

4. Η μέθοδος των ελαχίστων τετραγώνων

4.1. Θεωρητικό πλαίσιο. Έστω $\Lambda \in \mathbb{N}$ και δεδομένα $x, y \in \mathbb{R}^\Lambda$, όπου $x_j \neq x_i$ όταν $j \neq i$. Έστω ακόμα $m \in \mathbb{N}_0$ με $m < \Lambda$. Τότε ορίζουμε το σύνολο:

$$Z_m := \{z \in \mathbb{R}^\Lambda : \exists p \in \mathbb{P}^m \text{ τέτοιο ώστε : } z_\kappa = p(x_\kappa), \quad \kappa = 1, \dots, \Lambda\},$$

το οποίο είναι ένας υπόχωρος του γραμμικού χώρου $(\mathbb{R}^\Lambda, \mathbb{R})$ με διάσταση $N = m + 1$.

Διαλέγοντας μια βάση $(p^\ell)_{\ell=1}^{m+1}$ του γραμμικού χώρου $(\mathbb{P}^m, \mathbb{R})$, τα διανύσματα $(z^\ell)_{\ell=1}^{m+1} \subset Z_m$ με

$$z_i^\ell = p^\ell(x_i), \quad i = 1, \dots, \Lambda, \quad \ell = 1, \dots, m + 1,$$

αποτελούν μια βάση $\mathcal{B}_m$ του Z_m .

Εφοδιάζουμε το $(\mathbb{R}^\Lambda, \mathbb{R})$ με το συνηθισμένο εσωτερικό γινόμενο $(\cdot, \cdot)_2$, δηλ.

$$(v, w)_2 := \sum_{j=1}^{\Lambda} v_j w_j \quad \forall v, w \in \mathbb{R}^\Lambda,$$

το οποίο παράγει την Ευκλείδεια νόρμα

$$\|v\|_2 := \sqrt{(v, v)_2} = \left(\sum_{j=1}^{\Lambda} |v_j|^2 \right)^{\frac{1}{2}} \quad \forall v \in \mathbb{R}^\Lambda.$$

Το πρόβλημα των ελαχίστων τετραγώνων είναι η εύρεση της βέλτιστης προσέγγισης $y_m^* \in Z_m$ του y από τον Z_m ως προς τη νόρμα $\|\cdot\|_2$, κάτι που είναι εφικτό μέσω της απαίτησης:

$$(1) \quad (y_m^*, v)_2 = (y, v)_2 \quad \forall v \in Z_m.$$

Επειδή $y_m^* \in Z_m$, υπάρχουν $(a_\kappa)_{\kappa=1}^{m+1} \subset \mathbb{R}$ τέτοια ώστε:

$$y_m^* = \sum_{\kappa=1}^{m+1} a_\kappa z^\kappa$$

τα οποία προσδιορίζονται, μέσω της (1), ως λύση του ακόλουθου γραμμικού συστήματος:

$$(2) \quad \sum_{\kappa=1}^{m+1} a_\kappa (z^\kappa, z^j)_2 = (y, z^j)_2, \quad j = 1, \dots, N,$$

του οποίου ο πίνακας είναι συμμετρικός και θετικά ορισμένος. Όταν η βάση $\mathcal{B}_m$ είναι ορθοκανονική, τότε

$$(3) \quad a_\kappa = (y, z^\kappa)_2, \quad \kappa = 1, \dots, m + 1.$$

Σημείωση: Με την παραπάνω διαδικασία, βρίσκουμε πολυώνυμο $p_m^* \in \mathbb{P}^m$ το οποίο δίνεται από τον τύπο

$$(4) \quad p_m^*(x) = \sum_{\kappa=1}^{m+1} a_\kappa p^\kappa(x)$$

και λύνει το ακόλουθο πρόβλημα βελτιστοποίησης:

$$\sum_{j=1}^{\Lambda} |p_m^*(x_j) - y_j|^2 = \min_{q \in \mathbb{P}^m} \sum_{j=1}^{\Lambda} |q(x_j) - y_j|^2,$$

από το οποίο προέρχεται το όνομα μέθοδος ελαχίστων τετραγώνων.

Ιστορικό σημείωμα: Η πατρότητα της μεθόδου των ελαχίστων τετραγώνων αποδίδεται στον Gauss ο οποίος, χρησιμοποιώντας τη μέθοδο, κατάφερε το 1801 να υπολογίσει την τροχιά του αστεροειδούς Ceres χρησιμοποιώντας μερικές παρατηρήσεις της θέσης του και ο οποίος για ένα διάστημα δεν ήταν ορατός. Ο αστεροειδής Ceres βρίσκεται ανάμεσα στις τροχιές του Άρη και του Δία και έχει διάμετρο περίπου 945Km. Παρόλα αυτά η πρώτη δημοσίευση πάνω στη μέθοδο έγινε το 1806 από τον Legendre. Το πρόβλημα των ελαχίστων τετραγώνων καλείται συχνά πρόβλημα προσαρμογής των δεδομένων και προσπαθεί να ξεπεράσει το πρόβλημα των σφαλμάτων που υπάρχουν στις μετρήσεις κατά την αναζήτηση ή επαλήθευση ενός νόμου για την περιγραφή τους.

3.3. Εφαρμογή. Εφαρμόστε την παραπάνω μεθοδολογία για $\Lambda = 20$, $m = 1, 2$ ή 3 , και δεδομένα

1.01921	0.700761
-0.557546	0.793753
3.248364E-02	-0.819224
1.20935	1.50934
0.843566	1.05314
-0.983407	1.46828
0.626417	0.239290
0.305218	0.190248
-0.203555	-0.778884
-0.521226	0.728942
0.157991	-0.367680
1.49990	2.21791
1.62142	1.68701
1.47549	2.91171
1.79016	2.34913
1.02596	0.344372
-0.557240	1.02893
-0.951472	0.999800
1.96111	4.23217
1.00919	1.00686

όπου η πρώτη στήλη αντιστοιχεί στο διάνυσμα x και η δεύτερη στήλη στο διάνυσμα y .

Χρησιμοποιήστε τη συνήθη βάση του $\mathbb{P}^m$, δηλ. $\{1, t, \dots, t^m\}$ για να προσδιορίσετε μια βάση του Z_m . Στη συνέχεια προσδιορίστε το πολυώνυμο p_m^* του τύπου (4). Κάνετε γράφημα των παραπάνω σημείων μαζί με το πολυώνυμο p_m^* . Μπορείτε να χρησιμοποιήσετε την `np.linalg.cholesky` για να λύσετε το γραμμικό σύστημα.

Βιβλιογραφία: G. Hämmerlin - K.-H. Hoffmann, Numerical Mathematics, Springer-Verlag.

Γ. Ζουράρης