

6. ΜΕΘΟΔΟΣ ΤΩΝ ΕΛΑΧΙΣΤΩΝ ΤΕΤΡΑΓΩΝΩΝ

Θεωρούμε ένα γραμμικό σύστημα $m \times n$, με $m > n$

$$Av = b.$$

Ένα τέτοιο σύστημα λέγεται **υπερορισμένο** και είναι συνήθως *αδύνατο* (ειδικά όταν το m είναι πολύ μεγαλύτερο του n , $m \gg n$). Γράφουμε τότε

$$Av = b + e(v),$$

όπου $e(v)$ το **σφάλμα** (ή υπόλοιπο) για κάθε $v \in \mathbf{R}^n$. Το πρόβλημα που τίθεται είναι να ελαχιστοποιήσουμε, ως προς v , κάποια νόρμα του διανύσματος σφάλματος $e(v)$, δηλαδή την “απόσταση” μεταξύ Av και b . Στη **μέθοδο των Ελαχίστων Τετραγώνων**, ελαχιστοποιούμε τη νόρμα $\|Av - b\|_2$, ή ισοδύναμα, τη συνάρτηση

$$F(v) = \|e(v)\|_2^2 = \|Av - b\|_2^2.$$

Στις εφαρμογές (π.χ. στη μοντελοποίηση, ή προτυποποίηση), ο πίνακας A και το διάνυσμα b αντιστοιχούν συνήθως σε m μετρήσεις μεγεθών ενός φαινομένου, το v στις n παραμέτρους του μαθηματικού μοντέλου (που υποθέτουμε εδώ γραμμικό), τις οποίες θέλουμε να εκτιμήσουμε, και το $e(v)$ στα σφάλματα που οφείλονται στις μετρήσεις ή/και σε ενδεχόμενη μη γραμμικότητα του μοντέλου. Επειδή γίνονται πολλές μετρήσεις, έχουμε συνήθως $m \gg n$.

Θεώρημα 1. Υποθέτουμε ότι $m > n$ και ότι οι στήλες του πίνακα A είναι γραμμικά ανεξάρτητες. Τότε υπάρχει μοναδικό διάνυσμα $\bar{v}$ που ελαχιστοποιεί τη συνάρτηση F στον $\mathbf{R}^n$, το οποίο είναι η λύση του ομαλού συστήματος $n \times n$

$$(1) \quad A^T A \bar{v} = A^T b.$$

Επιπλέον, ο πίνακας $A^T A$ είναι συμμετρικός, και θετικά ορισμένος, δηλαδή ισχύει $v^T A^T A v > 0$, για κάθε $v \neq 0$.

Το διάνυσμα $\bar{v}$ καλείται **εκτίμηση με τα ελάχιστα τετράγωνα** και οι n εξισώσεις (1) καλούνται **κανονικές εξισώσεις**.

Απόδειξη. Θετούμε $\|\cdot\| := \|\cdot\|_2$. Έχουμε προφανώς

$$Av = \sum_{j=1}^n v_j A^j,$$

όπου A^j η στήλη j του πίνακα A . Από την ανεξαρτησία των στηλών του A προκύπτει άρα ότι

$$Av = 0 \Rightarrow v = 0 \quad (\text{ή ισοδύναμα } v \neq 0 \Rightarrow Av \neq 0).$$

Εξάλλου, ο πίνακας $A^T A$ είναι προφανώς συμμετρικός και έχουμε

$$A^T Av = 0 \Rightarrow v^T A^T Av = \|Av\|^2 = 0 \Rightarrow Av = 0 \Rightarrow v = 0,$$

που δείχνουν ότι ο πίνακας $A^T A$ είναι ομαλός και θετικά ορισμένος. Τώρα, έχουμε

$$\|Av - b\| - \|Au - b\| \leq \|A(v - u)\| \leq \|A\| \|v - u\|,$$

όπου $\|A\|$ η παραγόμενη φυσική νόρμα πίνακα, που δείχνει ότι η συνάρτηση $F^{1/2}$, και άρα η F , είναι συνεχής. Ας δείξουμε τώρα ότι

$$F(v) \rightarrow +\infty, \text{ όταν } \|v\| \rightarrow +\infty.$$

Έχουμε, θέτοντας $u = v / \|v\|$

$$[F(v)]^{1/2} = \|Av - b\| = \|v\| \left\| Au - \frac{b}{\|v\|} \right\| \geq \|v\| \left\| \|Au\| - \frac{\|b\|}{\|v\|} \right\|,$$

Έχουμε $u \neq 0$, αφού το u είναι μοναδιαίο, άρα $\|Au\| \neq 0$. Επομένως, η συνεχής και αυστηρά θετική συνάρτηση $u \rightarrow \|Au\|$ παίρνει την ελάχιστή της τιμή $\|A\bar{u}\| > 0$ σε ένα σημείο $\bar{u}$ του κλειστού και φραγμένου, άρα συμπαγούς, συνόλου (υπερσφαίρας)

$$S := \{u \mid \|u\| = 1\}.$$

Επιπλέον, $\|b\|/\|v\| \rightarrow 0$ όταν $\|v\| \rightarrow +\infty$. Συνεπώς, $F(v) \rightarrow +\infty$ όταν $\|v\| \rightarrow +\infty$. Έστω $w \in \mathbf{R}^n$. Υπάρχει άρα $r > 0$ τέτοιο ώστε

$$\|v\| > r \Rightarrow F(v) > F(w).$$

Επομένως, η ελαχιστοποίηση της F στον $\mathbf{R}^n$ ισοδυναμεί με την ελαχιστοποίηση της F στο συμπαγές σύνολο (κλειστή μπάλα)

$$B := \{v \mid \|v\| \leq r\},$$

στο οποίο η συνεχής συνάρτηση F έχει τουλάχιστο ένα σημείο ελαχίστου $\bar{v}$. Στο σημείο αυτό, η F ικανοποιεί τις αναγκαίες συνθήκες

$$\nabla F(\bar{v}) = 0,$$

δηλαδή

$$\frac{\partial F}{\partial x_k}(\bar{v}) = 0, \quad k = 1, \dots, n,$$

ή

$$\begin{aligned} \frac{\partial}{\partial x_k} \left[\sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} \bar{v}_j - b_i \right)^2 \right] &= \sum_{i=1}^m 2 \left(\sum_{j=1}^n a_{ij} \bar{v}_j - b_i \right) a_{ik} \\ &= 2 \sum_{i=1}^m a_{ik} \sum_{j=1}^n a_{ij} \bar{v}_j - 2 \sum_{i=1}^m a_{ik} b_i = 0, \quad k = 1, \dots, n. \end{aligned}$$

Οι εξισώσεις αυτές ισοδυναμούν με το γραμμικό σύστημα $n \times n$

$$A^T A \bar{v} = A^T b,$$

το οποίο έχει μια μοναδική λύση, αφού ο πίνακας $A^T A$ είναι ομαλός. Συνεπώς, το $\bar{v}$ είναι το μοναδικό σημείο ελαχίστου της F .

Το σύστημα (1) μπορεί να λυθεί π.χ. με τη μέθοδο απαλοιφής Gauss (συνήθως η διάσταση n δεν είναι πολύ μεγάλη). Θεωρητικά, μπορούμε να λύσουμε το σύστημα (1) ως προς $\bar{v}$ με αντιστροφή

$$\bar{v} = (A^T A)^{-1} A^T b.$$

Ο πίνακας $(A^T A)^{-1} A^T$ καλείται **ψευδοαντίστροφος** του A και ταυτίζεται με τον αντίστροφο A^{-1} του A στην περίπτωση όπου $m = n$.

Παράδειγμα. Έστω το γραμμικό μοντέλο, με παραμέτρους α και β

$$y = \alpha x + \beta,$$

και μετρήσεις

$$x_1, \dots, x_m \quad \text{και} \quad y_1, \dots, y_m.$$

Έχουμε εδώ

$$A = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \dots & \dots \\ x_m & 1 \end{bmatrix}, \quad b = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_m \end{bmatrix}, \quad v = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}.$$

Εφαρμόζοντας τη μέθοδο των Ελαχίστων Τετραγώνων, βρίσκουμε

$$A^T A = \begin{bmatrix} \sum_{i=1}^m x_i^2 & \sum_{i=1}^m x_i \\ \sum_{i=1}^m x_i & m \end{bmatrix}, \quad A^T b = \begin{bmatrix} \sum_{i=1}^m x_i y_i \\ \sum_{i=1}^m y_i \end{bmatrix},$$

και

$$\begin{aligned} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} &= (A^T A)^{-1} A^T b = \frac{1}{m \sum_{i=1}^m x_i^2 - (\sum_{i=1}^m x_i)^2} \begin{bmatrix} m & -\sum_{i=1}^m x_i \\ -\sum_{i=1}^m x_i & \sum_{i=1}^m x_i^2 \end{bmatrix} \begin{bmatrix} \sum_{i=1}^m x_i y_i \\ \sum_{i=1}^m y_i \end{bmatrix} \\ &= \frac{1}{m \sum_{i=1}^m x_i^2 - (\sum_{i=1}^m x_i)^2} \begin{bmatrix} m \sum_{i=1}^m x_i y_i - (\sum_{i=1}^m x_i)(\sum_{i=1}^m y_i) \\ -(\sum_{i=1}^m x_i)(\sum_{i=1}^m x_i y_i) + (\sum_{i=1}^m x_i^2)(\sum_{i=1}^m y_i) \end{bmatrix}. \end{aligned}$$