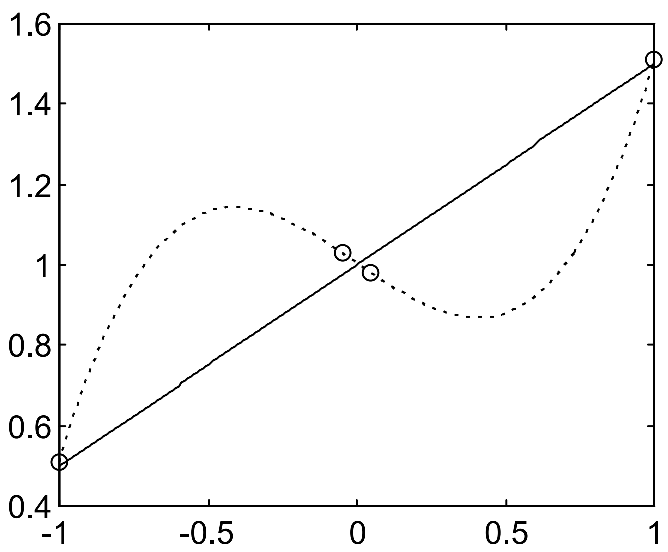
## Γραμμική Παλινδρόμηση

*Ας θεωρήσουμε το πρόβλημα υπολογισμού της τιμής μιας συνάρτησης σε ενδιάμεσα σημεία όταν δίνονται τα πειραματικά δεδομένα του πίνακα*

| $i$ | $x_i$ | $y_i$ |
|---|---|---|
| 1 | $-1$ | 0.51 |
| 2 | $-1/25$ | 1.03 |
| 3 | $1/20$ | 0.98 |
| 4 | 1 | 1.51 |

Όπως φαίνεται από τον πίνακα και όπως πολλές φορές υποθέτουμε (π.χ. $U = R \cdot i$ στον νόμο του *Ohm*) τα δεδομένα έχουν μια γραμμική σχέση.



Το μοναδικό πολυώνυμο 3ου βαθμού που παρεμβάλει τα δεδομένα είναι

$$p(x) = 1.007752 - 0.557799x + 0.002248x^2 + 1.057799x^3$$

Παρατηρώντας όμως το παραπάνω σχήμα, διαπιστώνουμε ότι μια ευθεία γραμμή π.χ. η

$$y(x) = 0.5x + 1$$

"ταιριάζει" καλύτερα στα δεδομένα του πίνακα.

Μια καλή αντιμετώπιση αυτού του είδους προβλημάτων θα ήταν να βρούμε τη "*βέλτιστη*" ευθεία η οποία θα μπορούσε να χρησιμοποιηθεί σαν μια προσεγγιστική συνάρτηση σε οποιοδήποτε ενδιάμεσο σημείο, ακόμα και όταν αυτή δεν συμφωνεί ακριβώς με τα δεδομένα.

Η μέθοδος των **Ελαχίστων Τετραγώνων** (*least squares*) υπολογίζει τη βέλτιστη προσεγγιστική ευθεία όταν το σφάλμα προσέγγισης είναι το άθροισμα των τετραγώνων των διαφορών μεταξύ των τιμών της ευθείας αυτής και των αντίστοιχων πειραματικών δεδομένων (data).

Για να γίνει εύκολα κατανοητή η μέθοδος των ελαχίστων τετραγώνων, θεωρούμε αρχικά τα δεδομένα του πίνακα, και τη γραμμική συνάρτηση

$$y(x) = ax + b$$

Αν με $y_i$ παραστήσουμε τις πειραματικές τιμές στα $x_i$, $i = 1,2,3,4$ και με

$$y(x_i) = ax_i + b$$

τις αντίστοιχες τιμές πάνω στην προσδιοριστέα προσεγγιστική ευθεία, τότε το σφάλμα της προσέγγισης μπορεί να εκφραστεί:

$$E(a,b) = \sum_{i=1}^{4} (y_i - (a \cdot x_i + b))^2.$$

Το πρόβλημα ανάγεται στον υπολογισμό των $a, b$ ώστε το σφάλμα να γίνει ελάχιστο. Η συνάρτηση σφάλματος είναι μία συνάρτηση δύο μεταβλητών Οι αναγκαίες συνθήκες για την ελαχιστοποίηση είναι:

$$\frac{\partial}{\partial a}E(a,b) = \frac{\partial}{\partial a}\sum_{i=1}^{4}\left(y_i - (a\cdot x_i + b)\right)^2 = 0,$$

$$\frac{\partial}{\partial b}E(a,b) = \frac{\partial}{\partial\ b}\sum_{i=1}^{4}(y_i - (a\cdot x_i + b))^2 = 0$$

Για το συγκεκριμένο παράδειγμα, η εξίσωση του σφάλματος γράφεται:

$$E(a,b) = \{0.51 - (-a + b)\}^2 + \{1.03 - (-0.04a + b)\}^2 +$$
$$+\{0.98 - (0.05a + b)\}^2 + \{1.51 - (a + b)\}^2$$

ή

$$E(a,b) = \frac{9123}{2000} - \frac{5039a}{2500} + \frac{20041a^2}{10000} - \frac{403b}{50} + \frac{ab}{50} + 4b^2.$$

Από την παραπάνω εξίσωση σφάλματος οι συνθήκες ελαχιστοποίησης δίνουν τις παρακάτω εξισώσεις

$$-2.0156 + 4.0082a + 0.02b = 0$$

$$-8.06 + 0.02a + 8.b = 0$$

Η λύση του γραμμικού συστήματος είναι:

$$a{\sim}0.497848, b{\sim}1.00625$$

Οπότε η βέλτιστη ευθεία των ελαχίστων τετραγώνων για τα δεδομένα του πίνακα είναι η

$$y(x) = 0.497848x + 1.00625$$

## Γραμμική προσέγγιση

Το γενικό πρόβλημα προσαρμογής της βέλτιστης ευθείας σε ένα σύνολο δεδομένων $(x_i, y_i)$ $i = 1,2,..,n$ με την μέθοδο ελαχίστων τετραγώνων, συνεπάγεται την ελαχιστοποίηση της παράστασης

$$E(a,b) = \sum_{i=1}^{n} (y_i - (a \cdot x_i + b))^2$$

Τούτο έχει σαν συνέπεια τις αναγκαίες συνθήκες

$$\frac{\partial}{\partial a} E(a,b) = \frac{\partial}{\partial a} \sum_{i=1}^{n} (y_i - (a \cdot x_i + b))^2 = 0,$$

$$\frac{\partial}{\partial b} E(a,b) = \frac{\partial}{\partial b} \sum_{i=1}^{n} (y_i - (a \cdot x_i + b))^2 = 0$$

ή

$$2 \sum_{i=1}^{n} (y_i - a x_i - b)(-x_i) = 0,$$

$$2 \sum_{i=1}^{n} (y_i - a x_i - b)(-1) = 0$$

Οι εξισώσεις αυτές απλοποιούνται στις λεγόμενες κανονικές εξισώσεις

$$a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i\, y_i,$$

$$a\sum_{i=1}^{n} x_i + nb = \sum_{i=1}^{n} y_i$$

Το παραπάνω σύστημα επιλυμένο ως προς *a* και *b* μας δίνει

$$a = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2},$$

$$b = \frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i y_i \sum_{i=1}^{n} x_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}.$$

# Παράδειγμα

Δίνονται τα σημεία

| i | $x_i$ | $y_i$ |
|---|---|---|
| 1 | 0.1 | 1.2 |
| 2 | 0.25 | 1.51 |
| 3 | 0.37 | 1.7 |
| 4 | 0.5 | 2 |
| 5 | 0.5 | 2.05 |
| 6 | 0.8 | 2.5 |
| 7 | 1.02 | 3 |
| 8 | 1.09 | 3.3 |
| 9 | 1.24 | 3.54 |
| 10 | 1.4 | 3.77 |

Διαμορφώνουμε τον παρακάτω πίνακα

| i | xi | yi | xi*yi | xi^2 |
|---|---|---|---|---|
| 1 | 0.1 | 1.2 | 0.12 | 0.01 |
| 2 | 0.25 | 1.51 | 0.3775 | 0.0625 |
| 3 | 0.37 | 1.7 | 0.629 | 0.1369 |
| 4 | 0.5 | 2 | 1 | 0.25 |
| 5 | 0.5 | 2.05 | 1.025 | 0.25 |
| 6 | 0.8 | 2.5 | 2 | 0.64 |
| 7 | 1.02 | 3 | 3.06 | 1.0404 |
| 8 | 1.09 | 3.3 | 3.597 | 1.1881 |
| 9 | 1.24 | 3.54 | 4.3896 | 1.5376 |
| 10 | 1.4 | 3.77 | 5.278 | 1.96 |
| | | | | |
| sums | 7.27 | 24.57 | 21.4761 | 7.0755 |

και

$$a = \frac{10 \cdot 21.4671 - 7.27 \cdot 24.57}{10 \cdot 7.0755 - 7.27^2} \approx 2.013568$$

$$b = \frac{7.0755 \cdot 24.57 - 21.4761 \cdot 7.27}{10 \cdot 7.0755 - 7.27^2} \approx 0.989481$$

Άρα η ζητούμενη ευθεία είναι: $y \approx 2.013568x + 0.989481$