολο αποτελούμενο από πέντε σημεία, επομένως το βέλτιστο πολυώνυμο που βρήκαμε είναι και βέλτιστο πολυώνυμο τρίτου βαθμού, δηλαδή $p^*(X_5) \in P_3$. Επίσης παρατηρούμε ότι η συνάρτηση $|x|$ είναι άρτια και έδωσε ως βέλτιστο πολυώνυμο ένα άρτιο πολυώνυμο. Αυτό συμβαίνει επειδή το σύνολο σημείων X_5, είναι συμμετρικό ως προς το 0. Ισχύει και, εδώ στη διακριτή περίπτωση, η παρατήρηση που κάναμε στη συνεχή. Μπορεί δηλαδή να αποδειχτεί γενικά ότι οι άρτιες συναρτήσεις ορισμένες σε συμμετρικό ως προς το 0 σύνολο σημείων δίνουν άρτιο βέλτιστο πολυώνυμο, ενώ οι περιττές, περιττό.

Θα δώσουμε εδώ έναν εναλλακτικό τρόπο εύρεσης του βέλτιστου πολυωνύμου χωρίς τη χρήση πολυωνύμων παρεμβολής αλλά με προσδιοριστέους συντελεστές. Έστω ότι $p_n^*(X_{n+2};x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_n x^n$ είναι το βέλτιστο πολυώνυμο της f, ορισμένης στο X_{n+2}, στον P_n. Τότε, σύμφωνα με τη θεωρία των εναλλασσόμενων σημείων θαι σχύει:

$$f(x_i) - p_n^*(X_{n+2};x_i) = \lambda(-1)^i \Leftrightarrow (-1)^i\lambda + p_n^*(X_{n+2};x_i) = f(x_i),\ i = 1,2,\ldots,n+2$$

όπου λ είναι το σφάλμα που πρέπει να προσδιοριστεί. Ισοδύναμα έχουμε:

$$(-1)^i\lambda + \alpha_0 + \alpha_1 x + \cdots + \alpha_n x^n = f(x_i),\ i = 1,2,\ldots,n+2.$$

Η σχέση όμως αυτή αποτελεί ένα $(n+2)\times(n+2)$ γραμμικό σύστημα $A\alpha = \mathrm{f}$ όπου

$$A = \begin{bmatrix} -1 & 1 & x_1 & \cdots & x_1^n \\ 1 & 1 & x_2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & & \vdots \\ (-1)^{n+2} & 1 & x_{n+2} & \cdots & x_{n+2}^n \end{bmatrix}, \quad \alpha = \begin{bmatrix} \lambda \\ \alpha_0 \\ \vdots \\ \alpha_n \end{bmatrix}, \quad \mathrm{f} = \begin{bmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_n) \end{bmatrix}$$

Με τη λύση του παραπάνω γραμμικού συστήματος βρίσκουμε συγχρόνως το σφάλμα λ και τους συντελεστές του βέλτιστου πολυωνύμου α_i, $i = 0, 1, \ldots, n$. Παρατηρούμε ότι το σύστημα αυτό είναι τύπου Vandermonde. Εδώ χρειάζεται μεγάλη προσοχή διότι, για σχετικά μεγάλα n, τα συστήματα αυτού του τύπου έχουν κακή κατάσταση, που σημαίνει ότι μικρά σφάλματα κατά τους υπολογισμούς επιφέρουν μεγάλα σφάλματα στη λύση.

Επιστρέφουμε τώρα στο Παράδειγμα 14 για να υπολογίσουμε το βέλτιστο πολυώνυμο $p_n^*(X_5 : x) = p_n^*(X_{5,5} : x)$ όπου $X_{5,5} = \{-\frac{1}{2}, 0, \frac{1}{2}, 1\}$ με τη μέθοδο προσδιοριστέων συντελεστών. Ο πίνακας τιμών της συνάρτησης για το $X_{5,5}$ είνα:

$$\begin{array}{c|cccc} x_i & -\frac{1}{2} & 0 & \frac{1}{2} & 1 \\ \hline f_i & \frac{1}{2} & 0 & \frac{1}{2} & 1 \end{array}.$$

Το 4×4 γραμμικό σύστημα είναι

$$\begin{bmatrix} -1 & 1 & -\frac{1}{2} & \frac{1}{4} \\ 1 & 1 & 0 & 0 \\ -1 & 1 & \frac{1}{2} & \frac{1}{4} \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \lambda \\ \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \\ 1 \end{bmatrix}.$$

Λύνουμε το σύστημα με τη μέθοδο απαλοιφής του Gauss

$$\begin{array}{cccccccc}
 & \underline{-1} & 1 & -\frac{1}{2} & \frac{1}{4} & & \frac{1}{2} \\
-1 & 1 & 1 & 0 & 0 & & 0 \\
1 & -1 & 1 & \frac{1}{2} & \frac{1}{4} & & \frac{1}{2} \\
-1 & 1 & 1 & 1 & 1 & & 1 \\
 & & & & & & \\
 & & \underline{2} & -\frac{1}{2} & \frac{1}{4} & & \frac{1}{2} \\
0 & & 0 & 1 & 0 & & 0 \\
1 & & 2 & \frac{1}{2} & \frac{5}{4} & & \frac{3}{2} \\
 & & & & & & \\
 & & & \underline{1} & 0 & & 0 \\
1 & & & 1 & 1 & & 1 \\
 & & & & & & \\
 & & & & \underline{1} & & 1
\end{array}.$$

$$\alpha^T = \begin{bmatrix} -\frac{1}{8} & \frac{1}{8} & 0 & 1 \end{bmatrix}$$

Βρήκαμε λοιπόν κι εδώ $\lambda = -\frac{1}{8}$ και $p_n^*(X_5 : x) = p_n^*(X_{5,5} : x) = x^2 + \frac{1}{8}$.

Η παραπάνω διαδικασία είναι αρκετά δαπανηρή σε χρόνο, αφού πρέπει να υπολογιστούν όλα τα λ για όλα τα σύνολα αποτελούμενα από $n+2$ σημεία, που είναι $\binom{m}{n+2}$ σε αριθμό. Ο αριθμός αυτός στη πράξη είναι πάρα πολύ μεγάλος, διότι συνήθως ο m είναι πολύ μεγαλύτερος του n, με αποτέλεσμα η παραπάνω διαδικασία να καθίσταται αναποτελεσματική. Για την αντιμετώπιση αυτού του προβλήματος επινοήθηκε ένας αλγόριθμος, όπου μεθοδικά μεταβαίνουμε από σύνολο σε σύνολο σημείων και που θα περιγραφεί στη συνέχεια.

2.3.2 Αλγόριθμος εναλλαγής σημείων για την εύρεση της βέλτιστης ομοιόμορφης προσέγγισης στο σύνολο X_m

Όπως έχουμε αναφέρει στην προηγούμενη παράγραφο, η εύρεση της ομοιόμορφης βέλτιστης προσέγγισης σε ένα σύνολο X_m ανάγεται στην εύρεση εκείνου του υποσυνόλου $X_{m,i}$, αποτελούμενο από $n+2$ σημεία, που δίνει το μεγαλύτερο σφάλμα. Ακόμη, για την εύρεση της βέλτιστης προσέγγισης συναρτήσεων σε ένα διάστημα $[a,b]$, στην πράξη θεωρούμε μια ομοιόμορφη διαμέριση X_m, m σημείων στο διάστημα αυτό, και προσεγγίζουμε τη βέλτιστη προσέγγιση στο $[a,b]$ με τη βέλτιστη προσέγγιση στο X_m. Όσο μεγαλύτερο είναι το m, τόσο καλύτερα προσεγγίζουμε τη βέλτιστη προσέγγιση στο $[a,b]$. Ο Remez ανέπτυξε έναν αλγόριθμο για την εύρεση της βέλτιστης ομοιόμορφης προσέγγισης στο διάστημα $[a,b]$, με όση ακρίβεια θέλουμε. Το πόσο μεγάλο θα είναι το m εξαρτάται από το μέτρο συνέχειας της f. Έχουν αποδειχτεί διάφορες προτάσεις για διάφορες κλάσεις συναρτήσεων.

Θα ασχοληθούμε όμως εδώ με το πως θα μπορέσουμε να βρούμε το υποσύνολο $X_{m,i}$, που δίνει τη βέλτιστη προσέγγιση, από το σύνολο X_m, χωρίς να χρειάζεται να υπολογίσουμε όλα τα $\lambda^{(i)}$, $\quad i = 1,2,\ldots,\begin{pmatrix} m \\ n+2 \end{pmatrix}$. Θα δώσουμε έναν αλγόριθμο μετάβασης από ένα υποσύνολο σημείων σε άλλο, με γνώμονα την αύξηση του σφάλματος.

Έστω ότι ξεκινάμε με ένα σύνολο $n+2$ σημείων του X_m που θα το συμβολίζουμε με $\{x_\sigma\}$ και θα το καλούμε σύνολο αναφοράς. Τα σύνολα αναφοράς τα συμβολίζουμε με ελληνικούς δείκτες, όπως $\{x_\sigma\}$, ενώ με x_σ θα συμβολίζουμε κάθε μια από τις τιμές x_i που ανήκει στο $\{x_\sigma\}$. Για απλότητα δηλαδή δε θα αλλάζουμε το δείκτη. Για παράδειγμα με το $\sum_\sigma f(x_\sigma)$ θα εννοούμε το άθροισμα όλων των $f(x)$ όπου $x \in \{x_\sigma\}$. Έστω $p_\sigma = p_n^*(f;\{x_\sigma\})$ η βέλτιστη προσέγγιση της f στο $\{x_\sigma\}$, τότε

$$|f(x_\sigma) - p_\sigma(x_\sigma)| = E_n(f;\{x_\sigma\}) = \rho_\sigma.$$

Έστω επίσης ότι

$$M = \max_{x\in X_m} |f(x) - p_\sigma(x)| = |f(x_j) - p_\sigma(x_j)|.$$

Αν $M = \rho_\sigma$ τότε το σύνολο αναφοράς $\{x_\sigma\}$ αποτελεί εναλλασσόμενο σύνολο σημείων για το X_m που σημαίνει ότι p_σ είναι η βέλτιστη προσέγγιση και για το X_m. Στην περίπτωση αυτή βρέθηκε η βέλτιστη προσέγγιση και τελειώνει ο αλγόριθμος. Αν $M > \rho_\sigma$ τότε $x_j \notin \{x_\sigma\}$ και θα πρέπει να βρούμε ένα άλλο σύνολο αναφοράς $\{x_\mu\}$. Το $\{x_\mu\}$ θα αποτελείται από το x_j που θα αντικαταστήσει ένα σημείο του $\{x_\sigma\}$ και από όλα τα υπόλοιπα σημεία του $\{x_\sigma\}$. Η αντικατάσταση θα γίνεται με τρόπο ώστε η συνάρτηση $e(x) = f(x) - p_\sigma(x)$ να εξακολουθεί να εναλλάσσει πρόσημο όταν το x κινείται στο σύνολο $\{x_\mu\}$. Άν το x_j

είναι ενδιάμεσο σημείο τότε θα αντικαταστήσει ένα από τα διπλανά σημεία που έχει το ίδιο πρόσημο με το $f(x) - p_\sigma(x)$. Άν το x_j είναι ακραίο σημείο τότε θα αντικαταστήσει το διπλανό του αν έχει το ίδιο πρόσημο ή το άλλο άκρο αν το πρόσημο με το διπλανό του είναι διαφορετικό. Υπολογίζουμε στη συνέχεια τη βέλτιστη προσέγγιση p_μ και το αντίστοιχο σφάλμα ρ_μ. Θα αποδείξουμε ότι

$$\rho_\mu > \rho_\sigma,$$

που σημαίνει ότι είμαστε σε καλό δρόμο αφού μεταβαίνουμε από μικρότερα σε μεγαλύτερα σφάλματα.

Από τη σχέση που μας δίνει το λ, έχουμε ότι

$$\begin{aligned}\lambda &= \frac{\sum_{i=1}^{n+2} \frac{f(x_i)}{\omega'(x_i)}}{\sum_{i=1}^{n+2} \frac{(-1)^i}{\omega'(x_i)}} = \sum_{i=1}^{n+2} \frac{\frac{1}{\omega'(x_i)}}{\sum_{j=1}^{n+2} \frac{(-1)^j}{\omega'(x_j)}} f(x_i) = \sum_{i=1}^{n+2} \frac{\frac{1}{(-1)^{n-i}|\omega'(x_i)|}}{\sum_{j=1}^{n+2} \frac{(-1)^j}{(-1)^{n-j}|\omega'(x_j)|}} f(x_i) \\ &= \sum_{i=1}^{n+2} \frac{\frac{1}{|\omega'(x_i)|}}{\sum_{j=1}^{n+2} \frac{1}{|\omega'(x_j)|}} (-1)^i f(x_i).\end{aligned}$$

Ονομάζουμε με

$$\lambda_i = \frac{\frac{1}{|\omega'(x_i)|}}{\sum_{j=1}^{n+2} \frac{1}{|\omega'(x_j)|}} > 0,$$

τότε το λ δίνεται από

$$\lambda = \sum_{i=1}^{n+2} \lambda_i (-1)^i f(x_i), \quad \text{με} \quad \sum_{i=1}^{n+2} \lambda_i = 1, \tag{2.19}$$

που σημαίνει ότι το σφάλμα λ της βέλτιστης προσέγγισης της συνάρτησης f, είναι ένας κυρτός συνδυασμός των τιμών $(-1)^i f(x_i)$ με συντελεστές (ή βάρη) λ_i. Ας σημειωθεί εδώ ότι οι συντελεστές λ_i εξαρτώνται μόνο από τις τιμές των συναρτήσεων ω' στα σημεία x_i και όχι από τις τιμές της συνάρτησης f, που σημαίνει ότι είναι σταθεροί εφόσον το σύνολο σημείων παραμένει το ίδιο. Αν στην παραπάνω σχέση αντί της f θέσουμε ένα πολυώνυμο $p \in P_n$, τότε η ποσότητα αυτή θα είναι μηδέν, δηλαδή

$$\sum_{i=1}^{n+2} \lambda_i (-1)^i p(x_i) = 0, \tag{2.20}$$

αφού θα αντικατοπτρίζει το σφάλμα της βέλτιστης προσέγγισης ενός πολυωνύμου $p \in P_n$ στον P_n. Η βέλτιστη προσέγγιση θα είναι το ίδιο το πολυώνυμο και επομένως το σφάλμα θα είναι μηδέν.

Στη συνέχεια θεωρούμε ως σύνολο αναφοράς το $\{x_\mu\}$ που προέκυψε από το $\{x_\sigma\}$ αφού έγινε η αλλαγή του σημείου. Αν εφαρμόσουμε τη (2.19) για την f το λ θα είναι το απόλυτο σφάλμα ρ_μ ή το αντίθετό του $-\rho_\mu$, επομένως

$$\rho_\mu = \left| \sum_\mu \lambda_\mu (-1)^\mu f(x_\mu) \right| = \left| \sum_\mu \lambda_\mu (-1)^\mu (f(x_\mu) - p_\sigma(x_\mu)) \right|, \tag{2.21}$$

όπου αφαιρέσαμε την ποσότητα $\left|\sum_\mu \lambda_\mu(-1)^\mu p_\sigma(x_\mu)\right|$ που είναι μηδέν λόγω της (2.20). Υπενθυμίζουμε ότι η αλλαγή σημείου έγινε κατά τρόπο ώστε το σφάλμα $f(x_\mu) - p_\sigma(x_\mu)$ να συνεχίσει να αλλάζει συνεχώς πρόσημο στο σύνολο αναφοράς $\{x_\mu\}$. Αυτό σημαίνει ότι η ποσότητα $(-1)^\mu(f(x_\mu) - p_\sigma(x_\mu))$ διατηρεί σταθερό πρόσημο καθώς το x_μ διατρέχει το σύνολο αναφοράς $\{x_\mu\}$. Επομένως το παραπάνω άθροισμα εντός των απόλυτων τιμών, είναι άθροισμα ομόσημων όρων αφού τα λ_μ είναι θετικά. Μετά από αυτή την παρατήρηση, η (2.21) γράφεται ως

$$\rho_\mu = \sum_\mu \lambda_\mu |f(x_\mu) - p_\sigma(x_\mu)|. \tag{2.22}$$

Επειδή τα σημεία του $\{x_\mu\}$ είναι και σημεία του $\{x_\sigma\}$ εκτός από ένα, έστω το x_j, ισχύει ότι

$$|f(x_\mu) - p_\sigma(x_\mu)| = \rho_\sigma, \quad \mu \neq j \ \text{ και } \ |f(x_j) - p_\sigma(x_j)| = M_\sigma.$$

Σύμφωνα με αυτά η (2.22) καταλήγει στην

$$\rho_\mu = \sum_{\mu\neq j} \lambda_\mu \rho_\sigma + \lambda_j M_\sigma = \sum_\mu \lambda_\mu \rho_\sigma + \lambda_j(M_\sigma - \rho_\sigma) = \rho_\sigma + \lambda_j(M_\sigma - \rho_\sigma) > \rho_\sigma,$$

αφού $\sum_\mu \lambda_\mu = 1$ και $M_\sigma > \rho_\sigma$. Η τελευταία σχέση απέδειξε το ζητούμενο.

Σύμφωνα με την παραπάνω ανάλυση, ο αλγόριθμος θα συγκλίνει σίγουρα στη βέλτιστη προσέγγιση αφού μεταβαίνουμε σε όλο και μεγαλύτερα σφάλματα, δε γνωρίζουμε όμως εκ των προτέρων των αριθμό των επαναλήψεων. Σίγουρα είναι ταχύτερος από εκείνον της διαδικασίας ελέγχου όλων των σφαλμάτων για όλα τα υποσύνολα που είναι σε πλήθος $\begin{pmatrix} m \\ n+2 \end{pmatrix}$. Έχουμε να παρατηρήσουμε ότι αν τύχει να ξεκινήσουμε από σύνολο αναφοράς που αντιστοιχεί σε σχετικά μεγάλο σφάλμα, τότε σύντομα θα φτάσουμε στη λύση. Αν όμως αυτό αντιστοιχεί σε σχετικά μικρό, τότε ίσως χρειαστούν πολλές επαναλήψεις. Γεννιέται λοιπόν το ερώτημα αν υπάρχει τρόπος να επιλέξουμε το καλύτερο δυνατό αρχικό σύνολο αναφοράς. Η απάντηση είναι θετική και σχετίζεται με την προσέγγιση ελαχίστων τετραγώνων που θα αναπτυχθεί στο επόμενο κεφάλαιο. Έχει αποδειχτεί ότι αν $q^* \in P_n$ είναι η προσέγγιση ελαχίστων τετραγώνων της f, ορισμένης στο X_m, στον P_n, τότε το σύνολο αναφοράς $\{x_\sigma\}$ για το οποίο η συνάρτηση σφάλμα $f(x_\sigma) - q^*(x_\sigma)$ εναλλάσσει πρόσημα, είναι ένα πολύ καλό αρχικό σύνολο αναφοράς για την εύρεση της βέλτιστης προσέγγισης με τον αλγόριθμο εναλλαγής σημείων. Απαιτείται βέβαια η εύρεση πρώτα της προσέγγισης ελαχίστων τετραγώνων, που όπως θα δούμε αποτελεί μια ευκολότερη και ταχύτερη διαδικασία. Δίνουμε στη συνέχεια δυο παραδείγματα:

Παράδειγμα 15 Δίνεται το σύνολο σημείων $X_5 = \{-2, -1, 0, 1, 2\}$. Να βρεθεί η βέλτιστη προσέγγιση δευτέρου βαθμού για τη συνάρτηση $f(x) = x^3$ με τον αλγόριθμο εναλλαγής σημείων.

Ο πίνακας τιμών της συνάρτησης είναι

x_i	-2	-1	0	1	2
f_i	-8	-1	0	1	8

.

Επιλέγουμε ως αρχικό σύνολο αναφοράς $\{x_\sigma\} = \{-2,-1,0,1\}$, αποτελούμενο από τα τέσσερα πρώτα σημεία. Προχωρούμε στη συνέχεια με τη γνωστή διαδικασία για την εύρεση του βέλτιστου πολυωνύμου.

$$\begin{array}{lll|lll} \omega'(-2) = & (-2+1)(-2)(-2-1) & =-6 & \omega'(0) = & 2\cdot 1(-1) & =-2 \\ \omega'(-1) = & (-1+2)(-1)(-1-1) & =2 & \omega'(1) = & (1+2)(1+1)1 & =6 \end{array}$$

$$\lambda_\sigma = \frac{\sum_{i=1}^4 \frac{f(x_i)}{\omega'(x_i)}}{\sum_{i=1}^4 \frac{(-1)^i}{\omega'(x_i)}} = \frac{\frac{-8}{-6}+\frac{-1}{2}+0+\frac{1}{6}}{\frac{-1}{-6}+\frac{1}{2}+\frac{-1}{-2}+\frac{1}{6}} = \frac{1}{\frac{4}{3}} = \frac{3}{4}.$$

$$\begin{array}{lll} p_\sigma(x) & = & p_3(x)-\lambda_\sigma l_3(x) = \sum_{i=1}^4 \frac{\omega(x)}{x-x_i}\frac{f(x_i)}{\omega'(x_i)} - \lambda_\sigma \sum_{i=1}^4 \frac{\omega(x)}{x-x_i}\frac{(-1)^i}{\omega'(x_i)} \\ & = & (x+1)x(x-1)\frac{-8}{-6}+(x+2)x(x-1)\frac{-1}{2}+(x+2)(x+1)x\frac{1}{6} \\ & - & \frac{3}{4}\Big((x+1)x(x-1)\frac{-1}{-6}+(x+2)x(x-1)\frac{1}{2}+(x+2)(x+1)(x-1)\frac{-1}{-2} \\ & + & (x+2)(x+1)x\frac{1}{6}\Big) = -\frac{3}{2}x^2+x+\frac{3}{4}. \end{array}$$

Υπολογίζουμε τώρα τα σφάλματα $e_\sigma(x_i)$ για όλα τα x_i που ανήκουν στο X_m, και ελέγχουμε τα πρόσημά των ώστε να γίνει η αλλαγή των σημείων.

$$\begin{array}{rclcrcl} e_\sigma(-2) & = & (-2)^3-(-\frac{3}{2}(-2)^2+(-2)+\frac{3}{4}) & = & -\frac{3}{4} & (-) & \\ e_\sigma(-1) & = & (-1)^3-(-\frac{3}{2}(-1)^2+(-1)+\frac{3}{4}) & = & \frac{3}{4} & (+) & \\ e_\sigma(0) & = & 0^3-(-\frac{3}{2}0^2+0+\frac{3}{4}) & = & -\frac{3}{4} & (-) & \\ e_\sigma(1) & = & 1^3-(-\frac{3}{2}1^2+1+\frac{3}{4}) & = & \frac{3}{4} & (+) & \text{έξοδος} \\ e_\sigma(2) & = & 2^3-(-\frac{3}{2}2^2+2+\frac{3}{4}) & = & \frac{45}{4} & (+) & \text{είσοδος} \end{array}$$

Πρέπει να παρατηρήσουμε εδώ ότι δε χρειαζόταν να υπολογιστούν τα σφάλματα για τα σημεία του $\{x_\sigma\}$. Είναι γνωστό από τη θεωρία ότι αυτά θα είναι ίσα με $\rho_\sigma = \frac{3}{4}$ με εναλλασσόμενα πρόσημα, ξεκινώντας με $-\lambda_\sigma = -\frac{3}{4}$. Τα υπολογίσαμε απλώς για να επιβεβαιωθεί η θεωρία. Το μέγιστο σφάλμα δίνεται στο επιπλέον σημείο 2 και είναι $M = \frac{45}{4}$ με θετικό πρόσημο. Θα εισαχθεί αυτό το σημείο σε αντικατάσταση του διπλανού του 1 που, το αντίστοιχο σφάλμα έχει το ίδιο πρόσημο. Έτσι το νέο σύνολο αναφοράς θα είναι το $\{x_\mu\} = \{-2,-1,0,2\}$. Στη συνέχεια ακολουθούμε την ίδια διαδικασία:

$$\begin{array}{lll|lll} \omega'(-2) = & (-2+1)(-2)(-2-2) & =-8 & \omega'(0) = & 2\cdot 1(-2) & =-4 \\ \omega'(-1) = & (-1+2)(-1)(-1-2) & =3 & \omega'(2) = & (2+2)(2+1)2 & =24 \end{array}$$

$$\lambda_\mu = \frac{\sum_{i=1}^4 \frac{f(x_i)}{\omega'(x_i)}}{\sum_{i=1}^4 \frac{(-1)^i}{\omega'(x_i)}} = \frac{\frac{-8}{-8}+\frac{-1}{3}+0+\frac{8}{24}}{\frac{-1}{-8}+\frac{1}{3}+\frac{-1}{-4}+\frac{1}{24}} = \frac{1}{\frac{3}{4}} = \frac{4}{3}.$$

$$\begin{array}{lll} p_\mu(x) & = & p_3(x)-\lambda_\mu l_3(x) = \sum_{i=1}^4 \frac{\omega(x)}{x-x_i}\frac{f(x_i)}{\omega'(x_i)} - \lambda_\mu \sum_{i=1}^4 \frac{\omega(x)}{x-x_i}\frac{(-1)^i}{\omega'(x_i)} \\ & = & (x+1)x(x-2)\frac{-8}{-8}+(x+2)x(x-2)\frac{-1}{3}+(x+2)(x+1)x\frac{8}{24} \\ & - & \frac{3}{4}\Big((x+1)x(x-2)\frac{-1}{-8}+(x+2)x(x-2)\frac{1}{3}+(x+2)(x+1)(x-2)\frac{-1}{-4} \\ & + & (x+2)(x+1)x\frac{1}{24}\Big) = -\frac{1}{3}x^2+\frac{10}{3}x+\frac{4}{3}. \end{array}$$

Δίνουμε τώρα τον πίνακα των σφαλμάτων $e_\mu(x_i)$ για όλα τα x_i που ανήκουν στο X_m, υπολογίζοντας μόνο το σφάλμα για το επιπλέον σημείο 1, αφού για τα σημεία του $\{x_\mu\}$ θα είναι $\frac{4}{3}$ με εναλλασσόμενα πρόσημα, ξεκινώντας με αρνητικό.

$$\begin{array}{lclcll}
e_\mu(-2) & & & = & -\frac{4}{3} & (-) & \\
e_\mu(-1) & & & = & \frac{4}{3} & (+) & \\
e_\mu(0) & & & = & -\frac{4}{3} & (-) & \text{έξοδος} \\
e_\mu(1) & = & 1^3-(-\frac{1}{3}1^2+\frac{10}{3}1+\frac{4}{3}) & = & -\frac{10}{3} & (-) & \text{είσοδος} \\
e_\mu(2) & & & = & \frac{4}{3} & (+) &
\end{array}$$

Φαίνεται καθαρά ότι το μέγιστο σφάλμα δίνεται στο σημείο 1 και είναι $M=\frac{10}{3}$ με αρνητικό πρόσημο. Θα εισαχθεί ξανά αυτό το σημείο σε αντικατάσταση του διπλανού του 0 που, το αντίστοιχο σφάλμα έχει το ίδιο πρόσημο. Το νέο σύνολο αναφοράς θα είναι το $\{x_\nu\}=\{-2,-1,1,2\}$ και συνεχίζεται η ίδια διαδικασία

$$\begin{array}{lll|lll}
\omega'(-2)= & (-2+1)(-2-1)(-2-2) & =-12 & \omega'(1)= & (1+2)(1+1)(1-2) & =-6 \\
\omega'(-1)= & (-1+2)(-1-1)(-1-2) & =6 & \omega'(2)= & (2+2)(2+1)(2-1) & =12
\end{array}$$

$$\lambda_\nu=\frac{\sum_{i=1}^4\frac{f(x_i)}{\omega'(x_i)}}{\sum_{i=1}^4\frac{(-1)^i}{\omega'(x_i)}}=\frac{\frac{-8}{-12}+\frac{-1}{6}+\frac{1}{-6}+\frac{8}{12}}{\frac{-1}{-12}+\frac{1}{6}+\frac{-1}{-6}+\frac{1}{12}}=\frac{1}{\frac{1}{2}}=2.$$

$$\begin{array}{rcl}
p_\nu(x) & = & p_3(x)-\lambda_\nu l_3(x)=\sum_{i=1}^4\frac{\omega(x)}{x-x_i}\frac{f(x_i)}{\omega'(x_i)}-\lambda_\nu\sum_{i=1}^4\frac{\omega(x)}{x-x_i}\frac{(-1)^i}{\omega'(x_i)} \\
& = & (x+1)(x-1)(x-2)\frac{-8}{-12}+(x+2)(x-1)(x-2)\frac{-1}{6} \\
& + & (x+2)(x+1)(x-2)\frac{1}{-6}+(x+2)(x+1)(x-1)\frac{8}{12} \\
& - & 2\Big((x+1)(x-1)(x-2)\frac{-1}{-12}+(x+2)(x-1)(x-2)\frac{1}{6} \\
& + & (x+2)(x+1)(x-2)\frac{-1}{-6}+(x+2)(x+1)(x-1)\frac{1}{12}\Big)=3x.
\end{array}$$

Ο πίνακα των σφαλμάτων $e_\nu(x_i)$ είναι:

$$\begin{array}{lclcrl}
e_\nu(-2) & & & = & -2 & (-) \\
e_\nu(-1) & & & = & 2 & (+) \\
e_\nu(0) & = & 0^3-3\times 0 & = & 0 & \\
e_\nu(1) & & & = & -2 & (-) \\
e_\nu(2) & & & = & 2 & (+)
\end{array}$$

Το μέγιστο σφάλμα είναι $M=\rho_\nu=2$, που σημαίνει ότι ο αλγόριθμος ολοκληρώθηκε. Το σύνολο αναφοράς $\{x_\nu\}$ θα είναι το σύνολο εναλλασσόμενων σημείων για το X_m που μας δίνει τη βέλτιστη ομοιόμορφη προσέγγιση

$$p_2^*(x)=3x.$$

Παρατηρούμε εδώ ότι το πολυώνυμο αυτό είναι πολυώνυμο πρώτου βαθμού, που σημαίνει ότι είναι και βέλτιστη προσέγγιση στον P_1. Ισχύει και εδώ η παρατήρηση που κάναμε και στην προσέγγιση συνεχών συναρτήσεων σε διάστημα καθώς και στη διακριτή περίπτωση, όπως και στο προηγούμενο παράδειγμα. Επειδή η συνάρτηση x^3 είναι περιττή και το σύνολο σημείων $X_5 = \{-2, -1, 0, 1, 2\}$ είναι συμμετρικό ως προς το μηδέν, η βέλτιστη προσέγγιση θα είναι περιττή συνάρτηση.

Θα δώσουμε εδώ ένα παράδειγμα όπου η συνάρτηση και το σύνολο σημείων είναι τα ίδια όπως προηγούμενα, αλλάζει όμως ο χώρος πολυωνύμων που θα είναι ο P_1. Φυσικά αναμένεται το ίδιο αποτέλεσμα $3x$, αλλά θα δούμε το πως λειτουργεί ο αλγόριθμος όταν παραμένουν δυο σημεία έξω από το σύνολο αναφοράς.

Παράδειγμα 16 Δίνεται το σύνολο σημείων $X_5 = \{-2, -1, 0, 1, 2\}$. Να βρεθεί τη βέλτιστη προσέγγιση πρώτου βαθμού για τη συνάρτηση $f(x) = x^3$ με τον αλγόριθμο εναλλαγής σημείων.

Ο πίνακας τιμών της συνάρτησης είναι

x_i	-2	-1	0	1	2
f_i	-8	-1	0	1	8

Επιλέγουμε ως αρχικό σύνολο αναφοράς $\{x_\sigma\} = \{-2, -1, 0\}$, αποτελούμενο από τα τρία πρώτα σημεία και ακολουθούμε τη γνωστή διαδικασία.

$$\omega'(-2) = (-2+1)(-2) = 2, \quad \omega'(-1) = (-1+2)(-1) = -1, \quad \omega'(0) = 2 \cdot 1 = 2$$

$$\lambda_\sigma = \frac{\sum_{i=1}^{3} \frac{f(x_i)}{\omega'(x_i)}}{\sum_{i=1}^{3} \frac{(-1)^i}{\omega'(x_i)}} = \frac{\frac{-8}{2} + \frac{-1}{-1} + 0}{\frac{-1}{2} + \frac{1}{-1} + \frac{1}{2}} = \frac{-3}{-2} = \frac{3}{2}.$$

$$\begin{aligned} p_\sigma(x) &= p_2(x) - \lambda_\sigma l_2(x) = \textstyle\sum_{i=1}^{3} \frac{\omega(x)}{x-x_i} \frac{f(x_i)}{\omega'(x_i)} - \lambda_\sigma \sum_{i=1}^{3} \frac{\omega(x)}{x-x_i} \frac{(-1)^i}{\omega'(x_i)} \\ &= (x+1)x\tfrac{-8}{2} + (x+2)x\tfrac{-1}{-1} - \tfrac{3}{2}\left((x+1)x\tfrac{-1}{2} + (x+2)x\tfrac{1}{-1} + (x+2)(x+1)\tfrac{-1}{2}\right) \\ &= 4x + \tfrac{3}{2}. \end{aligned}$$

Ο πίνακας των σφαλμάτων $e_\sigma(x_i)$ θα είναι:

$$\begin{array}{rcccll} e_\sigma(-2) &=& -\lambda_\sigma &=& -\frac{3}{2} & (-) \\ e_\sigma(-1) &=& \lambda_\sigma &=& \frac{3}{2} & (+) \\ e_\sigma(0) &=& -\lambda_\sigma &=& -\frac{3}{2} & (-) \quad \text{έξοδος} \\ e_\sigma(1) &=& 1^3 - (4 \cdot 1 + \frac{3}{2}) &=& -\frac{9}{2} & (-) \quad \text{είσοδος} \\ e_\sigma(2) &=& 2^3 - (4 \cdot 2 + \frac{3}{2}) &=& -\frac{3}{2} & (-) \end{array}$$

Το μέγιστο σφάλμα δίνεται στο σημείο 1 και είναι $M = \frac{9}{2}$ με αρνητικό πρόσημο. Θα εισαχθεί αυτό το σημείο σε αντικατάσταση του διπλανού του 0 που, το αντίστοιχο σφάλμα έχει το ίδιο πρόσημο. Έτσι το νέο σύνολο αναφοράς θα είναι το $\{x_\mu\} = \{-2, -1, 1\}$.

$$\omega'(-2) = (-2+1)(-2-1) = 3, \; \omega'(-1) = (-1+2)(-1-1) = -2, \; \omega'(1) = (1+2)(1+1) = 6$$

$$\lambda_\mu = \frac{\sum_{i=1}^{3} \frac{f(x_i)}{\omega'(x_i)}}{\sum_{i=1}^{3} \frac{(-1)^i}{\omega'(x_i)}} = \frac{\frac{-8}{3} + \frac{-1}{-2} + \frac{1}{6}}{\frac{-1}{3} + \frac{1}{-2} + \frac{-1}{6}} = \frac{-2}{-1} = 2.$$

$$\begin{aligned} p_\mu(x) &= p_2(x) - \lambda_\mu l_2(x) = \sum_{i=1}^{3} \frac{\omega(x)}{x-x_i} \frac{f(x_i)}{\omega'(x_i)} - \lambda_\mu \sum_{i=1}^{3} \frac{\omega(x)}{x-x_i} \frac{(-1)^i}{\omega'(x_i)} \\ &= (x+1)(x-1)\tfrac{-8}{3} + (x+2)(x-1)\tfrac{-1}{-2} + (x+2)(x+1)\tfrac{1}{6} \\ &- 2\left((x+1)(x-1)\tfrac{-1}{3} + (x+2)(x-1)\tfrac{1}{-2} + (x+2)(x+1)\tfrac{-1}{6}\right) = 3x. \end{aligned}$$

Ο πίνακας των σφαλμάτων $e_\mu(x_i)$ θα είναι:

$$\begin{array}{rcccrc} e_\mu(-2) &=& -\lambda_\mu &=& -2 & (-) \\ e_\mu(-1) &=& \lambda_\mu &=& 2 & (+) \\ e_\mu(0) &=& 0^3 - 3\cdot 0 &=& 0 & (-) \\ e_\mu(1) &=& -\lambda_\mu &=& -2 & (-) \\ e_\mu(2) &=& 2^3 - 3\cdot 2 &=& 2 & (+) \end{array}$$

Το μέγιστο σφάλμα είναι $M = \rho_\mu = 2$ και φτάσαμε στο αναμενόμενο αποτέλεσμα:

$$p_1^*(x) = 3x.$$

Προφανώς ο πίνακας των σφαλμάτων είναι ακριβώς ο ίδιος με τον τελευταίο πίνακα του προηγουμένου παραδείγματος και το σύνολο των εναλλασσόμενων σημείων αποτελείται από τέσσερα σημεία αφού είναι και βέλτιστη προσέγγιση στον P_2.

Θα δώσουμε εδώ έναν άλλο τρόπο υπολογισμού της βέλτιστης προσέγγισης, χωρίς την εύρεση των αντίστοιχων πολυωνύμων παρεμβολής p_2 και l_2, όταν πρόκειται για προσέγγιση στον P_1. Είναι ένας γεωμετρικός τρόπος που βασίζεται στο γεγονός ότι η ευθεία που ενώνει τα μέσα των δύο πλευρών ενός τριγώνου, απέχει από τις τρεις κορυφές του εξίσου και αφήνει δυο κορυφές στο ένα ημιεπίπεδο και μια στο άλλο. Απομένει λοιπόν να υπολογίσουμε την εξίσωση της ευθείας που ενώνει τα μέσα δυο πλευρών του τριγώνου, ώστε οι αποστάσεις από τις κορυφές να έχουν εναλλασσόμενο πρόσημο. Έστω ότι έχουμε το σύνολο σημείων $X_3 = \{x_1, x_2, x_3\}$, με $x_1 < x_2 < x_3$ και $f(x_1), f(x_2), f(x_3)$ είναι οι τιμές της συνάρτησης. Ορίζεται ένα τρίγωνο με κορυφές τα σημεία $(x_1, f(x_1))$, $(x_2, f(x_2))$ και $(x_3, f(x_3))$. Είναι προφανές ότι η ευθεία που ενώνει τα μέσα των πλευρών του τριγώνου που ορίζεται από τα δυο πρώτα και τα δυο δεύτερα σημεία, απέχει εξίσου από τις κορυφές με εναλλασσόμενο πρόσημο. Τα μέσα αυτά έχουν συντεταγμένες $\left(\frac{x_1+x_2}{2}, \frac{f(x_1)+f(x_2)}{2}\right)$ και $\left(\frac{x_2+x_3}{2}, \frac{f(x_2)+f(x_3)}{2}\right)$, αντίστοιχα. Κατασκευάζουμε στη συνέχεια την εξίσωση της ευθείας που ενώνει τα δυο αυτά σημεία:

$$y - \frac{f(x_1)+f(x_2)}{2} = \frac{\frac{f(x_2)+f(x_3)}{2} - \frac{f(x_1)+f(x_2)}{2}}{\frac{x_2+x_3}{2} - \frac{x_1+x_2}{2}} \left(x - \frac{x_1+x_2}{2}\right).$$

Λύνοντας ως προς y και θέτοντας $p_1^*(x)$ στη θέση του, παίρνουμε

$$p_1^*(x) = \frac{f(x_3)-f(x_1)}{x_3-x_1}\left(x - \frac{x_1+x_2}{2}\right) + \frac{f(x_1)+f(x_2)}{2}, \tag{2.23}$$

που αποτελεί έναν τύπο απευθείας εύρεσης του βέλτιστου πολυωνύμου.

Αν πάρουμε το σύνολο αναφοράς $\{x_\mu\} = \{-2, -1, 1\}$ του τελευταίου παραδείγματος για τη συνάρτηση x^3, ο τύπος αυτός δίνει:

$$p_1^*(x) = \frac{1-(-8)}{1-(-2)}\left(x - \frac{-2+(-1)}{2}\right) + \frac{-8+(-1)}{2} = 3x,$$

που είναι και το αναμενόμενο αποτέλεσμα.

Ασκήσεις

1. Να βρεθεί η βέλτιστη ομοιόμορφη προσέγγιση της συνάρτησης $f(x) = x + |x|$, ορισμένης στο διάστημα $[-1, 1]$, στον P_1. (Μπορείτε να χρησιμοποιήσετε τη γραφική παράσταση της συνάρτησης και ιδιότητες της βέλτιστης ομοιόμορφης προσέγγισης.)

2. Να βρεθεί η βέλτιστη ομοιόμορφη προσέγγιση της συνάρτησης $f(x) = \frac{1}{x}$, ορισμένης στο διάστημα $[1, 2]$, στον P_1 με τη μέθοδο προσδιοριστέων συντελεστών. (Θεωρήστε $p_1^*(x) = \alpha x + \beta$ και προσδιορίστε τα α και β. Να διατηρείτε κλάσματα και ριζικά στους υπολογισμούς.)

3. Να βρεθεί η βέλτιστη ομοιόμορφη προσέγγιση της $f \in X_4$, στον P_1, όπου

$$X_4 = \{-1, 0, 1, 2\} \quad \text{και} \quad \begin{array}{c|cccc} x_i & -1 & 0 & 1 & 2 \\ \hline f_i & 1 & 0 & 1 & 4 \end{array}.$$

4. Να βρεθεί η βέλτιστη ομοιόμορφη προσέγγιση της $f \in X_4$, στον P_1, όπου

$$X_4 = \{-2, 0, 1, 2\} \quad \text{και} \quad \begin{array}{c|cccc} x_i & -2 & 0 & 1 & 2 \\ \hline f_i & 0 & 2 & 0 & -2 \end{array}.$$

5. Να βρεθεί η βέλτιστη ομοιόμορφη προσέγγιση της $f \in X_5$, στον P_1, όπου

$$X_5 = \{-2, 1, 0, 1, 2\} \quad \text{και} \quad \begin{array}{c|ccccc} x_i & -2 & -1 & 0 & 1 & 2 \\ \hline f_i & 1 & -1 & 1 & 2 & 0 \end{array},$$

χρησιμοποιώντας τον αλγόριθμο της εναλλαγής σημείων και ξεκινώντας από $X_\sigma = \{-2, -1, 0\}$.

6. Να βρεθεί η βέλτιστη ομοιόμορφη προσέγγιση της $f \in X_4$, στον P_1 χρησιμοποιώντας τον αλγόριθμο της εναλλαγής σημείων, όπου

$$X_4 = \{0, 1, 2, 3\} \quad \text{και} \quad \begin{array}{c|cccc} x_i & 0 & 1 & 2 & 3 \\ \hline f_i & 1 & 1 & 2 & 4 \end{array}.$$

7. Να βρεθεί η βέλτιστη ομοιόμορφη προσέγγιση της $f \in X_5$, στον P_1, όπου
$$X_5 = \{-2, 1, 0, 1, 2\} \quad \text{και} \quad \begin{array}{c|ccccc} x_i & -2 & -1 & 0 & 1 & 2 \\ \hline f_i & 0 & 1 & -1 & 2 & -2 \end{array},$$
χρησιμοποιώντας τον αλγόριθμο της εναλλαγής σημείων και ξεκινώντας από $X_\sigma = \{-1, 0, 1\}$.

8. Να βρεθεί η βέλτιστη ομοιόμορφη προσέγγιση της συνάρτησης $f(x) = x^2$, ορισμένης στο διάστημα $[0, 1]$, στον P_1 χρησιμοποιώντας πολυώνυμα Chebyshev.

9. Να βρεθεί η βέλτιστη ομοιόμορφη προσέγγιση της $f \in X_4$, στον P_2 και στον P_1, όπου
$$X_4 = \{0, 1, 2, 3\} \quad \text{και} \quad \begin{array}{c|cccc} x_i & 0 & 1 & 2 & 3 \\ \hline f_i & -1 & 1 & -1 & 0 \end{array}.$$

10. Να βρεθεί η βέλτιστη προσέγγιση της συνάρτησης x^3, ορισμένης στο διάστημα $[a, b]$, στον P_3.

11. Ποιό από τα δυο πολυώνυμα $x^3 - \frac{1}{8}x$ και $x^2 - \frac{1}{8}$ δεν μπορεί να είναι η βέλτιστη προσέγγιση της συνάρτησης x^4, ορισμένης στο διάστημα $[-1, 1]$, στον P_3 και γιατί;

12. Να αποδειχτεί ότι η βέλτιστη ομοιόμορφη προσέγγιση της συνάρτησης $f(x) = \sin(x) + 3x^2 - 5x + 1$, ορισμένης στο διάστημα $[-2\pi, 2\pi]$, στον P_2, είναι $p_2^*(x) = 3x^2 - 5x + 1$.

13. Δίνεται ότι p_n^* είναι η βέλτιστη ομοιόμορφη προσέγγιση της συνάρτησης $f \in C[a, b]$ στον P_n. Να βρεθεί η βέλτιστη ομοιόμορφη προσέγγιση της συνάρτησης $f + q_n$ στον P_n, όπου $q_n \in P_n$.

Κεφάλαιο 3

ΠΡΟΣΕΓΓΙΣΗ ΕΛΑΧΙΣΤΩΝ ΤΕΤΡΑΓΩΝΩΝ

Το πρόβλημα της προσέγγισης ελαχίστων τετραγώνων για συνεχείς συναρτήσεις ορισμένες στο διάστημα $[-1,1]$ διατυπώνεται ως εξής:

Πρόβλημα 2 *Δοθείσας συνεχούς συνάρτησης f στο διάστημα $I=[-1,1]$ και μιας ολοκληρώσιμης κατά Riemann συνάρτησης βάρους w, θετικής στο I εκτός από πεπερασμένο σύνολο σημείων όπου $w(x)=0$, να βρεθεί πολυώνυμο $q_n^* \in P_n$ τέτοιο ώστε*

$$\|f-q_n^*\|_2 = \left[\int_{-1}^{1}[f(x)-q_n^*(x)]^2 w(x)dx\right]^{\frac{1}{2}} < \|f-p\|_2 \ \ \forall p \in P_n, \ \ p \neq q_n^*. \tag{3.1}$$

Το αντίστοιχο πρόβλημα για συναρτήσεις ορισμένες στο σύνολο σημείων $X_m=\{x_1,x_2,\ldots,x_m\}$ διατυπώνεται ως εξής:

Πρόβλημα 3 *Δοθείσας συνάρτησης f ορισμένης στο σύνολο σημείων $X_m=\{x_1,x_2,\ldots,x_m\}$ και μιας συνάρτησης βάρους w, θετικής στο X_m, να βρεθεί πολυώνυμο $q_n^* \in P_n$ τέτοιο ώστε*

$$\|f-q_n^*\|_2 = \left[\sum_{i=1}^{m}[f(x_i)-q_n^*(x_i)]^2 w(x_i)\right]^{\frac{1}{2}} < \|f-p\|_2 \ \ \forall p \in P_n, \ \ p \neq q_n^*. \tag{3.2}$$

Από τα γενικά θεωρήματα ύπαρξης της βέλτιστης προσέγγισης Θεώρημα 1 και μοναδικότητας Θεώρημα 3 της Εισαγωγής, προκύπτει ότι υπάρχει η προσέγγιση ελαχίστων τετραγώνων και είναι μοναδική, αφού ο χώρος P_n είναι πεπερασμένης διάστασης και η ευκλείδεια νόρμα $\|\cdot\|_2$ είναι αυστηρά κυρτή νόρμα. Από τη σχέση (3.1) μπορούμε να παρατηρήσουμε ότι η προσέγγιση ελαχίστων τετραγώνων υπάρχει και είναι μοναδική όχι μόνο για συνεχείς συναρτήσεις, αλλά και για συναρτήσεις που είναι ολοκληρώσιμες κατά Riemann σε σχέση με τη συνάρτηση βάρους w.

Η προσέγγιση ελαχίστων τετραγώνων μπορεί να ορισθεί κάλλιστα και σε οποιοδήποτε κλειστό διάστημα $[a,b]$, αρκεί να αντικαταστήσουμε με a και b τα -1 και 1, αντίστοιχα στον τύπο (3.1). Την ορίζουμε όμως στο διάστημα $I = [-1, 1]$ αφού κάθε συνάρτηση ορισμένη στο $[a, b]$ μπορεί να μετασχηματιστεί με το γραμμικό μετασχηματισμό

$$x = \frac{b-a}{2}t + \frac{b+a}{2}$$

στο διάστημα $[-1, 1]$. Βρίσκουμε τη συνάρτηση

$$g(t) = f\left(\frac{b-a}{2}t + \frac{b+a}{2}\right), \quad t \in [-1, 1],$$

προσδιορίζουμε την προσέγγιση ελαχίστων τετραγώνων της g και επανερχόμαστε στο διάστημα $[a, b]$, με τον αντίστροφο μετασχηματισμό

$$t = \frac{2}{b-a}x - \frac{b+a}{b-a}.$$

Προτιμούμε να εργασθούμε στο διάστημα $I = [-1, 1]$ επειδή αυτό διευκολύνει τους υπολογισμούς, όπως θα δούμε παρακάτω.

Απομένει λοιπόν να δούμε πως χαρακτηρίζεται η προσέγγιση ελαχίστων τετραγώνων και να δώσουμε αλγορίθμους για τον υπολογισμό της. Πριν προχωρήσουμε σε αυτό θα ορίσουμε το ευκλείδειο εσωτερικό γινόμενο που σχετίζεται μα την ευκλείδεια νόρμα $\|\cdot\|_2$ και θα δώσουμε κάποιες ιδιότητες αυτού που θα μας χρειαστούν παρακάτω.

Ορισμός 5 *Δοθέντων δυο συναρτήσεων f και g, συνεχών στο διάστημα $I = [-1, 1]$ και μιας συνάρτησης βάρους w, ορίζουμε ως εσωτερικό γινόμενο αυτών, την ποσότητα*

$$(f, g) = \int_{-1}^{1} f(x)g(x)w(x)dx. \tag{3.3}$$

Ο αντίστοιχος ορισμός για το εσωτερικό γινόμενο συναρτήσεων ορισμένων σε σύνολο σημείων, είναι.

Ορισμός 6 *Δοθέντων δυο συναρτήσεων f και g, ορισμένων στο σύνολο $X_m = \{x_1, x_2, \ldots, x_m\}$ και μιας συνάρτησης βάρους w, θετικής στο X_m, ορίζουμε ως εσωτερικό γινόμενο αυτών, την ποσότητα*

$$(f, g) = \sum_{i=1}^{m} f(x_i)g(x_i)w(x_i). \tag{3.4}$$

Πρέπει να παρατηρήσουμε εδώ ότι το ευκλείδειο εσωτερικό γινόμενο ορίζεται γενικότερα για μιγαδικές συναρτήσεις, εμείς όμως δώσαμε τον ορισμό που περιορίζεται στις πραγματικές συναρτήσεις επειδή με αυτές μόνο θα ασχοληθούμε. Θα δούμε εδώ ότι τα εσωτερικά γινόμενα, για τη συνεχή και διακριτή περίπτωση, που ορίσαμε έχουν τις ίδιες ακριβώς ιδιότητες. Ακόμη, θα δούμε παρακάτω ότι υπάρχει ένας θαυμάσιος δυϊσμός μεταξύ των προσεγγίσεων σε συνεχές διάστημα και σε σύνολο σημείων, σε βαθμό που οι αποδείξεις για τη μια περίπτωση είναι ακριβώς οι ίδιες και για την άλλη. Δίνουμε εδώ κάποιες ιδιότητες για το εσωτερικό γινόμενο:

(i) $(f, g) = (g, f)$
(ii) $(\lambda f, g) = (f, \lambda g) = \lambda(f, g), \quad \lambda \in I\!R.$
(iii) $(f + h, g) = (f, g) + (h, g), \quad (f, g + h) = (f, g) + (f, h)$ (γραμμικότητα).
(iv) $(f, f) = \|f\|_2^2.$

Θα λέμε ότι οι συναρτήσεις f και g είναι ορθογώνιες μεταξύ των αν $(f, g) = 0$. Παρατηρούμε ότι ο ορισμός αυτός της ορθογωνότητας συμπίπτει με την ορθογωνιότητα διανυσμάτων στον $I\!R^m$, όταν οι συναρτήσεις ορίζονται στο σύνολο σημείων X_m. Προχωρούμε, λοιπόν, στον χαρακτηρισμό της προσέγγισης ελαχίστων τετραγώνων.

3.1 Χαρακτηρισμός της προσέγγισης ελαχίστων τετραγώνων

Θεώρημα 14 *Έστω $f \in C(I)$ και w μια συνάρτηση βάρους ορισμένη στο I. Το πολυώνυμο q_n^* είναι η προσέγγιση ελαχίστων τετραγώνων της f, στο σύνολο πολυωνύμων P_n, ανν*

$$(f - q_n^*, p) = \int_{-1}^{1} (f(x) - q_n^*(x))p(x)w(x)dx = 0 \;\; \forall p \in P_n. \tag{3.5}$$

Απόδειξη: $(\Leftarrow)$ Έστω ότι ισχύει η σχέση (3.5) και ότι $p \in P_n$, τότε

$$\begin{aligned} \|f - p\|_2^2 &= \textstyle\int_{-1}^{1} (f(x) - p(x))^2 w(x)dx = (f - p, f - p) \\ &= ((f - q_n^*) + (q_n^* - p), (f - q_n^*) + (q_n^* - p)) \\ &= (f - q_n^*, f - q_n^*) + (q_n^* - p, q_n^* - p) + 2(f - q_n^*, q_n^* - p). \end{aligned}$$

Επειδή η σχέση (3.5) ισχύει για όλα τα $p \in P_n$ και επειδή $q_n^* - p \in P_n$ έχουμε ότι $(f - q_n^*, q_n^* - p) = 0$. Επομένως η παραπάνω σχέση γίνεται

$$\|f - p\|_2^2 = (f - q_n^*, f - q_n^*) + (q_n^* - p, q_n^* - p) = \|f - q_n^*\|_2^2 + \|q_n^* - p\|_2^2 \geq \|f - q_n^*\|_2^2,$$

όπου η ανισότητα γίνεται ισότητα μόνο όταν $q_n^* = p$. Αυτό σημαίνει ότι q_n^* είναι η προσέγγιση ελαχίστων τετραγώνων.

$(\Rightarrow)$ Έστω ότι q_n^* είναι η προσέγγιση ελαχίστων τετραγώνων της f και ότι δεν ισχύει η σχέση (3.5), δηλαδή υπάρχει $\bar{p} \in P_n$ για το οποίο ισχύει

$$(f - q_n^*, \bar{p}) = \int_{-1}^{1} (f(x) - q_n^*(x))\bar{p}(x)w(x)dx = a \neq 0.$$

Ορίζουμε με

$$b = (\bar{p}, \bar{p}) = \int_{-1}^{1} (\bar{p}(x))^2 w(x)dx > 0$$

και θέτουμε

$$\lambda = \frac{a}{b} \neq 0.$$

Στη συνέχεια θεωρούμε το πολυώνυμο $q = q_n^* + \lambda\bar{p} \in P_n$ και υπολογίζουμε τη $\|f - q\|_2^2$:

$$\begin{aligned}
\|f - q\|_2^2 &= \|f - q_n^* - \lambda\bar{p}\|_2^2 = (f - q_n^* - \lambda\bar{p}, f - q_n^* - \lambda\bar{p}) \\
&= (f - q_n^*, f - q_n^*) - 2(f - q_n^*, \lambda\bar{p}) + (\lambda\bar{p}, \lambda\bar{p}) \\
&= \|f - q_n^*\|_2^2 - 2\lambda(f - q_n^*, \bar{p}) + \lambda^2(\bar{p}, \bar{p}) = \|f - q_n^*\|_2^2 - 2\lambda a + \lambda^2 b \\
&= \|f - q_n^*\|_2^2 - \lambda^2 b < \|f - q_n^*\|_2^2.
\end{aligned}$$

Αυτό σημαίνει ότι το πολυώνυμο q αποτελεί καλύτερη προσέγγιση της f από ό,τι το q_n^*. Καταλήξαμε επομένως σε άτοπο αφού υποθέσαμε ότι το q_n^* είναι η προσέγγιση ελαχίστων τετραγώνων της f και η απόδειξη ολοκληρώθηκε. •

Το αντίστοιχο θεώρημα που χαρακτηρίζει την προσέγγιση ελαχίστων τετραγώνων της f, που ορίζεται στο σύνολο σημείων X_m, είναι το ακόλουθο.

Θεώρημα 15 *Έστω ότι η f ορίζεται στο σύνολο σημείων X_m και w είναι μια συνάρτηση βάρους ορισμένη στο X_m. Το πολυώνυμο q_n^* είναι η προσέγγιση ελαχίστων τετραγώνων της f, στο σύνολο πολυωνύμων P_n, ανν*

$$(f - q_n^*, p) = \sum_{i=1}^{m} (f(x_i) - q_n^*(x_i))p(x_i)w(x_i) = 0 \quad \forall p \in P_n. \tag{3.6}$$

Δε θα δώσουμε εδώ την απόδειξη γιατί είναι ακριβώς η ίδια με εκείνη του Θεωρήματος 14. Η απόδειξη του Θεωρήματος 14 δόθηκε υπό μορφή εσωτερικών γινομένων και όχι υπό μορφή ολοκληρωμάτων, ακριβώς για να αποτελεί απόδειξη και του Θεωρήματος 15. Μόνο σε ορισμένα σημεία δόθηκαν και κάποια ολοκληρώματα απλά για ευκολία του αναγνώστη, που θα μπορούσαν και αυτά να παραληφθούν. Αν αντικαταστήσουμε τα ολοκληρώματα με τα αντίστοιχα αθροίσματα, τότε η παραπάνω απόδειξη αποτελεί κάλλιστα απόδειξη του Θεωρήματος 15.

Έχουμε να παρατηρήσουμε εδώ ότι τα Θεωρήματα 14 και 15, ουσιαστικά μας λένε ότι το σφάλμα $f - q_n^*$ της προσέγγισης ελαχίστων τετραγώνων πρέπει να είναι ορθογώνιο προς ολόκληρο το χώρο P_n στον οποίο προσεγγίζουμε. Το ίδιο συμβαίνει και στον ευκλείδειο χώρο $\mathbb{R}^3$ στον οποίο ζούμε και έχουμε γεωμετρική άποψη. Όταν θέλουμε να προσεγγίσουμε ένα σημείο του χώρου με ένα σημείο σε ένα επίπεδο, φέρνουμε κάθετο από το σημείο στο επίπεδο. Ο ίδιος νόμος επομένως που ισχύει στον ευκλείδειο χώρο ο ίδιος ισχύει και στους χώρους συναρτήσεων όταν προσεγγίζουμε με την ευκλείδεια νόρμα. Αποδεικνύεται ότι η ορθογωνιότητα χαρακτηρίζει και την προσέγγιση όχι μόνο σε χώρους πολυωνύμων αλλά και άλλους υποχώρους όπως για παράδειγμα είναι ο υποχώρος των τριγωνομετρικών πολυωνύμων. Ακόμη αποδεικνύεται ότι η προσέγγιση και σε πιο αφηρημένους χώρους χαρακτηρίζεται από ορθογωνιότητα όταν η αντίστοιχη νόρμα παράγεται από εσωτερικό γινόμενο.

Στη συνέχεια θα δοθούν οι μέθοδοι και αλγόριθμοι για την εύρεση τη προσέγγισης ελαχίστων τετραγώνων.

3.2 Πολυώνυμο ελαχίστων τετραγώνων για συναρτήσεις σε συνεχές διάστημα

Θεωρούμε ότι η συνάρτηση f είναι συνεχής στο διάστημα $[-1, 1]$ και στόχος μας είναι να βρούμε το πολυώνυμο ελαχίστων τετραγώνων στον P_n, ως προς τη συνάρτηση βάρους w. Από τη χαρακτηριστική ιδιότητα ορθογωνιότητας (3.5) που δίνει το Θεώρημα 14, έχουμε:

$$(f - q_n^*, p) = 0 \Leftrightarrow (q_n^*, p) = (f, p) \quad \forall p \in P_n \tag{3.7}$$

ή σε μορφή ολοκληρωμάτων

$$\int_{-1}^{1} q_n^*(x)p(x)w(x)dx = \int_{-1}^{1} f(x)p(x)w(x)dx \quad \forall p \in P_n. \tag{3.8}$$

Απομένει λοιπόν να προσδιορίσουμε το πολυώνυμο q_n^* ώστε να ισχύει η σχέση (3.8) για όλα τα $p \in P_n$. Από πρώτη άποψη φαίνεται να είναι πολύ δύσκολο πρόβλημα διότι πρέπει να το αναζητήσουμε μέσα από ένα απειροσύνολο που είναι το P_n. Λόγω της γραμμικότητας που ισχύει για τα ολοκληρώματα ή γενικότερα για τα εσωτερικά γινόμενα, αρκεί να απαιτήσουμε να ισχύει η (3.8) μόνο για $n+1$ σε πλήθος πολυώνυμα που αποτελούν βάση του P_n. Επιλέγουμε μια βάση του χώρου P_n, εκφράζουμε το πολυώνυμο q_n^* ως γραμμικό συνδυασμό των πολυωνύμων της βάσης και απαιτούμε να ισχύει η (3.8) για όλα τα πολυώνυμα της βάσης. Κατασκευάζουμε έτσι ένα γραμμικό σύστημα $n + 1$ εξισώσεων με $n + 1$ αγνώστους, τους συντελεστές στο γραμμικό συνδυασμό. Η λύση αυτού του συστήματος δίνει και τη λύση του προβλήματός μας.

Η διαφορετικότητα των μεθόδων που αναπτύχθηκαν και θα δούμε στη συνέχεια συνίσταται στην επιλογή της βάσης. Θα εξετάσουμε, για απλότητα, σε όλο το παρόν κεφάλαιο μόνο την περίπτωση όπου $w(x) = 1$. Η περίπτωση για την οποία η συνάρτηση βάρους είναι διαφορετική από τη μονάδα εξετάζεται ακριβώς με τον ίδιο τρόπο, αρκεί να θεωρούμε και τη συνάρτηση w ως παράγοντα στα ολοκληρώματα.

3.2.1 Το Σύστημα των Κανονικών Εξισώσεων

Παίρνουμε ως βάση το σύνολο των μονωνύμων $\{1, x, x^2, \ldots, x^n\}$ και θεωρούμε ότι η προσέγγιση ελαχίστων τετραγώνων q_n^* εκφράζεται με το γραμμικό συνδυασμό

$$q_n^*(x) = \xi_0 + \xi_1 x + \xi_2 x^2 + \cdots + \xi_n x^n. \tag{3.9}$$

Αν θέσουμε $p(x) = x^i, \;\; i = 0, 1, 2, \ldots, n$, η σχέση (3.8) γίνεται

$$\int_{-1}^{1} x^i(\xi_0 + \xi_1 x + \xi_2 x^2 + \cdots + \xi_n x^n)dx = \int_{-1}^{1} x^i f(x)dx, \quad i = 0, 1, 2, \ldots, n$$

ή

$$\sum_{j=0}^{n} \left(\int_{-1}^{1} x^{i+j} dx \right) \xi_j = \int_{-1}^{1} x^i f(x)dx, \quad i = 0, 1, 2, \ldots, n. \tag{3.10}$$

Σε μορφή εσωτερικών γινομένων αυτή γίνεται:

$$\sum_{j=0}^{n}(x^i, x^j)\xi_j = (x^i, f), \quad i = 0, 1, 2, \ldots, n. \tag{3.11}$$

Οι σχέσεις αυτές δίνουν το λεγόμενο Σύστημα των Κανονικών Εξισώσεων:

$$\begin{bmatrix} a_{00} & a_{01} & a_{02} & \cdots & a_{0n} \\ a_{10} & a_{11} & a_{12} & \cdots & a_{1n} \\ a_{20} & a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n0} & a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} \xi_0 \\ \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix} = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}, \tag{3.12}$$

όπου

$$a_{ij} = (x^i, x^j) = \int_{-1}^{1} x^{i+j} dx, \quad i, j = 0, 1, \ldots, n \tag{3.13}$$

και

$$b_i = (x^i, f) = \int_{-1}^{1} x^i f(x) dx, \quad i = 0, 1, \ldots, n. \tag{3.14}$$

Το σύστημα αυτό μπορεί να γραφεί συνοπτικά υπό μορφή διανυσματικής εξίσωσης:

$$A\xi = b.$$

Με τη λύση του συστήματος των κανονικών εξισώσεων βρίσκουμε τους συντελεστές $\xi_i, \quad i = 0, 1, \ldots, n$, του γραμμικού συνδυασμού (3.9), που σημαίνει ότι υπολογίσαμε το πολυώνυμο ελαχίστων τετραγώνων q_n^*.

Ο αλγόριθμος επομένως συνίσταται στον υπολογισμό των στοιχείων του πίνακα A, στον υπολογισμό των στοιχείων του διανύσματος b και στη λύση ενός γραμμικού συστήματος $n + 1$ εξισώσεων με $n + 1$ αγνώστους. Παρατηρούμε ότι τα στοιχεία a_{ij} έχουν την ίδια τιμή όταν οι δείκτες i και j έχουν το ίδιο άθροισμα, που σημαίνει ότι τα στοιχεία του πίνακα A έχουν την ίδια τιμή κατά μήκος των διαγωνίων από κάτω αριστερά προς τα άνω δεξιά. έχουμε επομένως να υπολογίσουμε $2n + 1$ ολοκληρώματα για τον πίνακα A και $n + 1$ ολοκληρώματα για το διάνυσμα b. Οι πίνακες που έχουν τέτοια διάταξη των στοιχείων τους λέγονται πίνακες Hankel και έχουν κάποιες χαρακτηριστικές ιδιότητες. Δε θα ασχοληθούμε εδώ με τις ιδιότητες των πινάκων αυτών διότι αποτελεί αντικείμενο της Γραμμικής Άλγεβρας, θα αναφέρουμε όμως κάποια χαρακτηριστικά που επηρεάζουν τη λύση του προβλήματός μας. Αποδεικνύεται ότι οι πίνακες Hankel που παράγονται από προβλήματα ελαχίστων τετραγώνων έχουν πολύ μεγάλο δείκτη κατάστασης $\kappa(A) = \|A\|_2 \cdot \|A^{-1}\|_2$. Τότε λέμε ότι το σύστημα (3.12) έχει κακή κατάσταση, που σημαίνει ότι μικρά σφάλματα που υπεισέρχονται κατά την αποθήκευση των στοιχείων του συστήματος στον υπολογιστή ή κατά την εκτέλεση των ενδιάμεσων πράξεων, επιφέρουν μεγάλα σφάλματα στη λύση του συστήματος. Είναι τόσο μεγάλα τα σφάλματα, όταν το n είναι σχετικά μεγάλο, που η λύση που παίρνουμε δεν είναι αξιόπιστη. Αποδεικνύεται ότι για $n \geq 7$ δεν έχουμε αξιόπιστη λύση. Βλέπουμε λοιπόν ότι ενώ είναι απλός ο παραπάνω αλγόριθμος, στην πράξη δεν μπορεί να εφαρμοστεί και θα πρέπει να αναζητήσουμε άλλον όταν το n είναι σχετικά μεγάλο. Για πολύ μικρά n μπορεί να εφαρμοστεί κάλλιστα και να μας δώσει πολύ καλά αποτελέσματα. Ακολουθεί ένα παράδειγμα εφαρμογής του αλγορίθμου.

Παράδειγμα 17 Να βρεθεί η προσέγγιση ελαχίστων τετραγώνων της συνάρτησης $\sqrt{|x|}$ στον P_2, στο διάστημα $[-1, 1]$ με το σύστημα των κανονικών εξισώσεων.

Έχουμε να κατασκευάσουμε το 3×3 σύστημα των κανονικών εξισώσεων. Σύμφωνα με τους τύπους (3.13) και (3.14) υπολογίζουμε τα στοιχεία $a_{ij},\ i, j = 0, 1, 2$ και $b_i,\ i = 0, 1, 2$, αντίστοιχα:

$$a_{00} = (1, 1) = \int_{-1}^{1} 1dx = 2.$$

$$a_{01} = a_{10} = (1, x) = \int_{-1}^{1} xdx = 0,$$

αφού η συνάρτηση x είναι περιττή και το διάστημα ολοκλήρωσης είναι συμμετρικό ως προς το 0.

$$a_{02} = a_{11} = a_{20} = (1, x^2) = \int_{-1}^{1} x^2dx = \left[\frac{x^3}{3}\right]_{-1}^{1} = \frac{2}{3}.$$

$$a_{12} = a_{21} = (x, x^2) = \int_{-1}^{1} x^3dx = 0.$$

$$a_{22} = (x^2, x^2) = \int_{-1}^{1} x^4dx = \left[\frac{x^5}{5}\right]_{-1}^{1} = \frac{2}{5}.$$

$$b_0 = \left(1, \sqrt{|x|}\right) = \int_{-1}^{1} \sqrt{|x|}dx = 2\int_{0}^{1} x^{\frac{1}{2}}dx = 2\left[\frac{x^{\frac{3}{2}}}{\frac{3}{2}}\right]_{0}^{1} = \frac{4}{3}.$$

Εδώ θέσαμε $\int_{-1}^{1} \sqrt{|x|}dx = 2\int_0^1 x^{\frac{1}{2}}dx$, επειδή η συνάρτηση $\sqrt{|x|}$ είναι άρτια και το διάστημα ολοκλήρωσης είναι συμμετρικό ως προς το 0.

$$b_1 = \left(x, \sqrt{|x|}\right) = \int_{-1}^{1} x\sqrt{|x|}dx = 0,$$

επειδή η συνάρτηση $x\sqrt{|x|}$ είναι περιττή.

$$b_2 = \left(x^2, \sqrt{|x|}\right) = \int_{-1}^{1} x^2\sqrt{|x|}dx = 2\int_{0}^{1} x^{\frac{5}{2}}dx = 2\left[\frac{x^{\frac{7}{2}}}{\frac{7}{2}}\right]_{0}^{1} = \frac{4}{7}.$$

Επομένως, το σύστημα των κανονικών εξισώσεων είναι

$$\begin{bmatrix} 2 & 0 & \frac{2}{3} \\ 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \end{bmatrix} \begin{bmatrix} \xi_0 \\ \xi_1 \\ \xi_2 \end{bmatrix} = \begin{bmatrix} \frac{4}{3} \\ 0 \\ \frac{4}{7} \end{bmatrix}.$$

Παρατηρούμε ότι η δεύτερη εξίσωση λύνεται ανεξάρτητα από τις άλλες δυο και δίνει $\xi_1 = 0$. Οι άλλες δυο δίνουν το 2×2 σύστημα

$$\begin{bmatrix} 2 & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{5} \end{bmatrix} \begin{bmatrix} \xi_0 \\ \xi_2 \end{bmatrix} = \begin{bmatrix} \frac{4}{3} \\ \frac{4}{7} \end{bmatrix},$$

από το οποίο παίρνουμε τη λύση $\xi_0 = \frac{3}{7}$ και $\xi_2 = \frac{5}{7}$. Τελικά, το πολυώνυμο ελαχίστων τετραγώνων θα είναι

$$\frac{5}{7}x^2 + \frac{3}{7}.$$

Παρατηρούμε και εδώ, όπως είχαμε παρατηρήσει και στην ομοιόμορφη προσέγγιση, ότι επειδή η συνάρτηση $\sqrt{|x|}$ είναι άρτια και το διάστημα ολοκλήρωσης είναι συμμετρικό ως προς το 0, το πολυώνυμο ελαχίστων τετραγώνων που προέκυψε είναι και αυτό άρτιο.

3.2.2 Μέθοδοι Ορθογωνοποίησης

Για να υπερβούμε το πρόβλημα που έχουμε λόγω της κακής κατάστασης του συστήματος των κανονικών εξισώσεων, προσπαθούμε να κατασκευάσουμε μια βάση ορθογώνιων πολυωνύμων. Ας δούμε πρώτα πώς λύνεται το πρόβλημα ελαχίστων τετραγώνων αν έχουμε μια τέτοια βάση, το πώς θα την κατασκευάσουμε θα το δούμε στη συνέχεια.

Θεωρούμε την ορθογώνια βάση $\{P_0, P_1, P_2, \ldots, P_n\}$, όπου P_i είναι πολυώνυμο i βαθμού και

$$(P_i, P_j) = \int_{-1}^{1} P_i(x)P_j(x)dx = 0 \;\; \forall i \neq j, \;\; i, j = 1, 2, \ldots, n$$

και μόνο

$$(P_i, P_i) = \int_{-1}^{1} [P_i(x)]^2 dx > 0, \;\; i = 1, 2, \ldots, n.$$

Θεωρούμε επίσης ότι η προσέγγιση ελαχίστων τετραγώνων q_n^* εκφράζεται με το γραμμικό συνδυασμό

$$q_n^*(x) = \lambda_0 P_0(x) + \lambda_1 P_1(x) + \lambda_2 P_2(x) + \cdots + \lambda_n P_n(x). \tag{3.15}$$

Θέτουμε $p(x) = P_i(x), \;\; i = 0, 1, 2, \ldots, n$ και η σχέση (3.8) γίνεται

$$\int_{-1}^{1} P_i(x) \left(\lambda_0 P_0(x) + \lambda_1 P_1(x) + \cdots + \lambda_n P_n(x)\right) dx = \int_{-1}^{1} P_i(x) f(x) dx, \;\; i = 0, 1, 2, \ldots, n$$

ή σε μορφή εσωτερικών γινομένων

$$(P_i, (\lambda_0 P_0 + \lambda_1 P_1 + \cdots + \lambda_n P_n)) = (P_i, f), \;\; i = 0, 1, 2, \ldots, n.$$

Εφαρμόζοντας ιδιότητες εσωτερικών γινομένων, παίρνουμε

$$\sum_{j=0}^{n} \lambda_j (P_i, P_j) = (P_i, f), \;\; i = 0, 1, 2, \ldots, n$$

και λόγω της ορθογωνιότητας αυτή γίνεται

$$\lambda_i(P_i, P_i) = (P_i, f) \Leftrightarrow \lambda_i = \frac{(P_i, f)}{(P_i, P_i)}, \quad i = 0, 1, 2, \ldots, n. \tag{3.16}$$

Φτάσαμε λοιπόν στην εύρεση των συντελεστών λ_i και κατά συνέπεια στην εύρεση του πολυωνύμου ελαχίστων τετραγώνων, χωρίς να χρειάζεται να λύσουμε κάποιο σύστημα. Ουσιαστικά οι σχέσεις (3.16) αποτελούν ένα γραμμικό σύστημα, μόνο που ο αντίστοιχος πίνακας είναι διαγώνιος και η λύση του δίνεται με n διαιρέσεις και με πολύ μεγάλη ακρίβεια. Το τελικό συμπέρασμα είναι ότι, με τη χρήση κάποιας ορθογώνιας βάσης, καταφέρνουμε να βρούμε το πολυώνυμο ελαχίστων τετραγώνων με μεγάλη ακρίβεια αποφεύγοντας το σκόπελο του συστήματος των κανονικών εξισώσεων.

Πρέπει να παρατηρήσουμε εδώ ότι αν κάθε πολυώνυμο P_i, $i = 0, 1, \ldots, n$, της βάσης το πολλαπλασιάσουμε με ένα σταθερό παράγοντα c_i, τότε η ορθογωνιότητα δεν αλλοιώνεται. Το νέο σύνολο πολυωνύμων αποτελεί κι αυτό μια ορθογώνια βάση. Θα λέμε ότι μια ορθογώνια βάση είναι ανεξάρτητη ως προς σταθερό παράγοντα. Σε ορισμένα προβλήματα μας ενδιαφέρει η βάση να είναι κανονικοποιημένη, με την έννοια ότι η νόρμα κάθε στοιχείου της ισούται με τη μονάδα. Η κανονικοποίηση γίνεται εύκολα αν κάθε στοιχείο μιας ορθογώνιας βάσης το διαιρέσουμε με τη νόρμα του. Έστω η ορθογώνια βάση $\{P_0, P_1, P_2, \ldots, P_n\}$. Από αυτήν παίρνουμε την αντίστοιχη κανονικοποιημένη $\{\tilde{P}_0, \widetilde{P}_1, \widetilde{P}_2, \ldots, \widetilde{P}_n\}$, όπου

$$\widetilde{P}_i = \frac{P_i}{\|P_i\|_2}, \; i = 0, 1, \ldots, n.$$

Μια τέτοια βάση λέγεται ορθοκανονική. Αν θεωρήσουμε μια ορθοκανονική βάση για τη λύση του προβλήματος ελαχίστων τετραγώνων τότε η σχέση (3.16), που δίνει τα λ_i, απλοποιείται ακόμη περισσότερο και γινεται

$$\lambda_i = (\widetilde{P}_i, f), \quad i = 0, 1, 2, \ldots, n, \tag{3.17}$$

αφού $(\widetilde{P}_i, \widetilde{P}_i) = 1$.

Απομένει να δούμε κάποιες μεθόδους κατασκευής ορθογώνιας βάσης.

Gram-Schmidt Ορθογωνοποίηση

Κατά τη Gram-Schmidt ορθογωνοποίηση ξεκινούμε με μια οποιαδήποτε βάση γραμμικά ανεξάρτητων πολυωνύμων και έχοντας αυτή ως βάση αναφοράς προσπαθούμε να κατασκευάσουμε τη βάση ορθογώνιων πολυωνύμων. Συνήθως ξεκινούμε με τη βάση των μονωνύμων $\{1, x, x^2, \ldots, x^n\}$.

Αν υποθέσουμε ότι βρήκαμε την ορθογώνια βάση και αυτή είναι η $\{P_0, P_1, P_2, \ldots, P_n\}$, τότε το μονώνυμο x^i γράφεται ως γραμμικός συνδυασμός των πολυονωνύμων της βάσης ως εξής:

$$x^i = c_0 P_0(x) + c_1 P_1(x) + c_2 P_2(x) + \cdots + c_i P_i(x). \tag{3.18}$$

Θεωρώντας το εσωτερικό γινόμενο της σχέσης αυτής με το πολυώνυμο της βάσης P_j, με $j \leq i$, παίρνουμε

$$(x^i, P_j) = c_0(P_0, P_j) + c_1(P_1, P_j) + c_2(P_2, P_j) + \cdots + c_i(P_i, P_j)$$

και λόγω της ορθογωνιότητας καταλήγουμε

$$c_j(P_j, P_j) = (x^i, P_j) \Leftrightarrow c_j = \frac{(x^i, P_j)}{(P_j, P_j)}, \quad j = 0, 1, 2, \ldots, i. \tag{3.19}$$

Υποθέτουμε τώρα ότι έχουμε βρει τα πρώτα i στοιχεία της βάσης, $P_0, P_1, P_2, \ldots, P_{i-1}$ και θέλουμε να βρούμε το P_i. Επειδή οι ορθογώνιες βάσεις είναι ανεξάρτητες ως προς σταθερό παράγοντα, θεωρούμε ως P_i στη σχέση (3.18) το $c_i P_i$. Λύνοντας ως προς αυτό το P_i, αφού πρώτα αντικαταστήσουμε τους συντελεστές από την (3.19), παίρνουμε

$$P_i(x) = x^i - \frac{(x^i, P_0)}{(P_0, P_0)} P_0(x) - \frac{(x^i, P_1)}{(P_1, P_1)} P_1(x) - \cdots - \frac{(x^i, P_{i-1})}{(P_{i-1}, P_{i-1})} P_{i-1}(x). \tag{3.20}$$

Η σχέση (3.20) αποτελεί και τον αλγόριθμο Gram-Schmidt ορθογωνοποίησης, ο οποίος είναι πολύ απλός στην περιγραφή του. Ξεκινούμε αυθαίρετα με $P_0(x) = 1$, από την (3.20) για $i = 1$ βρίσκουμε το P_1, από την ίδια σχέση για $i = 2$ βρίσκουμε το P_2 και η ίδια διαδικασία συνεχίζεται μέχρι να βρούμε και το P_n. Ουσιαστικά με τον αλγόριθμο Gram-Schmidt ορθογωνοποίησης, αφού εκφράσουμε το x^i ως γραμμικό συνδυασμό των P_j, $j = 0, 1, 2, \ldots, i$, απαλείφουμε όλες τις συνιστώσες του x^i στις διευθύνσεις $P_0, P_1, P_2, \ldots, P_{i-1}$ και απομένει αναγκαστικά η διεύθυνση P_i. Ο ίδιος ακριβώς αλγόριθμος Gram-Schmidt ορθογωνοποίησης, με την ίδια φιλοσοφία, συναντάται και στη Γραμμική Άλγεβρα για την εύρεση ορθογώνιων βάσεων διανυσμάτων και για την κατασκευή ορθογώνιων πινάκων.

Θα δούμε στη συνέχεια την εφαρμογή του αλγορίθμου αυτού για την κατασκευή μιας βάσης $\{P_0, P_1, P_2, P_3\}$ στον χώρο P_3.

$$\begin{aligned}
P_0(x) &= 1, \\
P_1(x) &= x - \tfrac{(x,P_0)}{(P_0,P_0)} P_0(x) = x - \tfrac{(x,1)}{(1,1)} 1 = x - \tfrac{\int_{-1}^{1} x dx}{\int_{-1}^{1} 1 dx} 1 = x, \\
P_2(x) &= x^2 - \tfrac{(x^2,P_0)}{(P_0,P_0)} P_0(x) - \tfrac{(x^2,P_1)}{(P_1,P_1)} P_1(x) = x^2 - \tfrac{(x^2,1)}{(1,1)} 1 - \tfrac{(x^2,x)}{(x,x)} x \\
&= x^2 - \tfrac{\int_{-1}^{1} x^2 dx}{\int_{-1}^{1} 1 dx} 1 - \tfrac{\int_{-1}^{1} x^3 dx}{\int_{-1}^{1} x^2 dx} x = x^2 - \tfrac{1}{3}, \\
P_3(x) &= x^3 - \tfrac{(x^3,P_0)}{(P_0,P_0)} P_0(x) - \tfrac{(x^3,P_1)}{(P_1,P_1)} P_1(x) - \tfrac{(x^3,P_2)}{(P_2,P_2)} P_2(x) \\
&= x^3 - \tfrac{(x^3,1)}{(1,1)} 1 - \tfrac{(x^3,x)}{(x,x)} x - \tfrac{(x^3, x^2 - \frac{1}{3})}{(x^2 - \frac{1}{3}, x^2 - \frac{1}{3})} (x^2 - \tfrac{1}{3}) \\
&= x^3 - \tfrac{\int_{-1}^{1} x^3 dx}{\int_{-1}^{1} 1 dx} 1 - \tfrac{\int_{-1}^{1} x^4 dx}{\int_{-1}^{1} x^2 dx} x - \tfrac{\int_{-1}^{1} x^3 (x^2 - \frac{1}{3}) dx}{\int_{-1}^{1} (x^2 - \frac{1}{3})^2 dx} (x^2 - \tfrac{1}{3}) = x^3 - \tfrac{3}{5} x.
\end{aligned}$$

Με τη Gram-Schmidt ορθογωνοποίηση δημιουργούνται κάποια υπολογιστικά προβλήματα όταν το n είναι σχετικά μεγάλο και αυτό γίνεται επειδή για την εύρεση του P_i χρησιμοποιούνται όλα τα προηγούμενα πολυώνυμα. Αυτό έχει ως συνέπεια από τη μια να απαιτούνται πολύ υπολογισμοί και από την άλλη τα σφάλματα που δημιουργήθηκαν σε προηγούμενα πολυώνυμα να μεταβιβάζονται και να συσσωρεύονται σε επόμενα με αποτέλεσμα να έχουμε μεγάλα σφάλματα στο τελικό αποτέλεσμα. Προφανώς είναι πολύ πιο ακριβής μέθοδος από εκείνη του συστήματος των κανονικών εξισώσεων. Θα προσπαθήσουμε όμως να βρούμε κάποιον πιο ακριβή αλγόριθμο ορθογωνοποίησης.

Ορθογωνοποίηση με την αναδρομική σχέση τριών όρων

Στη μέθοδο αυτή χρησιμοποιούμε μια αναδρομική σχέση όπως και κατά την Gram-Schmidt ορθογωνοποίηση, μόνο που εδώ εμπλέκονται μόνο τρεις διαδοχικοί όροι στην αναδρομική σχέση. Συγκεκριμένα, παίρνουμε $P_0(x) = 1$, $P_1(x) = x$ που γνωρίζουμε ότι είναι ορθογώνια μεταξύ των και θεωρούμε την αναδρομική σχέση:

$$P_{k+1}(x) = (x - \alpha_k)P_k(x) - \beta_k P_{k-1}(x), \quad k = 1, 2, \ldots, n-1. \tag{3.21}$$

Στόχος μας είναι ο προσδιορισμός των α_k και β_k ώστε η βάση πολυωνύμων που παράγεται να είναι ορθογώνια. Για να το πετύχουμε θα χρησιμοποιήσουμε τέλεια επαγωγή. Υποθέτουμε ότι οι $k+1$ πρώτοι όροι $P_0, P_1, P_2, \ldots, P_k$ είναι ορθογώνια πολυώνυμα μεταξύ των και θα αποδείξουμε ότι με κατάλληλο προσδιορισμό των α_k και β_k και το P_{k+1} είναι ορθογώνιο με όλα τα προηγούμενα. Θεωρούμε το πολυώνυμο P_j, $j = 0, 1, 2, \ldots, k$ και υπολογίζουμε το εσωτερικό γινόμενο (P_{k+1}, P_j) διακρίνοντας τρεις περιπτώσεις:

1) $j < k-1$:

$$\begin{aligned} (P_{k+1}, P_j) &= ((x-\alpha_k)P_k - \beta_k P_{k-1}, P_j) = ((x-\alpha_k)P_k, P_j) - (\beta_k P_{k-1}, P_j) \\ &= (P_k, (x-\alpha_k)P_j) - \beta_k \left(P_{k-1}, P_j\right). \end{aligned}$$

Από την υπόθεση έχουμε ότι $(P_{k-1}, P_j) = 0$. Παρατηρούμε επίσης ότι το πολυώνυμο $(x-\alpha_k)P_j$ θα είναι πολυώνυμο το πολύ $k-1$ βαθμού. Το P_k υποθέσαμε ότι είναι πολυώνυμο της ορθογώνιας βάσης, επομένως θα είναι ορθογώνιο με οποιοδήποτε πολυώνυμο το πολύ $k-1$ βαθμού, αφού εκείνο θα μπορεί να γραφεί ως γραμμικός συνδυασμός των $P_0, P_1, P_2, \ldots, P_{k-1}$. Αυτό σημαίνει ότι $(P_k, (x-\alpha_k)P_j) = 0$, που αποδεικνύει ότι $(P_{k+1}, P_j) = 0$. Αποδείξαμε λοιπον ότι το P_{k+1} είναι ορθογώνιο με όλα τα P_j για $j < k-1$, για οποιαδήποτε α_k και β_k.

2) $j = k$:

$$\begin{aligned} (P_{k+1}, P_k) &= ((x-\alpha_k)P_k - \beta_k P_{k-1}, P_k) = (xP_k, P_k) - \alpha_k \left(P_k, P_k\right) - \beta_k \left(P_{k-1}, P_k\right) \\ &= (xP_k, P_k) - \alpha_k \left(P_k, P_k\right). \end{aligned}$$

Αν επιλέξουμε

$$\alpha_k = \frac{(xP_k, P_k)}{(P_k, P_k)}, \tag{3.22}$$

τότε παίρνουμε $(P_{k+1}, P_k) = 0$. Προσδιορίσαμε το α_k ώστε το πολυώνυμο P_{k+1} να είναι ορθογώνιο με το P_k.

3) $j = k - 1$:

$$\begin{aligned}(P_{k+1}, P_{k-1}) &= ((x-\alpha_k)P_k - \beta_k P_{k-1}, P_{k-1}) = (xP_k, P_{k-1}) - \alpha_k (P_k, P_{k-1}) \\ &- \beta_k (P_{k-1}, P_{k-1}) = (xP_k, P_{k-1}) - \beta_k (P_{k-1}, P_{k-1}).\end{aligned}$$

Επιλέγοντας

$$\beta_k = \frac{(xP_k, P_{k-1})}{(P_{k-1}, P_{k-1})}, \tag{3.23}$$

παίρνουμε $(P_{k+1}, P_{k-1}) = 0$. Προσδιορίσαμε λοιπόν το β_k ώστε το πολυώνυμο P_{k+1} να είναι ορθογώνιο με το P_{k-1}.

Έχουμε αποδείξει ότι η αναδρομική σχέση (3.21) αποτελεί ένα πολύ καλό εργαλείο για την κατασκευή βάσης ορθογώνιων πολυωνύμων. Θα αποδείξουμε εδώ ότι ο τύπος (3.23) που μας δίνει το β_k είναι ισοδύναμος με τον

$$\beta_k = \frac{(P_k, P_k)}{(P_{k-1}, P_{k-1})}. \tag{3.24}$$

Πραγματικά, ο αριθμητής του (3.23) μπορεί να γίνει $(xP_k, P_{k-1}) = (P_k, xP_{k-1})$. Το πολυώνυμο xP_{k-1} είναι ακριβώς k βαθμού και μπορεί να δοθεί ως γραμμικός συνδυασμός των πολυωνύμων της βάσης ως εξής:

$$xP_{k-1} = c_0P_0 + c_1P_1 + c_2P_2 + \cdots + c_kP_k.$$

Παρατηρούμε, από την αναδρομική σχέση και από το γεγονός ότι ξεκινούμε με $P_0(x) = 1$ και $P_1(x) = x$, ότι όλα τα πολυώνυμα της βάσης έχουν συντελεστή μεγιστοβαθμίου όρου τη μονάδα. Επίσης και το πολυώνυμο xP_{k-1} εχει συντελεστή μεγιστοβαθμίου όρου τη μονάδα. Επομένως, στο γραμμικό συνδυασμό θα έχουμε $c_k = 1$. Αντικαθιστώντας στο εσωτερικό γινόμενο, παίρνουμε

$$(P_k, xP_{k-1}) = c_0(P_k, P_0) + c_1(P_k, P_1) + c_2(P_k, P_2) + \cdots + (P_k, P_k) = (P_k, P_k),$$

που αποδεικνύει τον τύπο (3.24).

Μπορούμε να αποδείξουμε ότι ο τύπος (3.22) δίνει $\alpha_k = 0$ επειδή πήραμε ως διάστημα το $[-1, 1]$ που είναι συμμετρικό ως προς το μηδέν. Πραγματικά, ο αριθμητής του (3.22) δίνει

$$(xP_k, P_k) = \int_{-1}^{1} x[P_k(x)]^2 dx = 0,$$

επειδή η υπό ολοκλήρωση συνάρτηση είναι περιττή και το διάστημα ολοκλήρωσης είναι συμμετρικό ως προς το μηδέν. Σε μη συμμετρικά όμως διαστήματα αυτό δεν ισχύει.

Θα προσπαθήσουμε τώρα να υπολογίσουμε την ίδια βάση $\{P_0, P_1, P_2, P_3\}$, χρησιμοποιώντας την αναδρομική σχέση (3.21).

$$\begin{aligned}
P_0(x) &= 1,\\
P_1(x) &= x,\\
\beta_1 &= \tfrac{(P_1,P_1)}{(P_0,P_0)} = \tfrac{(x,x)}{(1,1)} = \tfrac{\frac{2}{3}}{2} = \tfrac{1}{3},\\
P_2(x) &= xP_1(x) - \beta_1 P_0(x) = x^2 - \tfrac{1}{3},\\
\beta_2 &= \tfrac{(P_2,P_2)}{(P_1,P_1)} = \tfrac{\left(x^2-\frac{1}{3},x^2-\frac{1}{3}\right)}{(x,x)} = \tfrac{\left(x^2,x^2\right)-2\frac{1}{3}(x^2,1)+\frac{1}{9}(1,1)}{\frac{2}{3}} = \tfrac{\frac{2}{5}-\frac{4}{9}+\frac{2}{9}}{\frac{2}{3}} = \tfrac{4}{15},\\
P_3(x) &= xP_2(x) - \beta_2 P_1(x) = x(x^2-\tfrac{1}{3}) - \tfrac{4}{15}x = x^3 - \tfrac{3}{5}x.
\end{aligned}$$

Πολυώνυμα Legendre

Τα πολυώνυμα Legendre είναι ένα σύνολο ορθογώνιων πολυωνύμων στο διάστημα $[-1,1]$ με πολύ καλές ιδιότητες. Είναι μια ειδική περίπτωση της γενικότερης κατηγορίας των ορθογώνιων πολυωνύμων Jacobi. Εκτός από τη Θεωρία Προσέγγισης, ευρεία εφαρμογή έχουν και στην Αριθμητική Ολοκλήρωση. Δε θα ασχοληθούμε όμως εδώ σε βάθος, απλώς θα τα παραθέσουμε ως εργαλείο για την προσέγγιση ελαχίστων τετραγώνων. Αυτά δίνονται από την αναδρομική σχέση τριών όρων

$$P_{k+1}(x) = \frac{2k+1}{k+1}xP_k(x) - \frac{k}{k+1}P_{k-1}(x), \quad k=1,2,\ldots, \tag{3.25}$$

με $P_0(x)=1$ και $P_1(x)=x$. Δίνουμε στη συνέχεια τους τέσσερις πρώτους όρους των πολυωνύμων Legendre.

$$\begin{aligned}
P_0(x) &= 1,\\
P_1(x) &= x,\\
P_2(x) &= \tfrac{2\cdot1+1}{1+1}xP_1(x) - \tfrac{1}{1+1}P_0(x) = \tfrac{3}{2}x^2 - \tfrac{1}{2},\\
P_3(x) &= \tfrac{2\cdot2+1}{2+1}xP_2(x) - \tfrac{2}{2+1}P_1(x) = \tfrac{5}{3}x\left(\tfrac{3}{2}x^2-\tfrac{1}{2}\right) - \tfrac{2}{3}x = \tfrac{5}{2}x^3 - \tfrac{3}{2}x,\\
P_4(x) &= \tfrac{2\cdot3+1}{3+1}xP_3(x) - \tfrac{3}{3+1}P_2(x) = \tfrac{7}{4}x\left(\tfrac{5}{2}x^3-\tfrac{3}{2}x\right) - \tfrac{3}{4}\left(\tfrac{3}{2}x^2-\tfrac{1}{2}\right) = \tfrac{35}{8}x^4 - \tfrac{15}{4}x^2 + \tfrac{3}{8}.
\end{aligned}$$

Παρατηρούμε ότι τα πολυώνυμα Legendre διαφέρουν από τα ορθογώνια πολυώνυμα που κατασκευάσαμε με τη μέθοδο Gram-Schmidt ή με την αναδρομική σχέση (3.21), μόνο ως προς σταθερό παράγοντα. Πραγματικά, αν διαιρέσουμε με το μεγιστοβάθμιο όρο τα πολυώνυμα Legendre, παίρνουμε ακριβώς τα πολυώνυμα που είχαμε κατασκευάσει. Για παράδειγμα, Το P_2 μας δίνει $x^2-\frac{1}{3}$ ενώ το P_3 δίνει $x^3-\frac{3}{5}x$.

Θα λύσουμε το ίδιο πρόβλημα του Παραδείγματος 17 χρησιμοποιώντας ορθογώνια πολυώνυμα.

Παράδειγμα 18 Να βρεθεί η προσέγγιση ελαχίστων τετραγώνων της συνάρτησης $\sqrt{|x|}$ στον P_2, στο διάστημα $[-1,1]$ με μέθοδο ορθογωνοποίησης.

Μας χρειάζεται το σύνολο τριών ορθογωνίων πολυωνύμων $\{P_0,P_1,P_2\}=\{1,x,x^2-\frac{1}{3}\}$, που έχουμε ήδη κατασκευάσει. Σύμφωνα με τη σχέση (3.15) το πολυώνυμο ελαχίστων τετραγώνων θα δίνεται από το γραμμικό συνδυασμό

$$q_n^*(x) = \lambda_0P_0(x) + \lambda_1P_1(x) + \lambda_2P_2(x)$$

όπου τα $\lambda_i,\ i=0,1,2$, θα δίνονται από τη σχέση (3.16). Επομένως, παίρνουμε

$$\lambda_0=\frac{(P_0,f)}{(P_0,P_0)}=\frac{\left(1,\sqrt{|x|}\right)}{(1,1)}=\frac{\int_{-1}^{1}\sqrt{|x|}dx}{2}=\frac{2\int_0^1\sqrt{x}dx}{2}=\frac{\frac{4}{3}}{2}=\frac{2}{3},$$

$$\lambda_1=\frac{(P_1,f)}{(P_1,P_1)}=\frac{\left(x,\sqrt{|x|}\right)}{(x,x)}=\frac{\int_{-1}^{1}x\sqrt{|x|}dx}{\frac{2}{3}}=0,$$

$$\begin{aligned}\lambda_2 &= \frac{(P_2,f)}{(P_2,P_2)}=\frac{\left(x^2-\frac{1}{3},\sqrt{|x|}\right)}{\left(x^2-\frac{1}{3},x^2-\frac{1}{3}\right)}=\frac{\left(x^2,\sqrt{|x|}\right)-\left(\frac{1}{3},\sqrt{|x|}\right)}{(x^2,x^2)-2\frac{1}{3}(x^2,1)+\frac{1}{9}(1,1)}\\ &= \frac{2\int_0^1 x^{\frac{5}{2}}dx-\frac{1}{3}2\int_0^1 x^{\frac{1}{2}}dx}{\int_{-1}^1 x^4dx-2\frac{1}{3}\int_{-1}^1 x^2dx+\frac{1}{9}\int_{-1}^1 1dx}=\frac{2\frac{2}{7}-\frac{1}{3}2\frac{2}{3}}{\frac{2}{5}-2\frac{1}{3}\frac{2}{3}+\frac{1}{9}2}=\frac{5}{7}.\end{aligned}$$

Τελικά, το πολυώνυμο ελαχίστων τετραγώνων θα είναι

$$q_n^*(x)=\lambda_0P_0(x)+\lambda_1P_1(x)+\lambda_2P_2(x)=\frac{2}{3}+\frac{5}{7}\left(x^2-\frac{1}{3}\right))=\frac{5}{7}x^2+\frac{3}{7}.$$

3.3 Πολυώνυμο ελαχίστων τετραγώνων για συναρτήσεις ορισμένες σε σύνολο σημείων

Θεωρούμε ότι η συνάρτηση f ορίζεται στο σύνολο σημείων $X_m=\{x_1,x_2,x_3,\ldots,x_m\}$ και οι τιμές της στα σημεία αυτά είναι $f_1,f_2,f_3,\ldots,f_m$. Όπως και στη συνεχή περίπτωση έτσι και εδώ, στόχος μας είναι να βρούμε το πολυώνυμο ελαχίστων τετραγώνων στον P_n, ως προς τη συνάρτηση βάρους w. Εδώ θα μπορούσαμε να μην πούμε τίποτε επιπλέον από αυτά που αναπτήξαμε στη συνεχή περίπτωση. Ισχύουν ακριβώς τα ίδια, υπάρχει ένας θαυμάσιος δυϊσμός όπως αναφέραμε και στην αρχή του Κεφαλαίου. Η μόνη διαφορά που υπάρχει, είναι ότι τα ολοκληρώματα της συνεχούς περίπτωσης αντικαθίστανται με αθροίσματα. Ακόμη, εδώ θα εργαζόμαστε κάθε φορά στο συγκεκριμένο σύνολο X_m, χωρίς να μπορούμε να το μετασχηματίσουμε, όπως κάναμε στη συνεχή περίπτωση με το $[-1,1]$.

Από τη χαρακτηριστική ιδιότητα ορθογωνιότητας (3.6) που δίνει το Θεώρημα 15, έχουμε και εδώ την ίδια σχέση (3.7), δηλαδή

$$(q_n^*,p)=(f,p)\ \ \forall p\in P_n.$$

Σε μορφή αθροισμάτων αυτή γίνεται

$$\sum_{k=1}^{m}q_n^*(x_k)p(x_k)w_k=\sum_{k=1}^{m}f_kp(x_k)w_k\ \ \forall p\in P_n. \tag{3.26}$$

Επιλέγουμε και εδώ μια βάση του χώρου P_n, εκφράζουμε το πολυώνυμο q_n^* ως γραμμικό συνδυασμό των πολυωνύμων της βάσης και απαιτούμε να ισχύει η (3.26) για όλα τα πολυώνυμα της βάσης. Θα εξετάσουμε και εδώ της αντίστοιχες περιπτώσεις, όπως και στη συνεχή, όπου η διαφορετικότητα οφείλεται στην επιλογή της βάσης. Θεωρούμε και εδώ ότι $w(x)=1$.

3.3.1 Το Σύστημα των Κανονικών Εξισώσεων

Όπως και στη συνεχή περίπτωση, παίρνουμε ως βάση το σύνολο των μονωνύμων $\{1, x, x^2, \ldots, x^n\}$ και θεωρούμε ότι η προσέγγιση ελαχίστων τετραγώνων q_n^* εκφράζεται με το γραμμικό συνδυασμό

$$q_n^*(x) = \xi_0 + \xi_1 x + \xi_2 x^2 + \cdots + \xi_n x^n. \tag{3.27}$$

Θέτουμε στη συνέχεια $p(x) = x^i$, $i = 0, 1, 2, \ldots, n$, και η σχέση (3.26) γίνεται

$$\sum_{k=1}^{m} x_k^i(\xi_0 + \xi_1 x_k + \xi_2 x_k^2 + \cdots + \xi_n x_k^n) = \sum_{k=1}^{m} x_k^i f_k, \quad i = 0, 1, 2, \ldots, n$$

ή

$$\sum_{j=0}^{n} \left(\sum_{k=1}^{m} x_k^{i+j} \right) \xi_j = \sum_{k=1}^{m} x_k^i f(x) dx, \quad i = 0, 1, 2, \ldots, n. \tag{3.28}$$

Σε μορφή εσωτερικών γινομένων παίρνουμε ακριβώς την (3.11). Οι παραπάνω σχέσεις δίνουν το Σύστημα των Κανονικών Εξισώσεων:

$$A\xi = b \Leftrightarrow \begin{bmatrix} a_{00} & a_{01} & a_{02} & \cdots & a_{0n} \\ a_{10} & a_{11} & a_{12} & \cdots & a_{1n} \\ a_{20} & a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n0} & a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} \xi_0 \\ \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix} = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}, \tag{3.29}$$

όπου

$$a_{ij} = (x^i, x^j) = \sum_{k=1}^{m} x_k^{i+j}, \quad i, j = 0, 1, \ldots, n \tag{3.30}$$

και

$$b_i = (x^i, f) = \sum_{k=1}^{m} x_k^i f_k, \quad i = 0, 1, \ldots, n. \tag{3.31}$$

Απομένει να λύσουμε το σύστημα των κανονικών εξισώσεων για να βρούμε τους συντελεστές του γραμμικού συνδυασμού (3.27) και κατά συνέπεια να υπολογίσουμε το πολυώνυμο ελαχίστων τετραγώνων q_n^*.

Γράφουμε τον πίνακα A και το διάνυσμα b στη μορφή των αθροισμάτων (3.30) και (3.31):

$$A = \begin{bmatrix} \sum_{k=1}^m 1 & \sum_{k=1}^m x_k & \sum_{k=1}^m x_k^2 & \cdots & \sum_{k=1}^m x_k^n \\ \sum_{k=1}^m x_k & \sum_{k=1}^m x_k^2 & \sum_{k=1}^m x_k^3 & \cdots & \sum_{k=1}^m x_k^{n+1} \\ \sum_{k=1}^m x_k^2 & \sum_{k=1}^m x_k^3 & \sum_{k=1}^m x_k^4 & \cdots & \sum_{k=1}^m x_k^{n+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^m x_k^n & \sum_{k=1}^m x_k^{n+1} & \sum_{k=1}^m x_k^{n+2} & \cdots & \sum_{k=1}^m x_k^{2n} \end{bmatrix}, b = \begin{bmatrix} \sum_{k=1}^m f_k \\ \sum_{k=1}^m x_k f_k \\ \sum_{k=1}^m x_k^2 f_k \\ \vdots \\ \sum_{k=1}^m x_k^n f_k \end{bmatrix}.$$

Εδώ παρατηρούμε ότι ο A και το b μπορούν να γραφούν στη μορφή

$$A = X^T X, \quad b = X^T \hat{f},$$

όπου X είνα ο $m \times (n+1)$ Vandermonde πίνακας

$$X = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ 1 & x_3 & x_3^2 & \cdots & x_3^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_m & x_m^2 & \cdots & x_m^n \end{bmatrix}$$

και $\hat{f}$ είναι το m-διάστατο διάνυσμα με τις τιμές της συνάρτησης ($\hat{f} = (f_1\ f_2\ f_3\ \cdots f_m)^T$). Επομένως, το σύστημα των κανονικών εξισώσεων γίνεται

$$X^T X\xi = X^T \hat{f} \Leftrightarrow X^T(X\xi - \hat{f}) = 0.$$

Το διάνυσμα $X\xi$ έχει ως συνιστώσες τις τιμές του ζητούμενου πολυωνύμου q_n^* στα σημεία του X_m, συνεπώς το διάνυσμα $X\xi - \hat{f}$ αποτελεί το διάνυσμα σφάλμα κατά την προσέγγιση ελαχίστων τετραγώνων. Είναι γνωστό από τη Γραμμική Άλγεβρα ότι το παραπάνω σύστημα κανονικών εξισώσεων, επιλύει το πρόβλημα ελαχιστοποίησης:

$$\min_{\xi \in \mathbb{R}^m} \|X\xi - \hat{f}\|_2. \tag{3.32}$$

Αυτό το πρόβλημα όμως είναι το πρόβλημα ελαχίστων τετραγώνων που θέσαμε εξαρχής. Διασταυρώσαμε με αυτόν τον τρόπο και από πλευράς της Γραμμικής Άλγεβρας ότι το σύστημα κανονικών εξισώσεων, λύνει το πρόβλημά μας. Βέβαια η Γραμμική Άλγεβρα λύνει το γενικότερο πρόβλημα ελαχιστοποίησης (3.32) για οποιοδήποτε πίνακα X και διάνυσμα $\hat{f}$ και όχι μόνο αυτά που προέρχονται από προσέγγιση συναρτήσεων με πολυώνυμα.

Ισχύουν και εδώ όλες οι παρατηρήσεις που δώσαμε στη συνεχή περίπτωση, ως προς την κατάσταση του συστήματος. Το σύστημα (3.29) έχει κακή κατάσταση και θα πρέπει να ψάξουμε για πιο ευσταθή αλγόριθμο διαμέσου μεθόδων ορθογωνοποίησης. Δίνουμε στη συνέχεια δυο παραδείγματα.

Παράδειγμα 19 Να βρεθεί η προσέγγιση ελαχίστων τετραγώνων στον P_2, της συνάρτησης $|x|$ στο σύνολο σημείων $X_5 = \{-1, -\frac{1}{2}, 0, \frac{1}{2}, 1\}$ με το σύστημα των κανονικών εξισώσεων.

Ο πίνακας τιμών της συνάρτησης είναι

x_k	-1	$-\frac{1}{2}$	0	$\frac{1}{2}$	1
f_k	1	$\frac{1}{2}$	0	$\frac{1}{2}$	1

Συνήθως, καταρτίζουμε ένα πίνακα με τις τιμές των δυνάμεων των x_k και με τις τιμές των δυνάμεων των x_k επί f_k που μας χρειάζονται, ώστε να βρίσκουμε τα αθροίσματα. Στο παράδειγμά μας ο πίνακας αυτός είναι:

1	x_k	x_k^2	x_k^3	x_k^4	f_k	x_kf_k	$x_k^2f_k$
1	-1	1	-1	1	1	-1	1
1	$-\frac{1}{2}$	$\frac{1}{4}$	$-\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{2}$	$-\frac{1}{4}$	$\frac{1}{8}$
1	0	0	0	0	0	0	0
1	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$
1	1	1	1	1	1	1	1
5	0	$\frac{5}{2}$	0	$\frac{17}{8}$	3	0	$\frac{9}{4}$

Το 3×3 σύστημα των κανονικών εξισώσεων θα είναι

$$\begin{bmatrix} 5 & 0 & \frac{5}{2} \\ 0 & \frac{5}{2} & 0 \\ \frac{5}{2} & 0 & \frac{17}{8} \end{bmatrix} \begin{bmatrix} \xi_0 \\ \xi_1 \\ \xi_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ \frac{9}{4} \end{bmatrix}.$$

Η δεύτερη εξίσωση λύνεται ανεξάρτητα και δίνει $\xi_1 = 0$. Οι άλλες δυο δίνουν το 2×2 σύστημα

$$\begin{bmatrix} 5 & \frac{5}{2} \\ \frac{5}{2} & \frac{17}{8} \end{bmatrix} \begin{bmatrix} \xi_0 \\ \xi_2 \end{bmatrix} = \begin{bmatrix} 3 \\ \frac{9}{4} \end{bmatrix},$$

από το οποίο παίρνουμε τη λύση $\xi_0 = \frac{6}{35}$ και $\xi_2 = \frac{6}{7}$. Τελικά, το πολυώνυμο ελαχίστων τετραγώνων θα είναι

$$q_2^*(x) = \frac{6}{7}x^2 + \frac{6}{35}.$$

Παρατηρούμε και εδώ, ότι επειδή η συνάρτηση $|x|$ είναι άρτια και το σύνολο σημείων X_m είναι συμμετρικό ως προς το 0, το πολυώνυμο ελαχίστων τετραγώνων που προέκυψε είναι και αυτό άρτιο.

Το δεύτερο παράδειγμα θα λύνει το ίδιο πρόβλημα χωρίς να είναι το X_m συμμετρικό ως προς το 0.

Παράδειγμα 20 Να βρεθεί η προσέγγιση ελαχίστων τετραγώνων στον P_2, της συνάρτησης $|x|$ στο σύνολο σημείων $X_4 = \{-1, 0, \frac{1}{2}, 1\}$ με το σύστημα των κανονικών εξισώσεων.

Οπίνακας τιμών της συνάρτησης είναι

x_k	-1	0	$\frac{1}{2}$	1
f_i	1	0	$\frac{1}{2}$	1

Ο αντίστοιχος πίνακας τιμών θα είναι είναι:

1	x_k	x_k^2	x_k^3	x_k^4	f_k	x_kf_k	$x_k^2f_k$
1	-1	1	-1	1	1	-1	1
1	0	0	0	0	0	0	0
1	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$
1	1	1	1	1	1	1	1
4	$\frac{1}{2}$	$\frac{9}{4}$	$\frac{1}{8}$	$\frac{33}{16}$	$\frac{5}{2}$	$\frac{1}{4}$	$\frac{17}{8}$

Το 3×3 σύστημα των κανονικών εξισώσεων θα είναι

$$\begin{bmatrix} 4 & \frac{1}{2} & \frac{9}{4} \\ \frac{1}{2} & \frac{9}{4} & \frac{1}{8} \\ \frac{9}{4} & \frac{1}{8} & \frac{33}{16} \end{bmatrix} \begin{bmatrix} \xi_0 \\ \xi_1 \\ \xi_2 \end{bmatrix} = \begin{bmatrix} \frac{5}{2} \\ \frac{1}{4} \\ \frac{17}{8} \end{bmatrix}.$$

Λύνοντας το σύστημα αυτό με ακριβή αριθμητική, διατηρώντας κλάσματα στους υπολογισμούς, παίρνουμε τη λύση $\xi_0 = \frac{6}{55}$, $\xi_1 = \frac{2}{55}$ και $\xi_2 = \frac{10}{11}$. Τελικά, το πολυώνυμο ελαχίστων τετραγώνων θα είναι

$$q_2^*(x) = \frac{10}{11}x^2 + \frac{2}{55}x + \frac{6}{55}.$$

3.3.2 Μέθοδοι Ορθογωνοποίησης

Θεωρούμε και εδώ την ορθογώνια βάση $\{P_0, P_1, P_2, \dots, P_n\}$, όπου P_i είναι πολυώνυμο i βαθμού και

$$(P_i, P_j) = \sum_{k=1}^{m} P_i(x_k)P_j(x_k) = 0 \quad \forall i \neq j, \quad i,j = 1,2,\dots,n$$

και μόνο

$$(P_i, P_i) = \sum_{k=1}^{m} [P_i(x_k)]^2 > 0, \quad i = 1,2,\dots,n.$$

Θεωρούμε επίσης ότι η προσέγγιση ελαχίστων τετραγώνων q_n^* εκφράζεται με το γραμμικό συνδυασμό

$$q_n^*(x) = \lambda_0 P_0(x) + \lambda_1 P_1(x) + \lambda_2 P_2(x) + \cdots + \lambda_n P_n(x). \tag{3.33}$$

Ακολουθώντας την ίδια διαδικασία όπως στη συνεχή περίπτωση, καταλήγουμε στη σχέση (3.16) που δίνει τα λ_i, δηλαδή

$$\lambda_i = \frac{(P_i, f)}{(P_i, P_i)}, \quad i = 0,1,2,\dots,n,$$

μόνο που τα εσωτερικά γινόμενα εδώ αντιστοιχούν σε αθροίσματα. Ισχύουν και εδώ οι παρατηρήσεις που έγιναν στη συνεχή περίπτωση, για το ότι δηλαδή μια ορθογώνια βάση είναι ανεξάρτητη από σταθερό παράγοντα.

Απομένει να δούμε κάποιες μεθόδους κατασκευής ορθογώνιας βάσης.

Gram-Schmidt Ορθογωνοποίηση

Κατά τη Gram-Schmidt ορθογωνοποίηση ξεκινούμε με τη βάση των μονωνύμων $\{1, x, x^2, \dots, x^n\}$. Ισχύει ακριβώς η ίδια θεωρία της συνεχούς περίπτωσης. Εργαζόμενοι με εσωτερικά γινόμενα καταλήγουμε στον ίδιο αλγόριθμο (3.20):

$$P_i(x) = x^i - \frac{(x^i, P_0)}{(P_0, P_0)}P_0(x) - \frac{(x^i, P_1)}{(P_1, P_1)}P_1(x) - \cdots - \frac{(x^i, P_{i-1})}{(P_{i-1}, P_{i-1})}P_{i-1}(x),$$

με $P_0(x) = 1$. Θα κατασκευάσουμε εδώ τις ορθογώνιες βάσεις $\{P_0, P_1, P_2\}$ στο χώρο P_2, που αντιστοιχούν στα σύνολα σημείων των παραδειγμάτων 19 και 20.

1) $X_5 = \{-1, -\frac{1}{2}, 0, \frac{1}{2}, 1\}$

$$\begin{aligned}
P_0(x) &= 1,\\
P_1(x) &= x - \tfrac{(x,P_0)}{(P_0,P_0)}P_0(x) = x - \tfrac{(x,1)}{(1,1)}1 = x - \tfrac{-1\cdot1+(-\frac{1}{2})\cdot1+0\cdot1+\frac{1}{2}\cdot1+1\cdot1}{1\cdot1+1\cdot1+1\cdot1+1\cdot1+1\cdot1}1 = x,\\
P_2(x) &= x^2 - \tfrac{(x^2,P_0)}{(P_0,P_0)}P_0(x) - \tfrac{(x^2,P_1)}{(P_1,P_1)}P_1(x) = x^2 - \tfrac{(x^2,1)}{(1,1)}1 - \tfrac{(x^2,x)}{(x,x)}x\\
&= x^2 - \tfrac{1\cdot1+\frac{1}{4}\cdot1+0\cdot1+\frac{1}{4}\cdot1+1\cdot1}{1\cdot1+1\cdot1+1\cdot1+1\cdot1+1\cdot1}1 - \tfrac{1\cdot(-1)+\frac{1}{4}\cdot(-\frac{1}{2})+0\cdot0+\frac{1}{4}\cdot\frac{1}{2}+1\cdot1}{(-1)\cdot(-1)+(-\frac{1}{2})\cdot(-\frac{1}{2})+0\cdot0+\frac{1}{2}\cdot\frac{1}{2}+1\cdot1}x = x^2 - \tfrac{1}{2}.
\end{aligned}$$

2) $X_4 = \{-1, 0, \frac{1}{2}, 1\}$

$$\begin{aligned}
P_0(x) &= 1,\\
P_1(x) &= x - \tfrac{(x,P_0)}{(P_0,P_0)}P_0(x) = x - \tfrac{(x,1)}{(1,1)}1 = x - \tfrac{-1\cdot1+0\cdot1+\frac{1}{2}\cdot1+1\cdot1}{1\cdot1+1\cdot1+1\cdot1+1\cdot1}1 = x - \tfrac{1}{8},\\
P_2(x) &= x^2 - \tfrac{(x^2,P_0)}{(P_0,P_0)}P_0(x) - \tfrac{(x^2,P_1)}{(P_1,P_1)}P_1(x) = x^2 - \tfrac{(x^2,1)}{(1,1)}1 - \tfrac{(x^2,x-\frac{1}{8})}{(x-\frac{1}{8},x-\frac{1}{8})}(x - \tfrac{1}{8})\\
&= x^2 - \tfrac{1\cdot1+0\cdot1+\frac{1}{4}\cdot1+1\cdot1}{1\cdot1+1\cdot1+1\cdot1+1\cdot1}1 - \tfrac{1\cdot(-\frac{9}{8})+0\cdot(-\frac{1}{8})+\frac{1}{4}\cdot\frac{3}{8}+1\cdot\frac{7}{8}}{(-\frac{9}{8})\cdot(-\frac{9}{8})+(-\frac{1}{8})\cdot(-\frac{1}{8})+\frac{3}{8}\cdot\frac{3}{8}+\frac{7}{8}\cdot\frac{7}{8}}(x - \tfrac{1}{8})\\
&= x^2 - \tfrac{9}{16} + \tfrac{1}{14}(x - \tfrac{1}{8}) = x^2 + \tfrac{1}{14}x - \tfrac{4}{7}.
\end{aligned}$$

Παρατηρούμε εδώ ότι το πρώτο παράδειγμα έδωσε αποτέλεσμα $P_1(x) = x$ και ένα άρτιο πολυώνυμο ως P_2, επειδή το σύνολο σημείων είναι συμμετρικό ως προς το μηδέν, ενώ το δεύτερο, που δεν ισχύει η συμμετρικότητα, δίνει διαφορετικό αποτέλεσμα. Ισχύει εδώ το γενικό συμπέρασμα ότι αν το σύνολο X_m είναι συμμετρικό ως προς το μηδέν, τότε το P_i είναι περιττό πολυώνυμο αν το i είναι περιττό και άρτιο αν το i είναι άρτιο.

Ορθογωνοποίηση με την αναδρομική σχέση τριών όρων

Ισχύει ακριβώς η θεωρία που αναπτύξαμε στη συνεχή περίπτωση. Η ορθογώνια βάση κατασκευάζεται από την αναδρομική σχέση των τριών όρων της (3.21):

$$P_{k+1}(x) = (x - \alpha_k)P_k(x) - \beta_k P_{k-1}(x), \quad k = 1, 2, \ldots, n-1,$$

όπου $P_0(x) = 1$ ενώ το $P_1(x)$ δίνεται από τη Gram-Schmidt ορθογωνοποίηση ή θεωρούμε ως $P_1(x) = x - \alpha_0$ και υπολογίζουμε το α_0 ώστε τα P_0 και P_1 να είναι ορθογώνια μεταξύ των. Η απαίτηση αυτή δίνει:

$$(P_1, P_0) = 0 \Leftrightarrow (x, 1) - \alpha_0(1, 1) \Leftrightarrow \alpha_0 = \frac{(x, 1)}{(1, 1)},$$

που είναι το αποτέλεσμα που πήραμε από τη Gram-Schmidt ορθογωνοποίηση. Παρατηρούμε εδώ ότι ο υπολογισμός του P_1, είναι και η μόνη διαφορά από τη συνεχή περίπτωση. Αυτό οφείλεται στο γεγονός ότι εκεί θεωρήσαμε το διάστημα $[-1, 1]$ που είναι συμμετρικό ως προς το μηδέν και δίνει πάντα ως αποτέλεσμα $P_1(x) = x$. Αν και εκεί θεωρούσαμε οποιοδήποτε διάστημα $[a, b]$, θα έπρεπε να κάνουμε αυτό που κάνουμε και εδώ για τον υπολογισμό του P_1.

Τα α_k και β_k δίνονται από τις σχέσεις (3.22) και (3.23) ή (3.24), αντίστοιχα. Έχουμε δηλαδή

$$\alpha_k = \frac{(xP_k, P_k)}{(P_k, P_k)}, \quad \beta_k = \frac{(xP_k, P_{k-1})}{(P_{k-1}, P_{k-1})} = \frac{(P_k, P_k)}{(P_{k-1}, P_{k-1})}.$$

Θα κατασκευάσουμε στη συνέχεια τις βάσεις $\{P_0, P_1, P_2\}$, που αντιστοιχούν στα δυο προηγούμενα παραδείγματα.

1) $X_5 = \{-1, -\frac{1}{2}, 0, \frac{1}{2}, 1\}$

$$\begin{aligned}
P_0(x) &= 1,\\
P_1(x) &= x,\\
\alpha_1 &= \tfrac{(xP_1,P_1)}{(P_1,P_1)} = \tfrac{(x^2,x)}{(x,x)} = \tfrac{1\cdot(-1)+\frac{1}{4}\cdot(-\frac{1}{2})+0\cdot 0+\frac{1}{4}\cdot\frac{1}{2}+1\cdot 1}{(-1)\cdot(-1)+(-\frac{1}{2})\cdot(-\frac{1}{2})+0\cdot 0+\frac{1}{2}\cdot\frac{1}{2}+1\cdot 1} = 0,\\
\beta_1 &= \tfrac{(P_1,P_1)}{(P_0,P_0)} = \tfrac{(x,x)}{(1,1)} = \tfrac{(-1)\cdot(-1)+(-\frac{1}{2})\cdot(-\frac{1}{2})+0\cdot 0+\frac{1}{2}\cdot\frac{1}{2}+1\cdot 1}{1\cdot 1+1\cdot 1+1\cdot 1+1\cdot 1+1\cdot 1} = \tfrac{1}{2},\\
P_2(x) &= (x-\alpha_1)P_1(x) - \beta_1 P_0(x) = x^2 - \tfrac{1}{2}.
\end{aligned}$$

2) $X_4 = \{-1, 0, \frac{1}{2}, 1\}$

$$\begin{aligned}
P_0(x) &= 1,\\
P_1(x) &= x - \tfrac{1}{8},\\
\alpha_1 &= \tfrac{(xP_1,P_1)}{(P_1,P_1)} = \tfrac{\left(x(x-\frac{1}{8}),x-\frac{1}{8}\right)}{\left(x-\frac{1}{8},x-\frac{1}{8}\right)} = \tfrac{\frac{9}{8}\cdot(-\frac{9}{8})+0\cdot(-\frac{1}{8})+\frac{3}{16}\cdot\frac{3}{8}+\frac{7}{8}\cdot\frac{7}{8}}{(-\frac{9}{8})\cdot(-\frac{9}{8})+(-\frac{1}{8})\cdot(-\frac{1}{8})+\frac{3}{8}\cdot\frac{3}{8}+\frac{7}{8}\cdot\frac{7}{8}} = -\tfrac{11}{56},\\
\beta_1 &= \tfrac{(P_1,P_1)}{(P_0,P_0)} = \tfrac{\left(x-\frac{1}{8},x-\frac{1}{8}\right)}{(1,1)} = \tfrac{(-\frac{9}{8})\cdot(-\frac{9}{8})+(-\frac{1}{8})\cdot(-\frac{1}{8})+\frac{3}{8}\cdot\frac{3}{8}+\frac{7}{8}\cdot\frac{7}{8}}{1\cdot 1+1\cdot 1+1\cdot 1+1\cdot 1+1\cdot 1} = \tfrac{35}{64},\\
P_2(x) &= (x-\alpha_1)P_1(x) - \beta_1 P_0(x) = \left(x+\tfrac{11}{56}\right)\left(x-\tfrac{1}{8}\right) - \tfrac{35}{64}1 = x^2 + \tfrac{1}{14}x - \tfrac{4}{7}.
\end{aligned}$$

Τέλος, θα λύσουμε τα προβλήματα των Παραδειγμάτων 19 και 20.

Παράδειγμα 21 Να βρεθεί η προσέγγιση ελαχίστων τετραγώνων στον P_2, της συνάρτησης $|x|$ στο σύνολο σημείων $X_5 = \{-1, -\frac{1}{2}, 0, \frac{1}{2}, 1\}$ με τη μέθοδο ορθογωνοποίησης.

Το σύνολο τριών ορθογώνιων πολυωνύμων είναι $\{P_0, P_1, P_2\} = \{1, x, x^2 - \frac{1}{2}\}$. Τα $\lambda_i,\ i = 0, 1, 2$, είναι

$$\begin{aligned}
\lambda_0 &= \tfrac{(P_0,f)}{(P_0,P_0)} = \tfrac{(1,|x|)}{(1,1)} = \tfrac{1\cdot 1+1\cdot\frac{1}{2}+1\cdot 0+1\cdot\frac{1}{2}+1\cdot 1}{1\cdot 1+1\cdot 1+1\cdot 1+1\cdot 1+1\cdot 1} = \tfrac{3}{5},\\
\lambda_1 &= \tfrac{(P_1,f)}{(P_1,P_1)} = \tfrac{(x,|x|)}{(x,x)} = \tfrac{(-1)\cdot 1+(-\frac{1}{2})\cdot\frac{1}{2}+0\cdot 0+\frac{1}{2}\cdot\frac{1}{2}+1\cdot 1}{(-1)\cdot(-1)+(-\frac{1}{2})\cdot(-\frac{1}{2})+0\cdot 0+\frac{1}{2}\cdot\frac{1}{2}+1\cdot 1} = 0,\\
\lambda_2 &= \tfrac{(P_2,f)}{(P_2,P_2)} = \tfrac{\left(x^2-\frac{1}{2},|x|\right)}{\left(x^2-\frac{1}{2},x^2-\frac{1}{2}\right)} = \tfrac{\frac{1}{2}\cdot 1+(-\frac{1}{4})\cdot\frac{1}{2}+(-\frac{1}{2})\cdot 0+(-\frac{1}{4})\cdot\frac{1}{2}+\frac{1}{2}\cdot 1}{\frac{1}{2}\cdot\frac{1}{2}+(-\frac{1}{4})\cdot(-\frac{1}{4})+(-\frac{1}{2})\cdot(-\frac{1}{2})+(-\frac{1}{4})\cdot(-\frac{1}{4})+\frac{1}{2}\cdot\frac{1}{2}} = \tfrac{6}{7}.
\end{aligned}$$

Τελικά, το πολυώνυμο ελαχίστων τετραγώνων θα είναι

$$q_n^*(x) = \lambda_0 P_0(x) + \lambda_1 P_1(x) + \lambda_2 P_2(x) = \frac{3}{5} + \frac{6}{7}\left(x^2 - \frac{1}{2}\right) = \frac{6}{7}x^2 + \frac{6}{35}.$$

Παράδειγμα 22 Να βρεθεί η προσέγγιση ελαχίστων τετραγώνων στον P_2, της συνάρτησης $|x|$ στο σύνολο σημείων $X_4 = \{-1, 0, \frac{1}{2}, 1\}$ με τη μέθοδο ορθογωνοποίησης.

Το σύνολο τριών ορθογώνιων πολυωνύμων είναι $\{P_0, P_1, P_2\} = \{1, x-\frac{1}{8}, x^2+\frac{1}{14}x-\frac{4}{7}\}$, και τα $\lambda_i,\ i = 0, 1, 2,$ θα είναι

$$\begin{aligned}
\lambda_0 &= \tfrac{(P_0,f)}{(P_0,P_0)} = \tfrac{(1,|x|)}{(1,1)} = \tfrac{1\cdot 1+1\cdot 0+1\cdot\frac{1}{2}+1\cdot 1}{1\cdot 1+1\cdot 1+1\cdot 1+1\cdot 1} = \tfrac{5}{8}, \\
\lambda_1 &= \tfrac{(P_1,f)}{(P_1,P_1)} = \tfrac{\left(x-\frac{1}{8},|x|\right)}{(x-\frac{1}{8},x-\frac{1}{8})} = \tfrac{(-\frac{9}{8})\cdot 1+(-\frac{1}{8})\cdot 0+\frac{3}{8}\cdot\frac{1}{2}+\frac{7}{8}\cdot 1}{(-\frac{9}{8})\cdot(-\frac{9}{8})+(-\frac{1}{8})\cdot(-\frac{1}{8})+\frac{3}{8}\cdot\frac{3}{8}+\frac{7}{8}\cdot\frac{7}{8}} = -\tfrac{1}{35}, \\
\lambda_2 &= \tfrac{(P_2,f)}{(P_2,P_2)} = \tfrac{\left(x^2+\frac{1}{14}x-\frac{4}{7},|x|\right)}{\left(x^2+\frac{1}{14}x-\frac{4}{7},x^2+\frac{1}{14}x-\frac{4}{7}\right)} = \tfrac{\frac{5}{14}\cdot 1+(-\frac{4}{7})\cdot 0+(-\frac{2}{7})\cdot\frac{1}{2}+\frac{1}{2}\cdot 1}{\frac{5}{14}\cdot\frac{5}{14}+(-\frac{4}{7})\cdot(-\frac{4}{7})+(-\frac{2}{7})\cdot(-\frac{2}{7})+\frac{1}{2}\cdot\frac{1}{2}} = \tfrac{10}{11}.
\end{aligned}$$

Τελικά, το πολυώνυμο ελαχίστων τετραγώνων θα είναι

$$\begin{aligned}
q_n^*(x) &= \lambda_0 P_0(x) + \lambda_1 P_1(x) + \lambda_2 P_2(x) = \tfrac{5}{8} - \tfrac{1}{35}\left(x - \tfrac{1}{8}\right) + \tfrac{10}{11}\left(x^2 + \tfrac{1}{14}x - \tfrac{4}{7}\right) \\
&= \tfrac{10}{11}x^2 + \tfrac{2}{55}x + \tfrac{6}{55}.
\end{aligned}$$

Βρήκαμε φυσικά τα ίδια αποτελέσματα με αυτά του συστήματος των κανονικών εξισώσεων.

Ασκήσεις

1. Δίνεται ότι οι συναρτήσεις $f_1, f_2 \in C[a, b]$ είναι ορθογώνιες μεταξύ των. Να αποδειχτεί η σχέση:
$$\|f_1 + f_2\|_2^2 = \|f_1\|_2^2 + \|f_2\|_2^2.$$

2. Δίνεται ότι οι συναρτήσεις $g_1, g_2 \in C[a, b]$ είναι κανονικοποιημένες ώστε $\|g_1\|_2 = \|g_2\|_2 = 1$. Να αποδειχτεί ότι οι συναρτήσεις $g_1 + g_2$ και $g_1 - g_2$ είναι ορθογώνιες μεταξύ των.

3. Να βρεθεί με οποιδήποτε τρόπο η προσέγγιση ελαχίστων τετραγώνων της $f(x) = x^3$, ορισμένης στο διάστημα $[0, 1]$, στον P_2.

4. Να βρεθεί με οποιδήποτε τρόπο η προσέγγιση ελαχίστων τετραγώνων της συνάρτησης $f(x) = \begin{cases} 1+x & , \quad x \in [-1, 0] \\ 1-x & , \quad x \in [0, 1] \end{cases}$, ορισμένης στο διάστημα $[-1, 1]$, στον P_2.

5. Να βρεθεί η προσέγγιση ελαχίστων τετραγώνων της $f(x) = e^x$, ορισμένης στο διάστημα $[-1, 1]$, στον P_2. (Να διατηρείτε κλάσματα και εκφράσεις του e στους υπολογισμούς).

6. Να βρεθεί η προσέγγιση ελαχίστων τετραγώνων της $f(x) = x^{\frac{1}{3}}$, ορισμένης στο διάστημα $[-1, 1]$, στον P_2 χρησιμοποιώντας το σύστημα των κανονικών εξισώσεων.

7. Να βρεθεί η προσέγγιση ελαχίστων τετραγώνων της $f(x) = \sqrt{|x|^3}$, ορισμένης στο διάστημα $[-1, 1]$, στον P_2 χρησιμοποιώντας ορθογώνια πολυώνυμα που παράγονται με την αναδρομική σχέση.

8. Να βρεθεί η προσέγγιση ελαχίστων τετραγώνων της $f(x) = x + |x|$, ορισμένης στο διάστημα $[-1, 1]$, στον P_2 χρησιμοποιώντας ορθογώνια πολυώνυμα Legendre.

9. Να βρεθεί η προσέγγιση ελαχίστων τετραγώνων της συνάρτησης x^4, ορισμένης στο σύνολο σημείων $X_4 = \{-2, -1, 1, 2\}$, στον P_2, χρησιμοποιώντας ορθογώνια πολυώνυμα που παράγονται με τη μέθοδο Gram Schmidt ορθογωνοποίησης.

10. Να βρεθεί η προσέγγιση ελαχίστων τετραγώνων της $f \in X_4$, στον P_2, όπου

$$X_4 = \{-1, 0, 1, 2\} \quad \text{και} \quad \begin{array}{c|cccc} x_i & -1 & 0 & 1 & 2 \\ \hline f_i & 1 & 0 & 1 & 2 \end{array},$$

χρησιμοποιώντας ορθογώνια πολυώνυμα που παράγονται με τη μέθοδο Gram Schmidt ορθογωνοποίησης.

11. Να βρεθεί η προσέγγιση ελαχίστων τετραγώνων της συνάρτησης $f \in X_5$, στον P_2, όπου

$$X_5 = \{-0, 1, 2, 3, 4\} \quad \text{και} \quad \begin{array}{c|ccccc} x_i & 0 & 1 & 2 & 3 & 4 \\ \hline f_i & 0 & 1 & -1 & 1 & -1 \end{array},$$

χρησιμοποιώντας ορθογώνια πολυώνυμα.

12. Να βρεθεί η προσέγγιση ελαχίστων τετραγώνων της $f \in X_5$, στον P_2, όπου

$$X_5 = \{-2, -1, 0, 1, 2\} \quad \text{και} \quad \begin{array}{c|ccccc} x_i & -2 & -1 & 0 & 1 & 2 \\ \hline f_i & 0 & 2 & 0 & 2 & 0 \end{array},$$

χρησιμοποιώντας ορθογώνια πολυώνυμα στο σύνολο X_5 που παράγονται από την αναδρομική σχέση.

13. Να βρεθεί με οποιδήποτε τρόπο η προσέγγιση ελαχίστων τετραγώνων της $f \in X_6$, στον P_2, όπου

$$X_6 = \{-2, -1, 0, 1, 2, 3\} \quad \text{και} \quad \begin{array}{c|cccccc} x_i & -2 & -1 & 0 & 1 & 2 & 3 \\ \hline f_i & 6 & 2 & 0 & 0 & 2 & 6 \end{array}.$$

Στη συνέχεια να βρεθεί η $\|\cdot\|_2$ του σφάλματος.

14. Δίνεται ότι q_n^* είναι η προσέγγιση ελαχίστων τετραγώνων της συνάρτησης $f \in C[a, b]$ στον P_n. Να βρεθεί η προσέγγιση ελαχίστων τετραγώνων της συνάρτησης $f + p_n$ στον P_n, όπου $p_n \in P_n$.

15. Δίνεται ότι q_1^* και q_2^* είναι οι προσεγγίσεις ελαχίστων τετραγώνων των συναρτήσεων $f_1 \in C[a, b]$ και $f_2 \in C[a, b]$ αντίστοιχα, στον P_n. Να βρεθεί η προσέγγιση ελαχίστων τετραγώνων της συνάρτησης $\alpha f_1 + \beta f_2$ στον P_n, όπου α και β είναι πραγματικές σταθερές.

Κεφάλαιο 4

ΠΑΡΕΜΒΟΛΗ ΜΕ ΣΥΝΑΡΤΗΣΕΙΣ SPLINE

Στην Πολυωνυμική Παρεμβολή κατά Lagrange ή με πεπερασμένες διαφορές, παρατηρούμε ότι όσο περισσότερα σημεία παίρνουμε, τόσο μεγαλύτερος θα είναι ο βαθμός του πολυωνύμου. Μας ενδιαφέρει να πάρουμε πολλά σημεία για καλύτερη ακρίβεια, αλλά τότε η μελέτη του προβλήματος γίνεται πιο πολύπλοκη αφού έχουμε να εργαστούμε με πολυώνυμα μεγάλου βαθμού. Ακόμη, και αυτή η ίδια η ακρίβεια δεν εξασφαλίζεται τελικά. Έχει αποδειχτεί ότι υπάρχουν συναρτήσεις όπου το σφάλμα της πολυωνυμικής παρεμβολής τείνει στο άπειρο όταν ο βαθμός του πολυωνύμου τείνει στο άπειρο. Αυτό μπορεί να εξηγηθεί γεωμετρικά από το γεγονός ότι τα πολυώνυμα μεγάλου βαθμού έχουν την τάση να ταλαντώνονται. Όσο πιο κοντά είναι τα σημεία παρεμβολής τόσο πιο έντονες είναι οι ταλαντώσεις με συνέπεια τα σφάλματα να μεγαλώνουν. Για να παρακαμφθεί αυτό το πρόβλημα έχει επινοηθεί η παρεμβολή με τμηματικά συνεχείς συναρτήσεις. Χωρίζουμε δηλαδή το διάσημα σε υποδιαστήματα και σε αυτά κάνουμε παρεμβολή με πολυώνυμα προκαθορισμένου βαθμού, αρκεί να εξασφαλίζεται η συνέχεια. Όπως θα δούμε παρακάτω, μπορούμε να κατασκευάσουμε τμηματικά συνεχείς συναρτήσεις απαιτώντας και κάποια ομαλότητα σε αυτές, δηλαδή να είναι συνεχείς και οι παράγωγοι μέχρι κάποιας τάξης.

4.1 Παρεμβολή με τμηματικά γραμμικές συναρτήσεις

Η πιο απλή από αυτές είναι η παρεμβολή με τμηματικά γραμμικές συναρτήσεις. Το μειονέκτημα αυτών είναι ότι δεν εξασφαλίζεται η ομαλότητα, για παράδειγμα στις τμηματικά γραμμικές συναρτήσεις δεν υπάρχει συνέχεια της πρώτης παραγώγου στην άκρη των υποδιαστημάτων. Η παρεμβολή Hermite εξασφαλίζει την ομαλότητα μέχρι και τις πρώτες παραγώγους.

Έστω $f \in C[a, b]$, την οποία επιθυμούμε να προσεγγίσουμε με τμηματικά συνεχείς συναρτήσεις. Θεωρούμε τη διαμέριση του διαστήματος $[a, b]$:

$$a = x_0 < x_1 < x_2 < \cdots < x_n = b.$$

Η τμηματικά συνεχής συνάρτηση που παρεμβάλεται στην f συνίσταται από τα πολυώνυμα παρεμβολής πρώτου βαθμού σε κάθε ένα υποδιάστημα $[x_i, x_{i+1}]$, $i = 0, 1, \ldots, n-1$. Είναι δηλαδή

$$p(x) = \frac{x - x_{i+1}}{x_i - x_{i+1}} f_i + \frac{x - x_i}{x_{i+1} - x_i} f_{i+1} = \frac{f_{i+1} - f_i}{h_i} x + \frac{x_{i+1} f_i - x_i f_{i+1}}{h_i}, \quad \begin{array}{c} x \in [x_i, x_{i+1}] \\ i = 0, 1, \ldots, n-1 \end{array},$$

όπου $h_i = x_{i+1} - x_i$, το μήκος του διαστήματος $[x_i, x_{i+1}]$.

Συμβολίζουμε με L_n το χώρο όλων των συνεχών τμηματικά γραμμικών συναρτήσεων που ορίζονται στο $[a, b]$ με διαμέριση τη $X_n = \{a = x_0, x_1, x_2, \ldots, x_n = b\}$. Η διάσταση αυτού του χώρου είναι $n + 1$ και υπολογίζεται εύκολα ως εξής. Η διάσταση των πολυωνύμων πρώτου βαθμού σε κάθε υποδιάστημα είναι 2. Αυτή πολλαπλασιάζεται επί n που είναι το πλήθος των υποδιαστημάτων και παίρνουμε διάσταση $2n$ του χώρου όλων των τμηματικά γραμμικών συναρτήσεων. Η απαίτηση της συνέχειας δίνει $n - 1$ περιορισμούς, έναν για κάθε ενδιάμεσο σημείο της διαμέρισης. Αφαιρώντας τους περιορισμούς, παίρνουμε $2n - (n - 1) = n + 1$ τη διάσταση του L_n. Ένας άλλος τρόπος υπολογισμού της διάστασης ενός χώρου βασίζεται στο γεγονός ότι όσες πληροφορίες απαιτούνται για τον υπολογισμό ενός στοιχείου του, τόση είναι και η διάστασή του. Στην περίπτωση του L_n, απαιτούνται οι $n+1$ τιμές της συνάρτησης σε όλα τα σημεία της διαμέρισης. Είναι εύκολο να δώσουμε και μια βάση του χώρου αυτού αλλά δε θα ασχοληθούμε περισσότερο.

4.2 Κυβική παρεμβολή Hermite

Έστω ότι $f \in C^1[a, b]$, όπου $C^1[a, b]$ είναι ο χώρος όλων των συνεχών συναρτήσεων στο $[a, b]$ με συνεχείς τις πρώτες παραγώγους στο (a, b). Θεωρούμε τη διάμεριση $a = x_0 < x_1 < x_2 < \cdots < x_n = b$. Έστω επίσης ότι f_i, f'_i, $i = 0, 1, \ldots, n$, είναι οι τιμές της f και της παραγώγου αντίστοιχα στα παραπάνω σημεία. Το γενικό πρόβλημα της παρεμβολής κατά Hermite, συνίσταται στο να βρούμε ένα πολυώνυμο h το οποίο θα έχει τις ίδιες τιμές με την f και τις ίδιες τιμές της παραγώγου της f σε όλα τα $n + 1$ σημεία παρεμβολής. Αποδεικνύεται ότι το πολυώνυμο παρεμβολής Hermite, είναι το πολύ $2n + 1$ βαθμού. Η διαδικασία υπολογισμού των πολυωνύμων αυτών είναι περισσότερο πολύπλοκη από εκείνη των πολυωνύμων παρεμβολής κατά Lagrange. Προκύπτει από την παρεμβολή κατά Lagrange χρησιμοποιώντας τη θεωρία των διηρημένων διαφορών. Δε θα ασχοληθούμε εδώ με το γενικό πρόβλημα, μας ενδιαφέρει κυρίως η παρεμβολή με τμηματικά κυβικές συναρτήσεις Hermite που θα αποτελέσει εργαλείο για τις κυβικές συναρτήσεις spline.

Η παρεμβολή με τμηματικά κυβικές συναρτήσεις Hermite, συνίσταται στην εύρεση της συνάρτησης h που αποτελείται από κυβικά πολυώνυμα $h_i(x)$ στο διαστήματα $[x_i, x_{i+1}]$, $i = 0, 1, \ldots, n - 1$, τέτοια ώστε

$$h(x_i) = f_i \;\text{ και }\; h'(x_i) = f'_i.$$

Λόγω του ορισμού της συνάρτησης h, εξασφαλίζεται και η συνέχεια της πρώτης παραγώγου αυτής στο (a,b), επομένως $h \in C^1[a,b]$.

Ας δούμε όμως τώρα πως θα υπολογίσουμε το κυβικό πολυώνυμο Hermite σε ένα απλό διάστημα $[a,b]$. Επιθυμούμε να προσδιορίσουμε ένα πολυώνυμο παρεμβολής στο $[a,b]$ έχοντας ως γνωστά τα $f(a)$, $f(b)$, $f'(a)$ και $f'(b)$. Η γνώση των παραγώγων σημαίνει ότι γνωρίζουμε τις κλίσεις στα σημεία a,b ή ισοδύναμα γνωρίζουμε την f σε γειτονικό σημείο του a, το $a+\epsilon$, και σε γειτονικό σημείο του b, το $b-\eta$, με $\epsilon,\eta>0$ να τείνουν στο 0. Θα θεωρήσουμε επομένως ότι έχουμε τα σημεία $a,\ a+\epsilon,\ b-\eta,\ b$. Προσδιορίζουμε το πολυώνυμο παρεμβολής κατά Lagrange και παίρνουμε όρια για $\epsilon,\ \eta>0$ να τείνουν στο 0. Τα σημεία παρεμβολής είναι 4 επομένως το πολυώνυμο παρεμβολής Lagrange θα είναι το πολύ τρίτου βαθμού. Αυτό θα είναι:

$$\begin{aligned} p_3(x) &= \tfrac{(x-a-\epsilon)(x-b+\eta)(x-b)}{-\epsilon(a-b+\eta)(a-b)} f(a) + \tfrac{(x-a)(x-b+\eta)(x-b)}{\epsilon(a+\epsilon-b+\eta)(a+\epsilon-b)} f(a+\epsilon) \\ &+ \tfrac{(x-a)(x-a-\epsilon)(x-b)}{(b-\eta-a)(b-\eta-a-\epsilon)(-\eta)} f(b-\eta) + \tfrac{(x-a)(x-a-\epsilon)(x-b+\eta)}{(b-a)(b-a-\epsilon)\eta} f(b). \end{aligned} \tag{4.1}$$

Θα προσδιορίσουμε, για απλότητα, τους δύο πρώτους όρους του παραπάνω αθροίσματος. Με την ίδια τεχνική θα προσδιοριστούν και οι δύο τελευταίοι οι οποίοι θα δώσουν συμμετρική έκφραση. Έστω s_1 το άθροισμα των δύο πρώτων όρων. Αυτό, μετά από κάποιες πράξεις, λαμβάνει τη μορφή:

$$s_1 = (x-a)(x-b+\eta)(x-b)\frac{\frac{f(a+\epsilon)}{(a+\epsilon-b+\eta)(a+\epsilon-b)} - \frac{f(a)}{(a-b+\eta)(a-b)}}{\epsilon} + \frac{(x-b+\eta)(x-b)}{(a-b+\eta)(a-b)} f(a).$$

Παρατηρούμε ότι όλα τα ϵ εμφανίζονται στο κλάσμα του πρώτου όρου. Διατηρούμε το η σταθερό και παίρνουμε όριο του κλάσματος αυτού για ϵ τείνοντος στο 0. Αυτό δίνει

$$\lim_{\epsilon\to 0} \frac{\frac{f(a+\epsilon)}{(a+\epsilon-b+\eta)(a+\epsilon-b)} - \frac{f(a)}{(a-b+\eta)(a-b)}}{\epsilon} = \left[\frac{f(a)}{(a-b+\eta)(a-b)}\right]',$$

αφού το παραπάνω όριο μας δίνει τον ορισμό της παραγώγου της συνάρτησης $\frac{f(a)}{(a-b+\eta)(a-b)}$ θεωρούμενης ως συνάρτησης του a. Έτσι, το όριο του s_1 για ϵ τείνοντος στο 0 λαμβάνει τη μορφή

$$\begin{aligned} \lim_{\epsilon\to 0} s_1 &= (x-a)(x-b+\eta)(x-b)\left[\tfrac{f(a)}{(a-b+\eta)(a-b)}\right]' + \tfrac{(x-b+\eta)(x-b)}{(a-b+\eta)(a-b)} f(a) \\ &= \tfrac{(x-a)(x-b+\eta)(x-b)}{(a-b+\eta)a-b)}\left[f'(a) - \tfrac{f(a)}{a-b+\eta} - \tfrac{f(a)}{a-b}\right] + \tfrac{(x-b+\eta)(x-b)}{(a-b+\eta)(a-b)} f(a). \end{aligned}$$

Στη συνέχεια θεωρούμε το όριο για $\eta>0$ να τείνει στο 0, οπότε μετά από πράξεις λαμβάνουμε

$$\lim_{\epsilon,\eta\to 0} s_1 = \left[\frac{(x-b)^2}{(b-a)^2} + 2\frac{(x-a)(x-b)^2}{(b-a)^3}\right] f(a) + \frac{(x-a)(x-b)^2}{(b-a)^2} f'(a).$$

Πρέπει να σημειώσουμε εδώ ότι το ίδιο αποτέλεσμα θα προκύψει αν θεωρήσουμε πρώτα το όριο ως προς η και στη συνέχεια ως προς ϵ. Επειδή τα ϵ και η αποτελούν μικρές μεταβολές σε διαφορετικά σημεία, στα a και b, αντίστοιχα, δεν επηρεάζεται το τελικό αποτέλεσμα με οποιαδήποτε σειρά και αν ληφθούν τα όρια. Το άθροισμα s_2 του τρίτου και τέταρτου όρου της (4.1) δίνει συμμετρική έκφραση ως προς a και b, που αν ακολουθήσουμε την ίδια διαδικασία λαμβάνουμε

$$\lim_{\epsilon,\eta\to 0} s_2 = \left[\frac{(x-a)^2}{(b-a)^2} - 2\frac{(x-b)(x-a)^2}{(b-a)^3}\right] f(b) + \frac{(x-b)(x-a)^2}{(b-a)^2} f'(b).$$

Όταν λέμε συμμετρική έκφραση ως προς a και b, εννοούμε ότι αν στον έναν τύπο εναλλάξουμε τους ρόλους των a και b, τότε προκύπτει ο άλλος. Παρατηρούμε ότι αυτό συμβαίνει στους δυο παραπάνω τύπους. Τελικά, η συνάρτηση Hermite h, που είναι το όριο του πολυωνύμου p_3, προκύπτει από το άθροισμα των δυο παραπάνω εκφράσεων.

$$\begin{aligned} h(x) &= \left[\tfrac{(x-b)^2}{(b-a)^2} + 2\tfrac{(x-a)(x-b)^2}{(b-a)^3}\right] f(a) + \tfrac{(x-a)(x-b)^2}{(b-a)^2} f'(a) \\ &+ \left[\tfrac{(x-a)^2}{(b-a)^2} - 2\tfrac{(x-b)(x-a)^2}{(b-a)^3}\right] f(b) + \tfrac{(x-b)(x-a)^2}{(b-a)^2} f'(b). \end{aligned} \tag{4.2}$$

Επιστρέφοντας στη διαμέριση $a = x_0 < x_1 < x_2 < \cdots < x_n = b$, μπορούμε να υπολογίσουμε την τμηματική συνάρτηση Hermite από την παραπάνω έκφραση αρκεί να αντικαταστήσουμε με x_i και x_{i+1} τα a και b, αντίστοιχα, για όλα τα $i = 0, 1, \ldots, n-1$.

$$\begin{aligned} h(x) &= \left[\tfrac{(x-x_{i+1})^2}{(\Delta x_i)^2} + 2\tfrac{(x-x_i)(x-x_{i+1})^2}{(\Delta x_i)^3}\right] f_i + \left[\tfrac{(x-x_i)(x-x_{i+1})^2}{(\Delta x_i)^2}\right] f'_i \\ &+ \left[\tfrac{(x-x_i)^2}{(\Delta x_i)^2} - 2\tfrac{(x-x_{i+1})(x-x_i)^2}{(\Delta x_i)^3}\right] f_{i+1} + \left[\tfrac{(x-x_{i+1})(x-x_i)^2}{(\Delta x_i)^2}\right] f'_{i+1}, \ x \in [x_i, x_{i+1}], \end{aligned} \tag{4.3}$$

όπου $\Delta x_i = x_{i+1} - x_i$.

Έστω ότι H_n είναι ο χώρος όλων των τμηματικών συναρτήσεων Hermite που ορίζονται στο $[a, b]$ με διαμέριση τη $X_n = \{a = x_0, x_1, x_2, \ldots, x_n = b\}$. Όπως και στην περίπτωση των τμηματικά γραμμικών συναρτήσεων, υπολογίζεται η διάσταση αυτού του χώρου. Η διάσταση των πολυωνύμων τρίτου βαθμού σε κάθε υποδιάστημα είναι 4. Αυτή πολλαπλασιάζεται επί n που είναι το πλήθος των υποδιαστημάτων και παίρνουμε διάσταση $4n$ του χώρου όλων των τμηματικά κυβικών συναρτήσεων. Η απαίτηση της συνέχειας και της συνέχειας της πρώτης παραγώγου δίνουν $2(n-1)$ περιορισμούς, δυο για κάθε ενδιάμεσο σημείο της διαμέρισης. Αφαιρώντας τους περιορισμούς, παίρνουμε $4n - 2(n-1) = 2n+2$ τη διάσταση του H_n. Αλλιώς, η διάσταση του H_n δίνεται από τον αριθμό των πληροφοριών που απαιτούνται για τον υπολογισμό της h. Αυτές είναι οι $n+1$ τιμές της συνάρτησης και οι $n+1$ τιμές της παραγώγου σε όλα τα σημεία της διαμέρισης, σύνολο $2n+2$ πληροφορίες.

Παράδειγμα 23 Να βρεθεί η κυβική συνάρτηση Hermite της συνάρτησης $f(x) = |x^5|$, στο διάστημα $[-1, 1]$, με σημεία διαμέρισης $x_0 = -1, \quad x_1 = 0, \quad x_2 = 1$.

Οι τιμές της συνάρτησης στα σημεία αυτά είναι

$$f_0 = f(-1) = |(-1)^5| = 1, \quad f_1 = f(0) = |0^5| = 0, \quad f_2 = f(1) = |1^5| = 1.$$

Η παράγωγος της f είναι

$$f'(x) = \begin{cases} [-x^5]' = -5x^4, & x \leq 0 \\ [x^5]' = 5x^4, & x \geq 0 \end{cases}$$

που δίνει τις τιμές στα σημεία διακριτοποίησης

$$f'_0 = f'(-1) = -5, \quad f'_1 = f'(0) = 0, \quad f'_2 = f'(1) = 5.$$

Έχουμε λοιπόν για τη συνάρτηση f τον πίνακα τιμών

x_i	-1	0	1
f_i	1	0	1
f'_i	-5	0	5

Υπολογίζουμε στη συνέχεια την h από τον τύπο (4.2) αντικαθιστώντας πρώτα με $a = -1,\ b = 0$ για το διάστημα $[-1, 0]$ και στη συνέχεια με $a = 0,\ b = 1$ για το διάστημα $[0, 1]$. Τελικά λαμβάνουμε

$$h(x) = \begin{cases} x^2 + 2(x+1)x^2 - 5(x+1)x^2 = -3x^3 - 2x^2, & x \in [-1, 0] \\ x^2 - 2(x-1)x^2 + 5(x-1)x^2 = 3x^3 - 2x^2, & x \in [0, 1] \end{cases}.$$

Παρατηρούμε στο παράδειγμα αυτό ότι το συνολικό διάστημα είναι συμμετρικό ως προς το μηδέν και τα σημεία διαμέρισης είναι συμμετρικά ως προς το μηδέν. Η συνάρτηση $f(x) = |x^5|$ είναι άρτια συνάρτηση, θα εξετάσουμε αν και η h που βρήκαμε είναι άρτια. Θεωρούμε $x > 0$, τότε αφού $-x < 0$ θα έχουμε $h(-x) = -3(-x)^3 - 2(-x)^2 = 3x^3 - 2x^2 = h(x)$. Στο ίδιο συμπέρασμα καταλήγουμε αν θεωρήσουμε $x < 0$. Ισχύει επομένως και στην παρεμβολή Hermite ο ίδιος νόμος που ισχύει και στην προσέγγιση, όπως είδαμε στα προηγούμενα κεφάλαια, σχετικά με τη συμμετρικότητα.

4.3 Παρεμβολή με κυβικές συναρτήσεις Spline

Θεωρούμε τη διαμέριση $X_n = \{x_0, x_1, x_2, \ldots, x_n\}$ όπου: $a = x_0 < x_1 < x_2 < \cdots < x_n = b$ και την f ορισμένη και συνεχή στο $[a, b]$. Έστω ότι $S(X_n)$ είναι το σύνολο όλων των συναρτήσεων $S(X_n; x) = s(x) \quad \in C^2[a, b]$, με την ιδιότητα το s να είναι ένα πολυώνυμο το πολύ τρίτου βαθμού σε κάθε υποδιάστημα $[x_i, x_{i+1}],\ i = 0, 1, \ldots, n-1$. Αυτές τις συναρτήσεις τις καλούμε κυβικές splines. Στόχος μας είναι να προσδιορίσουμε τις προϋποθέσεις με τις οποίες μπορούμε να παρεμβάλουμε μια συνάρτηση $f \in C[a, b]$ με μια κυβική spline. Αυτό επιτυγχάνεται με το ακόλουθο θεώρημα.

Θεώρημα 16 *Δοθέντων των αριθμών s'_0 και s'_n υπάρχει μια μοναδική κυβική συνάρτησηspline τέτοια ώστε*

$$s(f, X_n; x_i) = f_i$$

και

$$s'(f, X_n; a) = s'_0, \quad s'(f, X_n; b) = s'_n.$$

Απόδειξη: Θα αποδείξουμε το θεώρημα προσπαθώντας να κατασκευάσουμε μια τέτοια spline. Θα υποθέσουμε πρώτα ότι s'_i είναι οι τιμές της παραγώγου της κυβικής spline στους εσωτερικούς κόμβους $x_i,\ i = 1, 2, \ldots, n-1$. Στη συνέχεια θα υπολογίσουμε το πολυώνυμο παρεμβολής Hermite $s(x)$ με παραμέτρους τα s'_i. Τέλος, θα προσδιορίσουμε τα s'_i αφού απαιτήσαμε οι πλευρικές δεύτερες παράγωγοι στους κόμβους $x_i,\ i = 1, 2, \ldots, n-1$ να ταυτίζονται. Το πολυώνυμο παρεμβολής Hermite στα διαστήματα $[x_i, x_{i+1}], i = 0, 1, ..., n-1$ σύμφωνα με την (4.3) θα είναι

$$\begin{aligned} s(x) &= \left[\frac{(x-x_{i+1})^2}{(\Delta x_i)^2} + 2\frac{(x-x_i)(x-x_{i+1})^2}{(\Delta x_i)^3}\right] f_i + \left[\frac{(x-x_i)(x-x_{i+1})^2}{(\Delta x_i)^2}\right] s'_i \\ &+ \left[\frac{(x-x_i)^2}{(\Delta x_i)^2} - 2\frac{(x-x_{i+1})(x-x_i)^2}{(\Delta x_i)^3}\right] f_{i+1} + \left[\frac{(x-x_{i+1})(x-x_i)^2}{(\Delta x_i)^2}\right] s'_{i+1},\ x \in [x_i, x_{i+1}]. \end{aligned} \tag{4.4}$$

Παραγωγίζοντας δυο φορές αυτήν ως προς x λαμβάνουμε

$$\begin{aligned} s''(x) &= \left[\frac{2}{(\Delta x_i)^2} + 2\frac{6x-2(x_i+2x_{i+1})}{(\Delta x_i)^3}\right] f_i + \left[\frac{6x-2(x_i+2x_{i+1})}{(\Delta x_i)^2}\right] s'_i \\ &+ \left[\frac{2}{(\Delta x_i)^2} - 2\frac{6x-2(2x_i+x_{i+1})}{(\Delta x_i)^3}\right] f_{i+1} + \left[\frac{6x-2(2x_i+x_{i+1})}{(\Delta x_i)^2}\right] s'_{i+1}. \end{aligned} \tag{4.5}$$

Θέτοντας όπου x τα x_i και x_{i+1} λαμβάνουμε αντίστοιχα

$$\begin{aligned} s''(x_i) &= -\frac{6}{(\Delta x_i)^2} f_i + \frac{6}{(\Delta x_i)^2} f_{i+1} - \frac{4}{\Delta x_i} s'_i - \frac{2}{\Delta x_i} s'_{i+1} \\ &= \frac{6}{(\Delta x_i)^2}\Delta f_i - \frac{4}{\Delta x_i} s'_i - \frac{2}{\Delta x_i} s'_{i+1} \end{aligned} \tag{4.6}$$

και

$$\begin{aligned} s''(x_{i+1}) &= \frac{6}{(\Delta x_i)^2} f_i - \frac{6}{(\Delta x_i)^2} f_{i+1} + \frac{2}{\Delta x_i} s'_i + \frac{4}{\Delta x_i} s'_{i+1} \\ &= -\frac{6}{(\Delta x_i)^2}\Delta f_i + \frac{2}{\Delta x_i} s'_i + \frac{4}{\Delta x_i} s'_{i+1}, \end{aligned} \tag{4.7}$$

όπου θέσαμε $\Delta f_i = f_{i+1} - f_i$. Με βάση την (4.7) υπολογίζουμε την αριστερή πλευρική παράγωγο $s''(x_i)$ θεωρώντας ως διάστημα το $[x_{i-1}, x_i]$, μειώνοντας απλώς τους δείκτες κατά ένα. Προκύπτει έτσι ότι

$$s''(x_i) = -\frac{6}{(\Delta x_{i-1})^2}\Delta f_{i-1} + \frac{2}{\Delta x_{i-1}} s'_{i-1} + \frac{4}{\Delta x_{i-1}} s'_i. \tag{4.8}$$

Η (4.6) δίνει τη δεξιά πλευρική δεύτερη παράγωγο στο x_i ενώ η (4.8) την αριστερή στο ίδιο σημείο. Επειδή επιθυμούμε $s \in C^2[a, b]$ πρέπει να εξισώσουμε τις δυο πλευρικές δεύτερες παραγώγους, οπότε λαμβάνουμε την εξίσωση

$$\frac{6}{(\Delta x_i)^2}\Delta f_i - \frac{4}{\Delta x_i} s'_i - \frac{2}{\Delta x_i} s'_{i+1} = -\frac{6}{(\Delta x_{i-1})^2}\Delta f_{i-1} + \frac{2}{\Delta x_{i-1}} s'_{i-1} + \frac{4}{\Delta x_{i-1}} s'_i. \tag{4.9}$$

Πολλαπλασιάζουμε στη συνέχεια αυτήν κατά μέλη με την ποσότητα $\frac{\Delta x_{i-1}\Delta x_i}{2}$, μεταφέρουμε τους αγνώστους $s'_{i-1},\ s'_i$ και s'_{i+1} στο πρώτο μέλος και όλους τους γνωστούς όρους στο δεύτερο. Μετά από απλές πράξεις λαμβάνουμε

$$\Delta x_i s'_{i-1} + 2(\Delta x_{i-1} + \Delta x_i) s'_i + \Delta x_{i-1} s'_{i+1} = 3\left[\frac{\Delta x_i}{\Delta x_{i-1}}\Delta f_{i-1} + \frac{\Delta x_{i-1}}{\Delta x_i}\Delta f_i\right]. \tag{4.10}$$

Αυτή η εξίσωση ισχύει για όλα τα εσωτερικά σημεία, δηλαδή για όλα τα $i = 1, 2, ..., n-1$. Επομένως, καταλήγουμε σε ένα γραμμικό σύστημα $n-1$ εξισώσεων με $n-1$ αγνώστους, τα s'_i, $i = 1, 2, ..., n-1$. Για $i = 1$, ο πρώτος όρος του πρώτου μέλους της (4.10), $\Delta x_1 s'_0$, είναι γνωστός αφού $s'_0 = f'_0 = f'(a)$ από τις υποθέσεις του θεωρήματος, και μεταφέρεται στο δεύτερο μέλος. Το ίδιο συμβαίνει για $i = n-1$ και για τον τελευταίο όρο του πρώτου μέλους της (4.10), αφού $s'_n = f'_n = f'(b)$. Καταλήξαμε λοιπόν στο γραμμικό σύστημα

$$Av = q, \tag{4.11}$$

όπου

$$A = \begin{bmatrix} 2(\Delta x_0 + \Delta x_1) & \Delta x_0 & & & & \\ \Delta x_2 & 2(\Delta x_1 + \Delta x_2) & \Delta x_1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & & \Delta x_{n-1} & 2(\Delta x_{n-2} + \Delta x_{n-1}) \end{bmatrix},$$

$$v = \begin{bmatrix} s'_1 \\ s'_2 \\ \vdots \\ s'_{n-1} \end{bmatrix} \quad \text{και} \quad q = \begin{bmatrix} 3\left[\frac{\Delta x_1}{\Delta x_0}\Delta f_0 + \frac{\Delta x_0}{\Delta x_1}\Delta f_1\right] - \Delta x_1 f'(a) \\ 3\left[\frac{\Delta x_2}{\Delta x_1}\Delta f_1 + \frac{\Delta x_1}{\Delta x_2}\Delta f_2\right] \\ \vdots \\ 3\left[\frac{\Delta x_{n-1}}{\Delta x_{n-2}}\Delta f_{n-2} + \frac{\Delta x_{n-2}}{\Delta x_{n-1}}\Delta f_{n-1}\right] - \Delta x_{n-2} f'(b) \end{bmatrix}.$$

Ο πίνακας A είναι τριδιαγώνιος, δηλαδή έχει μη μηδενικά στοιχεία μόνο στη διαγώνιο και στις δυο γειτονικές δευτερεύουσες διαγωνίους. Είναι μη αρνητικός, αφού όλα τα στοιχεία του είναι θετικά ή μηδέν, και αυστηρά διαγώνια υπέρτερος, αφού κάθε διαγώνιο στοιχείο του είναι τουλάχιστον διπλάσιο από το άθροισμα των υπόλοιπων στοιχείων της γραμμής. Από τη γραμμική άλγεβρα γνωρίζουμε ότι αυτός είναι αντιστρέψιμος, οπότε υπάρχει μοναδική λύση. Υπολογίζοντας λοιπόν από το σύστημα αυτό τα s'_i, $i = 1, 2, ..., n-1$, έχουμε υπολογίσει και την κυβική συνάρτηση spline $s(f, X_n; x) = s(x)$ που είναι μοναδικό και το θεώρημα αποδείχτηκε.•

Η απόδειξη του θεωρήματος δίνει ουσιαστικά και τον αλγόριθμο κατασκευής της κυβικής spline. Από το γραμμικό σύστημα (4.10) υπολογίζουμε τα s'_i και στη συνέχεια από τη σχέση (4.4) υπολογίζουμε την κυβική συνάρτηση spline. Δίνουμε εδώ δύο παραδείγματα υπολογισμού της κυβικής συνάρτησης spline.

Παράδειγμα 24 Να βρεθεί η κυβική συνάρτηση spline της συνάρτησης $f(x) = |x|$, στο διάστημα $[-1, 1]$, με διαμέριση την $X_2 = \{-1, 0, 1\}$.

Ο πίνακας τιμών της συνάρτησης στη διαμέριση X_2 είναι

$$\begin{array}{c|ccc} x_i & -1 & 0 & 1 \\ \hline f_i & 1 & 0 & 1 \end{array}.$$

Η παράγωγος της f είναι

$$f'(x) = \left\{ \begin{array}{ll} [-x]' = -1, & x \leq 0 \\ [x]' = 1, & x \geq 0 \end{array} \right. .$$

Επομένως, οι τιμές της παραγώγου στα άκρα του διαστήματος είναι

$$f_0' = f'(-1) = -1, \quad f_2' = f'(1) = 1.$$

Έχουμε όλα τα δεδομένα και προχωρούμε στον υπολογισμό της κυβικής spline s. Επειδή $n = 2$ το γραμμικό σύστημα (4.11) εκφυλίζεται σε μια εξίσωση με άγνωστο το s_1', την

$$2(\Delta x_0 + \Delta x_1)s_1' = 3\left[\frac{\Delta x_1}{\Delta x_0}\Delta f_0 + \frac{\Delta x_0}{\Delta x_1}\Delta f_1\right] - \Delta x_1 f'(-1) - \Delta x_0 f'(1).$$

Αντικαθιστώντας τα

$$\Delta x_0 = \Delta x_1 = 1, \quad \Delta f_0 = -1, \quad \Delta f_1 = 1,$$

λαμβάνουμε

$$4s_1' = 0 \Leftrightarrow s_1' = 0.$$

Στη συνέχεια κάνουμε χρήση της έκφρασης (4.4) για τον υπολογισμό του s στα διαστήματα $[-1, 0]$ και $[0, 1]$. Τελικά λαμβάνουμε

$$s(x) = \left\{ \begin{array}{ll} (x^2 + 2(x+1)x^2)\,1 + (x+1)x^2(-1) = x^3 + 2x^2, & x \in [-1, 0] \\ (x^2 - 2(x-1)x^2)\,1 + (x-1)x^2 1 = -x^3 + 2x^2, & x \in [0, 1] \end{array} \right. .$$

Παρατηρούμε και στο παράδειγμα αυτό ότι επειδή το συνολικό διάστημα είναι συμμετρικό ως προς το μηδέν και τα σημεία διαμέρισης είναι συμμετρικά ως προς το μηδέν, η κυβική spline s βρέθηκε άρτια, αφού η συνάρτηση $f(x) = |x|$ είναι άρτια.

Είναι εύκολο να επιβεβαιώσουμε από το αποτέλεσμα ότι η συνάρτηση s ανήκει στο χώρο $C^2[-1, 1]$. Εδώ έχουμε να παρατηρήσουμε ότι ενώ η συνάρτηση $f(x) = |x|$ ανήκει απλά στον $C[-1, 1]$, δεν είναι καν παραγωγίσιμη, την προσεγγίσαμε με μια συνάρτηση που βρίσκεται στον $C^2[-1, 1]$, δηλαδή με μια αρκετά ομαλή συνάρτηση. Συνηθίζεται γενικά στη Θεωρία Προσέγγισης, να προσεγγίζουμε μη ομαλές συναρτήσεις με ομαλές, ώστε να μελετάται πιο εύκολα το πρόβλημα. Στη συνάρτηση αυτή δε θα μπορούσαμε να εφαρμόσουμε παρεμβολή με τμηματικές συναρτήσεις Hermite, όπως έγινε στο Παράδειγμα 23 για την $|x^5|$, επειδή δεν υπάρχει η παράγωγος στο σημείο 0.

Θα δούμε τώρα, στο επόμενο παράδειγμα, πως λειτουργεί ο αλγόριθμος για την ίδια συνάρτηση αλλά με περισσότερα σημεία στη διαμέριση.

Παράδειγμα 25 Να βρεθεί η κυβική συνάρτηση spline της συνάρτησης $f(x) = |x|$, στο διάστημα $[-1, 1]$, με διαμέριση την $X_2 = \{-1, -\frac{1}{2}, 0, \frac{1}{2}, 1\}$.

Τα δεδομένα του προβλήματος είναι

$$\begin{array}{c|ccccc} x_i & -1 & -\frac{1}{2} & 0 & \frac{1}{2} & 1 \\ \hline f_i & 1 & \frac{1}{2} & 0 & \frac{1}{2} & 1 \end{array}, f'(-1) = -1, \quad f'(1) = 1.$$

Το n στην περίπτωσή μας είναι 4 οπότε το γραμμικό σύστημα (4.11) δίνει το 3×3 σύστημα

$$\begin{bmatrix} 2(\Delta x_0 + \Delta x_1) & \Delta x_0 & 0 \\ \Delta x_2 & 2(\Delta x_1 + \Delta x_2) & \Delta x_1 \\ 0 & \Delta x_3 & 2(\Delta x_2 + \Delta x_3) \end{bmatrix} \begin{bmatrix} s_1' \\ s_2' \\ s_3' \end{bmatrix}$$

$$= \begin{bmatrix} 3\left[\frac{\Delta x_1}{\Delta x_0}\Delta f_0 + \frac{\Delta x_0}{\Delta x_1}\Delta f_1\right] - \Delta x_1 f'(-1) \\ 3\left[\frac{\Delta x_2}{\Delta x_1}\Delta f_1 + \frac{\Delta x_1}{\Delta x_2}\Delta f_2\right] \\ 3\left[\frac{\Delta x_3}{\Delta x_2}\Delta f_2 + \frac{\Delta x_2}{\Delta x_3}\Delta f_3\right] - \Delta x_2 f'(1) \end{bmatrix}.$$

Αντικαθιστώντας τα

$$\Delta x_0 = \Delta x_1 = \Delta x_2 = \Delta x_2 = \frac{1}{2}, \quad \Delta f_0 = \Delta f_1 = -\frac{1}{2}, \quad \Delta f_2 = \Delta f_3 = \frac{1}{2},$$

λαμβάνουμε

$$\begin{bmatrix} 2 & \frac{1}{2} & 0 \\ \frac{1}{2} & 2 & \frac{1}{2} \\ 0 & \frac{1}{2} & 2 \end{bmatrix} \begin{bmatrix} s_1' \\ s_2' \\ s_3' \end{bmatrix} = \begin{bmatrix} -\frac{5}{2} \\ 0 \\ \frac{5}{2} \end{bmatrix}.$$

Επιλύουμε το σύστημα αυτό με τη μέθοδο απαλοιφής του Gauss ή με οποιαδήποτε άλλη μέθοδο και λαμβάνουμε τη λύση

$$s_1' = -\frac{5}{4}, \quad s_2' = 0, \quad s_3' = \frac{5}{4}.$$

Στη συνέχεια από την (4.4) υπολογίζουμε το s σε όλα τα υποδιαστήματα. Στο διάστημα $[-1, -\frac{1}{2}]$ αυτή δίνει:

$$\begin{aligned} s(x) &= \left[\frac{\left(x+\frac{1}{2}\right)^2}{\left(\frac{1}{2}\right)^2} + 2\frac{(x+1)\left(x+\frac{1}{2}\right)^2}{\left(\frac{1}{2}\right)^3}\right] 1 + \frac{(x+1)\left(x+\frac{1}{2}\right)^2}{\left(\frac{1}{2}\right)^2}(-1) \\ &+ \left[\frac{(x+1)^2}{\left(\frac{1}{2}\right)^2} - 2\frac{\left(x+\frac{1}{2}\right)(x+1)^2}{\left(\frac{1}{2}\right)^3}\right] \frac{1}{2} + \frac{\left(x+\frac{1}{2}\right)(x+1)^2}{\left(\frac{1}{2}\right)^2}\left(-\frac{5}{4}\right) \\ &= -x^3 - \frac{5}{2}x^2 - 3x - \frac{1}{2}, \end{aligned}$$

στο διάστημα $[-\frac{1}{2}, 0]$:

$$\begin{aligned} s(x) &= \left[\frac{x^2}{\left(\frac{1}{2}\right)^2} + 2\frac{\left(x+\frac{1}{2}\right)x^2}{\left(\frac{1}{2}\right)^3}\right] \frac{1}{2} + \frac{\left(x+\frac{1}{2}\right)x^2}{\left(\frac{1}{2}\right)^2}\left(-\frac{5}{4}\right) \\ &= 3x^3 + \frac{7}{2}x^2, \end{aligned}$$

στο διάστημα $[0, \frac{1}{2}]$:

$$\begin{aligned} s(x) &= \left[\frac{x^2}{\left(\frac{1}{2}\right)^2} - 2\frac{\left(x-\frac{1}{2}\right)x^2}{\left(\frac{1}{2}\right)^3}\right]\frac{1}{2} + \frac{\left(x-\frac{1}{2}\right)x^2}{\left(\frac{1}{2}\right)^2}\frac{5}{4} \\ &= -3x^3 + \frac{7}{2}x^2, \end{aligned}$$

και τέλος, στο διάστημα $[\frac{1}{2}, 1]$:

$$\begin{aligned} s(x) &= \left[\frac{(x-1)^2}{\left(\frac{1}{2}\right)^2} + 2\frac{\left(x-\frac{1}{2}\right)(x-1)^2}{\left(\frac{1}{2}\right)^3}\right]\frac{1}{2} + \frac{\left(x-\frac{1}{2}\right)(x-1)^2}{\left(\frac{1}{2}\right)^2}\frac{5}{4} \\ &+ \left[\frac{\left(x-\frac{1}{2}\right)^2}{\left(\frac{1}{2}\right)^2} - 2\frac{(x-1)\left(x-\frac{1}{2}\right)^2}{\left(\frac{1}{2}\right)^3}\right]1 + \frac{(x-1)\left(x-\frac{1}{2}\right)^2}{\left(\frac{1}{2}\right)^2}1 \\ &= x^3 - \frac{5}{2}x^2 + 3x - \frac{1}{2}. \end{aligned}$$

Τελικά, η κυβική συνάρτηση spline s είναι

$$s(x) = \begin{cases} -x^3 - \frac{5}{2}x^2 - 3x - \frac{1}{2}, & x \in [-1, -\frac{1}{2}] \\ 3x^3 + \frac{7}{2}x^2, & x \in [-\frac{1}{2}, 0] \\ -3x^3 + \frac{7}{2}x^2, & x \in [0, \frac{1}{2}] \\ x^3 - \frac{5}{2}x^2 + 3x - \frac{1}{2}, & x \in [\frac{1}{2}, 1] \end{cases}.$$

Η κυβική συνάρτηση spline και στο παράδειγμα αυτό είναι άρτια, όπως το περιμέναμε.

4.3.1 Βάση του χώρου συναρτήσεων των κυβικών Splines

Ας δούμε όμως τώρα την παρεμβολή με κυβικές splines από την πλευρά των γραμμικών χώρων. Η διάσταση του χώρου $S(X_n)$ των κυβικών splines είναι $n + 3$. Η διάσταση των πολυωνύμων τρίτου βαθμού σε κάθε υποδιάστημα είναι 4. Αυτή πολλαπλασιάζεται επί n που είναι το πλήθος των υποδιαστημάτων και λαμβάνουμε διάσταση $4n$ του χώρου όλων των τμηματικά κυβικών συναρτήσεων. Η απαίτηση της συνέχειας, της συνέχειας της πρώτης παραγώγου και της συνέχειας της δεύτερης παραγώγου δίνουν $3(n-1)$ περιορισμούς, τρεις για κάθε ενδιάμεσο σημείο της διαμέρισης. Αφαιρώντας τους περιορισμούς, λαμβάνουμε $4n - 3(n-1) = n + 3$ τη διάσταση του $S(X_n)$. Αλλιώς, η διάσταση του $S(X_n)$ είναι τόση, όσες πληροφορίες απαιτούνται για τον υπολογισμό της s. Αυτές είναι οι $n + 1$ τιμές της συνάρτησης και οι 2 τιμές της παραγώγου στα ακραία σημεία a και b, σύνολο $n + 3$ πληροφοριών.

Θα προσπαθήσουμε τώρα να βρούμε μια βάση του χώρου αυτού. Θεωρούμε τις συναρτήσεις spline c_i τέτοιες ώστε

$$c_i(x_j) = \begin{cases} 0 & \text{για } j \neq i \\ 1 & \text{για } j = i \end{cases} \text{ και } \; c_i'(x_0) = c_i'(x_n) = 0, \quad i = 0, 1, \ldots, n, \tag{4.12}$$

τις οποίες ονομάζουμε θεμελιώδεις splines. Αυτές είναι $n + 1$ τον αριθμό, όσα και τα σημεία. Θα αποδείξουμε ότι το σύνολο $\{x, x^2, c_0(x), c_1(x), \ldots, c_n(x)\}$ αποτελεί βάση για τον $S(X_n)$. Πραγματικά, αν θεωρήσουμε το γραμμικό συνδυασμό

$$t(x) = \alpha x + \beta x^2 + \sum_{i=1}^{n} \alpha_i c_i(x),$$

είναι προφανές ότι $t \in S(X_n)$. Το $\sum_{i=1}^{n} \alpha_i c_i(x)$ ορίζει μια spline που προκύπτει από μια συνάρτηση με τιμές α_i στα σημεία $x_i, \ \ i = 0, 1, \ldots, n$, και 0 τις τιμές της παραγώγου στα άκρα. Τα α και β προσδιορίζονται ώστε να εξασφαλίζονται η συνθήκες $t'(a) = f'(a)$ και $t'(b) = f'(b)$. Αυτό σημαίνει ότι

$$\begin{array}{l} \alpha + 2\beta a = f'(a) \\ \alpha + 2\beta b = f'(b) \end{array} \Leftrightarrow \begin{array}{l} \alpha = \frac{bf'(a) - af'(b)}{b-a} \\ \beta = \frac{f'(b) - f'(a)}{2(b-a)} \end{array}. \tag{4.13}$$

Η συνάρτηση t λοιπόν που κατασκευάσαμε είναι μια κυβική spline. Οι τιμές της spline αυτής στα σημεία $x_i, \ \ i = 0, 1, \ldots, n$, θα είναι

$$t(x_i) = \alpha x_i + \beta x_i^2 + \alpha_i, \ \ i = 0, 1, \ldots, n.$$

Για να είναι αυτή, η spline που προκύπτει από τη συνάρτηση f, τα α και β θα δίνονται από τους παραπάνω τύπους ενώ τα α_i θα δίνονται από

$$\alpha x_i + \beta x_i^2 + \alpha_i = f_i \Leftrightarrow \alpha_i = f_i - \alpha x_i - \beta x_i^2, \ \ i = 0, 1, \ldots, n. \tag{4.14}$$

Αποδείχτηκε με αυτόν τον τρόπο ότι για κάθε συνάρτηση $f \in C[a, b]$ υπάρχει μοναδικός γραμμικός συνδυασμός των παραπάνω στοιχείων που δίνει τη συνάρτηση spline $s(f, X_n)$, επομένως τα στοιχεία αυτά αποτελούν βάση.

Θα δούμε στη συνέχεια το πως κατασκευάζεται μια τέτοια βάση χρησιμοποιώντας τη διαμέριση X_2 του Παραδείγματος 24.

Παράδειγμα 26 Να βρεθεί μια βάση του χώρου των κυβικών splines $S(X_2)$, όταν $X_2 = \{-1, 0, 1\}$. Στη συνέχεια να βρεθεί η κυβική spline για τη συνάρτηση $f(x) = |x|$ (Παράδειγμα 24), χρησιμοποιώντας τη βάση.

Σύμφωνα με τα παραπάνω, η βάση αυτή θα είναι η $\{x, x^2, c_0(x), c_1(x), c_2(x)\}$ όπου $c_i, \ i = 0, 1, 2$ είναι οι θεμελιώδεις splines όπως ορίστηκαν στην (4.12).
Η c_0, είναι η spline που παράγεται από τη συνάρτηση f με τιμές

$$\begin{array}{c|ccc} x_i & -1 & 0 & 1 \\ \hline f_i & 1 & 0 & 0 \end{array}, \quad f'(-1) = 0, \ f'(1) = 0.$$

Το γραμμικό σύστημα (4.11) εκφυλίζεται στην εξίσωση

$$4s_1' = -3 \Leftrightarrow s_1' = -\frac{3}{4}.$$

Σύμφωνα με την (4.4) λαμβάνουμε

$$c_0(x) = \begin{cases} (x^2 + 2(x+1)x^2)\,1 + x(x+1)^2\left(-\frac{3}{4}\right) = \frac{5}{4}x^3 + \frac{3}{2}x^2 - \frac{3}{4}x, & x \in [-1, 0] \\ x(x-1)^2\left(-\frac{3}{4}\right) = -\frac{3}{4}x^3 + \frac{3}{2}x^2 - \frac{3}{4}x, & x \in [0, 1] \end{cases}.$$

Η c_1, παράγεται από τη συνάρτηση f με τιμές

$$\begin{array}{c|ccc} x_i & -1 & 0 & 1 \\ \hline f_i & 0 & 1 & 0 \end{array}, \quad f'(-1) = 0, \ f'(1) = 0.$$

Το γραμμικό σύστημα (4.11) δίνει

$$4s_1' = 0 \Leftrightarrow s_1' = 0$$

και από την (4.4) λαμβάνουμε

$$c_1(x) = \begin{cases} \left((x+1)^2 - 2x(x+1)^2\right)1 = -2x^3 - 3x^2 + 1, & x \in [-1, 0] \\ \left((x-1)^2 - 2x(x-1)^2\right)1 = 2x^3 - 3x^2 + 1, & x \in [0, 1] \end{cases}.$$

Τέλος, η c_2 παράγεται από τη συνάρτηση f με τιμές

$$\begin{array}{c|ccc} x_i & -1 & 0 & 1 \\ \hline f_i & 0 & 0 & 1 \end{array}, \quad f'(-1) = 0,\ f'(1) = 0.$$

Το γραμμικό σύστημα (4.11) δίνει

$$4s_1' = 3 \Leftrightarrow s_1' = \frac{3}{4}$$

και από την (4.4) λαμβάνουμε

$$c_2(x) = \begin{cases} x(x+1)^2\left(\frac{3}{4}\right) = \frac{3}{4}x^3 + \frac{3}{2}x^2 + \frac{3}{4}x, & x \in [-1, 0] \\ \left(x^2 - 2(x-1)x^2\right)1 + x(x-1)^2\left(\frac{3}{4}\right) = -\frac{5}{4}x^3 + \frac{3}{2}x^2 + \frac{3}{4}x, & x \in [0, 1] \end{cases}.$$

Κατασκευάσαμε με τον τρόπο αυτό τη βάση του χώρου των κυβικών splines στο X_2. Στη συνέχεια θα υπολογίσουμε τους συντελεστές που δίνουν την κυβική συνάρτηση spline του Παραδείγματος 24 για τη συνάρτηση $f(x) = |x|$, ως γραμμικό συνδυασμό των στοιχείων της βάσης. Από την (4.13) προκύπτει ότι

$$\alpha = \frac{1f'(-1) - (-1)f'(1)}{2} = 0, \quad \beta = \frac{f'(1) - f'(-1)}{4} = \frac{1}{2}.$$

Η (4.14) στη συνέχεια δίνει

$$\alpha_0 = f_0 - \alpha x_0 - \beta x_0^2 = \frac{1}{2},\ \alpha_1 = f_1 - \alpha x_1 - \beta x_1^2 = 0, \quad \alpha_2 = f_2 - \alpha x_2 - \beta x_2^2 = \frac{1}{2}.$$

Επομένως, η κυβική spline θα δίνεται από το γραμμικό συνδυασμό

$$\begin{aligned} s(|x|, X_2; x) &= \tfrac{1}{2}x^2 + \tfrac{1}{2}c_0(x) + \tfrac{1}{2}c_2(x) \\ &= \begin{cases} \frac{1}{2}x^2 + \frac{1}{2}\left(\frac{5}{4}x^3 + \frac{3}{2}x^2 - \frac{3}{4}\right) + \frac{1}{2}\left(\frac{3}{4}x^3 + \frac{3}{2}x^2 + \frac{3}{4}x\right) \\ \frac{1}{2}x^2 + \frac{1}{2}\left(-\frac{3}{4}x^3 + \frac{3}{2}x^2 - \frac{3}{4}x\right) + \frac{1}{2}\left(-\frac{5}{4}x^3 + \frac{3}{2}x^2 + \frac{3}{4}x\right) \end{cases} \\ &= \begin{cases} x^3 + 2x^2, & x \in [-1, 0] \\ -x^3 + 2x^2, & x \in [0, 1] \end{cases}. \end{aligned}$$

Προφανώς, το αποτέλεσμα είναι το ίδιο με εκείνο του Παραδείγματος 24.

Ασκήσεις

1. Να βρεθεί η τμηματικά γραμμική συνάρτηση που προσεγγίζει τη συνάρτηση $f(x) = x + \frac{1}{x}$, στο σύνολο σημείων $X_4 = \{1, 2, 3, 4\}$.

2. Να βρεθεί η κυβική συνάρτηση Hermite στο $[0, 3]$ που προσεγγίζει τη συνάρτηση $f(x) = x^5 - 4x^4$, στο σύνολο σημείων $X_4 = \{0, 1, 2, 3\}$.

3. Να βρεθεί η κυβική συνάρτηση spline που προσεγγίζει τη συνάρτηση $f(x) = \frac{1}{x}$, στο σύνολο σημείων $X_4 = \{1, 2, 3, 4\}$.

4. Να βρεθεί η κυβική συνάρτηση spline που προσεγγίζει τη συνάρτηση f η οποία δίνεται στο σύνολο σημείων
$$X_4 = \{-2, -1, 1, 2\}, \quad \begin{array}{c|cccc} x_i & -2 & -1 & 1 & 2 \\ \hline f_i & -8 & -1 & 1 & 8 \end{array}.$$
και $f'(-2) = 12, f'(2) = 12$.

5. Να βρεθεί η κυβική συνάρτηση spline που προσεγγίζει τη συνάρτηση f, η οποία δίνεται στο σύνολο σημείων
$$X_4 = \{-1, 0, 1, 2\}, \quad \begin{array}{c|cccc} x_i & -1 & 0 & 1 & 2 \\ \hline f_i & 0 & 1 & 3 & 17 \end{array}$$
και $f'(-1) = 3, f'(2) = 24$.

6. Να βρεθεί η κυβική συνάρτηση spline που προσεγγίζει τη συνάρτηση f, η οποία δίνεται στο σύνολο σημείων
$$X_5 = \{-2, -1, 0, 1, 2\}, \quad \begin{array}{c|ccccc} x_i & -2 & -1 & 0 & 1 & 2 \\ \hline f_i & 0 & 3 & 0 & -3 & 0 \end{array}.$$
και $f'(-2) = 8, f'(2) = 8$.

7. Να βρεθεί η κυβική συνάρτηση spline που προσεγγίζει τη συνάρτηση f, η οποία δίνεται στο σύνολο σημείων
$$X_4 = \{-1, 0, 1, 2\}, \quad \begin{array}{c|cccc} x_i & -1 & 0 & 1 & 2 \\ \hline f_i & -1 & 0 & 0 & 1 \end{array}.$$
και $f'(-1) = 3, f'(2) = 3$.

8. Αν s είναι η κυβική συνάρτηση spline που προσεγγίζει τη συνάρτηση f στο σύνολο σημείων X_m, να βρεθεί η κυβική συνάρτηση spline που προσεγγίζει τη συνάρτηση $f + x^3$ στο ίδιο σύνολο σημείων.

Βιβλιογραφία

[1] E. W. Cheney, *An Introduction to Approximation Theory,* McGraw Hill Company, New York 1966.

[2] J. R. Rice, *The Approximation of Functions, Vols. 1, 2*, AddisonWesley, 1: 1964, 2:1968.

[3] Theodore J. Rivlin, *An Introduction to the Approximation of Functions,* Dover publications, Inc., New York 1981.

[4] Νικόλας Παπαμιχαήλ, *Εισαγωγή στη Θεωρία Συναρτήσεων Spline μιας μεταβλητής*, Εκδόσεις Πανεπιστημίου Πατρών, Πάτρα 2004.