

Βασικά Μαθηματικά

Γραμμική Άλγεβρα

Χρήστος Θωμαδάκης

Επίκουρος Καθηγητής, Τμήμα Στατιστικής, Οικονομικό Πανεπιστήμιο Αθηνών
cthomadak@aueb.gr

Δ.Π.Μ.Σ. «Βιοστατιστική και Επιστήμη Δεδομένων Υγείας»

- Θέμα: **Γραμμική Άλγεβρα.**
- Στόχος: να γνωρίσουμε τον συμβολισμό πινάκων που εμφανίζεται συνεχώς στη Στατιστική και τη Βιοστατιστική.
- Δεν θα ασχοληθούμε με τη Γραμμική Άλγεβρα ως αυτοτελή μαθηματική θεωρία.
- Θα δώσουμε έμφαση σε έννοιες που χρειαζόμαστε για να κατανοήσουμε τύπους όπως

$$Y = X\beta + \varepsilon.$$

Γιατί χρειαζόμαστε Γραμμική Άλγεβρα στη Βιοστατιστική

- Πολλά στατιστικά μοντέλα γράφονται πολύ πιο απλά με διανύσματα και πίνακες.
- Για παράδειγμα, στη γραμμική παλινδρόμηση συναντάμε:

$$Y = X\beta + \varepsilon,$$

$$X^T X, \quad X^T Y,$$

και

$$\hat{\beta} = (X^T X)^{-1} X^T Y.$$

- Στο τέλος του μαθήματος θέλουμε να μπορούμε να **διαβάσουμε και να ερμηνεύσουμε** αυτούς τους συμβολισμούς.

Από έναν αριθμό σε έναν πίνακα

- Ένας αριθμός μπορεί να περιγράψει μία μόνο μέτρηση, π.χ. CD4 ενός ασθενούς.
- Πολλές μετρήσεις της ίδιας μεταβλητής μπορούν να γραφτούν ως **διάνυσμα**.
- Πολλές μεταβλητές για πολλά άτομα μπορούν να οργανωθούν σε έναν **πίνακα**.

$$480 \longrightarrow \begin{pmatrix} 480 \\ 350 \\ 620 \end{pmatrix} \longrightarrow \begin{pmatrix} 52 & 1 & 480 \\ 61 & 0 & 350 \\ 43 & 1 & 620 \end{pmatrix}.$$

- Διανύσματα και πίνακες.
- Διαστάσεις και βασικές πράξεις.
- Πολλαπλασιασμός πινάκων.
- Ανάστροφος και αντίστροφος πίνακας.
- Σύντομη αναφορά σε γραμμικά συστήματα.
- Η μορφή $Y = X\beta + \varepsilon$ και η σύνδεση με τα ελάχιστα τετράγωνα.

Τι είναι ένα διάνυσμα

- Ένα **διάνυσμα** είναι μια διατεταγμένη συλλογή αριθμών.
- Ένα διάνυσμα-στήλη γράφεται, για παράδειγμα,

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

- Παράδειγμα: οι τιμές CD4 τριών ασθενών μπορούν να γραφτούν ως

$$x = \begin{pmatrix} 480 \\ 350 \\ 620 \end{pmatrix}.$$

- Το διάνυσμα

$$x = \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}$$

έχει 3 γραμμές και 1 στήλη: είναι 3×1 .

- Αν ανταλλάξουμε γραμμές και στήλες παίρνουμε

$$x^T = (2 \quad 4 \quad 6),$$

που είναι 1×3 .

- Το σύμβολο T θα το μελετήσουμε αναλυτικότερα στη συνέχεια.

Τι είναι ένας πίνακας

- Ένας πίνακας είναι μια ορθογώνια διάταξη αριθμών σε **γραμμές** και **στήλες**.
- Γράφουμε γενικά

$$A = (a_{ij}),$$

όπου a_{ij} είναι το στοιχείο στη γραμμή i και στη στήλη j .

$$A = \begin{pmatrix} 3 & 1 & 5 \\ 2 & 4 & 0 \end{pmatrix}.$$

- Ο πίνακας

$$A = \begin{pmatrix} 3 & 1 & 5 \\ 2 & 4 & 0 \end{pmatrix}$$

έχει 2 γραμμές και 3 στήλες.

- Λέμε ότι ο A είναι πίνακας 2×3 .

- Η σειρά είναι πάντοτε:

γραμμές $\times$ στήλες.

- Ποιο είναι το στοιχείο a_{23} ;

Ο πίνακας δεδομένων στη Βιοστατιστική

- Σε ένα συνηθισμένο dataset:
 - κάθε **γραμμή** αντιστοιχεί σε ένα άτομο,
 - κάθε **στήλη** αντιστοιχεί σε μία μεταβλητή.

$$X = \begin{pmatrix} 52 & 1 & 480 \\ 61 & 0 & 350 \\ 43 & 1 & 620 \\ 38 & 0 & 510 \end{pmatrix}$$

Ηλικία

Θεραπεία

CD4

Ένα dataset ως πίνακας

Άτομο	Ηλικία	Θεραπεία	CD4
1	52	1	480
2	61	0	350
3	43	1	620
4	38	0	510

- Οι τρεις αριθμητικές μεταβλητές μπορούν να θεωρηθούν ως οι στήλες ενός πίνακα.
- Αυτή είναι η βασική ιδέα πίσω από τον πίνακα σχεδιασμού X στα στατιστικά μοντέλα.

$$A = \begin{pmatrix} 7 & -2 & 4 \\ 2 & 5 & 0 \\ 1 & 8 & 3 \end{pmatrix}.$$

- Το a_{12} βρίσκεται στη γραμμή 1 και στη στήλη 2, άρα $a_{12} = -2$.
- Το a_{31} βρίσκεται στη γραμμή 3 και στη στήλη 1, άρα $a_{31} = 1$.
- Ερώτηση: ποια είναι τα a_{23} και a_{33} ;

- Ένας πίνακας λέγεται **τετραγωνικός** όταν έχει ίδιο αριθμό γραμμών και στηλών.

$$A = \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 2 \\ 3 & 4 & 1 \\ 0 & 5 & 2 \end{pmatrix}.$$

- Ο A είναι 2×2 , ενώ ο B είναι 3×3 .
- Η έννοια του αντίστροφου και της ορίζουσας αφορά τετραγωνικούς πίνακες.

- Ο **μηδενικός πίνακας** έχει όλα τα στοιχεία ίσα με 0.
- Ο **μοναδιαίος πίνακας** έχει 1 στην κύρια διαγώνιο και 0 αλλού.

$$0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

- Ο μοναδιαίος πίνακας παίζει για τους πίνακες τον ρόλο που παίζει ο αριθμός 1 στους πραγματικούς αριθμούς.

- Ένας **διαγώνιος πίνακας** έχει μηδενικά όλα τα στοιχεία εκτός της κύριας διαγωνίου.
- Ένας πίνακας λέγεται **συμμετρικός** όταν

$$A^T = A.$$

$$D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \quad S = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}.$$

- Συμμετρικοί πίνακες εμφανίζονται πολύ συχνά στη Στατιστική, π.χ. σε πίνακες διακύμανσης–συνδιακύμανσης.

- Προσθέτουμε τα αντίστοιχα στοιχεία:

$$\begin{pmatrix} 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 4 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \end{pmatrix}.$$

- Τα δύο διανύσματα πρέπει να έχουν την ίδια διάσταση.

- Η πρόσθεση πινάκων γίνεται στοιχείο προς στοιχείο.

$$\begin{pmatrix} 3 & -1 \\ 8 & -7 \end{pmatrix} + \begin{pmatrix} 4 & -6 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 7 & -7 \\ 10 & -4 \end{pmatrix}.$$

- Η πρόσθεση ορίζεται μόνο για πίνακες με **ίδιες διαστάσεις**.

- Αν c είναι πραγματικός αριθμός, τότε πολλαπλασιάζουμε κάθε στοιχείο του πίνακα με c .

$$2 \begin{pmatrix} 3 & 2 \\ -1 & 4 \end{pmatrix} = \begin{pmatrix} 6 & 4 \\ -2 & 8 \end{pmatrix}.$$

Εσωτερικό γινόμενο δύο διανυσμάτων

- Για δύο διανύσματα ίδιας διάστασης,

$$x^T y = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

- Παράδειγμα:

$$(1 \quad 2 \quad 3) \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = 1 \cdot 4 + 2 \cdot 5 + 3 \cdot 6 = 32.$$

- Το αποτέλεσμα είναι ένας **αριθμός**.

Άθροισμα τετραγώνων σε συμβολισμό πινάκων

- Av

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

τότε

$$x^T x = x_1^2 + x_2^2 + \cdots + x_n^2.$$

- Άρα το $x^T x$ είναι ένας σύντομος τρόπος να γράψουμε ένα **άθροισμα τετραγώνων**.
- Θα το ξαναδούμε όταν γράψουμε τα ελάχιστα τετράγωνα ως $e^T e$.

Πότε ορίζεται το γινόμενο πινάκων

- Αν ο A έχει διαστάσεις $m \times n$ και ο B έχει διαστάσεις $n \times k$, τότε το γινόμενο AB ορίζεται.
- Το αποτέλεσμα έχει διαστάσεις $m \times k$:

$$(m \times n)(n \times k) = (m \times k).$$

- Ο **εσωτερικός αριθμός** των διαστάσεων πρέπει να συμφωνεί.

Έλεγχος διαστάσεων πριν από τον υπολογισμό

- Το γινόμενο

$$(3 \times 2)(2 \times 1)$$

ορίζεται και δίνει πίνακα 3×1 .

- Το γινόμενο

$$(2 \times 3)(2 \times 2)$$

δεν ορίζεται, επειδή $3 \neq 2$.

- Πριν κάνουμε οποιονδήποτε υπολογισμό, ελέγχουμε πρώτα τις διαστάσεις.

- Έστω $A : 4 \times 3$, $B : 3 \times 2$, $C : 2 \times 4$.
- Ποια από τα παρακάτω γινόμενα ορίζονται;

$$AB, \quad BA, \quad BC, \quad CA.$$

- Όπου το γινόμενο ορίζεται, ποια είναι η διάσταση του αποτελέσματος;

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 \cdot 5 + 2 \cdot 6 \\ 3 \cdot 5 + 4 \cdot 6 \end{pmatrix} = \begin{pmatrix} 17 \\ 39 \end{pmatrix}.$$

- Κάθε στοιχείο του αποτελέσματος προκύπτει από **γραμμή επί στήλη**.
- Οι διαστάσεις είναι

$$(2 \times 2)(2 \times 1) = (2 \times 1).$$

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 0 \\ 1 & 5 \end{pmatrix}.$$

$$AB = \begin{pmatrix} 1 \cdot 2 + 2 \cdot 1 & 1 \cdot 0 + 2 \cdot 5 \\ 3 \cdot 2 + 4 \cdot 1 & 3 \cdot 0 + 4 \cdot 5 \end{pmatrix} = \begin{pmatrix} 4 & 10 \\ 10 & 20 \end{pmatrix}.$$

- Κάθε στοιχείο είναι πάλι αποτέλεσμα **γραμμής επί στήλη**.

Το γινόμενο πινάκων δεν είναι αντιμεταθετικό

- Για πραγματικούς αριθμούς ισχύει $ab = ba$.
- Για πίνακες γενικά **δεν** ισχύει:

$$AB \neq BA.$$

$$AB = \begin{pmatrix} 4 & 10 \\ 10 & 20 \end{pmatrix}, \quad BA = \begin{pmatrix} 2 & 4 \\ 16 & 22 \end{pmatrix}.$$

- Ακόμη και όταν και τα δύο γινόμενα ορίζονται, δεν είναι συνήθως ίσα.

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 0 \\ 1 & 4 \end{pmatrix}, \quad x = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

- Υπολογίστε $A + B$.
- Υπολογίστε $2A$.
- Υπολογίστε Ax .
- Υπολογίστε AB και BA . Είναι ίσα;

Ο ανάστροφος ενός πίνακα

- Ο **ανάστροφος** ενός πίνακα A προκύπτει αν οι γραμμές του γίνουν στήλες.
- Συμβολίζεται με

$$A^T.$$

- Το στοιχείο a_{ij} του A γίνεται στοιχείο a_{ji} του A^T .

Παράδειγμα ανάστροφου πίνακα

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \implies A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

- Ο A είναι 2×3 .
- Ο A^T είναι 3×2 .
- Άρα ο ανάστροφος ανταλλάσσει και τις **διαστάσεις** του πίνακα.

- Αν εφαρμόσουμε δύο φορές την αναστροφή:

$$(A^T)^T = A.$$

- Για άθροισμα πινάκων:

$$(A + B)^T = A^T + B^T.$$

- Για γινόμενο πινάκων:

$$(AB)^T = B^T A^T.$$

- Στο τελευταίο αποτέλεσμα **αντιστρέφεται η σειρά** των πινάκων.

Γιατί μας ενδιαφέρει το X^T

- Αν ο πίνακας X έχει διαστάσεις $n \times p$, τότε

$$X^T : p \times n.$$

- Επομένως:

$$X^T X : (p \times n)(n \times p) = p \times p,$$

και αν $Y : n \times 1$,

$$X^T Y : (p \times n)(n \times 1) = p \times 1.$$

- Αυτές είναι ακριβώς οι ποσότητες που θα εμφανιστούν στα ελάχιστα τετράγωνα.

Ο πίνακας $X^T X$ είναι συμμετρικός

- Θυμόμαστε ότι ένας πίνακας A είναι συμμετρικός όταν $A^T = A$.
- Για τον $X^T X$:

$$(X^T X)^T = X^T (X^T)^T = X^T X.$$

- Άρα ο $X^T X$ είναι πάντα συμμετρικός.
- Αυτό είναι ένα από τα πολλά σημεία όπου η Γραμμική Άλγεβρα εμφανίζεται φυσικά στη Στατιστική.

Τι είναι ο αντίστροφος ενός πίνακα

- Για έναν αριθμό $a \neq 0$, ο αντίστροφός του είναι $1/a$, επειδή

$$a \frac{1}{a} = 1.$$

- Αντίστοιχα, ο αντίστροφος ενός πίνακα A συμβολίζεται με A^{-1} και ικανοποιεί

$$AA^{-1} = A^{-1}A = I.$$

- Ο A^{-1} δεν είναι ο ανάστροφος A^T : πρόκειται για διαφορετικές έννοιες.

Δεν έχουν όλοι οι πίνακες αντίστροφο

- Για να έχει ένας πίνακας αντίστροφο πρέπει πρώτα να είναι **τετραγωνικός**.
- Αλλά ούτε κάθε τετραγωνικός πίνακας είναι αντιστρέψιμος.
- Για να ελέγξουμε αν ένας τετραγωνικός πίνακας είναι αντιστρέψιμος, χρησιμοποιούμε την **ορίζουσα**.

Ορίζουσα πίνακα 2×2

- Για

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

η ορίζουσα είναι

$$\det(A) = ad - bc.$$

- Στο μάθημα θα περιοριστούμε στον υπολογισμό ορίζουσας 2×2 .

- Ένας τετραγωνικός πίνακας A είναι αντιστρέψιμος όταν

$$\det(A) \neq 0.$$

- Αν

$$\det(A) = 0,$$

τότε ο A^{-1} δεν υπάρχει.

- Στη Στατιστική αυτή η συνθήκη εμφανίζεται, για παράδειγμα, όταν θέλουμε να αντιστρέψουμε τον $X^T X$.

- Αν

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

και $ad - bc \neq 0$, τότε

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

- Πρώτα:

$$\det(A) = 2 \cdot 1 - 1 \cdot 1 = 1.$$

- Επομένως ο αντίστροφος υπάρχει και

$$A^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}.$$

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I.$$

- Ο πολλαπλασιασμός με τον αντίστροφο επαναφέρει τον μοναδιαίο πίνακα.
- Αυτός είναι και ο βασικός έλεγχος ότι υπολογίσαμε σωστά τον A^{-1} .

Άσκηση στον αντίστροφο πίνακα

- Δίνεται

$$A = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}.$$

- Υπολογίστε την $\det(A)$.
- Ελέγξτε αν υπάρχει ο A^{-1} .
- Αν υπάρχει, υπολογίστε τον αντίστροφο και επαληθεύστε ότι $AA^{-1} = I$.

- Το σύστημα

$$\begin{aligned}2x - y &= 0, \\ -x + 2y &= 3\end{aligned}$$

μπορεί να γραφτεί ως

$$\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \end{pmatrix}.$$

- Συνοπτικά: $Ax = b$.

Λύση γραμμικού συστήματος με τον αντίστροφο

- Αν A^{-1} υπάρχει και

$$Ax = b,$$

τότε πολλαπλασιάζουμε και τα δύο μέλη από αριστερά με A^{-1} :

$$A^{-1}Ax = A^{-1}b.$$

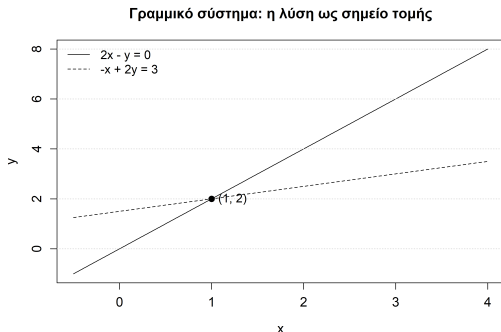
- Επειδή $A^{-1}A = I$, παίρνουμε

$$x = A^{-1}b.$$

Η λύση ενός συστήματος ως σημείο τομής

- Κάθε γραμμική εξίσωση δύο μεταβλητών παριστάνει μια ευθεία.
- Η κοινή λύση του συστήματος είναι το σημείο όπου οι δύο ευθείες τέμνονται.
- Στο παράδειγμα η λύση είναι

$$(x, y) = (1, 2).$$



- Στο πρώτο τρίωρο είδαμε το απλό γραμμικό μοντέλο

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i.$$

- Για n άτομα έχουμε n τέτοιες εξισώσεις.
- Με τη Γραμμική Άλγεβρα μπορούμε να τις γράψουμε όλες μαζί ως

$$Y = X\beta + \varepsilon.$$

- Οι n παρατηρούμενες τιμές της εξαρτημένης μεταβλητής γράφονται ως

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}.$$

- Το Y έχει διάσταση

$$n \times 1.$$

- Για απλή γραμμική παλινδρόμηση ο πίνακας X είναι

$$X = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}.$$

- Η πρώτη στήλη αποτελείται από 1 ώστε να περιλαμβάνεται ο σταθερός όρος β_0 .
- Ο X έχει διάσταση $n \times 2$.

- Οι παράμετροι του μοντέλου γράφονται ως

$$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}.$$

- Το β έχει διάσταση 2×1 .
- Ο β_0 είναι ο σταθερός όρος και ο β_1 η κλίση.

Τι είναι το γινόμενο $X\beta$

$$X\beta = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} \beta_0 + \beta_1 x_1 \\ \beta_0 + \beta_1 x_2 \\ \vdots \\ \beta_0 + \beta_1 x_n \end{pmatrix}.$$

- Άρα το $X\beta$ περιέχει τις μέσες τιμές που προβλέπει το μοντέλο για όλα τα άτομα.

$$Y = X\beta + \varepsilon$$

- Οι διαστάσεις είναι

$$(n \times 1) = (n \times p)(p \times 1) + (n \times 1).$$

- Ο Y περιέχει τις παρατηρήσεις.
- Ο X περιέχει τις επεξηγηματικές μεταβλητές.
- Το β περιέχει τους συντελεστές του μοντέλου.
- Το ε περιέχει τα τυχαία σφάλματα.

Παράδειγμα με ηλικία και θεραπεία

- Έστω ότι θέλουμε να περιγράψουμε μια απόκριση Y με βάση την ηλικία και τη θεραπεία.

$$X = \begin{pmatrix} 1 & 45 & 0 \\ 1 & 62 & 1 \\ 1 & 51 & 1 \\ 1 & 38 & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_0 \\ \beta_{\text{age}} \\ \beta_{\text{trt}} \end{pmatrix}.$$

- Η πρώτη στήλη αντιστοιχεί στον σταθερό όρο.
- Η δεύτερη στην ηλικία και η τρίτη στη θεραπεία.

Κατηγορική μεταβλητή ως 0 και 1

- Μια κατηγορική μεταβλητή με δύο κατηγορίες μπορεί να κωδικοποιηθεί αριθμητικά.
- Για παράδειγμα:

$\text{Treatment} = 0$ για ομάδα ελέγχου,

$\text{Treatment} = 1$ για ενεργό θεραπεία.

- Έτσι η μεταβλητή μπορεί να αποτελέσει μία στήλη του πίνακα X .
- Η ίδια λογική επεκτείνεται σε περισσότερες κατηγορίες με περισσότερες στήλες.

- Αφού εκτιμήσουμε το β , οι προβλεπόμενες τιμές γράφονται

$$\hat{Y} = X\hat{\beta}.$$

- Το $\hat{Y}$ είναι διάνυσμα $n \times 1$, όπως και το Y .
- Κάθε στοιχείο $\hat{Y}_i$ είναι η τιμή που προβλέπει το μοντέλο για το άτομο i .

- Το κατάλοιπο είναι

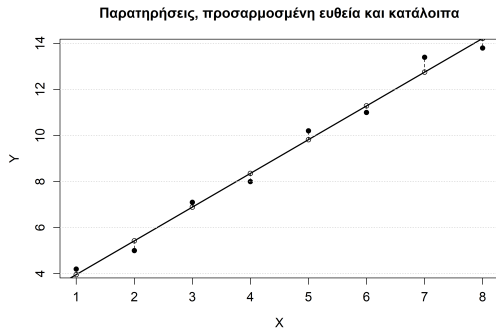
$$e_i = Y_i - \hat{Y}_i.$$

- Σε μορφή διανύσματος :

$$\mathbf{e} = \mathbf{Y} - \hat{\mathbf{Y}}.$$

- Επειδή $\hat{\mathbf{Y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$:

$$\mathbf{e} = \mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}.$$



Το άθροισμα τετραγώνων των καταλοίπων

- Θυμόμαστε ότι για ένα διάνυσμα e :

$$e^T e = \sum_{i=1}^n e_i^2.$$

- Επομένως το άθροισμα τετραγώνων των καταλοίπων γράφεται

$$e^T e.$$

- Επειδή $e = Y - X\beta$:

$$e^T e = (Y - X\beta)^T (Y - X\beta).$$

Η μέθοδος των ελαχίστων τετραγώνων

- Η μέθοδος των ελαχίστων τετραγώνων επιλέγει το β που ελαχιστοποιεί

$$S(\beta) = (Y - X\beta)^T(Y - X\beta).$$

- Πρόκειται για την ίδια ιδέα που είδαμε στο πρώτο τρίωρο:

$$\sum_{i=1}^n (Y_i - \hat{Y}_i)^2.$$

- Η Γραμμική Άλγεβρα απλώς γράφει το ίδιο άθροισμα πολύ πιο συμπαγώς.

- Από το πρώτο τρίωρο γνωρίζουμε ότι για να βρούμε ένα ελάχιστο θέτουμε την παράγωγο ίση με 0.
- Για τη συνάρτηση

$$S(\beta) = (Y - X\beta)^T(Y - X\beta),$$

η συνθήκη ελαχιστοποίησης οδηγεί στις **κανονικές εξισώσεις**:

$$X^T X \hat{\beta} = X^T Y.$$

- Δεν χρειάζεται να απομνημονεύσουμε την παραγωγή. Μας ενδιαφέρει κυρίως η μορφή του αποτελέσματος.

Η λύση των ελαχίστων τετραγώνων

- Αν ο $X^T X$ είναι αντιστρέψιμος, από

$$X^T X \hat{\beta} = X^T Y$$

παίρνουμε

$$\hat{\beta} = (X^T X)^{-1} X^T Y.$$

- Εδώ εμφανίζονται μαζί όλες οι έννοιες που είδαμε:
 - ανάστροφος,
 - γινόμενο πινάκων,
 - αντίστροφος,
 - έλεγχος διαστάσεων.

Έλεγχος διαστάσεων στον τύπο των ελαχίστων τετραγώνων

- Αν $X : n \times p$ και $Y : n \times 1$, τότε

$$X^T : p \times n,$$

$$X^T X : p \times p,$$

$$(X^T X)^{-1} : p \times p,$$

$$X^T Y : p \times 1.$$

- Επομένως

$$\hat{\beta} : (p \times p)(p \times 1) = p \times 1,$$

όπως πρέπει.

- Έστω τρεις παρατηρήσεις (x_i, Y_i) :

$$(x, Y) = (1, 3), (2, 5), (3, 7).$$

$$X = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}, \quad Y = \begin{pmatrix} 3 \\ 5 \\ 7 \end{pmatrix}.$$

- Ποια είναι η διάσταση του X^T ;
- Ποια είναι η διάσταση του $X^T X$;
- Ποια είναι η διάσταση του $X^T Y$;

$$X^T X = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 6 \\ 6 & 14 \end{pmatrix}.$$

$$X^T Y = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 3 \\ 5 \\ 7 \end{pmatrix} = \begin{pmatrix} 15 \\ 34 \end{pmatrix}.$$

- Για το προηγούμενο παράδειγμα:

$$(X^T X)^{-1} = \frac{1}{6} \begin{pmatrix} 14 & -6 \\ -6 & 3 \end{pmatrix}.$$

- Άρα

$$\hat{\beta} = (X^T X)^{-1} X^T Y = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

- Η προσαρμοσμένη ευθεία είναι

$$\hat{Y} = 1 + 2x.$$

- Στην R οι βασικές πράξεις γράφονται πολύ κοντά στη μαθηματική σημειογραφία:
- `t (X)` για τον ανάστροφο X^T .
- `X %*% Y` για γινόμενο πινάκων.
- `solve (A)` για τον αντίστροφο A^{-1} .
- Επομένως ο τύπος

$$\hat{\beta} = (X^T X)^{-1} X^T Y$$

μπορεί να υπολογιστεί ως

```
solve (t (X) %*% X) %*% t (X) %*% Y.
```

Έλεγχος με τη γραμμική παλινδρόμηση στην R

- Μπορούμε να συγκρίνουμε το αποτέλεσμα των πράξεων πινάκων με τη συνάρτηση `lm()`.
- Για τα ίδια δεδομένα:

`coef(lm(y ~ x))`

δίνει τους ίδιους συντελεστές με τον τύπο

$$(X^T X)^{-1} X^T Y.$$

- Τα στατιστικά πακέτα αυτοματοποιούν τέτοιες πράξεις Γραμμικής Άλγεβρας.

Τι πρέπει να κρατήσουμε

- Οι **διαστάσεις** καθορίζουν αν μια πράξη πινάκων επιτρέπεται.
- Ο πολλαπλασιασμός πινάκων γίνεται με **γραμμή επί στήλη**.
- Ο ανάστροφος A^T ανταλλάσσει γραμμές και στήλες.
- Ο αντίστροφος A^{-1} ικανοποιεί $AA^{-1} = I$ και δεν υπάρχει για κάθε πίνακα.
- Τα γραμμικά μοντέλα γράφονται πολύ συμπαγώς ως

$$Y = X\beta + \varepsilon.$$

Μπορούμε τώρα να διαβάσουμε τον τύπο

$$\hat{\beta} = (X^T X)^{-1} X^T Y$$

- Τι είναι το X^T ;
- Γιατί ορίζεται το $X^T X$;
- Τι σημαίνει $(X^T X)^{-1}$;
- Ποια είναι η διάσταση του $\hat{\beta}$;

- 1 Αν A είναι 3×5 , ποια είναι η διάσταση του A^T ;
- 2 Αν $A : 4 \times 3$ και $B : 3 \times 2$, ποια είναι η διάσταση του AB ; Ορίζεται το BA ;
- 3 Για

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix},$$

υπάρχει ο A^{-1} ;

- 4 Αν $X : n \times p$ και $Y : n \times 1$, ποια είναι η διάσταση του $X^T Y$;