

# Ανασκόπηση στοιχείων

## ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ

Τζούβελη Παρασκευή



# Γραμμική Άλγεβρα

# Τανυστές

- Τανυστές μηδενικής τάξης - Βαθμωτά (Scalars)
  - ένας μόνο αριθμός ( $s \in \mathfrak{R}$ )

- Τανυστές πρώτης τάξης - Διανύσματα (Vectors)
  - μονοδιάστατος πίνακας όπου  $x_1 \dots x_n \in \mathfrak{R}$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{bmatrix}$$

- Τανυστές δεύτερης τάξης - Πίνακες (Matrices)
  - δισδιάστατοι πίνακες, με  $a_{11} \dots a_{nn} \in \mathfrak{R}$

και

$$A \in \mathfrak{R}^{m \times n}, A = \begin{bmatrix} a_{11} & a_{12} & a_{1n} \\ \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & a_{mn} \end{bmatrix}$$

- Τανυστές μεγαλύτερης από 2ης τάξης - Τανυστές (Tensors)
  - πολυδιάστατοι πίνακες π.χ. 3ης τάξης όπου  $b_{111} \dots b_{kmn} \in \mathfrak{R}$

και  $B \in \mathfrak{R}^{k \times m \times n}$

# Ανάστροφο διάνυσμα $\mathbf{x}^T$ και πίνακας $\mathbf{A}^T$

- Ορίζουμε τον ανάστροφο (transpose) του διανύσματος  $\mathbf{x}$  και συμβολίζουμε  $\mathbf{x}^T$ , ως τον πίνακα γραμμή  $N$  θέσεων με στοιχεία:

$$\mathbf{x}^T = [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \dots \quad \mathbf{x}_n]$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{bmatrix}$$

- Ορίζουμε τον ανάστροφος πίνακα του πίνακα  $\mathbf{A} = (a_{ij}) \in M_{m \times n}$  και συμβολίζουμε  $\mathbf{A}^T$ , ως τον πίνακα που προκύπτει από τον  $\mathbf{A}$  με εναλλαγή μεταξύ γραμμών και στηλών του:

$$(\mathbf{A}^T)_{i,j} = (\mathbf{A})_{j,i} \quad , \quad \mathbf{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{bmatrix} \quad \mathbf{A}^T = \begin{bmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \end{bmatrix}$$

Κύρια διαγώνιος

# Ανάστροφος πίνακας $A^T$

Παραδείγματα

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \dots & \dots & \dots & \dots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{bmatrix}_{m \times n}$$

$$C = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 5 & -6 \\ -4 & 3 & 0 \\ -1 & 8 & 9 \end{bmatrix}_{4 \times 3}$$

$$A^T = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,m} \\ a_{2,1} & a_{2,2} & \dots & a_{2,m} \\ \dots & \dots & \dots & \dots \\ a_{n,1} & a_{n,2} & \dots & a_{n,m} \end{bmatrix}_{n \times m}$$

$$C^T = \begin{bmatrix} 1 & 2 & -4 & -1 \\ 0 & 5 & 3 & 8 \\ -1 & -6 & 0 & 9 \end{bmatrix}_{3 \times 4}$$

Ιδιότητες

$$\begin{aligned} (A^T)^T &= A \\ (A+B)^T &= A^T + B^T \\ (AB)^T &= B^T A^T \end{aligned}$$



# Τετραγωνικοί Πίνακες

Ένας πίνακας  $A$  είναι τετραγωνικού τύπου  $n$  όταν έχει διάσταση  $n \times n$

π.χ.  $A = \begin{pmatrix} 5 & 2 \\ 4 & 6 \end{pmatrix}$  είναι τετραγωνικού τύπου 2

## Διαγώνιοι Πίνακες

Ένας πίνακας  $A \in \mathbb{R}^{n \times n}$  ονομάζεται διαγώνιος πίνακας, αν και μόνο αν  $a_{ij} = 0$  όταν  $i \neq j$   
δηλαδή όλα τα στοιχεία έξω από την κύρια διαγώνιο είναι 0

π.χ.  $A = \begin{pmatrix} 5 & 0 \\ 0 & -3 \end{pmatrix}$

# Μοναδιαίος Πίνακας Τύπου $n$ , $I_n$

Ένας διαγώνιος πίνακας  $I \in \mathbb{R}^{n \times n}$  ονομάζεται μοναδιαίος πίνακας αν και μόνο αν

$$(I_n)_{ij} = \delta_{ij} = \begin{cases} 1, & \text{για } i=j \\ 0, & \text{για } i \neq j \end{cases}$$

## Ίχνος Τετραγωνικού Πίνακας

Έστω  $A \in \mathbb{R}^{n \times n}$ , τότε το ίχνος του  $A$  ορίζεται ως :  $\text{tr}(A) = \text{trace}(A) : \sum_{i=1}^n a_{ii} = a_{11} + a_{22} + \dots + a_{nn}$

δηλαδή ίσο με το άθροισμα των στοιχείων της κύριας διαγωνίου.



# Κάτω Τριγωνικός Πίνακας

Ένας πίνακας  $A \in \mathbb{R}^{n \times n}$  ονομάζεται κάτω τριγωνικός αν και μόνο αν  $a_{ij} = 0$  όταν  $i < j$ , δηλαδή όλα τα στοιχεία πάνω από την κύρια διαγώνιο είναι 0

$$\begin{pmatrix} 3 & 0 & 0 \\ 4 & 6 & 0 \\ 5 & 7 & 8 \end{pmatrix}$$

# Άνω Τριγωνικός Πίνακας

Ένας πίνακας  $A \in \mathbb{R}^{n \times n}$  ονομάζεται άνω τριγωνικός αν και μόνο αν  $a_{ij} = 0$  όταν  $i > j$ , δηλαδή όλα τα στοιχεία κάτω από την κύρια διαγώνιο είναι 0

$$\begin{pmatrix} 3 & 4 & 5 \\ 0 & 6 & 7 \\ 0 & 0 & 8 \end{pmatrix}$$





# Αντίστροφος πίνακας

Ο πίνακας  $A \in \mathbb{R}^{n \times n}$  είναι αντιστρέψιμος εάν υπάρχει ένας πίνακας  $B \in \mathbb{R}^{n \times n}$  έτσι ώστε  $B \cdot A = I$  και  $A \cdot B = I$ , όπου  $I$  είναι  $n \times n$  μοναδιαίος πίνακας.

Υπάρχει το πολύ ένας τέτοιος  $B$  και λέγεται αντίστροφος του  $A$  και συμβολίζεται με  $A^{-1}$   
:  $A \cdot A^{-1} = A^{-1} \cdot A = I$

→ Οπότε, για να είναι ο πίνακας  $A \in \mathbb{R}^{n \times n}$  αντιστρέψιμος θα πρέπει η ορίζουσά του είναι διάφορη του 0 :  $\det A \neq 0$

## Ιδιότητες

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

$$(AB)^{-1} = B^{-1}A^{-1}$$

$$(A+B)^{-1} \neq A^{-1} + B^{-1}$$



# Αντίστροφος πίνακας

Παράδειγμα 1.1 (Υπαρξη αντιστρόφου πίνακα 2x2) Θεωρούμε έναν πίνακα  $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in \mathbb{R}^{2 \times 2}$ . Αν τον πολλαπλασιάσουμε με τον πίνακα  $B = \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} \in \mathbb{R}^{2 \times 2}$  τότε θα έχουμε  $AB = (a_{11}a_{22} - a_{12}a_{21})\mathbf{I}$   
 Τότε ο αντίστροφος  $A^{-1}$  του πίνακα θα είναι ο

$$A^{-1} = \frac{1}{(a_{11}a_{22} - a_{12}a_{21})} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} \quad (1.1)$$

αν και μόνο αν  $(a_{11}a_{22} - a_{12}a_{21}) \neq 0$  (όπου  $(a_{11}a_{22} - a_{12}a_{21})$  είναι η ορίζουσα του πίνακα  $A$ ). Μπορούμε να χρησιμοποιούμε την ορίζουσα ενός πίνακα για να αποφανθούμε αν αυτός είναι αντιστρέψιμος.

- Εάν  $A = [a]$  τότε  $A^{-1} = [1/a]$
- Εάν  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$  και  $ad - bc \neq 0$  τότε  $A^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$



# Συμμετρικοί πίνακες

- Ένας τετραγωνικός πίνακας  $S \in \mathbb{R}^{n \times n}$  ο οποίος είναι ίσος με τον ανάστροφό του, δηλαδή  $S = S^T$ , είναι ένας συμμετρικός πίνακας.

π.χ. εάν  $A$  είναι ένας πίνακας μέτρησης αποστάσεων, με  $A_{i,j}$  που δίνει την απόσταση από το σημείο  $i$  στο σημείο  $j$ , τότε  $A_{i,j} = A_{j,i}$  επειδή οι συναρτήσεις απόστασης είναι συμμετρικές.

$$\begin{pmatrix} 3 & 5 \\ 5 & -1 \end{pmatrix}$$

$$\begin{pmatrix} 7 & -5 & 2 \\ -5 & 3 & 4 \\ 2 & 4 & 8 \end{pmatrix}$$

$$\begin{pmatrix} 5 & 6 & -4 & -3 \\ 6 & 7 & 2 & 9 \\ -4 & 2 & 1 & 5 \\ -3 & 9 & 5 & 6 \end{pmatrix}$$



# Αντισυμμετρικοί πίνακες

- Αν αντίθετα, ο τετραγωνικός πίνακας  $S \in \mathbb{R}^{n \times n}$  ο οποίος είναι ίσος με τον αρνητικό του ανάστροφο, δηλαδή  $S = S^T$  τότε λέμε ότι ο  $S$  είναι ένας **αντισυμμετρικός** πίνακας.

Δηλαδή:  $(A)^{ij} = -(A)^{ji}, \forall i, j$

Παραδείγματα

$$\begin{pmatrix} 0 & 3 \\ -3 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 5 & 2 \\ -5 & 0 & 4 \\ -2 & -4 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 6 & -4 & 3 \\ -6 & 0 & 2 & 9 \\ 4 & -2 & 0 & -5 \\ -3 & -9 & 5 & 0 \end{pmatrix}$$



# Hermitian πίνακες

- Σε μιγαδικούς πίνακες, η συμμετρία συχνά αντικαθιστάται από την έννοια Hermitian πίνακες  $A^H$ , που ικανοποιεί  $A^* = \overline{A^T} = A$  όπου ο αστερίσκος υποδηλώνει το συζυγή ανάστροφο ενός πίνακα, δηλαδή, η μετάθεση του συζυγή του  $A$ :  $a_{ij} = \overline{a_{ji}}$

Παράδειγμα

$$A: \begin{bmatrix} \overline{a} & \overline{b+ic} \\ \overline{b-ic} & \overline{d} \end{bmatrix} \quad A^T: \begin{bmatrix} a & b-ic \\ b+ic & d \end{bmatrix} \quad \overline{A^T} \begin{bmatrix} a & b+ic \\ b-ic & d \end{bmatrix}$$



# Πράξεις πινάκων

- Πρόσθεση πινάκων

Αν οι πίνακες έχουν το ίδιο μέγεθος μπορούμε να προσθέσουμε τα αντίστοιχα στοιχεία τους:  $C=A+B$  όπου  $C_{i,j} = A_{i,j} + B_{i,j}$

- Πίνακας με βαθμωτό ή διάνυσμα

Προσθέτουμε/πολλαπλασιάζουμε το βαθμωτό με κάθε στοιχείο του πίνακα:

$$\mathbf{D} = a \cdot \mathbf{B} + c \text{ όπου } D_{i,j} = a \cdot B_{i,j} + c$$

- Πολλαπλασιασμός πινάκων

Προϋπόθεση: Οι στήλες του A ίσες με τις γραμμές του B

$$C = A * B, C_{i,j} = \sum_k (A_{i,k} \cdot B_{k,j}) \quad C(m \times p) = A(m \times n) * B(n \times p)$$

- Γινόμενο στοιχείο προς στοιχείο (Element-wise product ή Hadamard product)

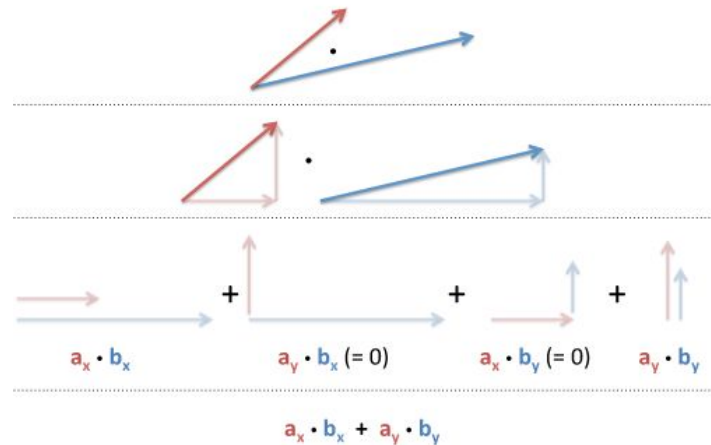
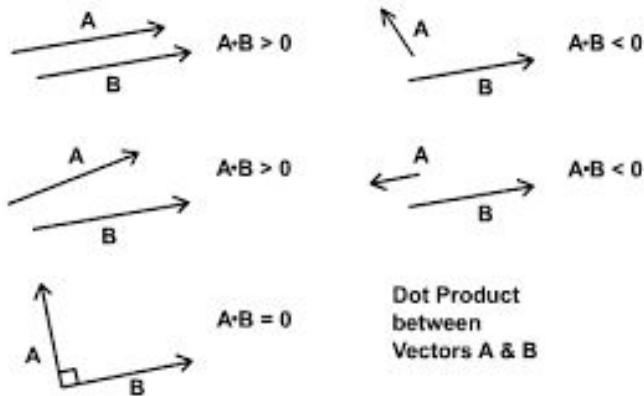
$$A \cdot B = a_{i,j} \cdot b_{i,j}$$

# Εσωτερικό Γινόμενο Διανυσμάτων

Η ποσότητα  $x^T y$  είναι το εσωτερικό γινόμενο των διανυσμάτων  $x$  και  $y$  του  $\mathcal{R}^n$

Αν τα  $x, y$  είναι ορθογώνια,  $x^T y = 0$

$$\begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix} = [x_1 \quad x_2 \quad \dots \quad x_n] * \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix} = x_1 * y_1 + x_2 * y_2 \dots + x_n * y_n$$





# Γραμμική Ανεξαρτησία Διανυσμάτων

Εάν ο τετριμμένος συνδυασμός είναι ο μόνος που παράγει το μηδέν δηλαδή

$$c_1 u_1 + c_2 u_2 + \dots + c_n u_n = 0 \text{ συμβαίνει μόνον όταν } c_1 = c_2 = \dots = c_n = 0,$$

τότε τα διανύσματα  $u_1, u_2, \dots, u_n$  είναι γραμμικώς ανεξάρτητα.

$$\text{Πχ. } \begin{bmatrix} 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix} = [1 \quad 1] * \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 0$$

Αν αυτό δεν συμβαίνει, τότε είναι γραμμικώς εξαρτημένα και κάποιο από αυτά είναι γραμμικός συνδυασμός των υπολοίπων.

$$\begin{aligned} c_1 - c_2 + 2c_3 &= 0 \\ c_2 + 3/2c_3 &= 0 \end{aligned}$$

$$u_1 = \begin{bmatrix} 1 \\ 1 \\ 3 \\ 3 \end{bmatrix}, u_2 = \begin{bmatrix} -1 \\ 1 \\ -1 \\ -3 \end{bmatrix}, u_3 = \begin{bmatrix} 2 \\ 5 \\ 9 \\ 6 \end{bmatrix} \rightarrow c_1 \begin{bmatrix} 1 \\ 1 \\ 3 \\ 3 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 1 \\ -1 \\ -3 \end{bmatrix} + c_3 \begin{bmatrix} 2 \\ 5 \\ 9 \\ 6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & 3/2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$





# Γραμμική Ανεξαρτησία Διανυσμάτων

**Άσκηση** Εξετάστε αν τα διανύσματα  $(1,1,0,0)$ ,  $(1,0,1,0)$ ,  $(0,0,1,1)$ ,  $(0,1,0,1)$  είναι γραμμικά ανεξάρτητα ή όχι και ελέγξτε εάν το διάνυσμα  $(0,0,0,1)$  βρίσκεται στο χώρο που παράγουν.

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \xrightarrow{\substack{R_2 - R_1 \\ (-)}} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \xrightarrow{\substack{R_2 \cdot (-1) \\ (-)}} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \xrightarrow{\substack{R_4 - R_3 \\ (-)}} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = U, \quad \text{άρα } r(A) = [\text{πλήθος μη-μηδεν. γραμμών του } U]$$

$$\Rightarrow r(A) = 3 \rightarrow \text{Πλήθος βασικών μεταβλητών}$$

και  $\dim N(A) = n - r(A) = 4 - 3 = 1 \neq 0 \rightarrow$  Πλήθος ελευθέρων μεταβλητών  
 άρα  $\exists c \neq 0$  τ.ω.  $Ac = 0$  και τα διανύσματα είναι γραμμικά εξαρτημένα



# Γραμμική Ανεξαρτησία Διανυσμάτων

**Άσκηση** Εξετάστε αν τα διανύσματα  $(1,1,0,0)$ ,  $(1,0,1,0)$ ,  $(0,0,1,1)$ ,  $(0,1,0,1)$  είναι γραμμικά ανεξάρτητα ή όχι και ελέγξτε εάν το διάνυσμα  $(0,0,0,1)$  βρίσκεται στο χώρο που παράγουν.

Το διάνυσμα  $(0,0,0,1)$  βρίσκεται στο χώρο που παράγουν αν  $\exists$   
 $\lambda_1, \lambda_2, \lambda_3, \lambda_4 \in \mathbb{R}$  τ.ω. 
$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \lambda_1 \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + \lambda_3 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} + \lambda_4 \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

$$\Leftrightarrow \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

η τελευταία γραμμή δίνει  $0\lambda_1 + 0\lambda_2 + 0\lambda_3 + 0\lambda_4 = 1$ , άρα το σύστημα είναι ασύμβατο και επομένως,

$$(0,0,0,1) \notin \langle (1,1,0,0), (1,0,1,0), (0,0,1,1), (0,1,0,1) \rangle$$

Αν δεν υπάρχει λύση, τότε λέμε ότι το σύστημα είναι **ασύμβατο** (inconsistent).



# Γραμμική Ανεξαρτησία Διανυσμάτων

Ένα σύνολο  $n$  διανυσμάτων του  $\mathcal{R}^m$  είναι κατ' ανάγκην εξαρτημένο, όταν  $n > m$

Κάθε διάνυσμα  $u$  του διανυσματικού χώρου  $V$  μπορεί να εκφραστεί ως γραμμικός συνδυασμός του  $w$ :  $u = c_1 w_1 + \dots + c_l w_l$  για κάποιους συντελεστές  $c_i$

Έστω διανύσματα που ξεκινούν από την αρχή ενός 3D χώρου:

2 εξαρτημένα διανύσματα  $\Rightarrow$  Περιέχονται στην ίδια ευθεία

3 εξαρτημένα διανύσματα  $\Rightarrow$  Περιέχονται στο ίδιο επίπεδο



# Βάση και Διάσταση Διανυσματικού Χώρου

Βάση ενός διανυσματικού χώρου είναι ένα σύνολο διανυσμάτων που έχει ταυτόχρονα τις δύο ιδιότητες:

1. Είναι γραμμικώς ανεξάρτητο
2. Παράγει τον χώρο.

Όταν ένας τετραγωνικός πίνακας είναι αντιστρέψιμος, τότε οι στήλες του είναι ανεξάρτητες και αποτελούν μία βάση του  $\mathcal{R}^n$ .

Αν  $u_1, \dots, u_m$  και  $w_1, \dots, w_n$  είναι δύο βάσεις του ίδιου διανυσματικού χώρου  $V$ , τότε  $m=n$

Το  $m$  (ή το  $n$ ) εκφράζει τους “βαθμούς ελευθερίας” του χώρου και ονομάζεται διάσταση του  $V$

→  $\dim(V) = [\text{μέγιστο πλήθος γραμμικώς ανεξάρτητων διανυσμάτων του } V]$

π.χ.  $\dim(\mathcal{R}^2)=2$ ,  $\dim(\mathcal{R}^3)=3$



# Βάση και Διάσταση Διανυσματικού Χώρου

Να βρεθεί μία βάση και η διάσταση του χώρου

$$V = \{(x, y, z) \in \mathbb{R}^3 \mid x - y + 2z = 0\}$$

$$x - y + 2z = 0 \rightarrow x = y - 2z, y, z \in \mathbb{R}$$

$$V = \{(x, y, z) \in \mathbb{R}^3 \mid x = y - 2z\} = \{(y - 2z, y, z) \mid y, z \in \mathbb{R}\}$$

$$= \{(y, y, 0) + (-2z, 0, z) \mid y, z \in \mathbb{R}\}$$

$$= \{y(1, 1, 0) + z(-2, 0, 1) \mid y, z \in \mathbb{R}\}$$

$$= \langle (1, 1, 0), (-2, 0, 1) \rangle$$

$$\mathbf{e}_1, \mathbf{e}_2$$

Οδηγοί

Για να αποτελέσουν βάση του χώρου αυτά τα 2 διανύσματα θα πρέπει να είναι γραμμικώς ανεξάρτητα:

$$\begin{bmatrix} 1 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} \xrightarrow{1st*2+2nd} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}$$

# Ορθογώνιοι πίνακες

- Ένας τετραγωνικός πίνακας ονομάζεται ορθογώνιος πίνακας εάν και μόνο αν τα  $n$  διανύσματα-στήλες (ή -γραμμές) του αποτελούν ορθοκανονικό σύστημα του χώρου διάστασης  $n \times n$

Π.χ. (ταυτοτικός)

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

(μετάθεσης)

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

(περιστροφή)

$$\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

- Ισοδύναμα, ένας τετραγωνικός πίνακας  $A$  είναι ορθογώνιος αν η μετάθεσή του είναι ίση με τον αντίστροφό του:  $A^T = A^{-1}$   
και ισχύει  $A^T A = A A^T = I$

# Αντίστροφος πίνακας

Παράδειγμα

δ) Αν  $A \in \mathbb{R}^{n \times n}$  είναι αντιστρέψιμος, τότε:  $(A^{-1})^T = (A^T)^{-1}$

$$\text{π.χ. } A = \begin{bmatrix} 1 & -2 \\ -3 & 5 \end{bmatrix} \Rightarrow A^{-1} = \frac{1}{5-6} \begin{bmatrix} 5 & 2 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} -5 & -2 \\ -3 & -1 \end{bmatrix} \Rightarrow (A^{-1})^T = \begin{bmatrix} -5 & -3 \\ -2 & -1 \end{bmatrix}$$

$$\text{και } A^T = \begin{bmatrix} 1 & -3 \\ -2 & 5 \end{bmatrix} \Rightarrow (A^T)^{-1} = \frac{1}{-1} \begin{bmatrix} 5 & 3 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} -5 & -3 \\ -2 & -1 \end{bmatrix}$$



# Ορθογώνιοι πίνακες

- Η ορίζουσα οποιουδήποτε ορθογώνιου πίνακα είναι είτε +1 (κάνει μια καθαρή περιστροφή), ή -1 (είναι μια καθαρή αντανάκλαση, ή μια σύνθεση της αντανάκλασης και της περιστροφής).
- Εάν  $\mathbf{A} \in \mathbb{R}^{n \times n}$  είναι ορθογώνιος πίνακας, τότε  $\mathbf{A}^T \mathbf{A} = \mathbf{I}_n$

Έπεται ότι:

$$\det(\mathbf{A}^T \mathbf{A}) = \det(\mathbf{I}_n) = 1,$$

$$\det(\mathbf{A}^T) \det(\mathbf{A}) = 1,$$

$$\det(\mathbf{A}) = \pm 1$$



# Σύστημα Γραμμικών Εξισώσεων

$\mathbf{A} \in \mathbb{R}^{m \times n}$ , γνωστός πίνακας,

$\mathbf{b} \in \mathbb{R}^m$ , διάνυσμα,

$\mathbf{x} \in \mathbb{R}^n$ , διάνυσμα με τις άγνωστες μεταβλητές

$$\mathbf{Ax} = \mathbf{b}$$

Αναλύεται σε:

$$\begin{array}{l}
 A_{1,:} \mathbf{x} = b_1 \\
 A_{2,:} \mathbf{x} = b_2 \\
 \dots\dots\dots \\
 A_{m,:} \mathbf{x} = b_m
 \end{array}
 \quad \text{ή} \quad
 \begin{array}{l}
 A_{1,1}x_1 + A_{1,2}x_2 + \dots + A_{1,n}x_n = b_1 \\
 A_{2,1}x_1 + A_{2,2}x_2 + \dots + A_{2,n}x_n = b_2 \\
 \dots\dots\dots \\
 A_{m,1}x_1 + A_{m,2}x_2 + \dots + A_{m,n}x_n = b_m
 \end{array}$$

Μπορεί να έχει καμία λύση, πολλές λύσεις, ακριβώς μία λύση (πολ/σμό με ανάστροφο)

# Σύστημα Γραμμικών Εξισώσεων



$$Ax = b$$

Σκέψη :

Οι στήλες του  $A$  ορίζουν διαφορετικές κατευθύνσεις που μπορούμε να “κινήσουμε” στο χώρο από το αρχικό σημείο  $0$  για να προσεγγίσουμε το  $b$ .

$$Ax = \sum_i x_i A_{:,i} \rightarrow \text{Γραμμικός συνδυασμός}$$

Το εύρος (span) είναι το σύνολο όλων των αρχικών διανυσμάτων να ληφθούν με

ενός των γραφών  $v = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$    $w = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$   μάτων τορούν των

some combinations of  $v$  and  $w$



# Σύστημα Γραμμικών Εξισώσεων

$$\mathbf{Ax} = \mathbf{b}$$



Έχει λύση η  $\mathbf{Ax} = \mathbf{b}$ ;

Θα ελέγχουμε αν το  $\mathbf{b}$  είναι στο εύρος των στηλών του  $\mathbf{A}$ :

Για να μπορεί το σύστημα  $\mathbf{Ax} = \mathbf{b}$  να έχει λύση για όλες τις τιμές του  $\mathbf{b} \in \mathcal{R}_m$ , απαιτούμε ο χώρος των στηλών του  $\mathbf{A}$  να είναι ο  $\mathcal{R}_m$

Πρέπει να έχει ακριβώς  $m$  γραμμικές ανεξάρτητες στήλες, όχι τουλάχιστον  $m$ .



Προκειμένου η μήτρα  $A$  να έχει αντίστροφο, πρέπει επιπλέον να διασφαλίσουμε ότι η εξίσωση έχει το πολύ μία λύση για κάθε τιμή του  $b$ .

(Διαφορετικά, υπάρχουν περισσότεροι από έναν τρόποι παραμετροποίησης κάθε λύσης)

Για να έχει αντίστροφο η μήτρα  $A$  θα πρέπει:

- να είναι τετράγωνη, δηλαδή, απαιτούμε ότι  $m = n$  και
- όλες οι στήλες πρέπει να είναι γραμμικώς ανεξάρτητες

Μια τετράγωνη μήτρα με γραμμικά εξαρτημένες στήλες είναι γνωστή ως μοναδική.

Εάν η  $A$  δεν είναι τετράγωνη, αλλά μοναδική, μπορεί ακόμα να επιλυθεί η εξίσωση.

→ Ωστόσο, δεν μπορούμε να χρησιμοποιήσουμε τη μέθοδο της αντιστροφής της μήτρας για να βρούμε τη λύση.



Λύση: Αντιστροφή της μήτρας

$$Ax = b$$

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

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

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

# Τριγωνική παραγοντοποίηση $A=LU$

Αν δεν απαιτούνται εναλλαγές γραμμών, ο αρχικός πίνακας μπορεί να γραφεί ως γινόμενο  $A=LU$

$$L : \text{κάτω τριγωνικός} \quad L = \begin{bmatrix} 1 & 0 & \dots & 0 \\ \lambda_{2,1} & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \lambda_{n,1} & \lambda_{n,2} & \dots & 1 \end{bmatrix} \quad U = \begin{bmatrix} u_{1,1} & u_{1,2} & \dots & u_{1,n} \\ 0 & u_{2,2} & \dots & u_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & u_{m,n} \end{bmatrix}$$

$U$  : άνω τριγωνικός

- Εμφανίζεται μετά τη διαδοχική απαλοιφή και πριν την ανάδρομη αντικατάσταση.
- Τα διαγώνια στοιχεία του είναι οι οδηγοί

π.χ.

$$A = \begin{bmatrix} 3 & 1 \\ 4 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{4}{3} & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 0 & \frac{2}{3} \end{bmatrix} = LU$$

$$A = \begin{bmatrix} -5 & 1 & -3 & 4 \\ 8 & -7 & 3 & 2 \\ -3 & -6 & -1 & -1 \\ 0 & 0 & 3 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{8}{5} & 1 & 0 & 0 \\ \frac{3}{5} & \frac{33}{27} & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} -5 & 1 & -3 & 4 \\ 0 & -\frac{27}{5} & -\frac{9}{5} & \frac{42}{5} \\ 0 & 0 & 3 & -\frac{41}{3} \\ 0 & 0 & 0 & \frac{68}{3} \end{bmatrix} = LU$$

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 7 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{bmatrix} = LU$$



# Εύρεση $A^{-1}$ : Μέθοδος των Gauss-Jordan

Έστω η εξίσωση :  $AA^{-1}=I$  .

Εάν θεωρηθεί στήλη προς στήλη αυτή η εξίσωση προσδιορίζει τις στήλες του  $A^{-1}$

$$Ax_1 = e_1, Ax_2 = e_2, Ax_3 = e_3$$

όπου

$$[e_1 \quad e_2 \quad e_3] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$[A \quad e_1 \quad e_2 \quad e_3] = [U \quad L]$$

$$[U \quad L] = [I \quad A^{-1}]$$

# Εύρεση $A^{-1}$ : Μέθοδος των Gauss-Jordan

$$A = \begin{pmatrix} 9 & 3 & 4 \\ 4 & 3 & 4 \\ 1 & 1 & 1 \end{pmatrix} \quad (A \mid I_3) = \begin{pmatrix} 9 & 3 & 4 & | & 1 & 0 & 0 \\ 4 & 3 & 4 & | & 0 & 1 & 0 \\ 1 & 1 & 1 & | & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 9 & 3 & 4 & | & 1 & 0 & 0 \\ 0 & 5/3 & 20/9 & | & -4/9 & 1 & 0 \\ 0 & 2/3 & 5/9 & | & -1/9 & 0 & 1 \end{pmatrix}$$

$$\begin{aligned} [A \ e_1 \ e_2 \ e_3] &= [U \ L] \\ [U \ L] &= [I \ A^{-1}] \end{aligned}$$

$$= \begin{pmatrix} 9 & 0 & 0 & | & 9/5 & -9/5 & 0 \\ 0 & 5/3 & 20/9 & | & -4/9 & 1 & 0 \\ 0 & 0 & -1/3 & | & 1/15 & -2/5 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 9 & 0 & 0 & | & 9/5 & -9/5 & 0 \\ 0 & 5/3 & 0 & | & 0 & -5/3 & 20/3 \\ 0 & 0 & -1/3 & | & 1/15 & -2/5 & 1 \end{pmatrix}$$

Διαιρούμε όλα τα στοιχεία της  $i$ -γραμμής του πίνακα εκ δεξιών της διακεκομμένης γραμμής με το μη μηδενικό στοιχείο της  $i$ -γραμμής του διαγώνιου πίνακα και παίρνουμε:

$$A^{-1} = \begin{pmatrix} 1/5 & -1/5 & 0 \\ 0 & -1 & 4 \\ -1/5 & 6/5 & -3 \end{pmatrix}$$



# Παραγοντοποίηση $A=LDU$

Ο  $L$  παραμένει ίδιο όπως στην  $A=LU$

Ο πίνακας  $D$  είναι διαγώνιος και περιέχει την διαγώνιο του πίνακα  $U$  της παραγοντοποίησης  $LU$  (στοιχεία οδηγών):

Ο νέος άνω τριγωνικός πίνακας  $U$  προκύπτει από τον πίνακα  $U$  της παραγοντοποίησης  $LU$  διαιρώντας κάθε στοιχείο του με το στοιχείο της διαγωνίου (δηλ. τον οδηγό) της ίδιας γραμμής

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 7 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1/2 \end{bmatrix} = LDU$$

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 7 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 3 & 6 \\ 0 & 0 & 4 \end{pmatrix} = LU$$

# Νόρμες (Norms)

Μέτρηση μεγέθους ενός διανύσματος  $\mathbf{x}$

$$L^p = \|\mathbf{x}\|_p = \left( \sum |x_i|^p \right)^{1/p} \quad p \in \mathbb{R}, p \geq 1$$

Η νόρμα είναι μία συνάρτηση  $f$  που ικανοποιεί τις ακόλουθες ιδιότητες:

$$f(\mathbf{x}) = 0 \Rightarrow \mathbf{x} = \mathbf{0}$$

$$f(\mathbf{x} + \mathbf{y}) \leq f(\mathbf{x}) + f(\mathbf{y}) \text{ (the **triangle inequality**)}$$

$$\forall \alpha \in \mathbb{R}, f(\alpha \mathbf{x}) = |\alpha| f(\mathbf{x})$$



# $L^2$ νόρμα (Ευκλείδεια νόρμα)

- Για  $p=2 \rightarrow$  η ευκλείδεια απόσταση από την αρχή μέχρι το σημείο που προσδιορίζεται από το  $x$ .
- Δηλώνεται συχνά ως  $\|x\|$ , με το δείκτη 2 να παραλείπεται.
- Συνήθως υπολογίζουμε το τετράγωνο της  $L^2$  νόρμας, απλά ως  $x^T x$
- Υπολογιστικά, συχνά, το τετράγωνο της  $L^2$  νόρμα μπορεί να αυξάνεται πολύ αργά κοντά το σημείο που προσδιορίζεται από το  $x$   
 $\rightarrow$  απαγορευτικό υπολογιστικά



# L<sup>1</sup> νόρμα

- Συχνά είναι σημαντικό να γίνεται διάκριση μεταξύ στοιχείων που είναι ακριβώς μηδενικά και αυτών που είναι μικρά αλλά μη μηδενικά.

$$\|x\|_1 = \sum_i |x_i|$$

- Κάθε φορά που ένα στοιχείο του  $x$  μετακινείται κατά  $\epsilon$  μακριά από το 0, η L<sup>1</sup> αυξάνεται κατά  $\epsilon$
- Υποκαθιστά το πλήθος των μη μηδενικών τιμών.

Μερικές φορές μετράμε το μέγεθος του διανύσματος υπολογίζοντας τον αριθμό των μη-μηδενικών στοιχείων (Λανθασμένη ορολογία το νόρμα L<sup>0</sup>).



# L<sup>∞</sup> νόρμα (Max νόρμα)

Υπολογίζει την απόλυτη τιμή του μεγαλύτερου στοιχείου του διανύσματος  $x$

$$\|x\|_{\infty} = \max_i |x_i|$$

## Υπολογισμός μεγέθους μήτρας

Νόρμα Frobenius

$$\|A\|_F = \sqrt{\sum_{i,j} A_{i,j}^2}$$

Είναι το ανάλογο της L<sup>2</sup> ενός διανύσματος.

## Εσωτερικό γινόμενο με χρήση νόρμας

$$x^T y = \|x\|_2 \|y\|_2 \cos\theta \quad \text{όπου } \theta \text{ η γωνία μεταξύ } x \text{ και } y$$



# Διαγώνιοι πίνακες

Ένας πίνακας  $D$  είναι διαγώνιος εάν και μόνο εάν  $D_{i,j} = 0$  για όλα τα  $i \neq j$

$diag(v)$  : τετραγωνικός διαγώνιος πίνακας,  
το διάνυσμα  $v$  περιέχει τις τιμές της διαγωνίου

Ο πολ/σμός του με άλλο πίνακα είναι υπολογιστικά αποδοτικός.

$$diag(v)x = v \odot x$$

Αντίστροφος :  $\exists$  αν τα στοιχεία της διαγωνίου είναι μη μηδενικά

$$diag(v)^{-1} = diag([1/v_1, \dots, 1/v_n])$$

- Σε πολλές περιπτώσεις, αλγόριθμοι μηχανικής μάθησης περιορίζουν πίνακες να είναι διαγώνιοι, κερδίζοντας υπολογιστικό κόστος και χρόνο.
- Είναι δυνατή η κατασκευή διαγώνιου πίνακα

# Ιδιοτιμές και ιδιοδιανύσματα

## Επίλυση γραμμικών συστημάτων

Επινοήθηκαν για επίλυση διαφορικών εξισώσεων :

Η λύση θα είναι της μορφής :

$$\frac{dy}{dt} = Ay$$

$$y(t) = e^{\lambda t} x$$

$$\Rightarrow \lambda e^{\lambda t} x = A e^{\lambda t} x \Rightarrow \underline{Ax = \lambda x}$$

Ιδιοτιμή      Ιδιοδιάνυσμα

- Ο αριθμός  $\lambda$  είναι ιδιοτιμή του  $A$  όταν και μόνον όταν ισχύει  $\det(A - \lambda I) = 0$
- Αυτή είναι η χαρακτηριστική εξίσωση και σε κάθε λύση της  $\lambda$  αντιστοιχεί ένα ιδιοδιάνυσμα  $x$ :

$$(A - \lambda I)x = 0 \quad \text{ή} \quad Ax = \lambda x$$

# Βήματα επίλυσης του προβλήματος ιδιοτιμών



1. Υπολογισμός της ορίζουσας του
2. Εύρεση των ριζών αυτού του πολυωνύμου  $A - \lambda I \rightarrow$  ιδιοτιμές
3. Για κάθε ιδιοτιμή, επίλυση του συστήματος  $(A - \lambda I)x = 0 \rightarrow$  ιδιοδιανύσματα

π.χ. Διαφορική εξίσωση

$$\frac{dy_1}{dt} = 4y_1 - 5y_2$$

$$\frac{dy_2}{dt} = 2y_1 - 3y_2$$

$$t = 0 : y_1 = 8, y_2 = 5$$

$$A = \begin{bmatrix} 4 & -5 \\ 2 & -3 \end{bmatrix}$$

$$1., 2. \quad |A - \lambda I| = \begin{vmatrix} 4 & -5 \\ 2 & -3 \end{vmatrix} - \lambda \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = \begin{vmatrix} 4-\lambda & -5-0 \\ 2-0 & -3-\lambda \end{vmatrix} = (4-\lambda)(-3-\lambda) + 10 = -(4-\lambda)(3+\lambda) + 10 = (\lambda-4)(\lambda+3) - 10 = \lambda^2 - 4\lambda + 3\lambda - 12 + 10 = \lambda^2 - \lambda - 2 \rightarrow \lambda_1 = -1, \lambda_2 = 2$$

3.

$$(A - \lambda_1 I)x = \begin{bmatrix} 5 & -5 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \rightarrow x_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$(A - \lambda_2 I)x = \begin{bmatrix} 2 & -5 \\ 2 & -5 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \rightarrow x_2 = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$$

Ειδικές λύσεις :

$$u = e^{\lambda_1 t} x_1 = e^{-t} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$u = e^{\lambda_2 t} x_2 = e^{2t} \begin{bmatrix} 5 \\ 2 \end{bmatrix}$$

Γενική λύση:

$$u = c_1 e^{\lambda_1 t} x_1 + c_2 e^{\lambda_2 t} x_2$$

$$\text{Χρήση αρχικών συνθηκών} \rightarrow \begin{cases} 8 = c_1 + 5c_2 \\ 5 = c_1 + 2c_2 \end{cases}$$

$$u = 3e^{-t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + e^{2t} \begin{bmatrix} 5 \\ 2 \end{bmatrix}$$





# Ιδιοτιμές και ιδιοδιανύσματα

Εξίσωση-κλειδί:  $Ax = \lambda x$

- Οι ιδιοτιμές παριστάνουν το βηματισμό στο χρόνο (magnitude)
- Τα ιδιοδιανύσματα παριστάνουν τις “κανονικές καταστάσεις” του συστήματος και επιδρούν ανεξάρτητα το ένα από το άλλο (direction)
- Μπορούμε να παρακολουθήσουμε τη συμπεριφορά κάθε ιδιοδιανύσματος και να συνδυάσουμε αυτές τις κανονικές καταστάσεις για να βρούμε μια λύση

→ Διαγωνοποίηση

$$\left. \begin{aligned} \frac{dy}{dt} &= Ay \\ y(t) &= e^{\lambda t} x \end{aligned} \right\}$$

$$\Rightarrow \lambda e^{\lambda t} x = A e^{\lambda t} x \Rightarrow Ax = \lambda x$$

Eigenvector Matrix

$$\Rightarrow A \begin{bmatrix} x_1 & x_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 x_1 & \lambda_2 x_2 \end{bmatrix}$$

Eigenvalue Matrix

$$\Rightarrow A \begin{bmatrix} x_1 & x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

$$\Rightarrow AV = V\Lambda$$

$$A^2 = ;$$

Ο πίνακας ιδιοτιμών υψωμένος στο τετράγωνο

$$\Rightarrow AVV^{-1} = V\Lambda V^{-1}$$

$$A^2 = V\Lambda V^{-1}V\Lambda V^{-1} = V\Lambda I\Lambda V^{-1} = V\Lambda^2 V^{-1}$$

$$\Rightarrow AI = V\Lambda V^{-1}$$

$$A^n = ;$$

Ίδιος πίνακας ιδιοδιανυσμάτων όπως ο A

$$\Rightarrow \boxed{A = V\Lambda V^{-1}}$$

$$\boxed{A^n = V\Lambda^n V^{-1}}$$

Ο πίνακας ιδιοτιμών υψωμένος στην n



# Πίνακες, ιδιοτιμές, ιδιοδιανύσματα

Συμμετρικός πίνακας

$$S = S^T$$

→ πραγματικές ιδιοτιμές

→ ορθογώνια ιδιοδιανύσματα

Ανάστροφος πίνακας

$$A^T = -A$$

→ φανταστικές ιδιοτιμές

→ ορθογώνια μιγαδικά ιδιοδιανύσματα

Ορθογώνιος πίνακας

$$Q^T Q = I$$

→ για όλες τις ιδιοτιμές ισχύει:  $|\lambda| = 1$

→ ορθογώνια μιγαδικά ιδιοδιανύσματα

Ένας πίνακας ονομάζεται **ιδιάζων (singular)** ανν οποιαδήποτε από τις ιδιοτιμές του είναι μηδέν.



# Πίνακες, ιδιοτιμές, ιδιοδιανύσματα

## Ορισμένος πίνακας

- Ένας συμμετρικός  $n \times n$  πίνακας ονομάζεται θετικά (αρνητικά) ορισμένος, εάν για όλα τα μη μηδενικά διανύσματα  $x \in \mathbb{R}^n$  το  $Q(x) = x^T A x$  παίρνει μόνο θετικές (αρνητικές τιμές).
- Εάν ένας συμμετρικός  $n \times n$  πίνακας παίρνει μόνο θετικές (αρνητικές) ονομάζεται γνησίως-θετικός (αρνητικός)  $\rightarrow x^T A x = 0 \Rightarrow x = 0$
- Εάν ένας συμμετρικός  $n \times n$  πίνακας παίρνει θετικές (αρνητικές) ή μηδενικές ονομάζεται ημι-θετικός (αρνητικός)  $\rightarrow \forall x, x = x^T A x \geq 0$
- Εάν ένας συμμετρικός  $n \times n$  δεν είναι ούτε θετικός ούτε αρνητικός τότε ο πίνακας είναι αόριστος.



# Παραγοντοποίηση

$$A=LU \quad (\text{Απαλοιφή})$$

$$A=QR \quad (\text{Gram-Schmidt})$$

$$S=Q\Lambda Q^T \quad (\text{Symmetric: } [\sigma_1 \ \dots \ \sigma_n] \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} \sigma_1^T \\ \vdots \\ \sigma_n^T \end{bmatrix} )$$

$$A=X\Lambda X^{-1}$$

$$A=U\Sigma V^T \quad (\text{Ορθογώνιος} \times \text{Διαγώνιος} \times \text{Ορθογώνιος})$$



# Trace Operator

Δίνει το άθροισμα όλων των διαγωνίων τιμών ενός πίνακα  $Tr(\mathbf{A}) = \sum_i A_{i,i}$

Frobenius νόρμα του A:  $\|\mathbf{A}\|_F = \sqrt{Tr(\mathbf{A}\mathbf{A}^T)}$

## ΙΔΙΟΤΗΤΕΣ

$$Tr(\mathbf{A}) = Tr(\mathbf{A}^T)$$

$Tr(\mathbf{A}\mathbf{B}\mathbf{C}) = Tr(\mathbf{C}\mathbf{A}\mathbf{B}) = Tr(\mathbf{B}\mathbf{C}\mathbf{A})$ , αν το επιτρέπουν οι διαστάσεις των πινάκων ή

$$Tr\left(\prod_{i=1}^n \mathbf{F}^{(i)}\right) = Tr\left(\mathbf{F}^{(n)} \prod_{i=1}^n \mathbf{F}^{(i)}\right)$$

Ένα βαθμωτό είναι το δικό του ίχνος:  $a = Tr(\mathbf{a})$

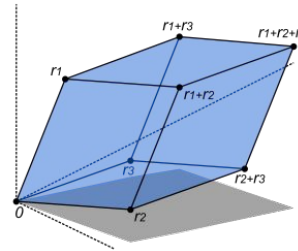
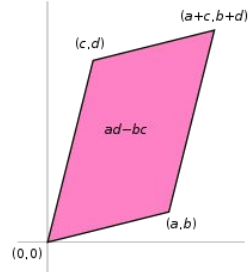
Παράδειγμα

$$Tr \begin{bmatrix} a_{00} & a_{01} & a_{02} \\ a_{10} & a_{11} & a_{12} \\ a_{20} & a_{21} & a_{22} \end{bmatrix} = a_{00} + a_{11} + a_{22}$$

# Ορίζουσα (Determinant)

$\det(A)$ : αντιστοιχεί πίνακα σε βαθμωτό

- Ισούνται με το γινόμενο όλων των ιδιοτιμών του πίνακα
- Γεωμετρική ερμηνεία
  - Η απόλυτη τιμή της ορίζουσας δίνει την κλίμακα με την οποία το εμβαδόν ή ο όγκος (ή μιας μεγαλύτερης διάστασης αναλογία) πολλαπλασιάζεται με τον σχετικό γραμμικό μετασχηματισμό,
  - το πρόσημό της δείχνει αν ο μετασχηματισμός διατηρεί τον προσανατολισμό.



Συνοπτικά, η ορίζουσα παρέχει γεωμετρική πληροφορία σχετικά με τη διάταξη και την κατεύθυνση των διανυσμάτων στο επίπεδο.



# Ορίζουσα Πίνακα: Γεωμετρική ερμηνεία

Έστω ότι έχουμε 2 διανύσματα στο επίπεδο  $v=[a \ c]$  και  $w=[b \ d]$ .

Ο πίνακας που σχηματίζεται από αυτά τα δύο διανύσματα είναι:  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

Η ορίζουσα  $\det(A)=ad-cb$ , είναι επίσης η περιοχή του παραλληλογράμμου που σχηματίζεται από τα δύο διανύσματα.

- Αν η ορίζουσα είναι θετική, τότε τα δύο διανύσματα δείχνουν προς ίδια κατεύθυνση και η περιοχή του παραλληλογράμμου είναι θετική.
- Αν είναι αρνητική, τα δύο διανύσματα δείχνουν προς αντίθετες κατευθύνσεις και η περιοχή είναι αρνητική.
- Αν η ορίζουσα είναι μηδέν, τότε τα δύο διανύσματα είναι γραμμικά εξαρτημένα, και η περιοχή του παραλληλογράμμου είναι μηδενική.

π.χ.  $A_{2 \times 2}$ ,  $\det(A)=-2$  : όταν εφαρμόζεται στην περιοχή ενός επιπέδου με πεπερασμένο εμβαδόν, θα μετασχηματιστεί σε μια περιοχή με το διπλάσιο εμβαδόν, ενώ αντιστρέφει τον προσανατολισμό της.

$\det(A)=0$  : ο χώρος συστέλλεται

$\det(A)=1$  : ο μετασχηματισμός διατηρεί τον όγκο

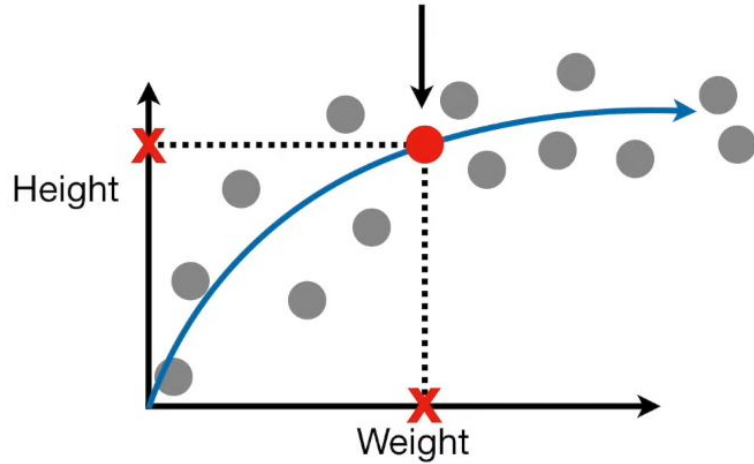


# Πιθανότητες και Μηχανική Μάθηση

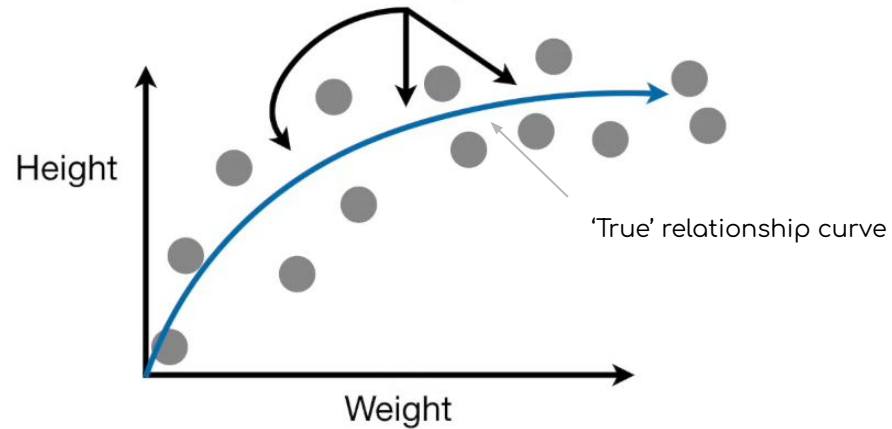
Pattern Recognition and Machine Learning, C.  
Bishop, Κεφάλαιο 1

# Bias & Variance

Ideally, we would know the exact mathematical formula that describes the relationship between weight and height...

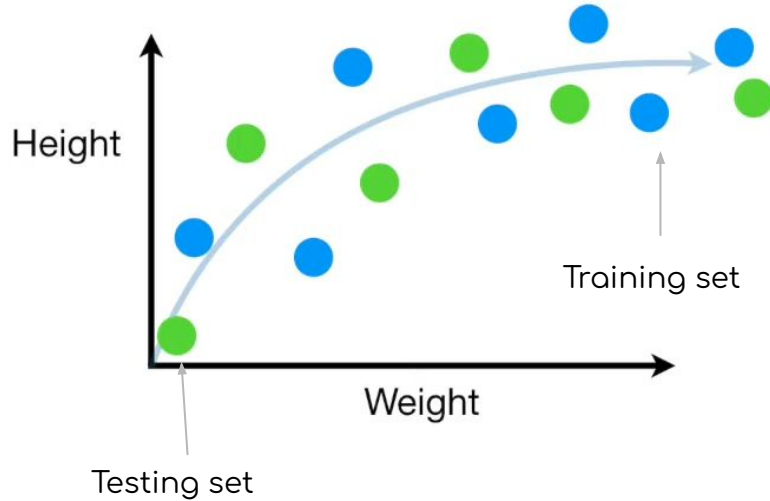


...but, in this case, we don't know the formula, so we're going to use two machine learning methods to approximate this relationship.

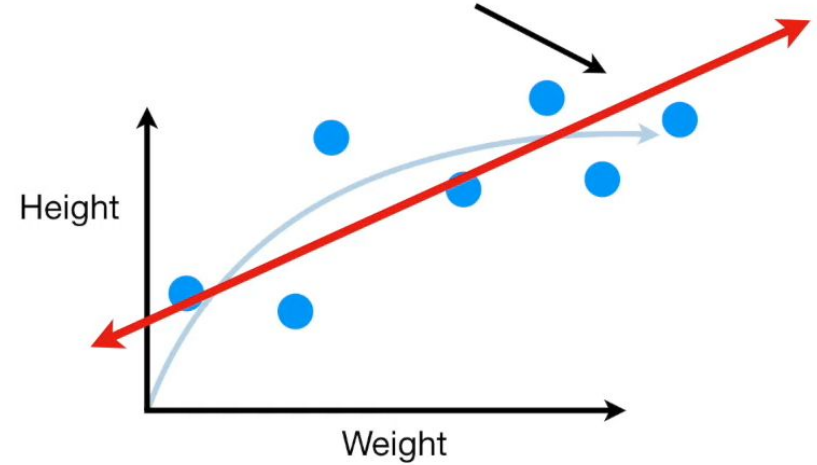


# Bias & Variance

The first thing we do is split the data into two sets, one for training the machine learning algorithms and one for testing them.



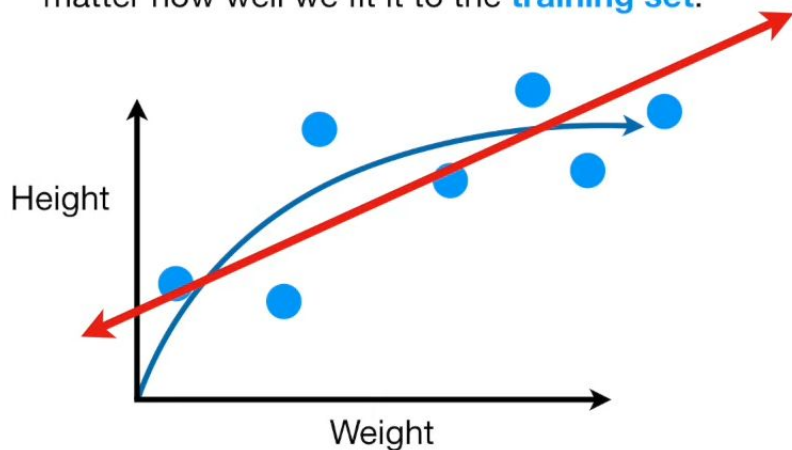
The first machine learning algorithm that we will use is Linear Regression (aka “Least Squares”).



**NOTE:** The **Straight Line** doesn't have the flexibility to accurately replicate the arc in the “true” relationship.

# Bias & Variance

Thus, the **Straight Line** will never capture the true relationship between weight and height, no matter how well we fit it to the **training set**.



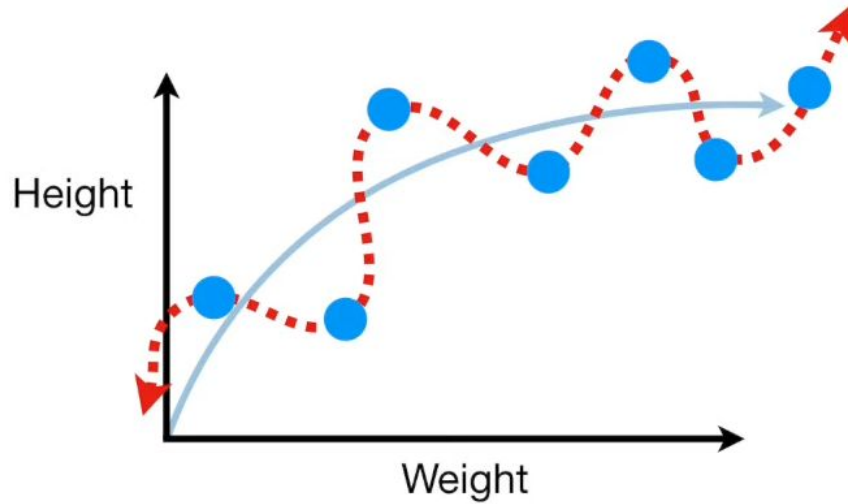
Because the **Straight Line** can't be curved like the “true” relationship, it has a relatively large amount of bias.

The inability for a machine learning method (like linear regression) to capture the true relationship is called **bias**.

# Bias & Variance

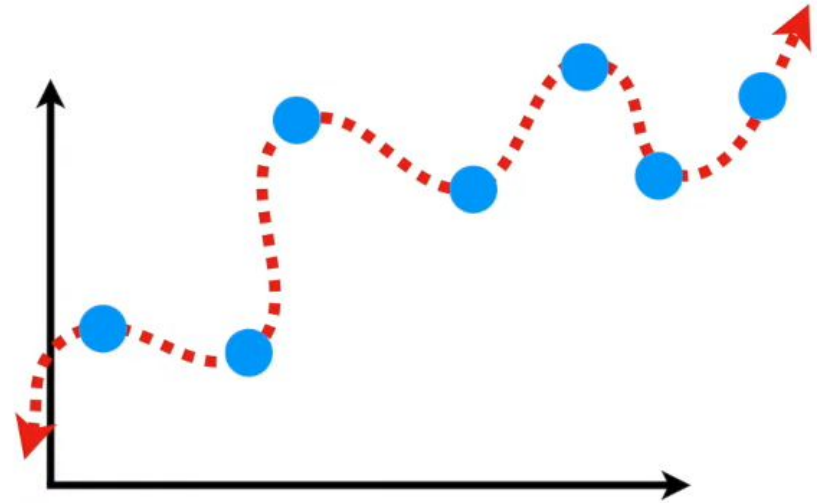
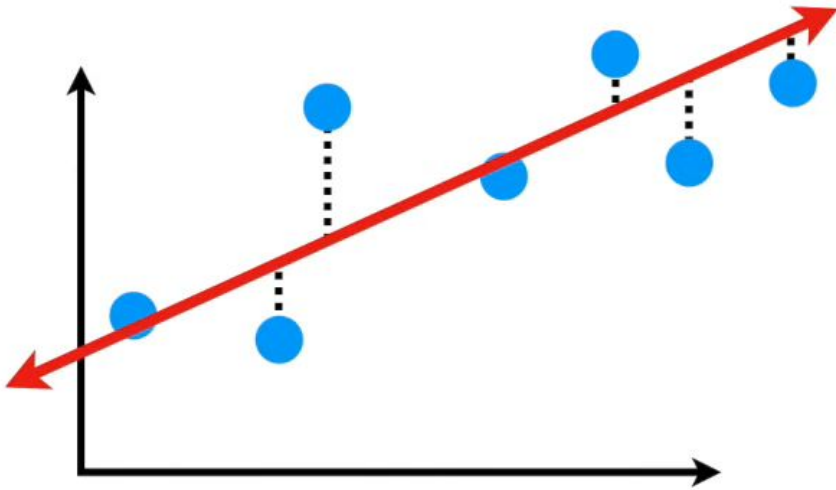


Because the **Squiggly Line** can handle the arc in the true relationship between weight and height, it has very little **bias**.



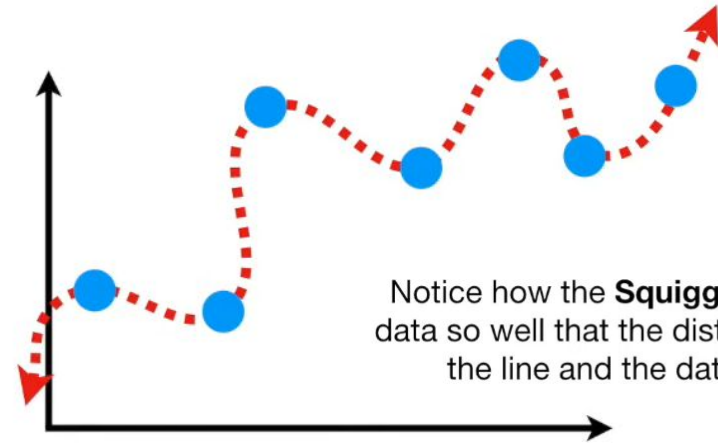
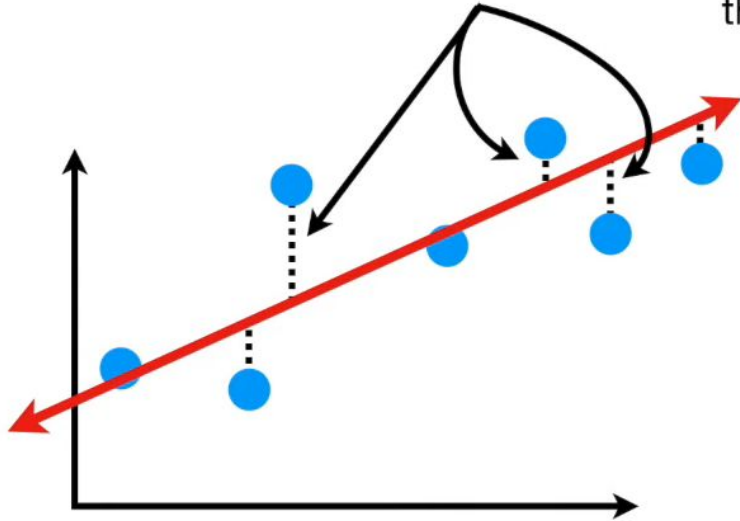
# Bias & Variance

We can compare how well the **Straight Line** and the **Squiggly Line** fit the **training set** by calculating their sums of squares.



# Bias & Variance

In other words, we measure the distances from the fit lines to the data, square them and add them up.

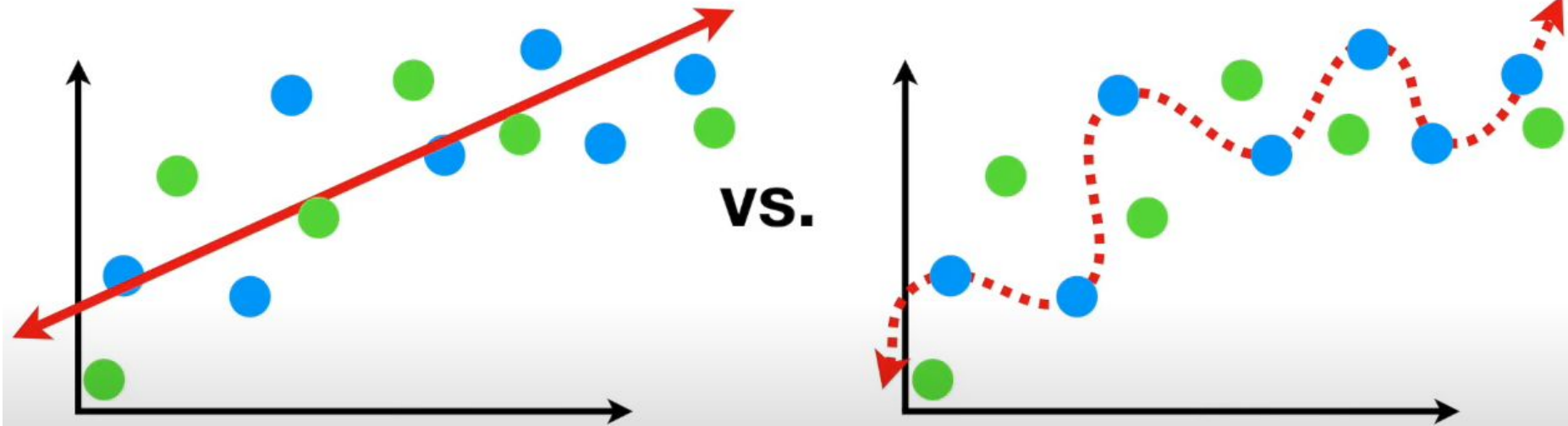


Notice how the **Squiggly Line** fits the data so well that the distances between the line and the data are all 0.

(psst. they are squared so that negative distances do not cancel out the positive distances)...

# Bias & Variance

We also have a **testing set**...

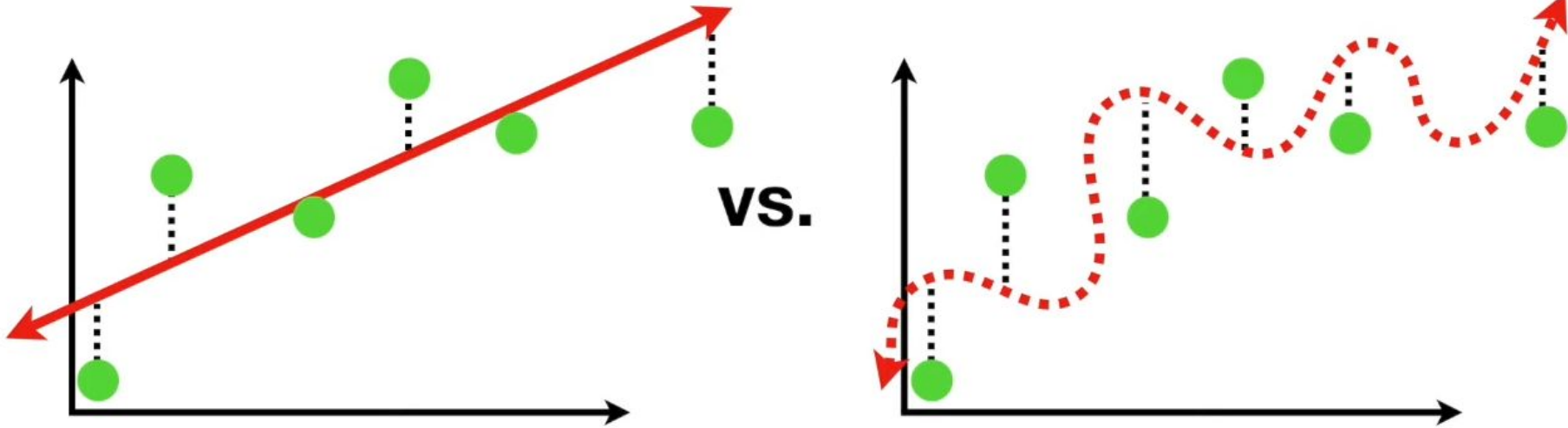




# Bias & Variance

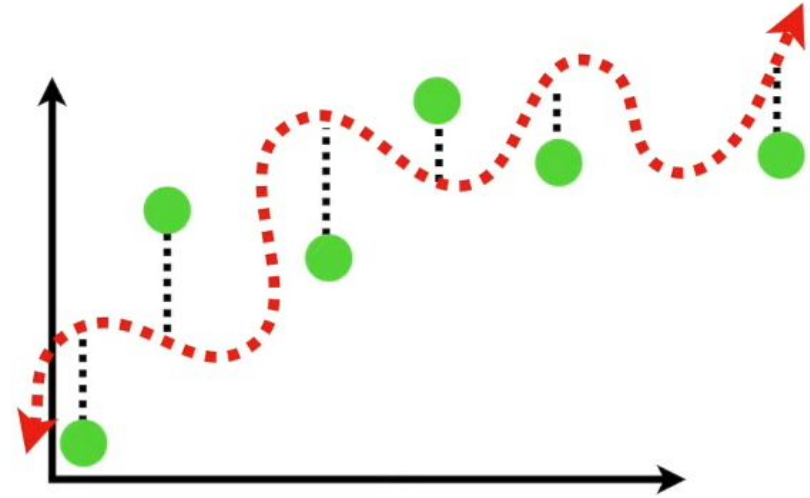
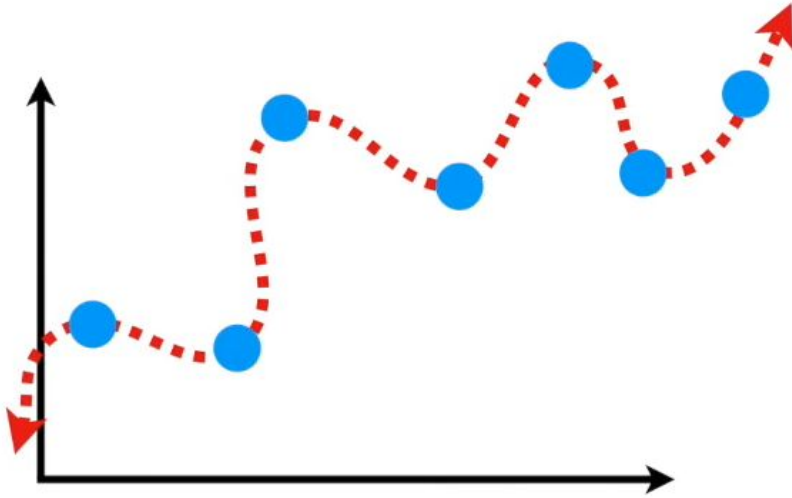


Now let's calculate the Sums of Squares for the **testing set.**



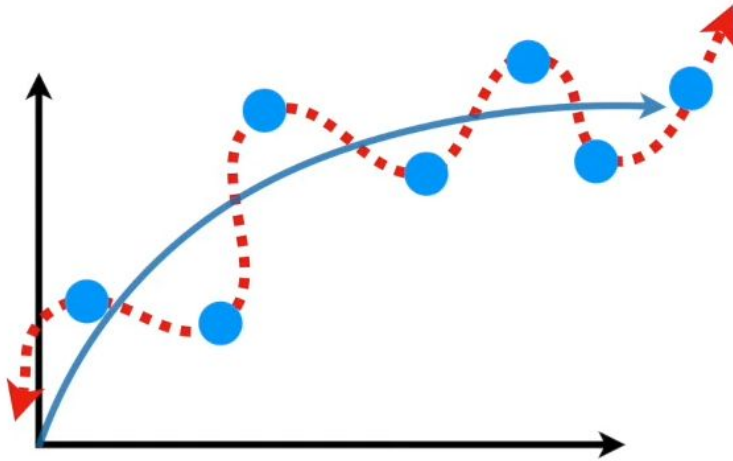
# Bias & Variance

In Machine Learning lingo, the difference in fits between data sets is called **Variance**.

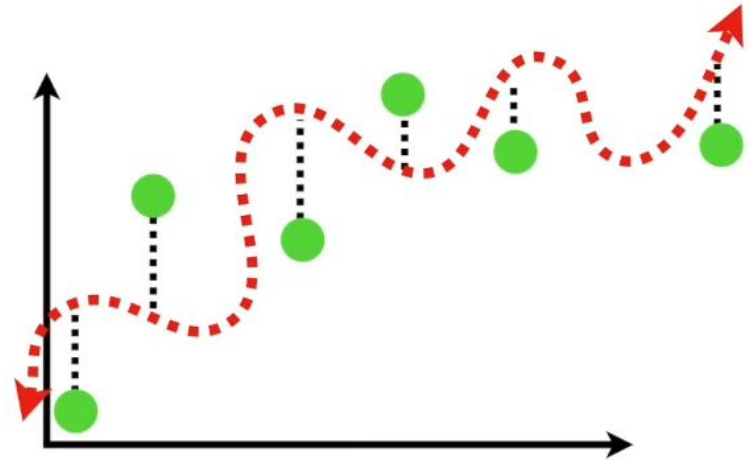


# Bias & Variance

The **Squiggly Line** has **low bias**, since it is flexible and can adapt to the curve in the relationship between weight and height...

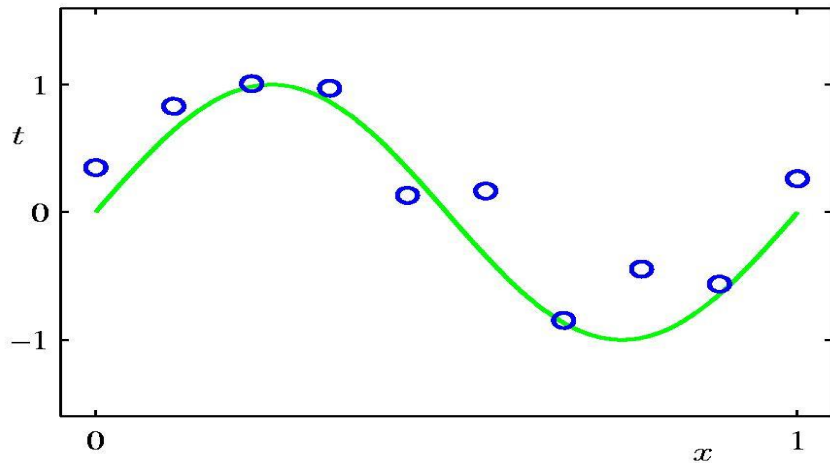


...but the **Squiggly Line** has **high variability**, because it results in vastly different Sums of Squares for different datasets.



In other words, it's hard to predict how well the **Squiggly Line** will perform with future data sets. It might do well sometimes, and other times it might do terribly.

# Polynomial Curve Fitting



- Δημιουργία συνόλου δεδομένων από  $\sin(2\pi x)$
- Προσθήκη τυχαίου θορύβου (Gaussian,  $\sigma=0.3$ )

Σύνολο τιμών εκπαίδευσης με  $N=10$  σημεία δεδομένων :

- παρατηρήσεις της μεταβλητής εισόδου που ισαπέχουν στο  $[0,1]$

$$\mathbf{x} = (x_1, \dots, x_N)^T.$$

- παρατηρήσεις αντίστοιχης μεταβλητής στόχου (προσθήκη τυχαίου Gaussian θορύβου  $\sigma=0.3$  στις αντίστοιχες τιμές του  $x$ ).

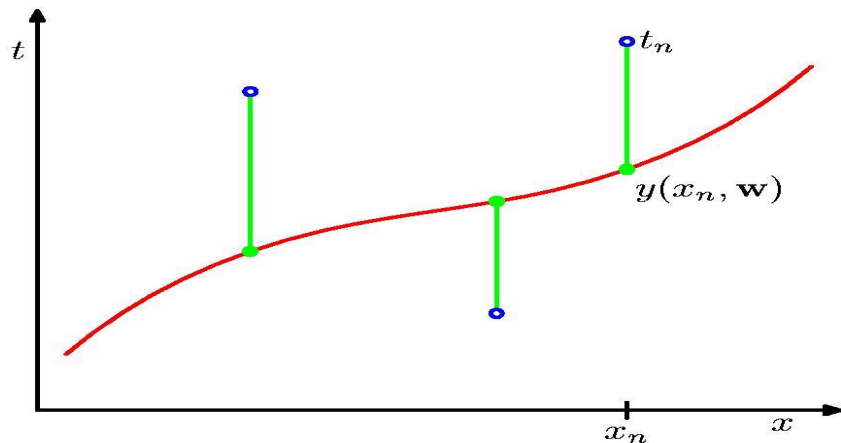
$$\mathbf{t} = (t_1, \dots, t_N)^T.$$

Γενίκευση από ένα πεπερασμένο σύνολο δεδομένων  $\rightarrow$  προσέγγιση με πολυωνυμική συνάρτηση:

$$y(x, \mathbf{w}) = w_0 + w_1x + w_2x^2 + \dots + w_Mx^M = \sum_{j=0}^M w_jx^j$$

Γραμμικό μοντέλο: μη γραμμική συνάρτηση ως προς  $x$  αλλά γραμμική ως προς  $w$ .

# Sum-of-Squares Error Function



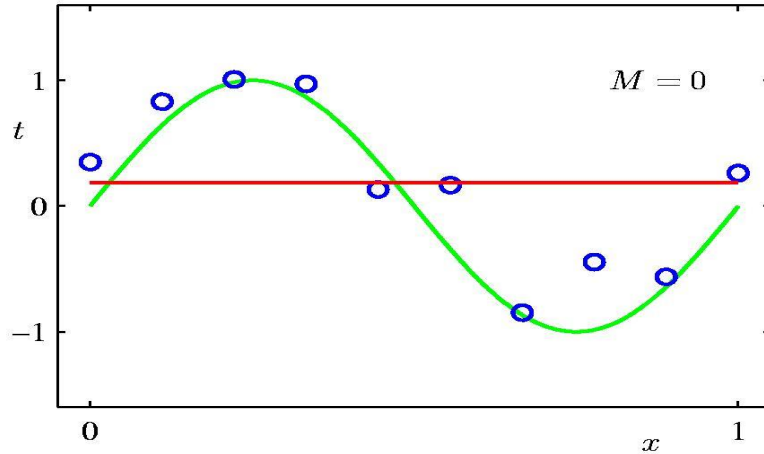
Οι τιμές του  $\mathbf{w}$  θα καθοριστούν με προσαρμογή του πολυωνύμου στα δεδομένα εκπαίδευσης

→ Ελαχιστοποίηση συνάρτησης σφάλματος.

$$E(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N \{y(x_n, \mathbf{w}) - t_n\}^2$$

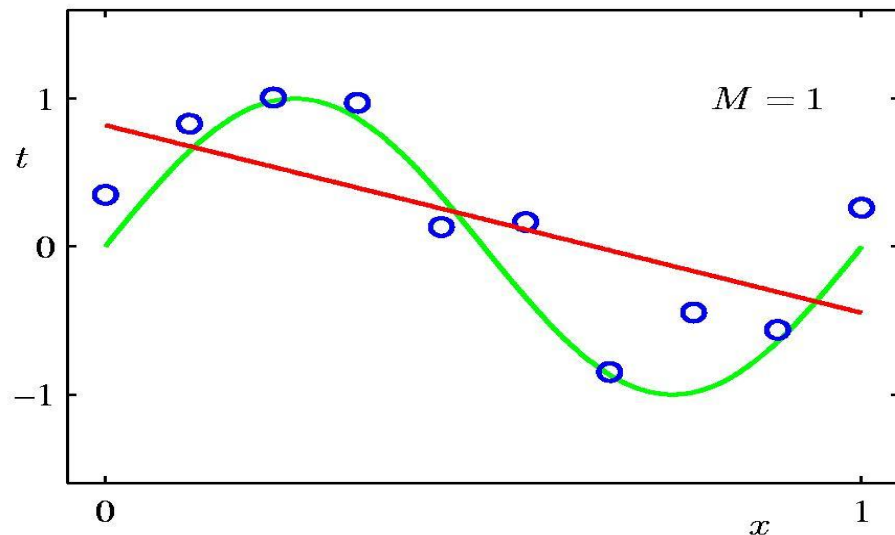
- τετραγωνική συνάρτηση των συντελεστών  $\mathbf{w}$ ,
- οι παράγωγοί της είναι γραμμικές ως προς  $\mathbf{w}$
- ελαχιστοποίηση της  $E(\mathbf{w})$  θα έχει μοναδική λύση, την  $\mathbf{w}^*$
- Το πολυώνυμο που προκύπτει δίνεται από την  $y(x, \mathbf{w}^*)$

# 0<sup>th</sup> Order Polynomial



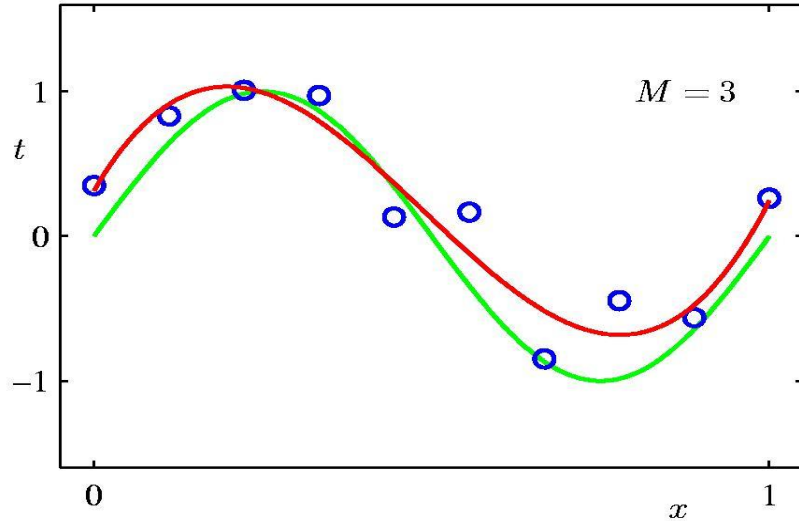
- Ανεπαρκείς προσαρμογές στα δεδομένα
- Κακή αναπαράσταση της  $\sin(2\pi x)$

# 1<sup>st</sup> Order Polynomial



- Ανεπαρκείς προσαρμογές στα δεδομένα
- Κακή αναπαράσταση της  $\sin(2\pi x)$

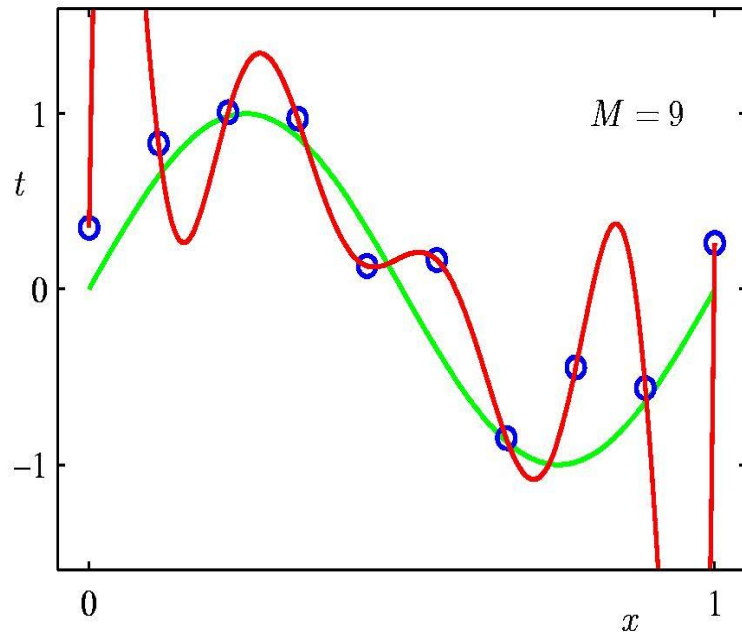
# 3<sup>rd</sup> Order Polynomial



- Καλύτερη αναπαράσταση της  $\sin(2\pi x)$



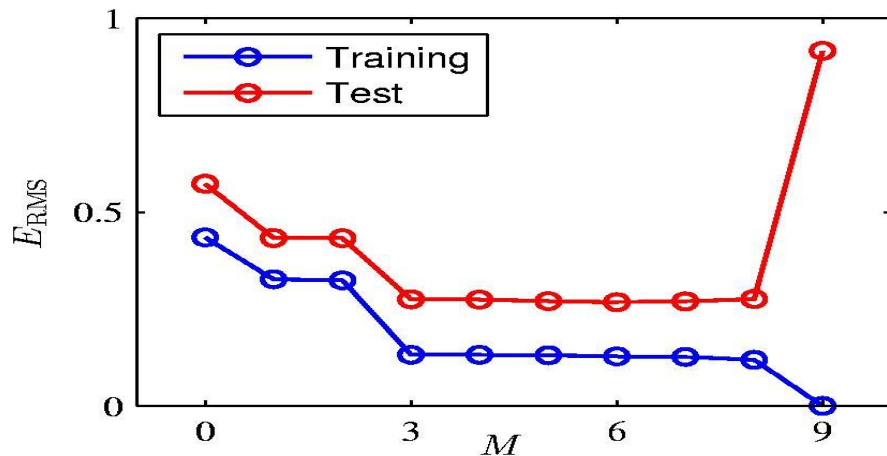
# 9<sup>th</sup> Order Polynomial



- Εξαιρετική προσαρμογή στα δεδομένα
- $E(w^*)=0$ , το πολυώνυμο περνά από κάθε σημείο.
- Πάλλεται έντονα η καμπύλη
- Πολύ χαμηλή ποιότητα αναπαράστασης της συνάρτησης  $\sin(2\pi x)$

→ ΥΠΕΡΕΚΠΑΙΔΕΥΣΗ  
(Over-fitting)

# Over-fitting



- Γράφημα RMS για το σύνολο εκπαίδευσης (training) και για ανεξάρτητο σύνολο δοκιμών για διάφορα  $M$
- Το RMS του δοκιμαστικού συνόλου αποτελεί ένα μέτρο για το πόσο καλά κάνουμε την πρόβλεψη των τιμών του  $t$  για τις νέες παρατηρήσεις δεδομένων του  $x$
- $M=9$ , training  $\rightarrow$  RMS=0  $\rightarrow$  πολυώνυμο με 10 βαθμούς ελευθερίας που αντιστοιχούν στα 10  $w(w_0, \dots, w_9)$   $\rightarrow$  ακριβή προσαρμογή στα 10 σημεία του συνόλου εκπαίδευσης

Root-Mean-Square (RMS) Error:  $E_{RMS} = \sqrt{2E(\mathbf{w}^*)/N}$



# Polynomial Coefficients

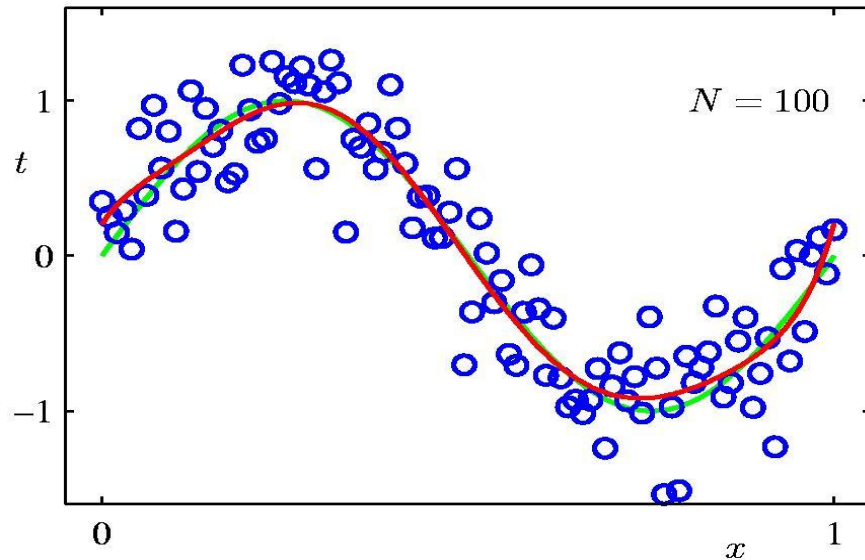
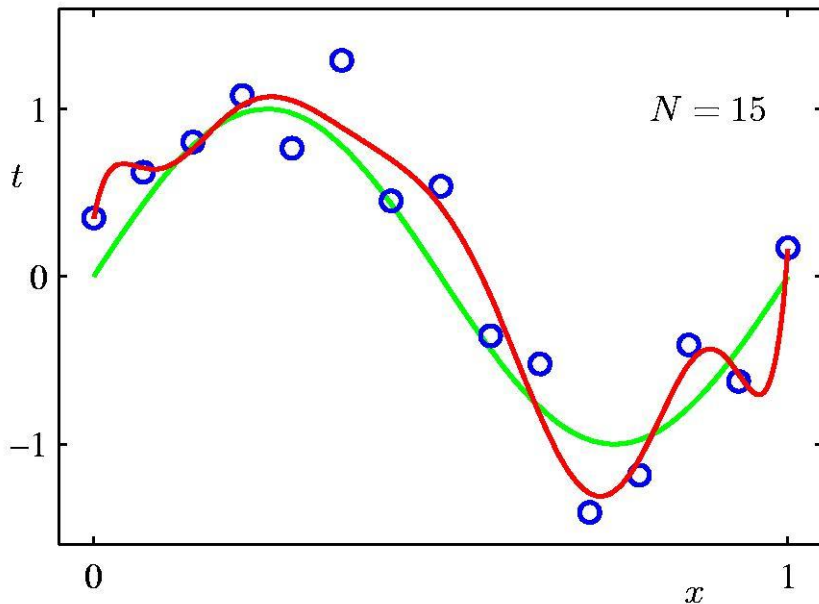
	$M = 0$	$M = 1$	$M = 3$	$M = 9$
$w_0^*$	0.19	0.82	0.31	0.35
$w_1^*$		-1.27	7.99	232.37
$w_2^*$			-25.43	-5321.83
$w_3^*$			17.37	48568.31
$w_4^*$				-231639.30
$w_5^*$				640042.26
$w_6^*$				-1061800.52
$w_7^*$				1042400.18
$w_8^*$				-557682.99
$w_9^*$				125201.43

↑  $M \rightarrow$  ↑  $w$

→  $w$  εξαιρετικά προσαρμοσμένα στα δεδομένα στο τυχαίο θόρυβο των τιμών στόχου  $t$

# Data Set Size:

## 9<sup>th</sup> Order Polynomial



$N \uparrow \rightarrow$  πιο σύνθετο το μοντέλο που μπορεί να προσαρμοστεί στα δεδομένα.

# Regularization

Penalize large coefficient values

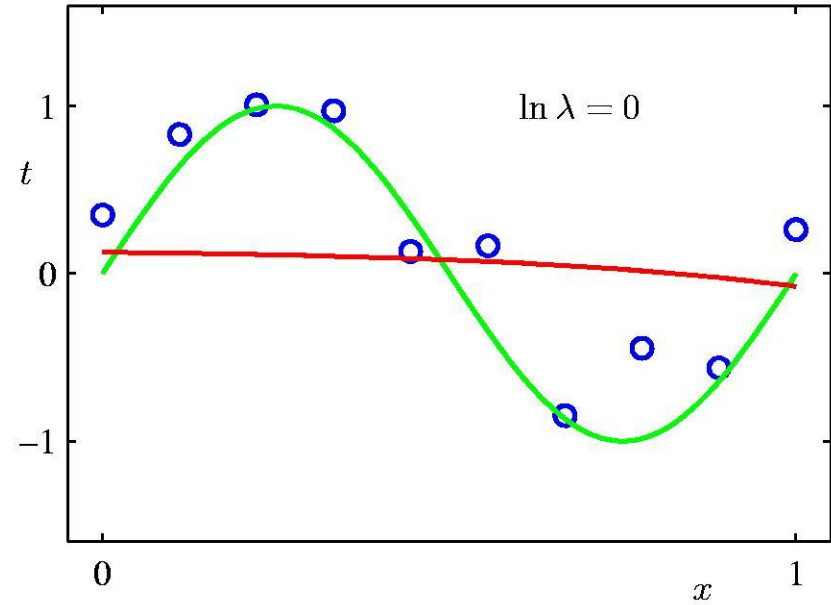
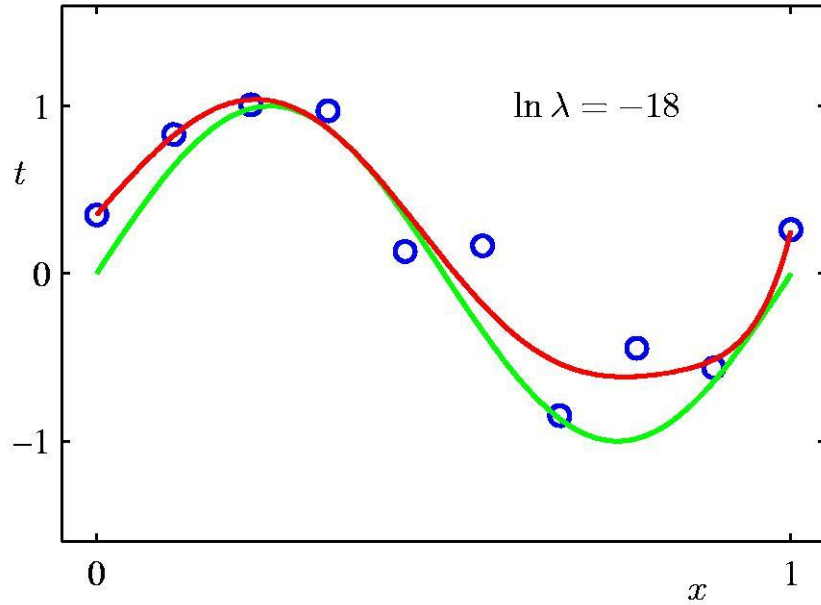
$$\tilde{E}(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N \{y(x_n, \mathbf{w}) - t_n\}^2 + \frac{\lambda}{2} \|\mathbf{w}\|^2$$

Όπου:

- $\|\mathbf{w}\|^2 = \mathbf{w}^T \mathbf{w} = w_0^2 + \dots + w_M^2$
- το  $\lambda$  καθορίζει τη σχετική σημασία του όρου κανονικοποίησης σε σύγκριση με όρο του τετραγωνικού σφάλματος

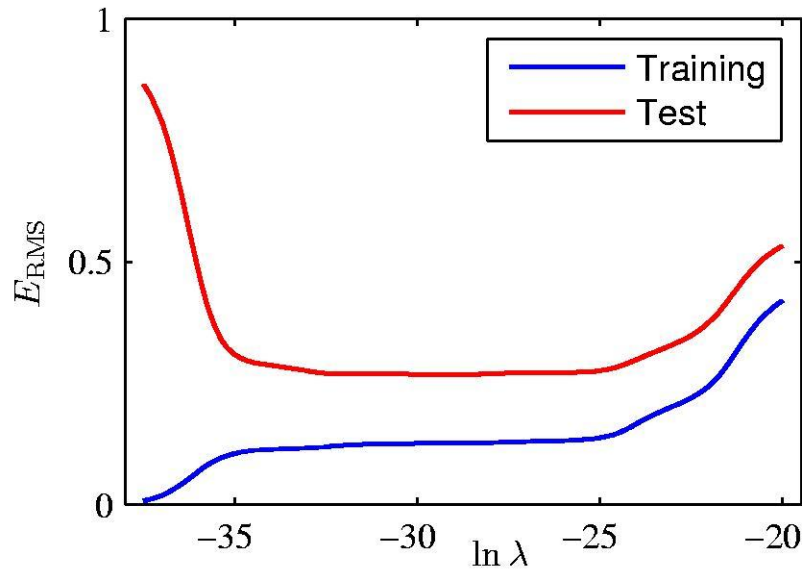


# Regularization - M=9



# Regularization: $E_{\text{RMS}}$ VS $\ln \lambda$

Το  $\lambda$  καθορίζει τώρα την πολυπλοκότητα του μοντέλου και συνεπώς καθορίζει το βαθμό της υπερεκπαίδευσης.

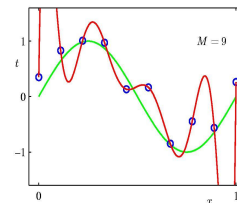


# Polynomial Coefficients

	$\ln \lambda = -\infty$	$\ln \lambda = -18$	$\ln \lambda = 0$
$w_0^*$	0.35	0.35	0.13
$w_1^*$	232.37	4.74	-0.05
$w_2^*$	-5321.83	-0.77	-0.06
$w_3^*$	48568.31	-31.97	-0.05
$w_4^*$	-231639.30	-3.89	-0.03
$w_5^*$	640042.26	55.28	-0.02
$w_6^*$	-1061800.52	41.32	-0.01
$w_7^*$	1042400.18	-45.95	-0.00
$w_8^*$	-557682.99	-91.53	0.00
$w_9^*$	125201.43	72.68	0.01

Μοντέλο χωρίς κανονικοποίηση:

$$\ln \lambda = -\infty$$

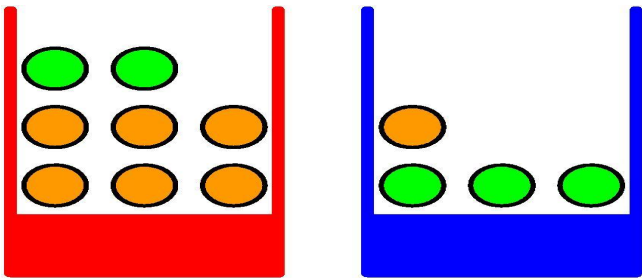


Όταν η τιμή του  $\lambda$  αυξάνει το τυπικό μέγεθος των συντελεστών μειώνεται



# Probability Theory - προσέγγιση

## Apples and Oranges



Βασική έννοια στην αναγνώριση προτύπων:

Αβεβαιότητα

→ θορύβος στις μετρήσεις

→ πεπερασμένο σύνολο δεδομένων

Θεωρία πιθανοτήτων

→ ποσοτικοποίηση της αβεβαιότητας

→ χειρισμός της αβεβαιότητας

Θεωρία πιθανοτήτων + θεωρία αποφάσεων →

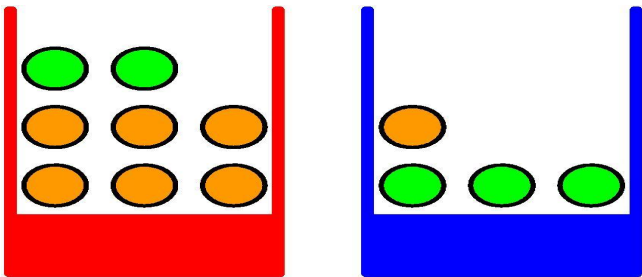
Βέλτιστες Προβλέψεις

### Πείραμα

1. Τυχαία επιλογή δοχείου
2. Τυχαία επιλογή φρούτου
3. Επανατοποθέτηση φρούτου

# Probability Theory - προσέγγιση

## Apples and Oranges



→ Επανάληψη του πειράματος πολλές φορές

→ Έστω 40% επιλέχθηκε το red Box και 60% το blue Box

> Τυχαία μεταβλητή:  $B = \{ r , b \}$

> Τυχαία μεταβλητή:  $F = \{ a , o \}$

$$p(B = r) = 4/10$$

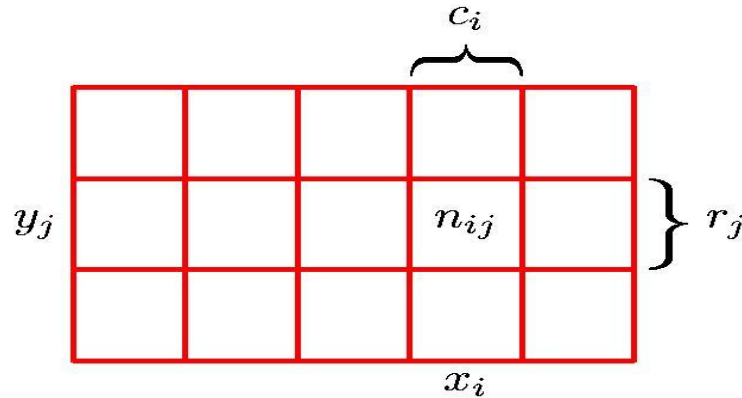
$$p(B = b) = 6/10$$

Πιθανά ερωτήματα:

- Ποια είναι η συνολική πιθανότητα να επιλέξουμε μήλο;
- Αν έχουμε επιλέξει ένα πορτοκάλι, ποια είναι η πιθανότητα να είναι από το μπλε δοχείο;

Κανόνες  
αθροίσματος  
και γινομένου  
πιθανοτήτων

# Probability Theory



$X = x_i, i=\{1..M\}$  (Boxes)

$Y = y_j, j=\{1..L\}$  (Fruits)

$N$ : συνολικές δοκιμές

$n_{ij}$ = πλήθος δοκιμών όπου  $X = x_i$  και  $Y = y_j$

$c_i$ = πλήθος δοκιμών όπου  $X = x_i$  ανεξάρτητα του  $Y$

$r_j$ = πλήθος δοκιμών όπου  $Y = y_j$  ανεξάρτητα του  $X$

Joint Probability (Από κοινού πιθανότητα)

$$p(X = x_i, Y = y_j) = \frac{n_{ij}}{N}$$

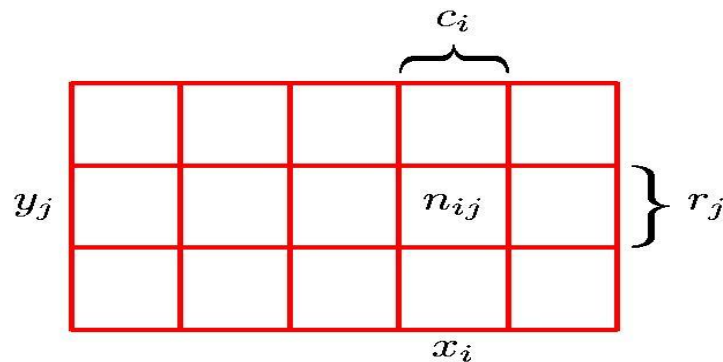
Marginal Probability

$$p(X = x_i) = \frac{c_i}{N}$$

Conditional Probability

$$p(Y = y_j | X = x_i) = \frac{n_{ij}}{c_i}$$

# Probability Theory



## Sum Rule

$$\begin{aligned}
 p(X = x_i) &= \frac{c_i}{N} = \frac{1}{N} \sum_{j=1}^L n_{ij} \\
 &= \sum_{j=1}^L p(X = x_i, Y = y_j)
 \end{aligned}$$

## Product Rule

$$\begin{aligned}
 p(X = x_i, Y = y_j) &= \frac{n_{ij}}{N} = \frac{n_{ij}}{c_i} \cdot \frac{c_i}{N} \\
 &= p(Y = y_j | X = x_i) p(X = x_i)
 \end{aligned}$$



# The Rules of Probability

Sum Rule

$$p(X) = \sum_Y p(X, Y)$$

Product Rule

$$p(X, Y) = p(Y|X)p(X)$$



# Bayes' Theorem

$$p(X) = \sum_Y p(X|Y)p(Y)$$

$$p(Y|X) = \frac{p(X|Y)p(Y)}{p(X)}$$

posterior  $\propto$  likelihood  $\times$   
prior

$X = x_i, i=\{1..M\}$  (Boxes)

$Y = y_j, j=\{1..L\}$  (Fruits)

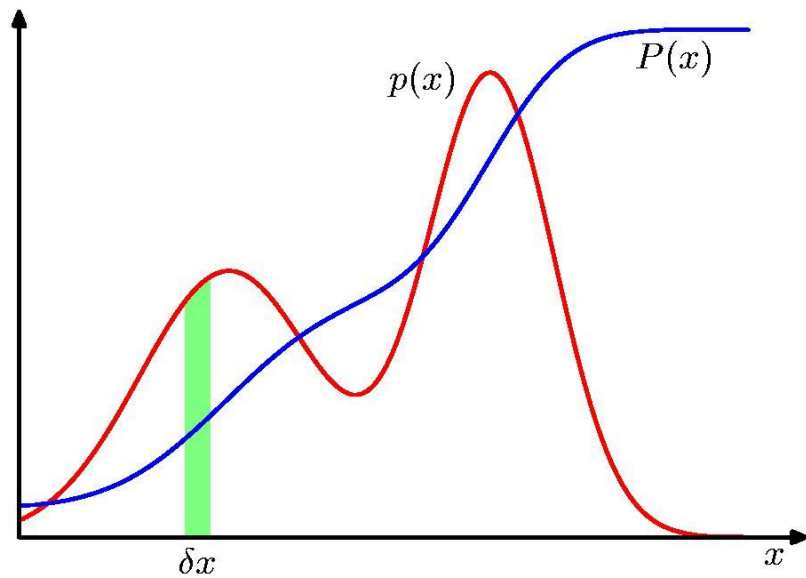
Ποιο κουτί επιλέχθηκε πριν μάθουμε ποιο φρούτο επιλέχθηκε;

→ prior: η πιθανότητα είναι γνωστή πριν διαλέξουμε το φρούτο.

Μόλις διαλέξουμε φρούτο, έστω το πορτοκάλι, με το θεώρημα του Bayes, μπορούμε να υπολογίζουμε την  $p(B|F)$

→ posterior : η πιθανότητα  $p(B|F)$  που λαμβάνουμε αφού έχουμε παρατηρήσει το F.

# Probability Densities



$$p(x) \geq 0$$

Για διακριτές μεταβλητές →

Πυκνότητα Πιθανότητας

$$p(x \in (a, b)) = \int_a^b p(x) dx$$

$$P(z) = \int_{-\infty}^z p(x) dx$$

$$\int_{-\infty}^{\infty} p(x) dx = 1$$

# Expectations

---

Προσδοκία: η μέση τιμή μιας συνάρτησης  $f(x)$  κάτω από κατανομή πιθανότητας  $p(x)$


Διακριτή  
Κατανομή

$$\mathbb{E}[f] = \sum_x p(x)f(x)$$

Συνεχείς  
Μεταβλητές

$$\mathbb{E}[f] = \int p(x)f(x) dx$$

Δεσμευμένη  
Προσδοκία

$$\mathbb{E}_x[f|y] = \sum_x p(x|y)f(x)$$


Conditional Expectation  
(discrete)

$$\mathbb{E}[f] \simeq \frac{1}{N} \sum_{n=1}^N f(x_n)$$

Approximate Expectation  
(discrete and continuous)

---



# Variances and Covariances

---

**Διακύμανση της  $f(x)$**  : μέτρο μεταβλητότητας που υπάρχει στην  $f(x)$  γύρω από τη μέση τιμή της  $E\{f(x)\}$

$$\text{var}[f] = \mathbb{E} \left[ (f(x) - \mathbb{E}[f(x)])^2 \right] = \mathbb{E}[f(x)^2] - \mathbb{E}[f(x)]^2$$

Συνδιακύμανση 2 τυχαίων μεταβλητών  $x, y$ : εκφράζει την έκταση που τα  $x$  και  $y$  μεταβάλλονται από

κοινού

$$\begin{aligned} \text{cov}[x, y] &= \mathbb{E}_{x,y} [\{x - \mathbb{E}[x]\} \{y - \mathbb{E}[y]\}] \\ &= \mathbb{E}_{x,y} [xy] - \mathbb{E}[x]\mathbb{E}[y] \end{aligned}$$

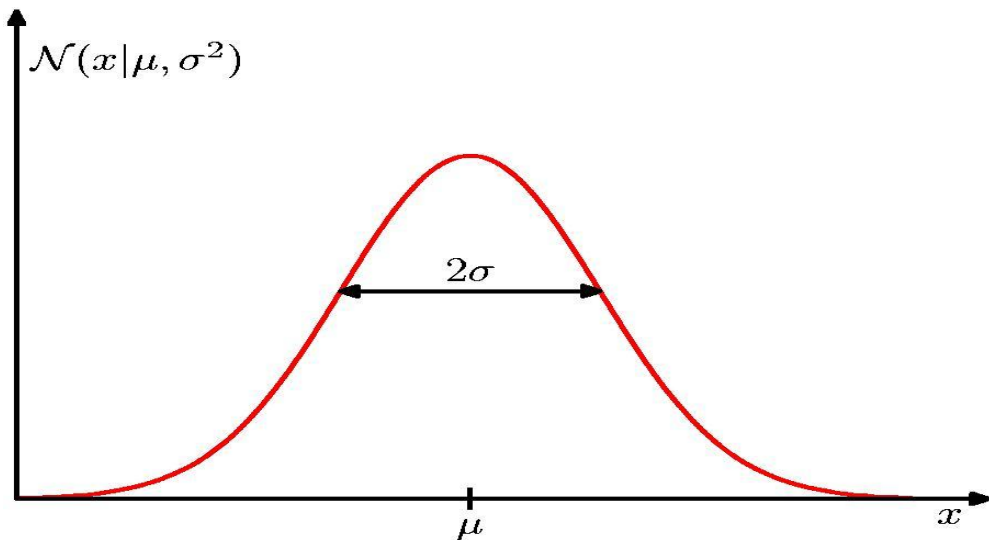
$$\begin{aligned} \text{cov}[\mathbf{x}, \mathbf{y}] &= \mathbb{E}_{\mathbf{x},\mathbf{y}} [\{\mathbf{x} - \mathbb{E}[\mathbf{x}]\} \{\mathbf{y}^T - \mathbb{E}[\mathbf{y}^T]\}] \\ &= \mathbb{E}_{\mathbf{x},\mathbf{y}} [\mathbf{x}\mathbf{y}^T] - \mathbb{E}[\mathbf{x}]\mathbb{E}[\mathbf{y}^T] \end{aligned}$$

# The Gaussian Distribution

---

$$\mathcal{N}(x|\mu, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left\{-\frac{1}{2\sigma^2}(x - \mu)^2\right\}$$

$x$ : μία μόνο πραγματική μεταβλητή  
 $\mu$ : μέσος,  $\sigma$ : τυπική απόκλιση



- Ικανοποιεί την ανισότητα

$$\mathcal{N}(x|\mu, \sigma^2) > 0$$

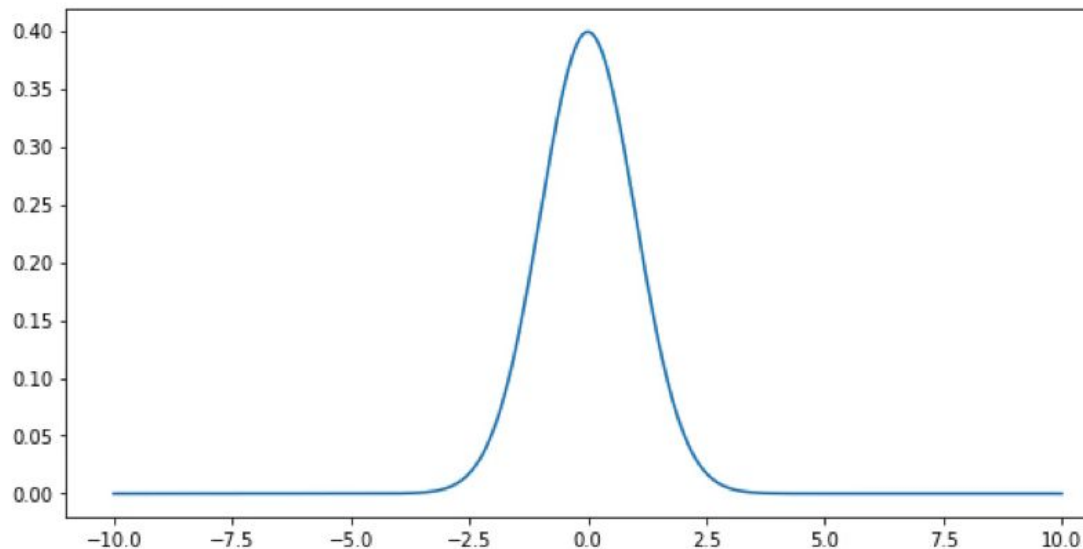
- Κανονικοποιημένη

$$\int_{-\infty}^{\infty} \mathcal{N}(x|\mu, \sigma^2) dx = 1$$

# Flashback - Normal Distribution

Density

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right)$$



# Gaussian Mean and Variance

---

$$\mathbb{E}[x] = \int_{-\infty}^{\infty} \mathcal{N}(x|\mu, \sigma^2) x \, dx = \mu$$

: Μέση τιμή του  $x$  ή μέσος

$$\mathbb{E}[x^2] = \int_{-\infty}^{\infty} \mathcal{N}(x|\mu, \sigma^2) x^2 \, dx = \mu^2 + \sigma^2$$

: Ροπή 2ης τάξης

$$\text{var}[x] = \mathbb{E}[x^2] - \mathbb{E}[x]^2 = \sigma^2$$

: Διακύμανση

---



# Εκτίμηση παραμέτρων

- **Mean**

$$\mu = \mathbf{E}[x] \text{ hence } \hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i$$

- **Variance**

$$\sigma^2 = \mathbf{E} [(x - \mu)^2] \text{ hence } \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{\mu})^2$$

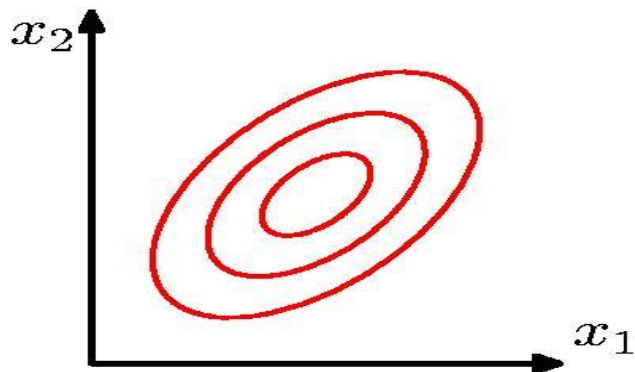
# The Multivariate Gaussian

---

Ορίζεται σε ένα  $D$ -διαστάσεων διάνυσμα  $\mathbf{x}$  συνεχών μεταβλητών

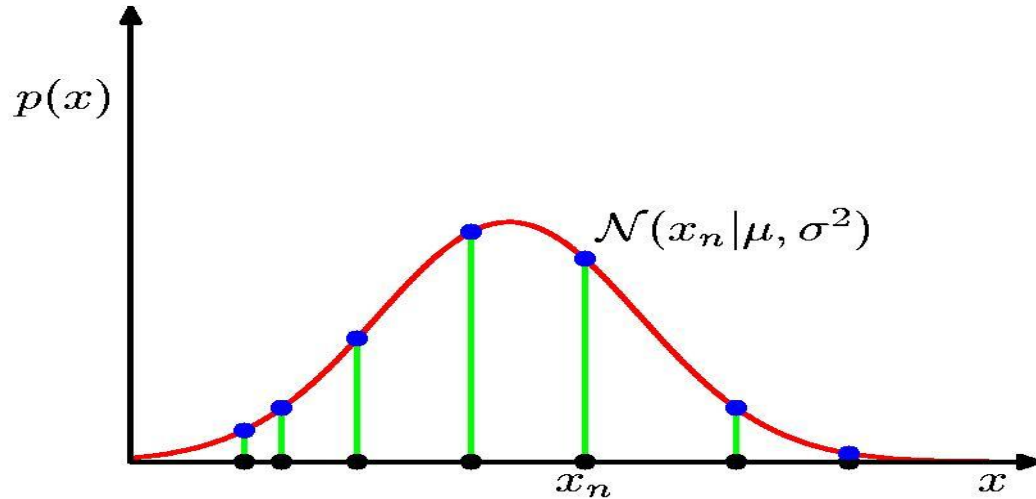
$$\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{D/2}} \frac{1}{|\boldsymbol{\Sigma}|^{1/2}} \exp \left\{ -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}) \right\}$$

$D \times D$  μητρώο  $\boldsymbol{\Sigma}$  συνδιακύμανσης  
 $|\boldsymbol{\Sigma}|$ : ορίζουσα του  $\boldsymbol{\Sigma}$



# Gaussian Parameter Estimation

---



Likelihood function

$$p(\mathbf{x} | \mu, \sigma^2) = \prod_{n=1}^N \mathcal{N}(x_n | \mu, \sigma^2)$$

$$\mathbf{x} = (x_1, \dots, x_N)^T$$

( $x$ : i..i.d. Independent and identically distributed random variables)

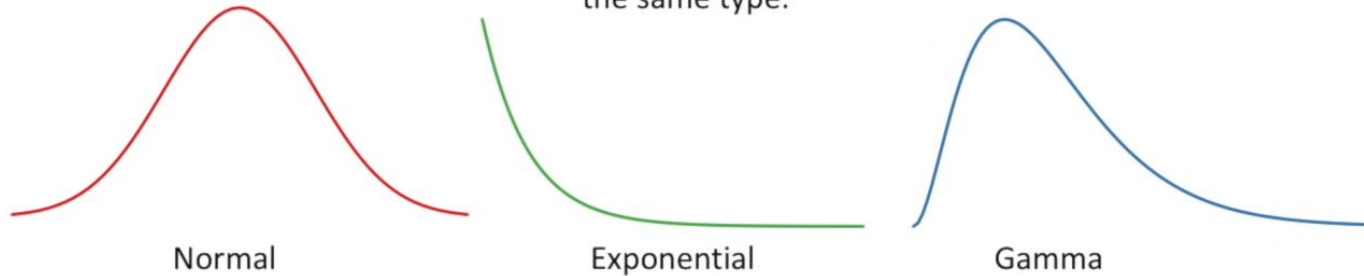
Άγνωστα:  $\mu, \sigma^2$

---

# Maximum (Log) Likelihood

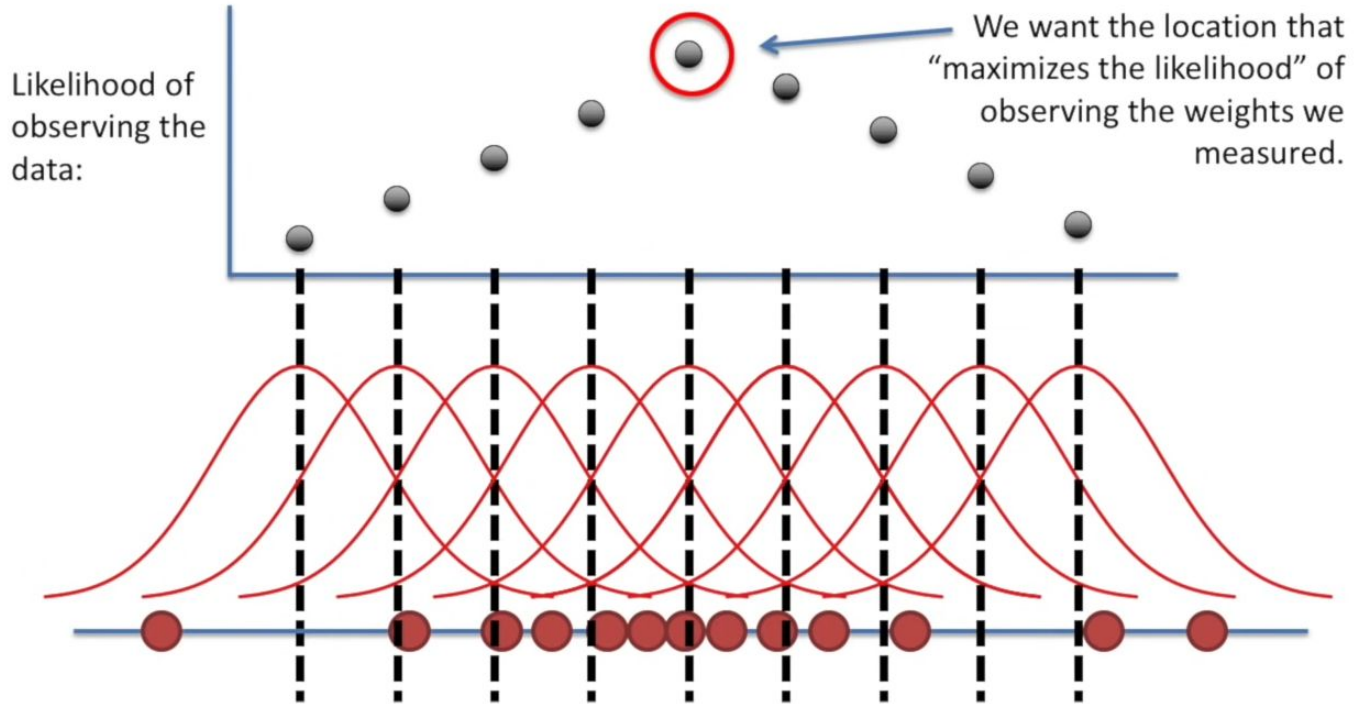
---

The reason you want to fit a distribution to your data is it can be easier to work with and it is also more general - it applies to every experiment of the same type.

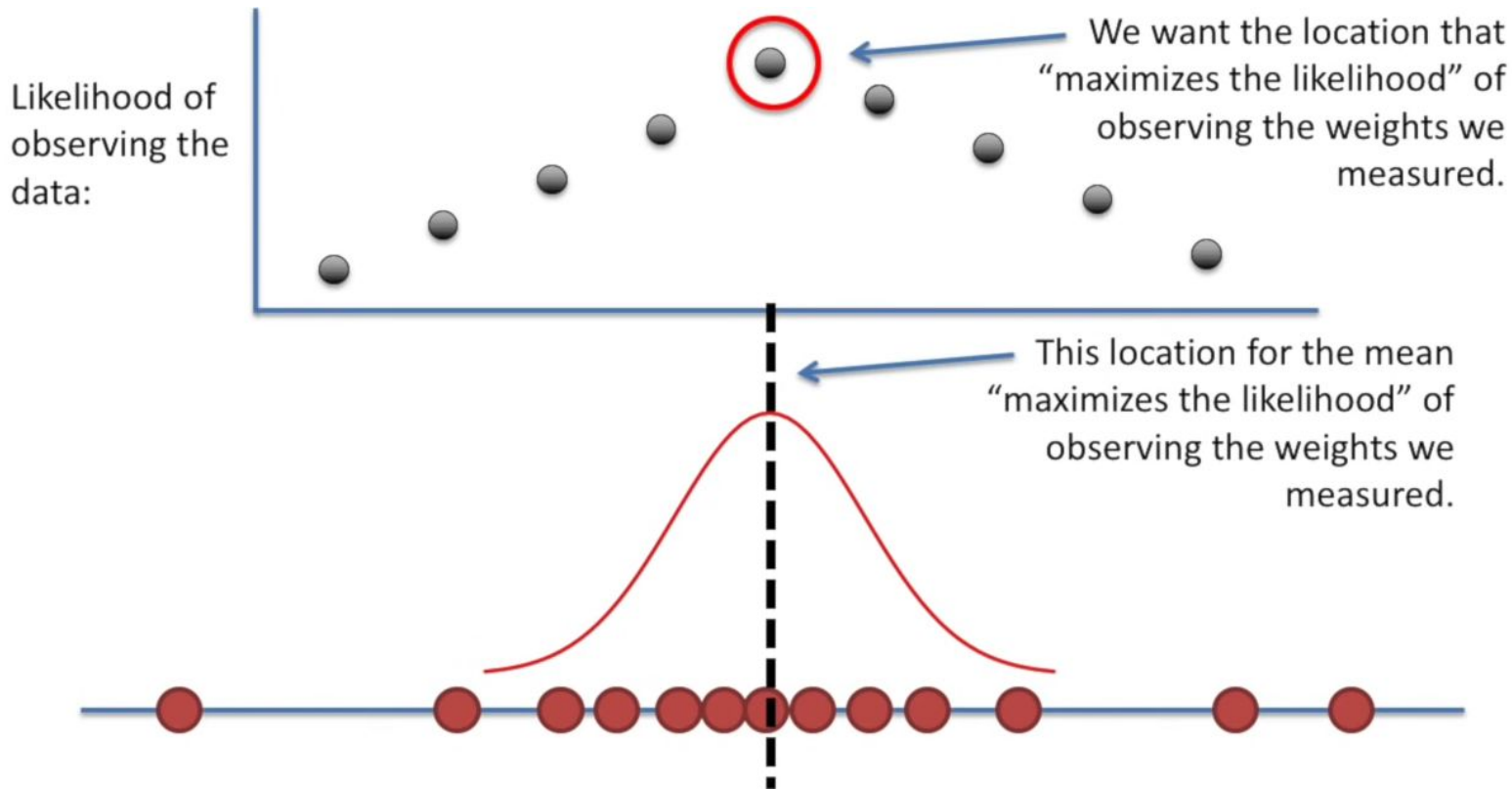




# Maximum (Log) Likelihood



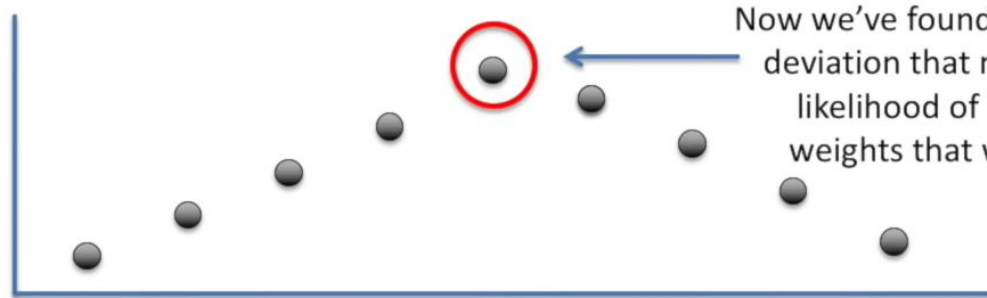
# Maximum (Log) Likelihood



# Maximum (Log) Likelihood

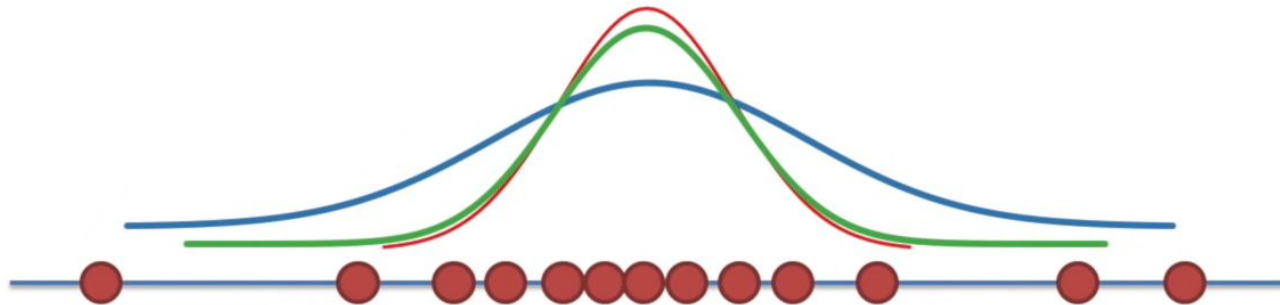
---

Likelihood of observing the data:



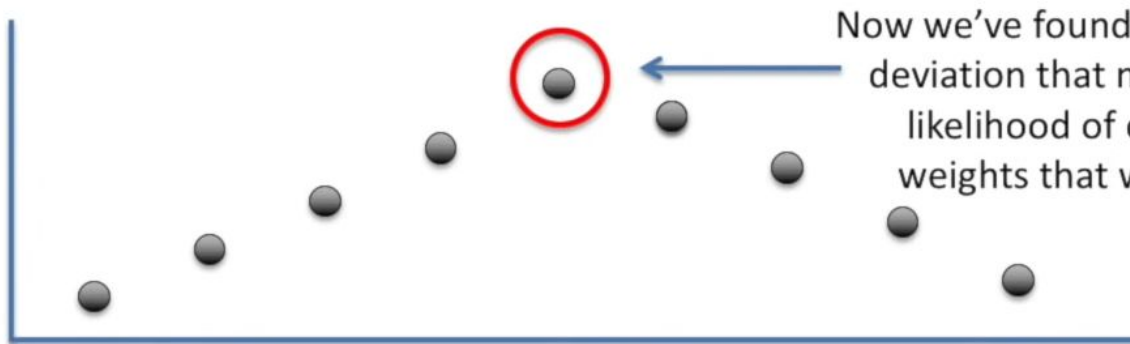
Now we've found the standard deviation that maximizes the likelihood of observing the weights that we measured.

Standard Deviation



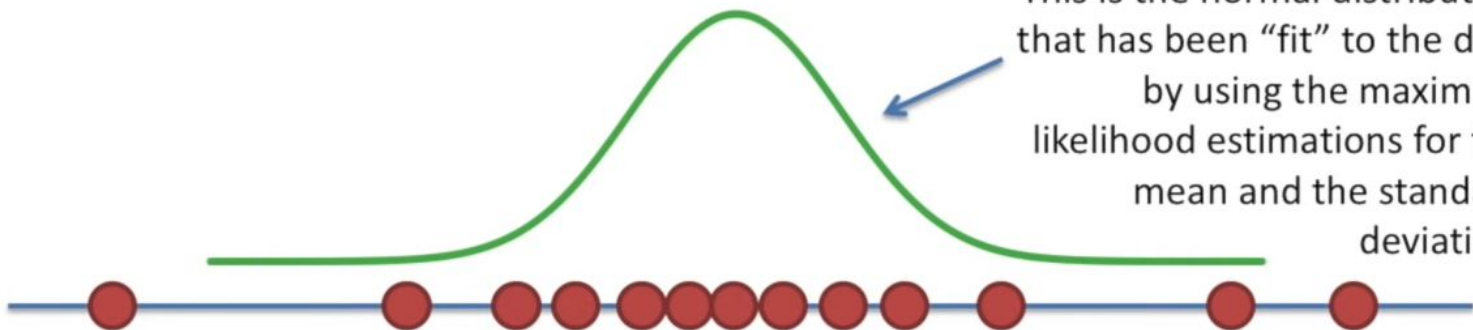
# Maximum (Log) Likelihood

Likelihood of observing the data:



Now we've found the standard deviation that maximizes the likelihood of observing the weights that we measured.

Standard Deviation



This is the normal distribution that has been "fit" to the data by using the maximum likelihood estimations for the mean and the standard deviation.



# Likelihood

- Observe some data  $X = \{x_1, \dots, x_n\}$
- Assume that the data is drawn from a Gaussian

$$p(X; \mu, \sigma^2) = \prod_{i=1}^n p(x_i; \mu, \sigma^2) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i - \mu)^2}{2\sigma^2}\right)$$

- **Fitting parameters is maximizing  $p(X; \mu, \sigma^2)$  wrt.  $\mu, \sigma^2$**   
(maximize likelihood that data was generated by model)
- **Practical simplification**

$$\underset{\mu, \sigma^2}{\text{maximize}} p(X; \mu, \sigma^2) \iff \underset{\mu, \sigma^2}{\text{minimize}} -\log p(X; \mu, \sigma^2)$$



# Maximum Likelihood

- Estimate parameters by finding ones that explain the data

$$\underset{\mu, \sigma^2}{\text{minimize}} -\log p(X; \mu, \sigma^2)$$

- **Decompose likelihood**

$$-\log p(X; \mu, \sigma^2) = \sum_{i=1}^n \left( \frac{1}{2} \log(2\pi\sigma^2) + \frac{1}{2\sigma^2} (x_i - \mu)^2 \right) = \frac{n}{2} \log(2\pi\sigma^2) + \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2$$

Minimize  $\mu = \frac{1}{n} \sum_{i=1}^n x_i$

$$p(X; \mu, \sigma^2) = \prod_{i=1}^n p(x_i; \mu, \sigma^2) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i - \mu)^2}{2\sigma^2}\right)$$



# Maximum Likelihood

- Estimating the variance

$$\frac{n}{2} \log(2\pi\sigma^2) + \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2$$

- Take derivatives with respect to it

$$\partial_{\sigma^2} [\cdot] = \frac{n}{2\sigma^2} - \frac{1}{2\sigma^4} \sum_{i=1}^n (x_i - \mu)^2 = 0$$

$$\implies \sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2$$

# Maximum (Log) Likelihood

---

$$p(\mathbf{x}|\mu, \sigma^2) = \prod_{n=1}^N \mathcal{N}(x_n|\mu, \sigma^2) \quad \longrightarrow \quad \text{Μέσο δείγμα}$$

$$\ln p(\mathbf{x}|\mu, \sigma^2) = -\frac{1}{2\sigma^2} \sum_{n=1}^N (x_n - \mu)^2 - \frac{N}{2} \ln \sigma^2 - \frac{N}{2} \ln(2\pi) \quad \xrightarrow{\text{μεγιστοποίηση}} \quad \mu_{\text{ML}} = \frac{1}{N} \sum_{n=1}^N x_n$$

Δειγματική διακύμανση σε σχέση με το μέσο δείγμα

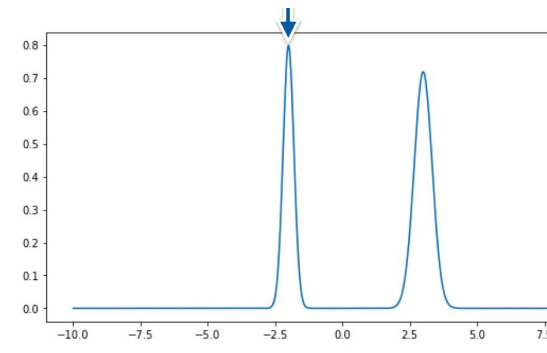
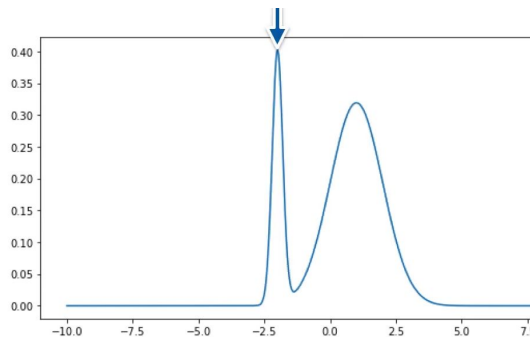
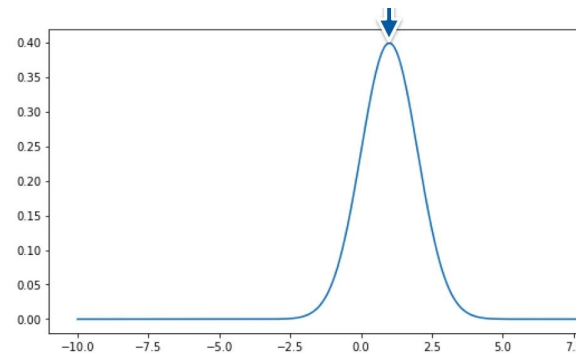
$$\xrightarrow{\text{μεγιστοποίηση}} \quad \sigma_{\text{ML}}^2 = \frac{1}{N} \sum_{n=1}^N (x_n - \mu_{\text{ML}})^2$$

(-) : υποεκτιμά συστηματικά τη διακύμανση της κατανομής - bias (μεροληψία)

---



# Maximum Likelihood Estimation



Κάτι μένει εκτός... χρησιμοποιώντας το MLE

- Data - ‘student didn’t do homework’
- Possible parameters
  - ‘dog ate homework’
  - ‘abducted by aliens’
  - ‘too lazy’
  - ‘sick grandmother’
- **All parameters explain the data.**



# Maximize a posteriori Estimation

- Posterior Probability

$$p(w | X) \propto p(X | w)p(w)$$

hence  $-\log p(w | X) = -\log p(X | w) - \log p(w) + c$

penalty

- Maximum a Posteriori Estimation

$$\underset{w}{\text{minimize}} -\log p(X; \mu, \sigma^2) - \log p(w)$$

- No homework example

$p(\text{'no homework'} | \text{explanation}) = 1$  (all explanations work)

lazy student	grandma sick	dog ate it	alien abduc
0.8	0.19	0.0099	0.0001



Τι σχέση έχει το regression με το MLE;



# Regression

- Recall optimization problem

$$\underset{w}{\text{minimize}} \sum_{i=1}^n (y_i - f(x_i, w))^2 + \text{penalty}(w)$$

$-\log p(w)$

Does the model work?

Additive Gaussian Noise

- Data generation model

$$y_i = f(x_i, w) + \epsilon_i \text{ where } \epsilon_i \sim \mathcal{N}(0, \sigma^2)$$

- Gaussian Prior  $p(w)$  hence  $-\log p(w) = \frac{1}{2\sigma^2} \|w\|^2 + \text{const}$



- Maximum a posteriori

$$\underset{w}{\text{minimize}} \quad -\log p(w | X, Y)$$

$$\Leftrightarrow \underset{w}{\text{minimize}} \quad \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - f(x_i, w))^2 + \frac{1}{2\bar{\sigma}^2} \|w\|^2 + \text{const}$$

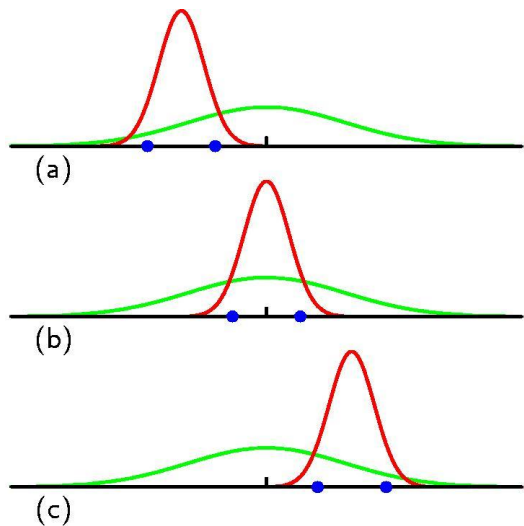
$$\Leftrightarrow \underset{w}{\text{minimize}} \quad \frac{1}{2n} \sum_{i=1}^n (y_i - f(x_i, w))^2 + \frac{\lambda}{2} \|w\|^2$$

# Properties of $\mu_{\text{ML}}$ and $\sigma_{\text{ML}}^2$

$$\mathbb{E}[\mu_{\text{ML}}] = \mu$$

$$\mathbb{E}[\sigma_{\text{ML}}^2] = \left(\frac{N-1}{N}\right) \sigma^2$$

$$\begin{aligned} \tilde{\sigma}^2 &= \frac{N}{N-1} \sigma_{\text{ML}}^2 \\ &= \frac{1}{N-1} \sum_{n=1}^N (x_n - \mu_{\text{ML}})^2 \end{aligned}$$



- Πράσινη καμπύλη: πραγματική Γκαουσιανή κατανομή
- Κόκκινη καμπύλη: Ληφθείσες Γκαουσιανές κατανομές
- Μπλε κουκίδες: Σημεία δεδομένων

→ Σωστός μέσος όρος

→ Η διακύμανση υποτιμημένη κατά  $(N-1)/N$ :  
μετρίεται σε σχέση με το μέσο δείγμα και όχι με  
τον πραγματικό μέσο  $\tilde{\sigma}^2$

→ η παράμετρο της μεταβλητότητας  
είναι αμερόληπτη (unbiased)

$N \rightarrow \infty$  η λύση της Maximum Likelihood  
= πραγματική διακύμανσης της κατανομής που  
παράγει τα δεδομένα



Ανάλυση  
Κύριων  
Συνιστωσών

Principal  
Component  
Analysis



# Ανάλυση Κύριων Συνιστωσών: παράδειγμα

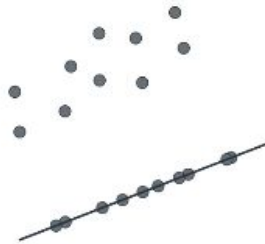
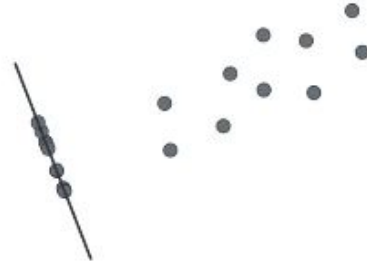
Taking a picture



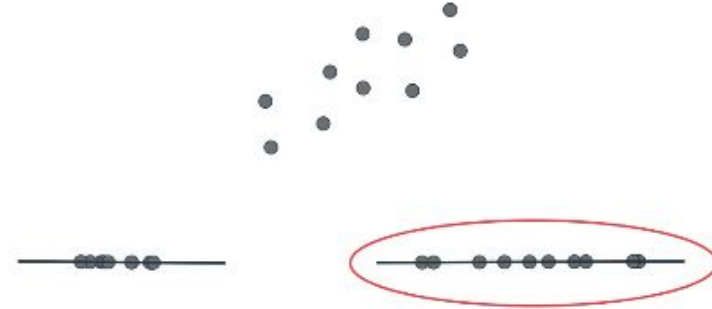
Taking a picture



Dimensionality Reduction



Dimensionality Reduction



Ποια είναι η ιδανική γραμμή για προβολή των δεδομένων (όπου αυτά θα είναι όσο πιο διακριτά γίνεται);

# Housing Data



## Dataset' features

Size  
Number of rooms  
Number of bathrooms  
Schools around  
Crime rate

Διάνυσμα 5 χαρακτηριστικών  $\xrightarrow{\text{Μείωση διαστάσεων}}$  Διάνυσμα 2 χαρακτηριστικών

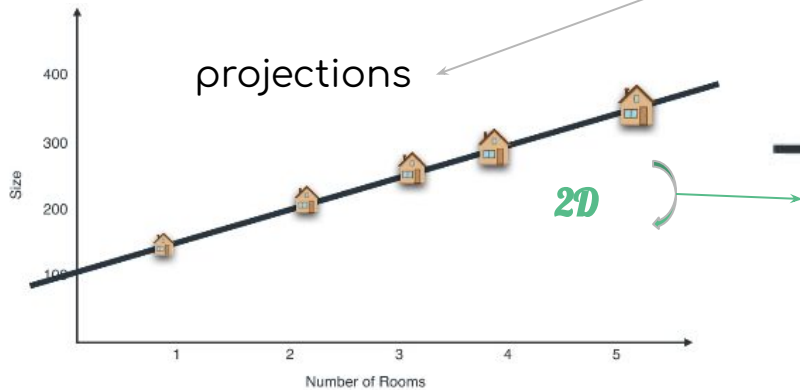
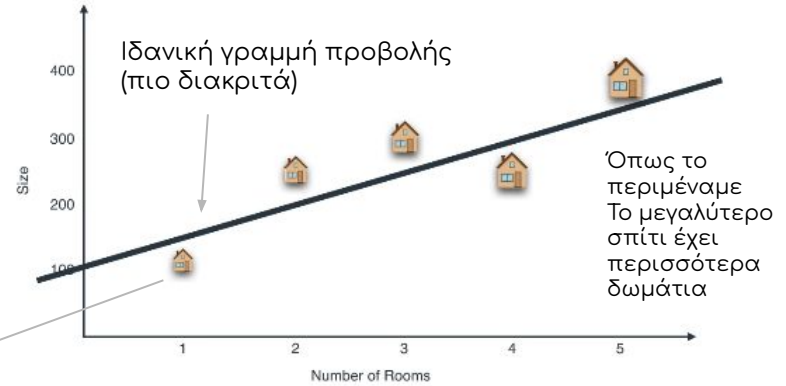
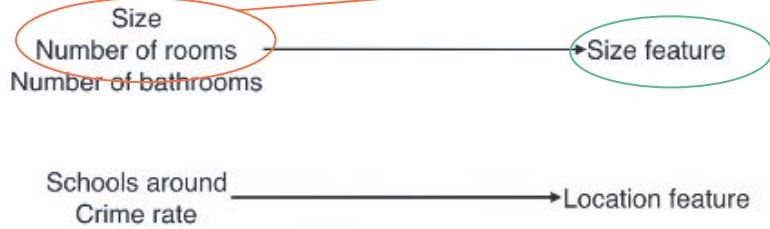
Size  
Number of rooms  $\longrightarrow$  Size feature  
Number of bathrooms

Schools around  $\longrightarrow$  Location feature  
Crime rate

# Ανάλυση Κύριων Συνιστωσών: παράδειγμα

Θα δούμε πώς πάμε από το χώρο των 2 χαρακτηριστικών (2D) στο 1D

## Housing Data



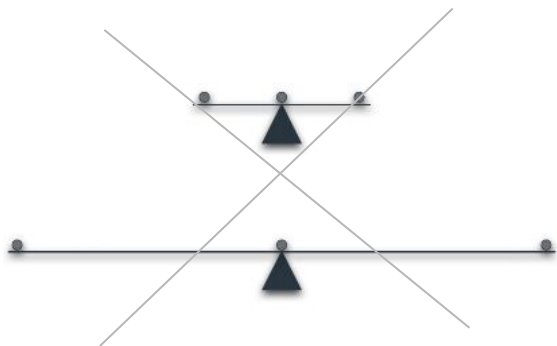
# Ανάλυση Κύριων Συνιστωσών: παράδειγμα

σημείο  
ισορροπίας

## Mean



$$\text{Mean} = \frac{1+2+6}{3} = 3$$

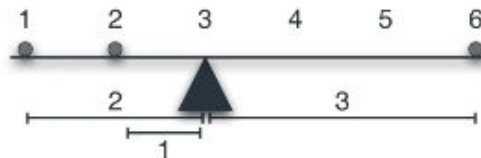


Ίδιο  
mean  
value

## Variance

$$\text{Variance} = \frac{1^2+0^2+1^2}{3} = 2/3$$

$$\text{Variance} = \frac{5^2+0^2+5^2}{3} = 50/3$$



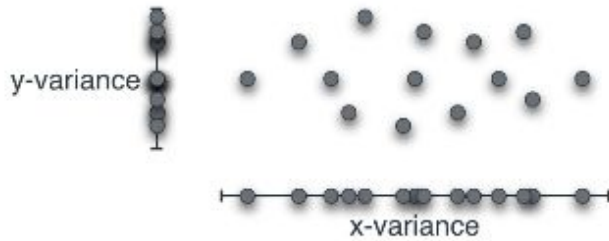
$$\text{Variance} = \frac{2^2+1^2+3^2}{3} = 14/3$$

Παίρνω την  
απόσταση κάθε  
σημείου από το  
κέντρο.

# Ανάλυση Κύριων Συνιστωσών: παράδειγμα

Εξετάζουμε τη διακύμανση στο χώρο  
→ εξετάζουμε και τις δύο διαστάσεις

## Variance?



## Variance?

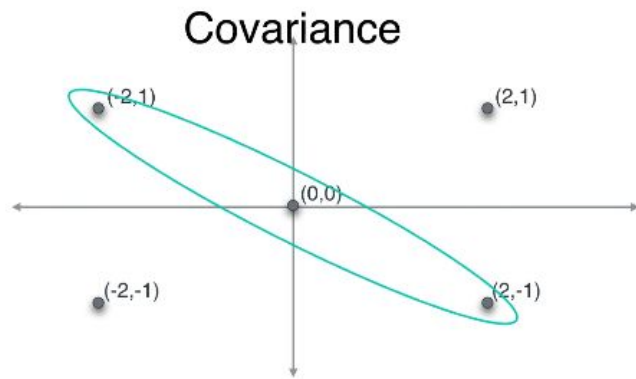
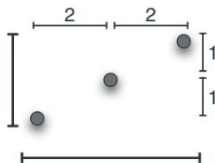
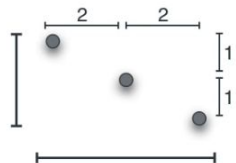


$$\text{x-variance} = \frac{2^2 + 0^2 + 2^2}{3} = 8/3$$

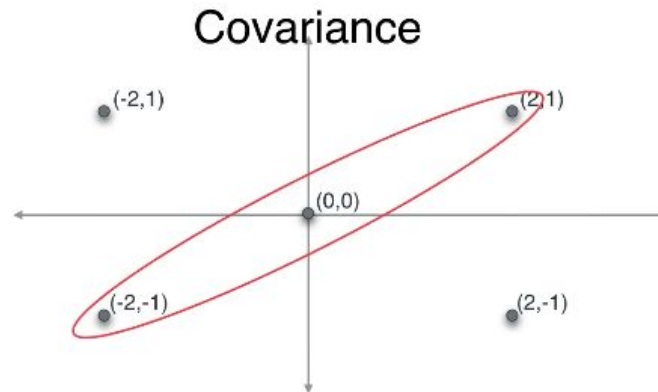
$$\text{y-variance} = \frac{1^2 + 0^2 + 1^2}{3} = 2/3$$

- Έχω δύο διαφορετικά σύνολα
  - έχουν την ίδια διακύμανση στις προβολές  $x$  και  $y$

# Ανάλυση Κύριων Συνιστωσών: παράδειγμα

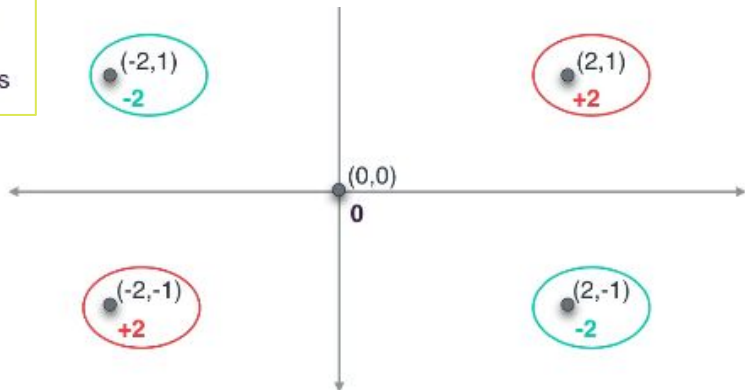


Διαφορά;

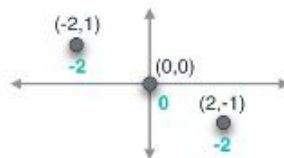


# Ανάλυση Κύριων Συνιστωσών: παράδειγμα

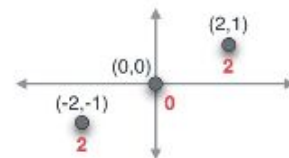
Product  
of  
ordinates



## Covariance



$$\text{covariance} = \frac{(-2) + 0 + (-2)}{3} = -4/3$$

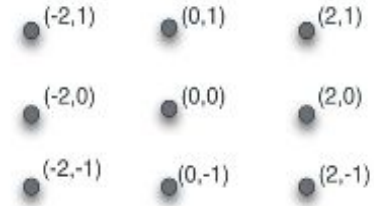
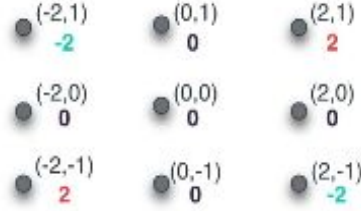
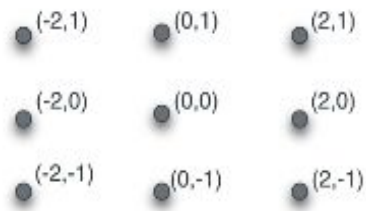


$$\text{covariance} = \frac{2 + 0 + 2}{3} = 4/3$$

# Ανάλυση Κύριων Συνιστωσών: παράδειγμα

Ας υπολογίσουμε το covariance των σημείων

## Covariance



$$\text{covariance} = \frac{-2 + 0 + 2 + 0 + 0 + 0 + 2 + 0 + -2}{9} = 0$$

Είναι 0 και αυτό φαίνεται λογικό αφού δεν φαίνεται να υπάρχει κάποια θετική ή αρνητική συσχέτιση



# Ανάλυση Κύριων Συνιστωσών: παράδειγμα



## Covariance



negative  
covariance



covariance zero  
(or very small)

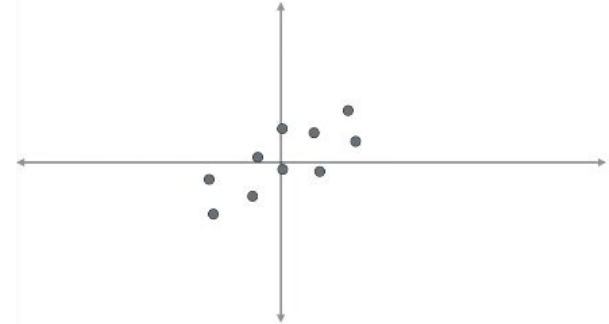
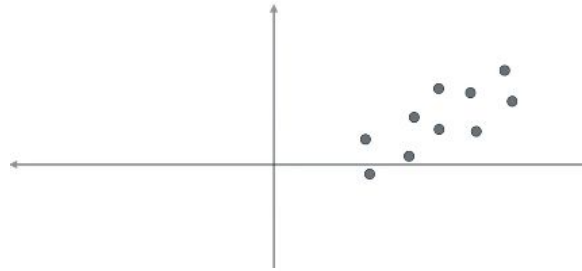


positive  
covariance

# Ανάλυση Κύριων Συνιστωσών: παράδειγμα

Πώς θα βρούμε την τέλεια προβολή;

Θα ορίσουμε σύστημα αξόνων -> θα βρούμε το σημείο όπου τα δεδομένα ισορροπούν

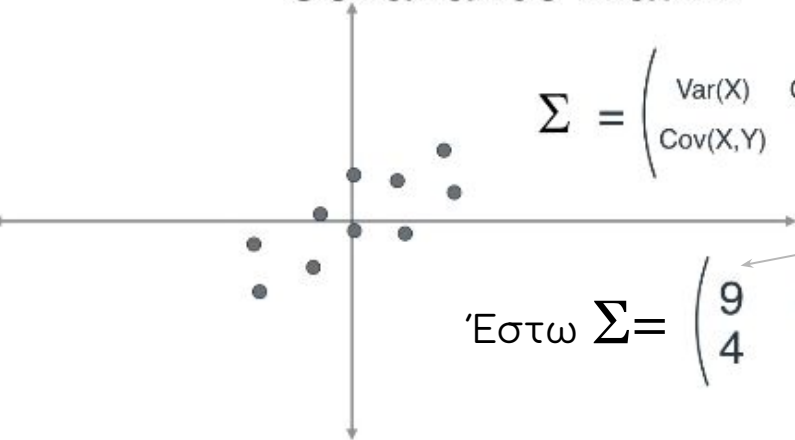


Covariance matrix

$$\Sigma = \begin{pmatrix} \text{Var}(X) & \text{Cov}(X,Y) \\ \text{Cov}(X,Y) & \text{Var}(Y) \end{pmatrix}$$

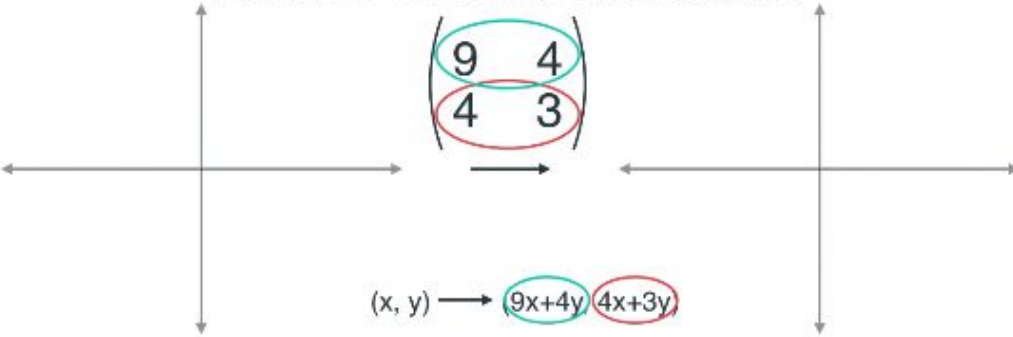
Έστω  $\Sigma = \begin{pmatrix} 9 & 4 \\ 4 & 3 \end{pmatrix}$

Μεγάλη διακύμανση στο x άξονα σε αντίθεση με τον y

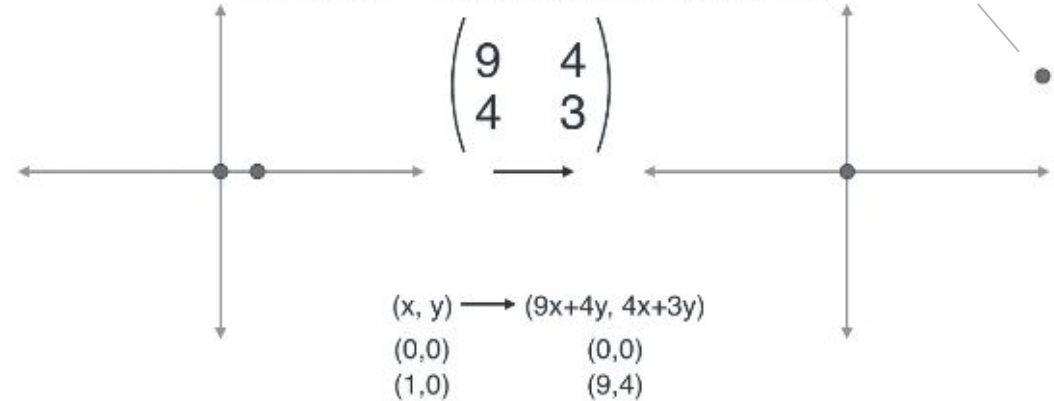


# Ανάλυση Κύριων Συνιστωσών: παράδειγμα

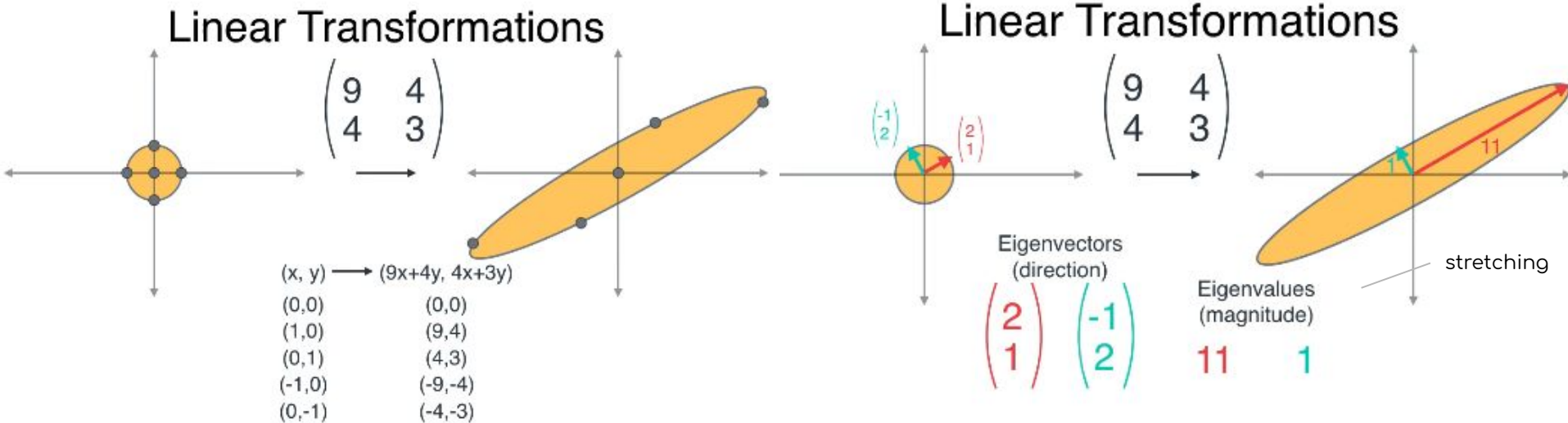
## Linear Transformations



## Linear Transformations



# Ανάλυση Κύριων Συνιστωσών: παράδειγμα



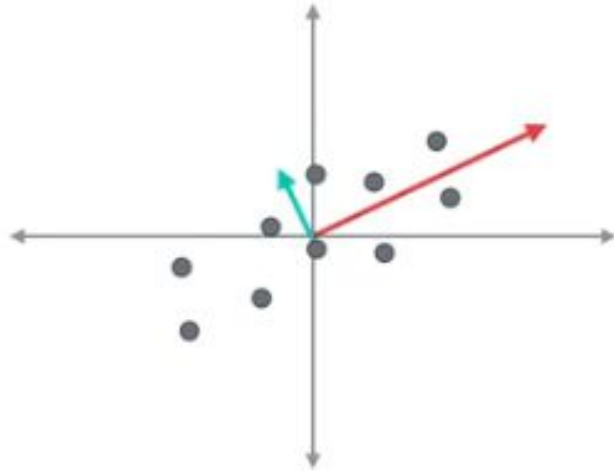
$$\begin{vmatrix} x-9 & -4 \\ -4 & x-3 \end{vmatrix} = (x-9)(x-3) - (-4)(-4) = x^2 - 12x + 11 = (x-11)(x-1)$$

$$\begin{pmatrix} 9 & 4 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = 11 \begin{pmatrix} u \\ v \end{pmatrix} \quad \begin{pmatrix} 9 & 4 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = 1 \begin{pmatrix} u \\ v \end{pmatrix}$$

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$



## Principal Component Analysis (PCA)



$$\Sigma = \begin{pmatrix} 9 & 4 \\ 4 & 3 \end{pmatrix}$$

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

Eigenvectors  
(direction)

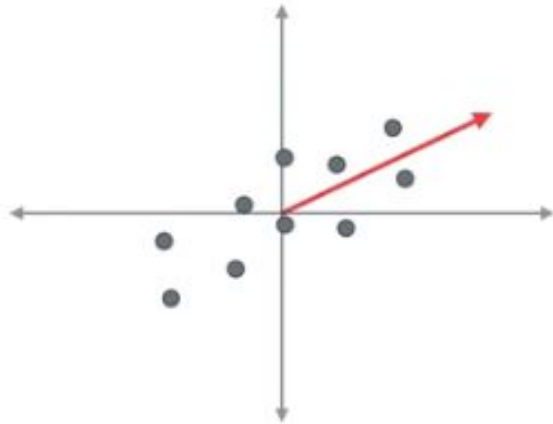
11

1

Eigenvalues  
(magnitude)



## Principal Component Analysis (PCA)



$$\Sigma = \begin{pmatrix} 9 & 4 \\ 4 & 3 \end{pmatrix}$$

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Eigenvectors  
(direction)

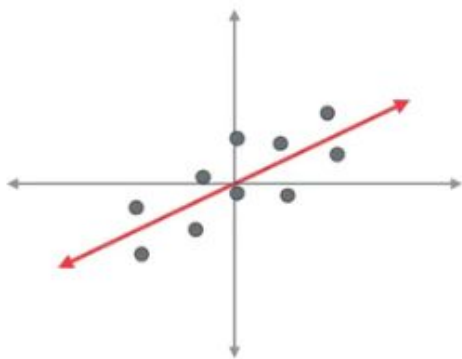
11

Eigenvalues  
(magnitude)



# Ανάλυση Κύριων Συνιστωσών: παράδειγμα

## Principal Component Analysis (PCA)



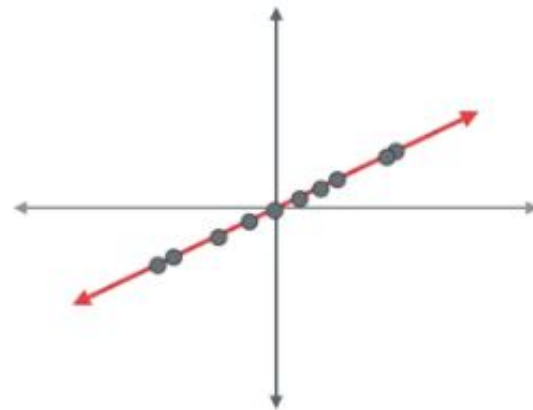
$$\Sigma = \begin{pmatrix} 9 & 4 \\ 4 & 3 \end{pmatrix}$$

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

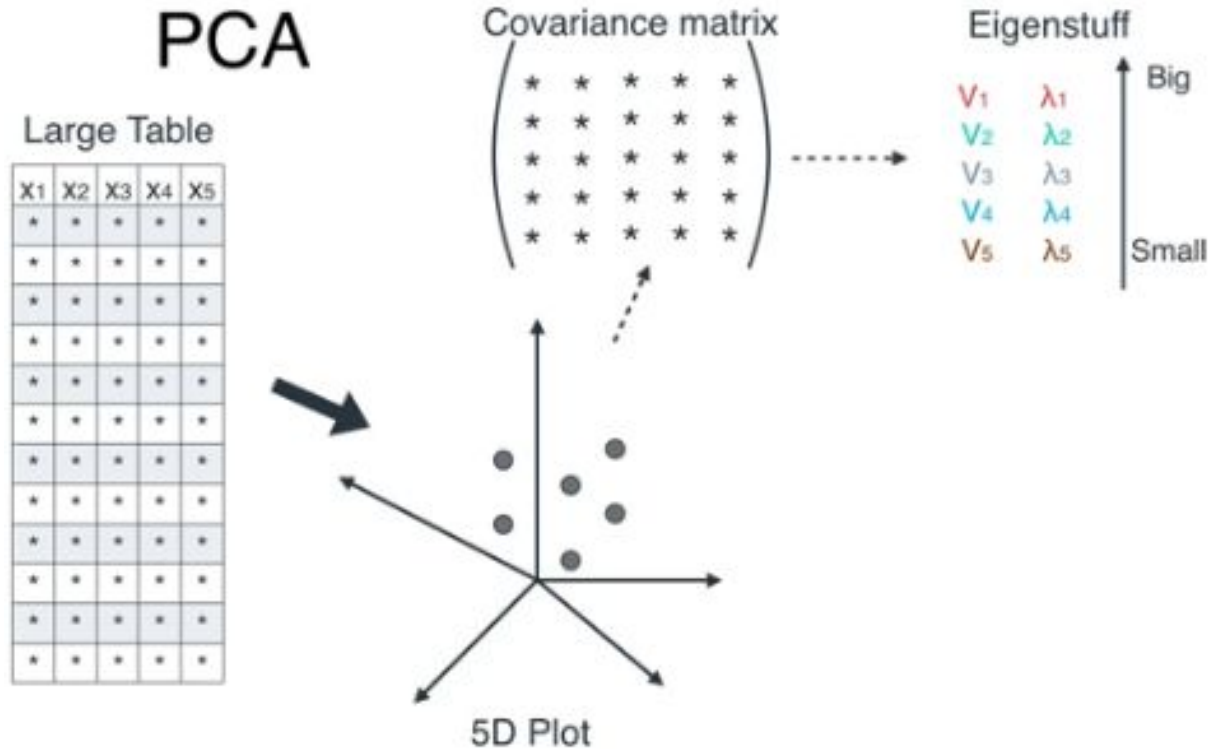
Eigenvectors  
(direction)

$$11$$

Eigenvalues  
(magnitude)

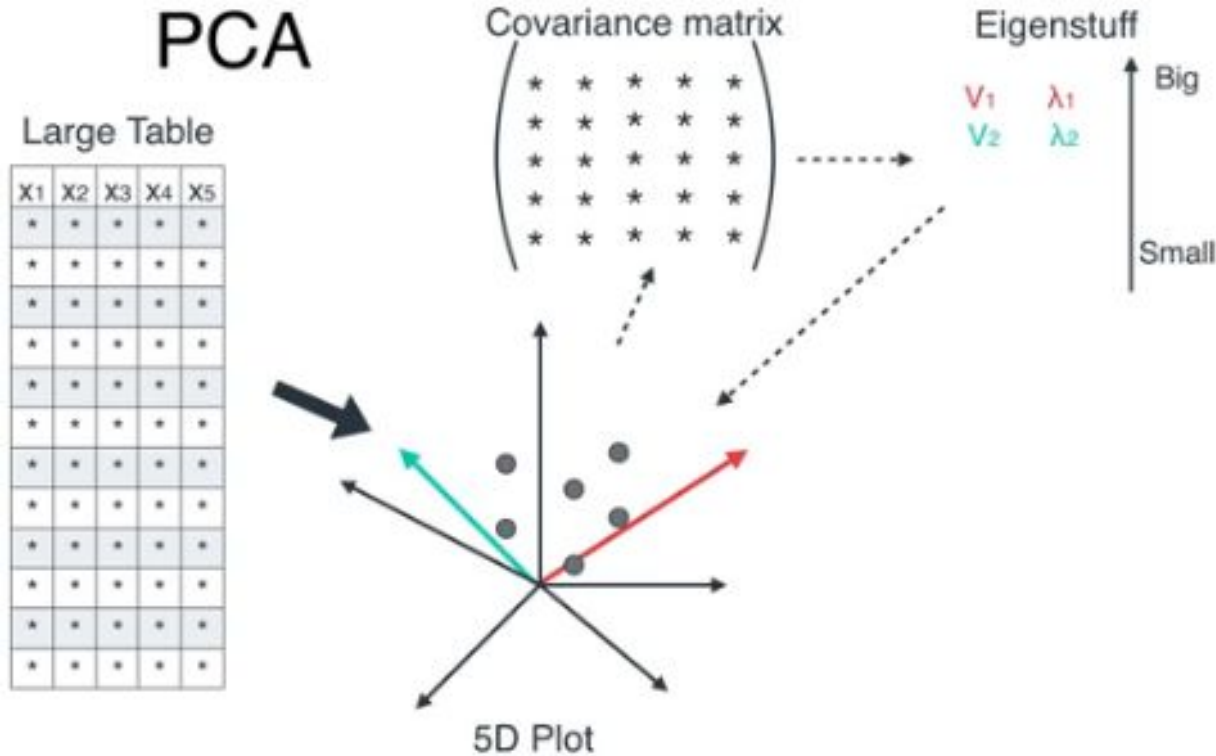


# Ανάλυση Κύριων Συνιστωσών: παράδειγμα

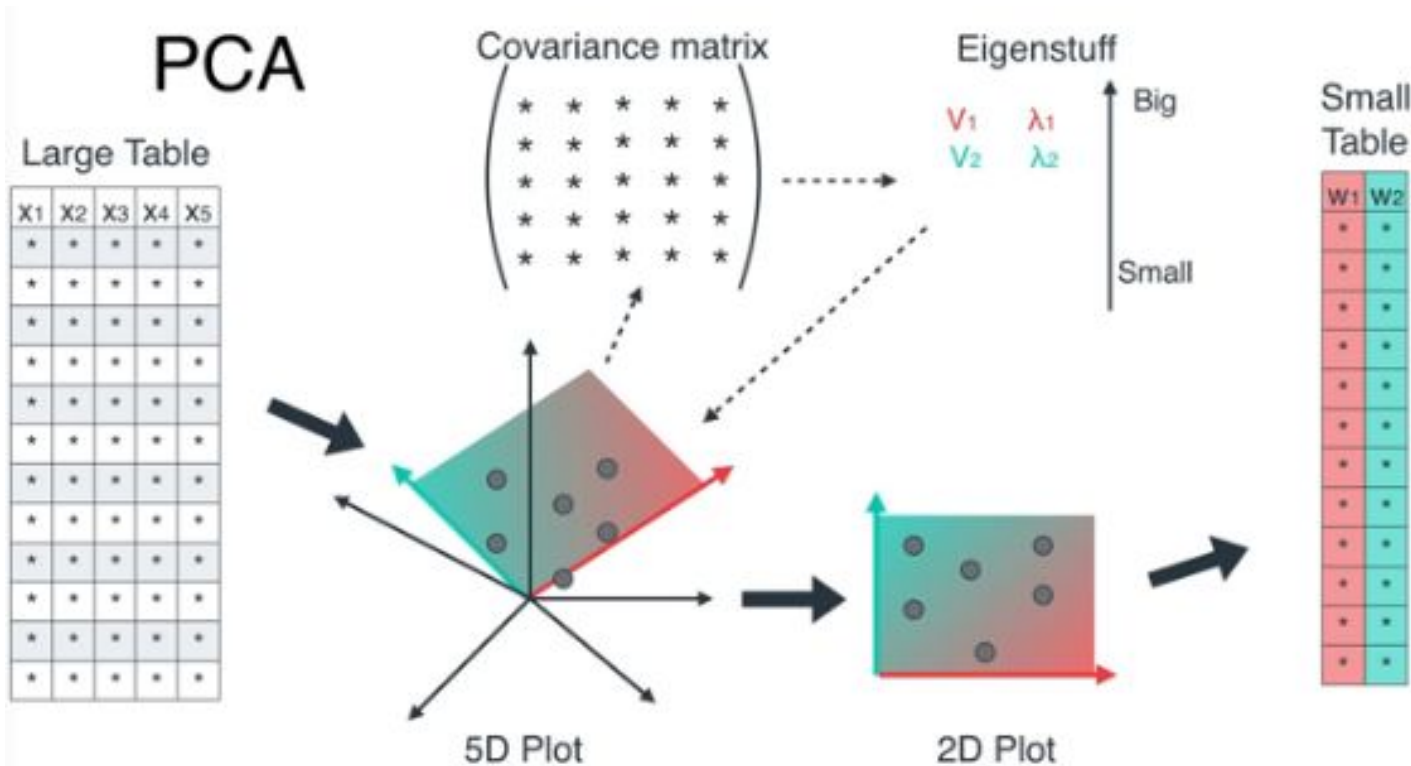




# Ανάλυση Κύριων Συνιστωσών: παράδειγμα



# Ανάλυση Κύριων Συνιστωσών: παράδειγμα





# Ανάλυση Κύριων Συνιστωσών (Principal Components Analysis)

Έστω  $x = (x^{(1)}, \dots, x^{(m)}) \in \mathcal{R}^n$

Συμπύεση: αποθήκευση των σημείων με τρόπο που απαιτεί λιγότερη μνήμη, αλλά μπορεί να χάσει κάποια ακρίβεια.

→ αναπαράσταση σε μικρότερες διαστάσεις, χωρίς να χάνουμε πολύ σε ακρίβεια

$\forall x^{(i)} \in \mathcal{R}^n$  θα υπολογίσουμε το  $c^{(i)} \in \mathcal{R}^l$

Αν  $l < n$ , η αποθήκευση των  $c$  θα απαιτήσει λιγότερη μνήμη από την αποθήκευση των  $x$ .

Αναζητούμε τη συνάρτηση κωδικοποίησης  $f(x) = c$   
και τη συνάρτηση αποκωδικοποίησης  $x \approx g(f(c))$



# Ανάλυση Κύριων Συνιστωσών

Το PCA ορίζεται από την επιλογή της συνάρτησης αποκωδικοποίησης.

$$\text{Έστω, } g(\mathbf{c}) = D\mathbf{c}, \quad D \in \mathbb{R}^{n \times l}$$

Ο υπολογισμός του βέλτιστου  $\mathbf{c}$  για αυτόν τον αποκωδικοποιητή μπορεί να είναι ένα δύσκολο πρόβλημα.

→ Για να παραμείνει εύκολο το πρόβλημα της κωδικοποίησης,  
ο PCA περιορίζει τις στήλες του D ώστε να είναι ορθογώνιες μεταξύ τους.

(ο D δεν είναι ακόμα από τεχνικής απόψεως "ορθογώνιος πίνακας" εκτός και αν  $l = n$ )

# Ανάλυση Κύριων Συνιστωσών

→ Πρέπει να δημιουργήσουμε το βέλτιστο  $c^*$  για κάθε  $x$

Να ελαχιστοποιηθεί η απόσταση μεταξύ του σημείου εισόδου  $x$  και της ανακατασκευής του,  $g(c^*)$  → χρήση νόρμας  $L^2$



$$c^* = \underset{c}{\operatorname{argmin}} \|x - g(c)\|_2$$

Επιλέγουμε το τετράγωνο της  $L^2$  νόρμας επειδή είναι μονοτονική και αύξουσα συνάρτηση για θετικές τιμές. Οπότε:

$$c^* = \underset{c}{\operatorname{argmin}} \|x - g(c)\|_2^2$$

Οπότε αρκεί να ελαχιστοποιήσουμε το  $(x - g(c))^T (x - g(c))$



$$c^* = \underset{c}{\operatorname{argmin}} - 2x^T g(c) + g(c)^T g(c)$$

$$c^* = \underset{c}{\operatorname{argmin}} - 2x^T Dc + c^T D^T Dc$$

$$= \underset{c}{\operatorname{argmin}} - 2x^T Dc + c^T I_1 c$$

$$= \underset{c}{\operatorname{argmin}} - 2x^T Dc + c^T c$$

$$\nabla c(-2x^T Dc + c^T c) = 0$$

$$- 2D^T x + 2c = 0$$

$$c = D^T x$$

Μπορούμε να κωδικοποιήσουμε το  $x$  χρησιμοποιώντας απλώς έναν πίνακα !

Κωδικοποιητής :  $f(x) = D^T x$ .

Ανασυγκρότηση PCA :  $r(x) = g(f(x)) = DD^T x$



# Ανάλυση Κύριων Συνιστωσών

Πρέπει να επιλέξουμε τον πίνακα κωδικοποίησης  $D$ .

→ επανεξετάζουμε την ιδέα της ελαχιστοποίησης της απόστασης  $L^2$  μεταξύ του  $x$  και του  $c$

Δεδομένου ότι θα χρησιμοποιήσουμε τον ίδιο πίνακα  $D$  για να αποκωδικοποιήσουμε όλα τα σημεία, δεν μπορούμε πλέον να θεωρούμε τα σημεία μεμονωμένα.

Αντ' αυτού, πρέπει να ελαχιστοποιήσουμε την νόρμα Frobenius του πίνακα σφαλμάτων, λαμβάνοντας υπόψη όλες τις διαστάσεις και όλα τα σημεία.

$$D^* = \underset{D}{\operatorname{argmin}} \sqrt{\sum_{i,j} (x_j^{(i)} - r(x^{(i)})_j)^2}$$

υπό την προϋπόθεση ότι  $D^T D = I_l$

# Ανάλυση Κύριων Συνιστωσών

Εύρεση του  $D^*$

$l = 1$  :  $D \rightarrow$  vector  $d$  οπότε το πρόβλημα μειώνεται σε  $r(x) = DD^T x$

$$d^* = \underset{d}{\operatorname{argmin}} \sum_i \left\| (x_i^{(i)} - dd^T x^{(i)}) \right\|_2^2$$

$$d^* = \underset{d}{\operatorname{argmin}} \sum_i \left\| (x_i^{(i)} - d^T x^{(i)} d) \right\|_2^2$$

$$d^* = \underset{d}{\operatorname{argmin}} \sum_i \left\| (x_i^{(i)} - x^{(i)T} dd) \right\|_2^2$$

υπό την προϋπόθεση ότι

$$\|d\|_2 = 1$$



# Ανάλυση Κύριων Συνιστωσών

$$d^* = \underset{d}{\operatorname{argmin}} \left\| (X - Xdd^T) \right\|_F^2$$

$$d^T d = 1$$

υπό την προϋπόθεση ότι

$$\underset{d}{\operatorname{argmax}} \operatorname{Tr}(d^T X^T X d)$$

Πρόβλημα βελτιστοποίησης → χρήση eigendecomposition.

Συγκεκριμένα, η βέλτιστη  $d$  δίνεται από τον ιδιοδιάνυσμα του  $X^T X$  που αντιστοιχεί στη μεγαλύτερη ιδιοτιμή.

Αυτή η παράγωγος είναι ειδική για την περίπτωση του  $l = 1$  και ανακτά μόνο το πρώτο κύριο συστατικό.

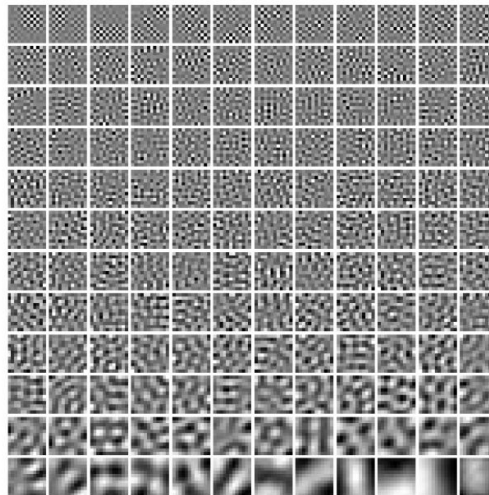
Γενικότερα, όταν θέλουμε να ανακτήσουμε μια βάση των κύριων συστατικών, η μήτρα  $D$  δίνεται από τα ιδιοδιανύσματα που αντιστοιχούν στις μεγαλύτερες

# Εφαρμογές : Συμπύεση εικόνας

PCA example — original



PCA example — PCA basis



PCA example — 90% compressed



PCA example — 50% compressed

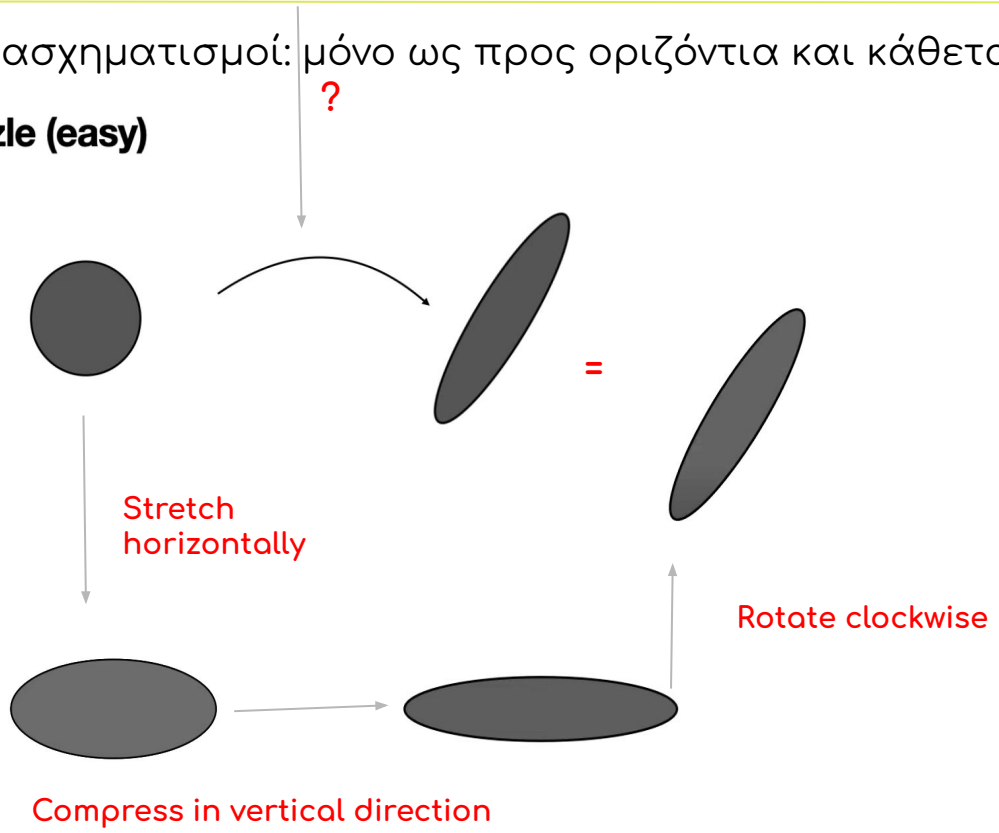




# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)

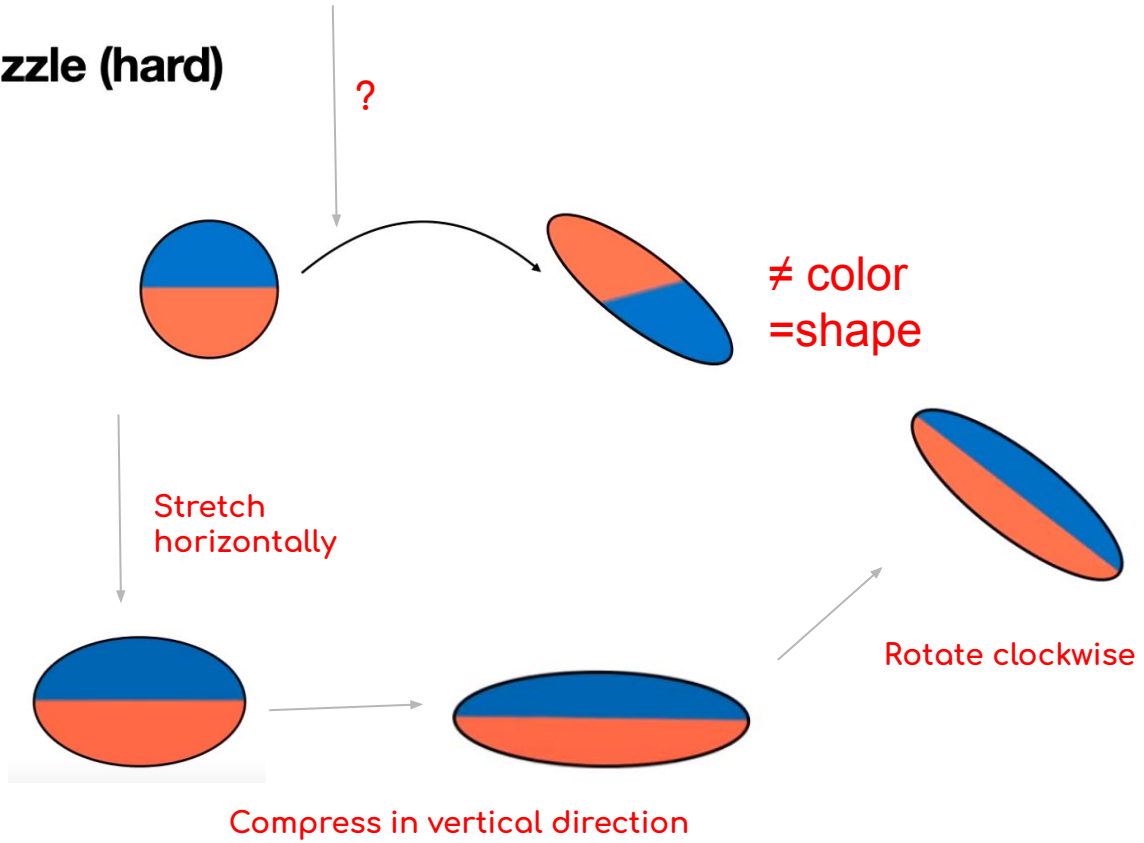
Μετασχηματισμοί: μόνο ως προς οριζόντια και κάθετα κατεύθυνση, όχι υπό άλλη γωνία

**Puzzle (easy)**



# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)

**Puzzle (hard)**

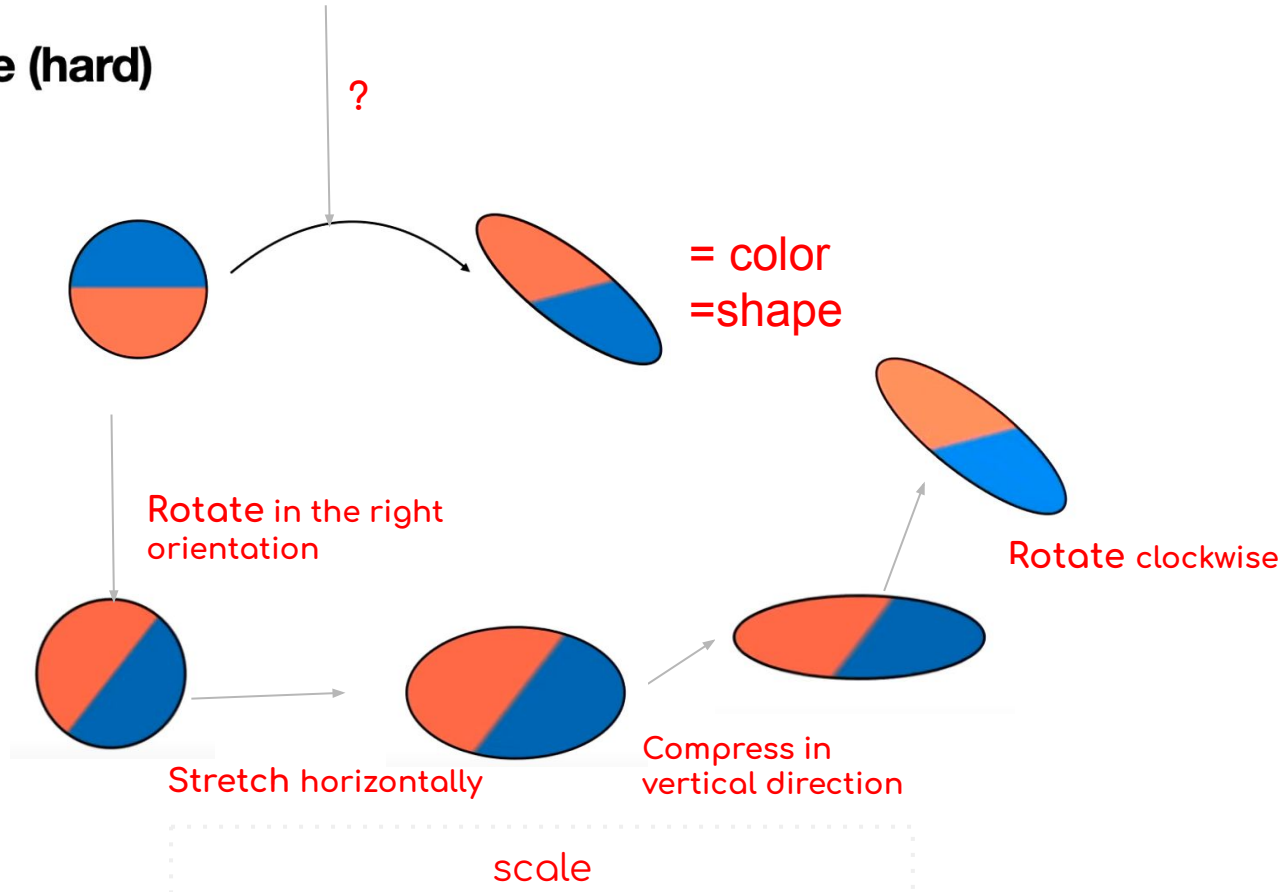


# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)

## Puzzle (hard)

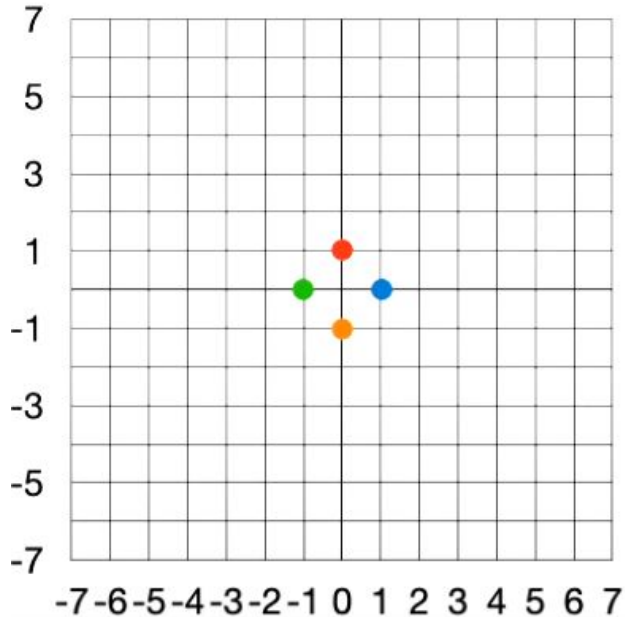
### Λύση

Με περιστροφή, κλιμάκωση και περιστροφή μπορούμε να μιμηθούμε κάθε γραμμικό μετασχηματισμό

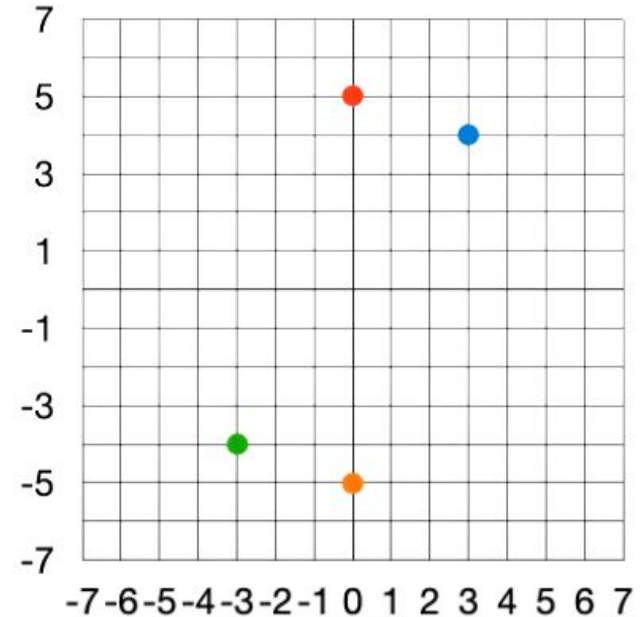


# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)

$$A = \begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix}$$



$(p, q)$	$(3p+0q, 4p+5q)$
$(1, 0)$	$(3, 4)$
$(0, 1)$	$(0, 5)$
$(-1, 0)$	$(-3, -4)$
$(0, -1)$	$(0, -5)$

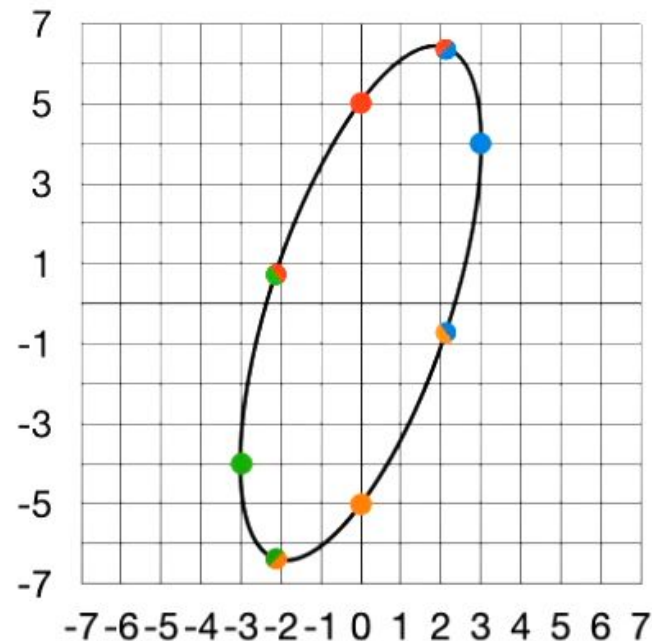
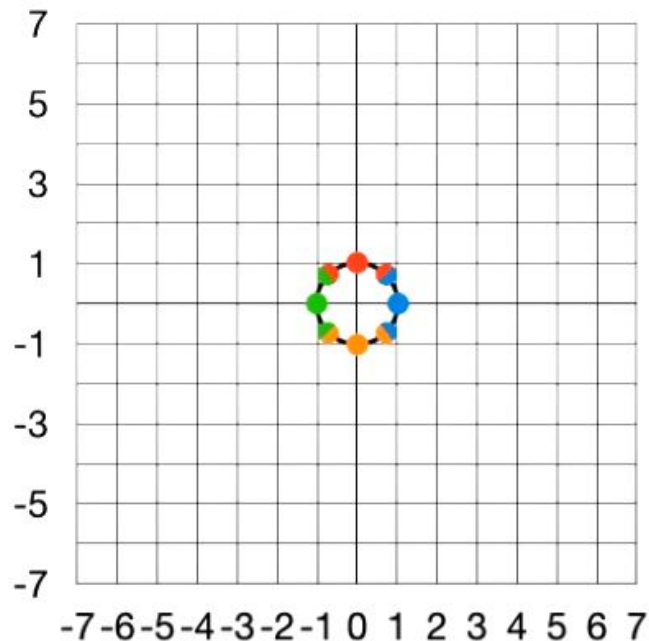


# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)

Γραμμικός μετασχηματισμός

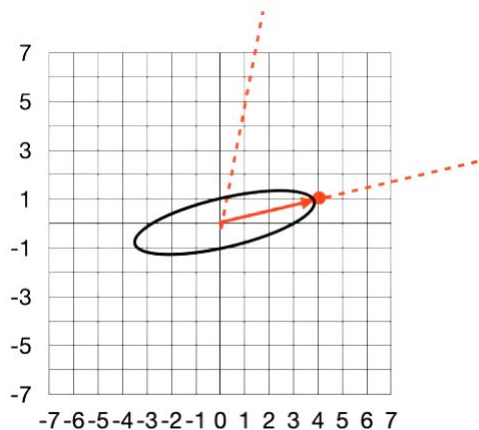
$$A = \begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix}$$

$(p,q)$	$(3p+0q, 4p+5q)$
$(1,0)$	$(3, 4)$
$(0,1)$	$(0, 5)$
$(-1,0)$	$(-3, -4)$
$(0,-1)$	$(0, -5)$

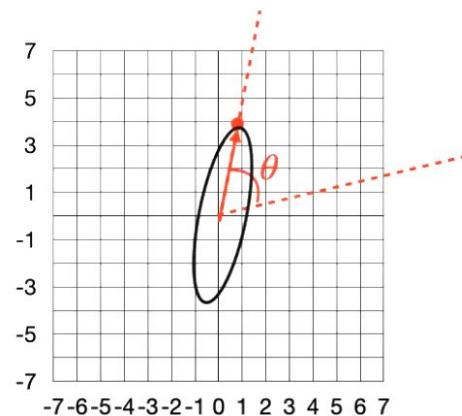


# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)

## Πίνακας περιστροφής



$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$



$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

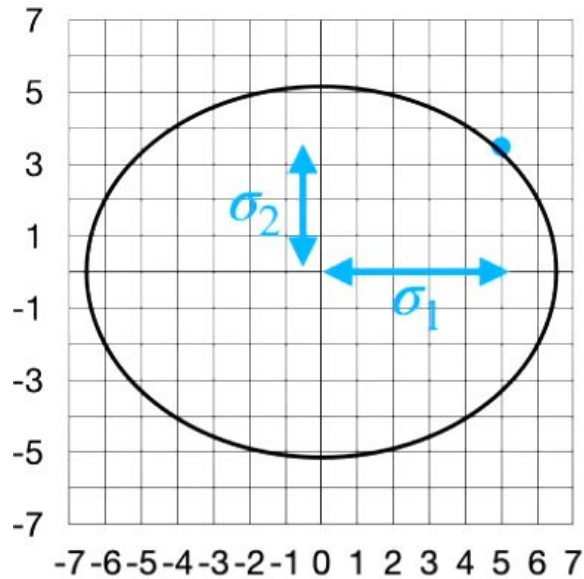




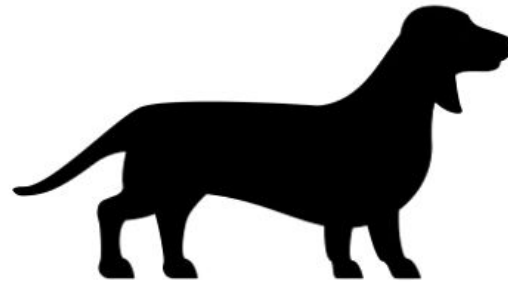
# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)



Πίνακας κλιμάκωσης

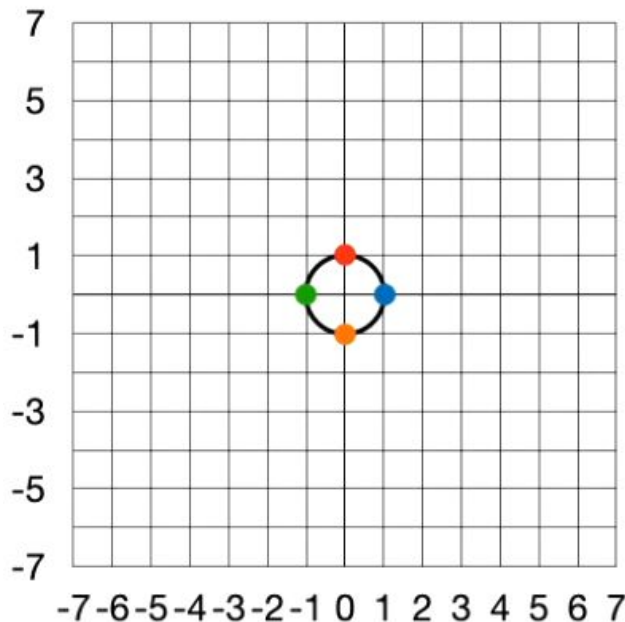


$$\begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}$$



# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)

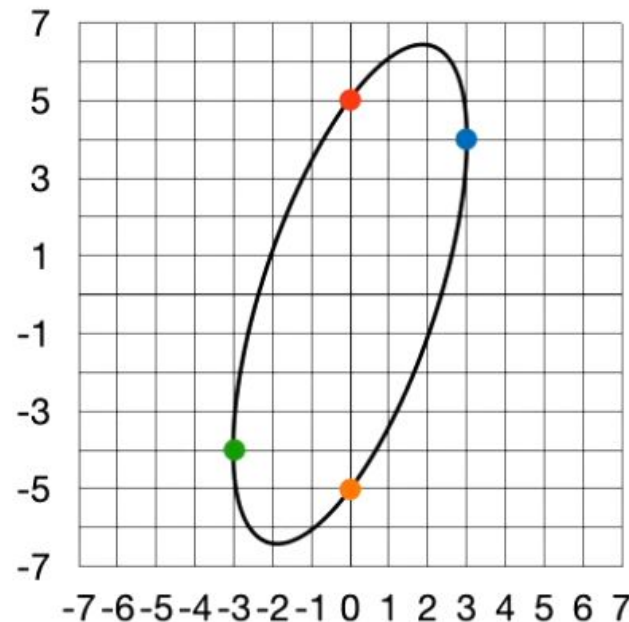
$$A = \begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix}$$



↻  $\begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$

↕  $\begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}$

↻  $\begin{bmatrix} \cos(\phi) & \sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{bmatrix}$



# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)

**WolframAlpha** computational intelligence.

singular value decomposition [[3,0],[4,5]]

Extended Keyboard Upload

Input:

singular value decomposition  $\begin{pmatrix} 3 & 0 \\ 4 & 5 \end{pmatrix}$

Result:

$M = U \cdot \Sigma \cdot V^T$

where

$M = \begin{pmatrix} 3 & 0 \\ 4 & 5 \end{pmatrix}$

$U = \begin{pmatrix} \frac{1}{\sqrt{10}} & -\frac{3}{\sqrt{10}} \\ \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{pmatrix}$

$\Sigma = \begin{pmatrix} 3\sqrt{5} & 0 \\ 0 & \sqrt{5} \end{pmatrix}$

$V = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$

$$A = U \Sigma V^T$$

$$\begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \begin{bmatrix} \cos(\phi) & \sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{bmatrix}$$

```
from numpy.linalg import svd
A = np.array([[3,0],[4,5]])
svd(A)
```

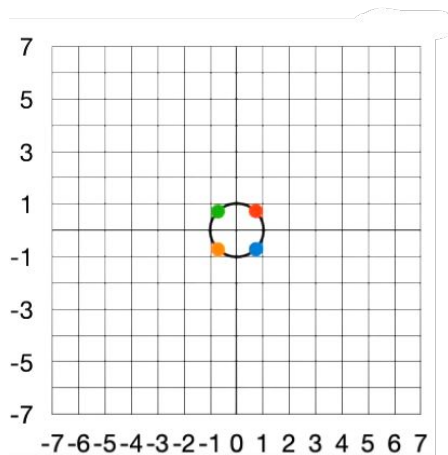
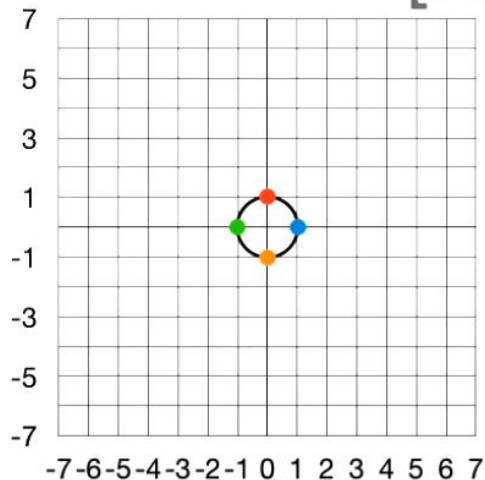
```
(array([[ -0.31622777, -0.9486833 ],
        [ -0.9486833 ,  0.31622777 ]]),
 array([6.70820393, 2.23606798]),
 array([[ -0.70710678, -0.70710678 ],
        [ -0.70710678,  0.70710678 ]]))
```

# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)

$$A = U\Sigma V^\dagger$$

$$\begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix} = \begin{bmatrix} 0.316 & -0.949 \\ 0.949 & 0.316 \end{bmatrix} \begin{bmatrix} 6.708 & 0 \\ 0 & 2.236 \end{bmatrix} \begin{bmatrix} 0.7071 & 0.7071 \\ -0.7071 & 0.7071 \end{bmatrix}$$

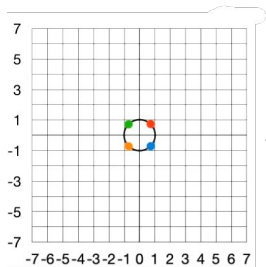
$$\begin{bmatrix} 1/\sqrt{10} & -3/\sqrt{10} \\ 3/\sqrt{10} & 1/\sqrt{10} \end{bmatrix} \begin{bmatrix} 3\sqrt{5} & 0 \\ 0 & \sqrt{5} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$



Rotation of  $\theta = -\frac{\pi}{4} = -45^\circ$

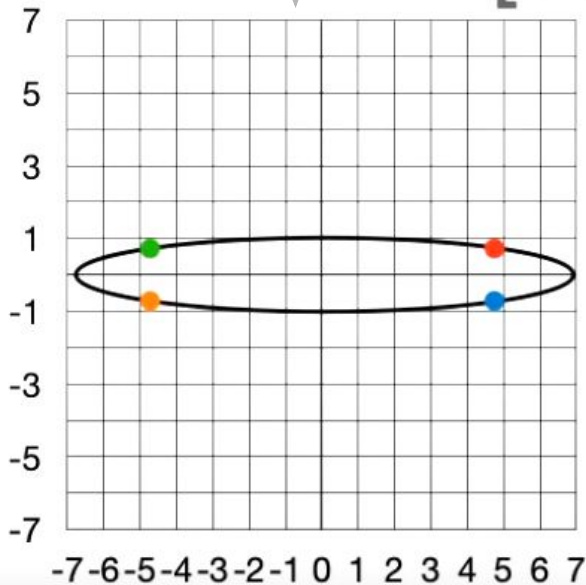


# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)



$$A = U\Sigma V^\dagger$$

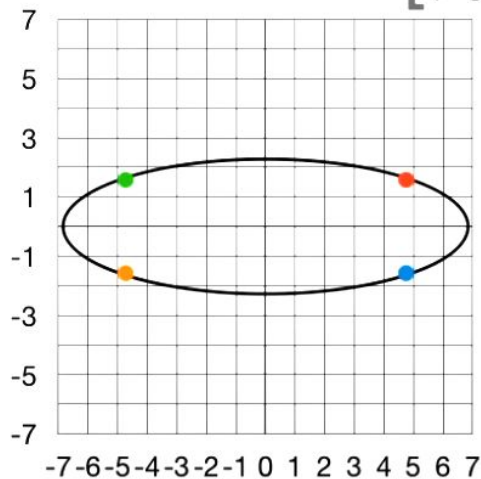
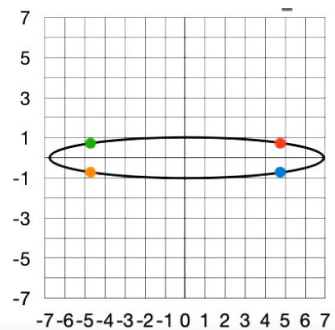
$$\begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix} = \begin{bmatrix} 0.316 & -0.949 \\ 0.949 & 0.316 \end{bmatrix} \begin{bmatrix} 6.708 & 0 \\ 0 & 2.236 \end{bmatrix} \begin{bmatrix} 0.7071 & 0.7071 \\ -0.7071 & 0.7071 \end{bmatrix}$$



$$\begin{bmatrix} 1/\sqrt{10} & -3/\sqrt{10} \\ 3/\sqrt{10} & 1/\sqrt{10} \end{bmatrix} \begin{bmatrix} 3\sqrt{5} & 0 \\ 0 & \sqrt{5} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

Horizontal scaling by  $3\sqrt{5}$

# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)



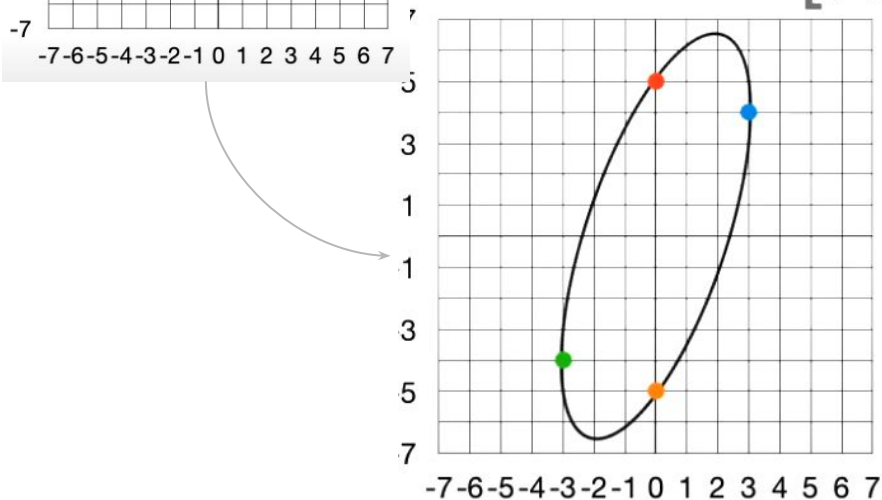
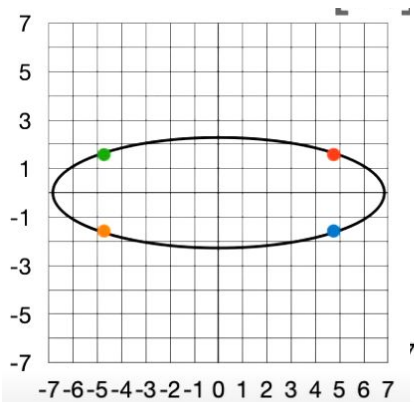
$$\begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix} = \begin{bmatrix} 0.316 & -0.949 \\ 0.949 & 0.316 \end{bmatrix} \begin{bmatrix} 6.708 & 0 \\ 0 & 2.236 \end{bmatrix} \begin{bmatrix} 0.7071 & 0.7071 \\ -0.7071 & 0.7071 \end{bmatrix}$$

$$\begin{bmatrix} 1/\sqrt{10} & -3/\sqrt{10} \\ 3/\sqrt{10} & 1/\sqrt{10} \end{bmatrix} \begin{bmatrix} 3\sqrt{5} & 0 \\ 0 & \sqrt{5} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

Vertical scaling by  $\sqrt{5}$



# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)

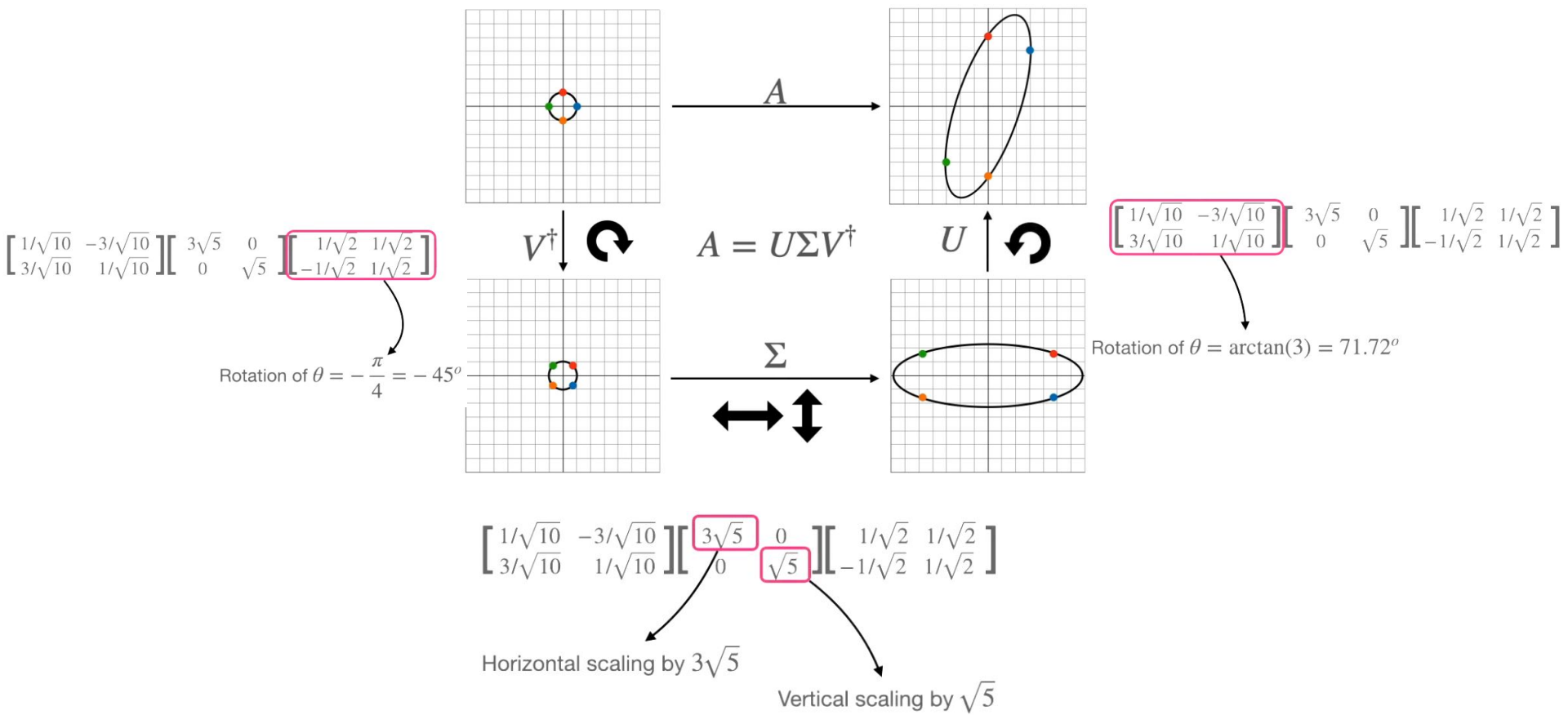


$$\begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix} = \begin{bmatrix} 0.316 & -0.949 \\ 0.949 & 0.316 \end{bmatrix} \begin{bmatrix} 6.708 & 0 \\ 0 & 2.236 \end{bmatrix} \begin{bmatrix} 0.7071 & 0.7071 \\ -0.7071 & 0.7071 \end{bmatrix}$$

$$\begin{bmatrix} 1/\sqrt{10} & -3/\sqrt{10} \\ 3/\sqrt{10} & 1/\sqrt{10} \end{bmatrix} \begin{bmatrix} 3\sqrt{5} & 0 \\ 0 & \sqrt{5} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

Rotation of  $\theta = \arctan(3) = 71.72^\circ$

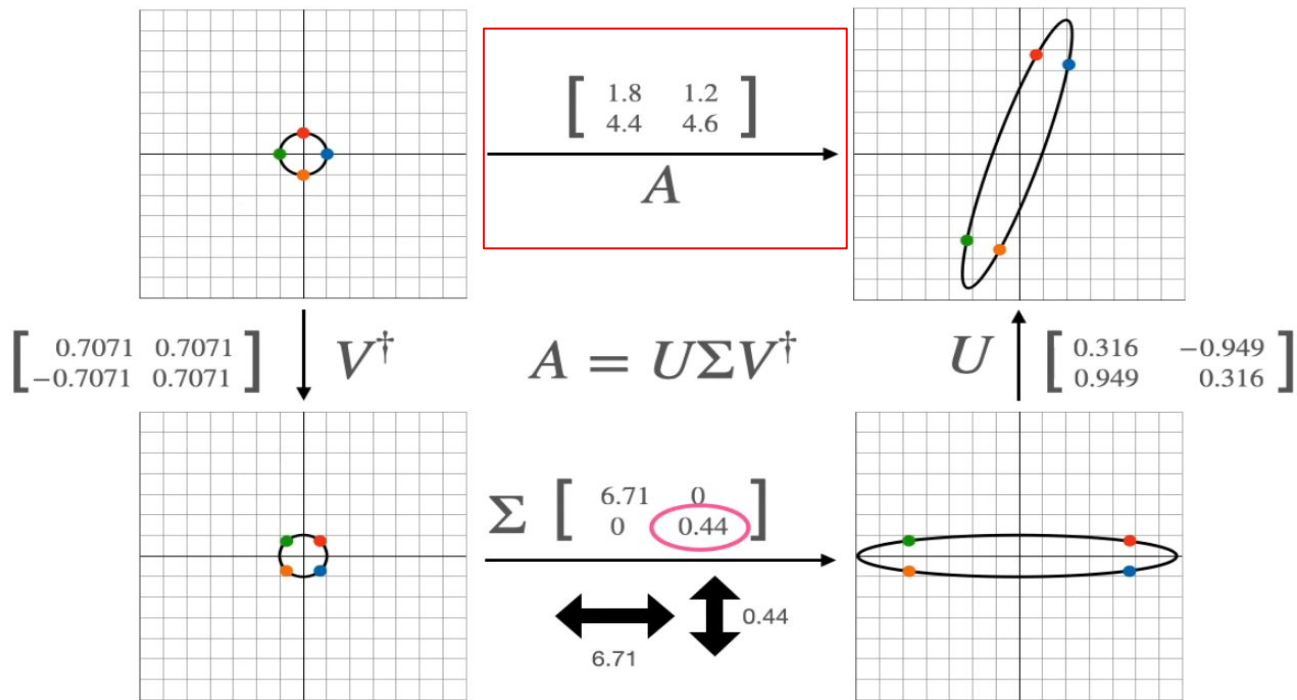
# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)



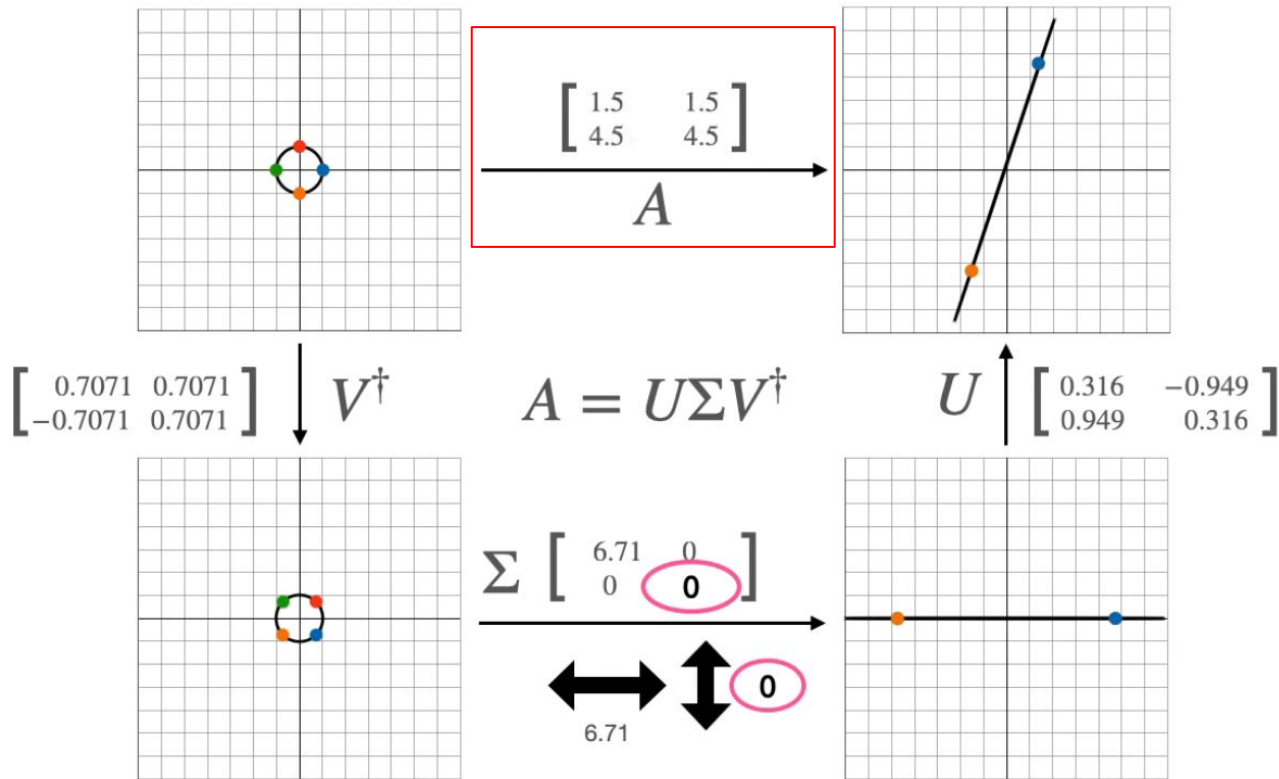


# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)

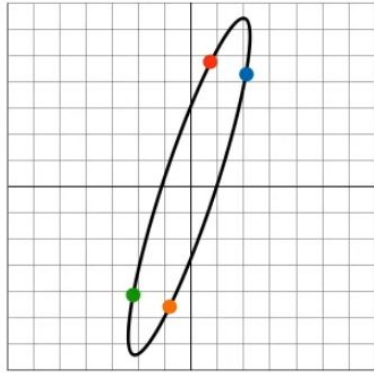
Μείωση διαστάσεων



# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)

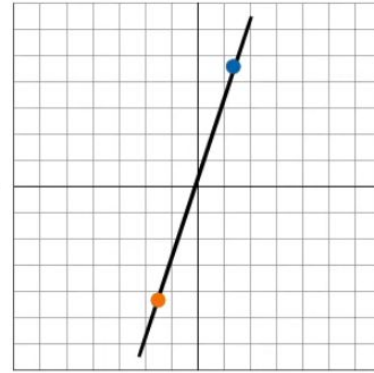


# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)



$$\begin{bmatrix} 1.8 & 1.2 \\ 4.4 & 4.6 \end{bmatrix}$$

Rank 2



$$\begin{bmatrix} 1.5 & 1.5 \\ 4.5 & 4.5 \end{bmatrix}$$

Rank 1

# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)



Πίνακας με rank=1

	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
<b>1</b>	1	2	3	4
<b>-1</b>	-1	-2	-3	-4
<b>2</b>	2	4	6	8
<b>10</b>	10	20	30	40



<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
<b>-1</b>	<b>-2</b>	<b>-3</b>	<b>-4</b>
<b>2</b>	<b>4</b>	<b>6</b>	<b>8</b>
<b>10</b>	<b>20</b>	<b>30</b>	<b>40</b>

=

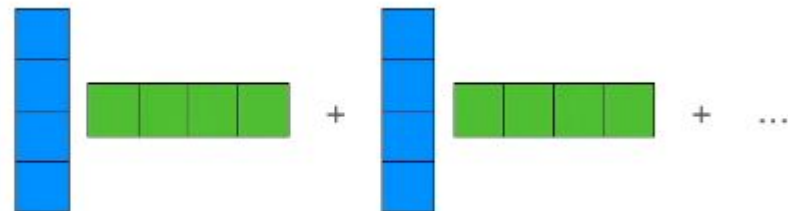
<b>1</b>			
<b>-1</b>			
<b>2</b>	<b>1</b>	<b>2</b>	<b>3</b>
<b>10</b>	<b>2</b>	<b>4</b>	<b>8</b>

# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)

Πίνακας με  $\text{rank} > 1$ : προσέγγιση όπως λειτουργεί για  $\text{rank} = 1$

3	1	4	1
5	9	2	6
5	3	5	8
9	7	9	3

=

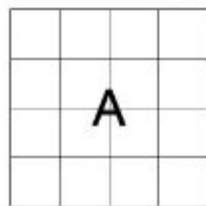


+ ...

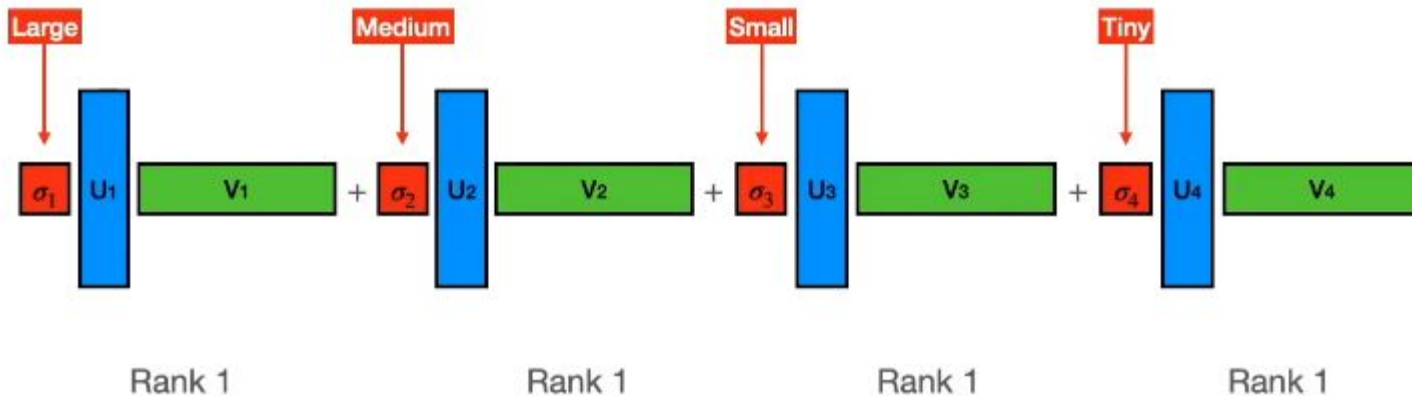
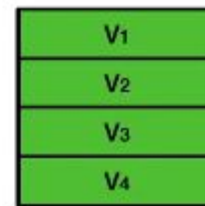
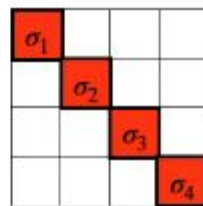
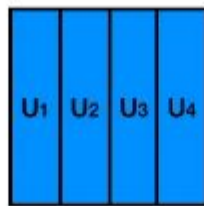
# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)



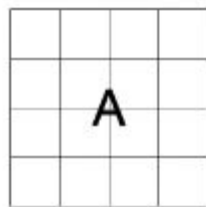
$$A = U \Sigma V^T$$



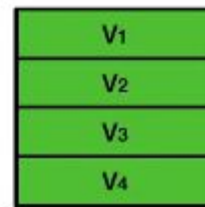
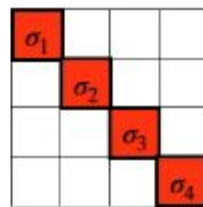
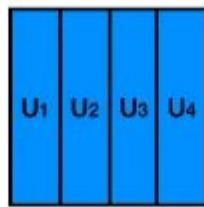
=



# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)



=



Large

$\sigma_1$



Rank 1

Medium

$\sigma_2$



Rank 1

Small

$\sigma_3$



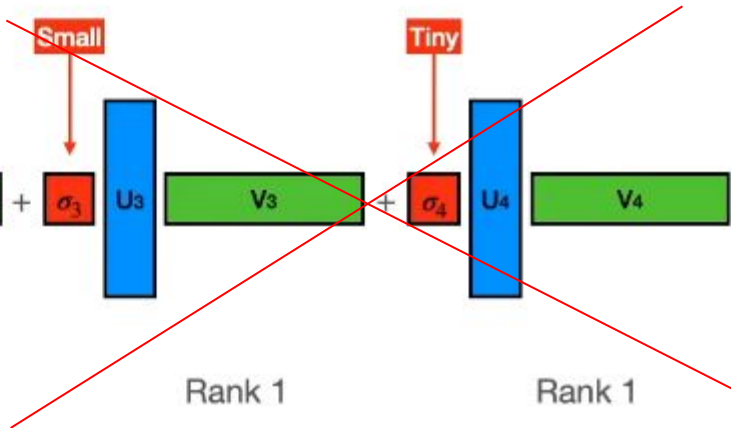
Rank 1

Tiny

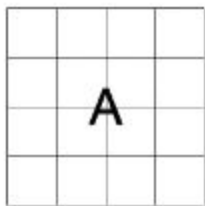
$\sigma_4$



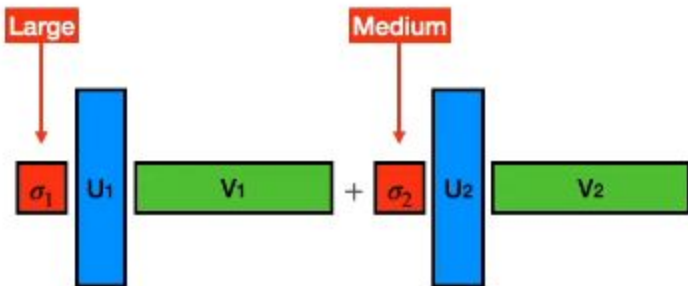
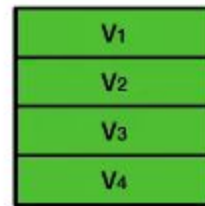
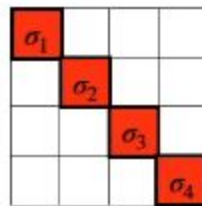
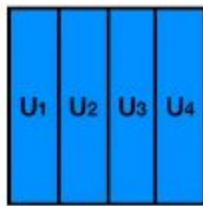
Rank 1



# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)



=



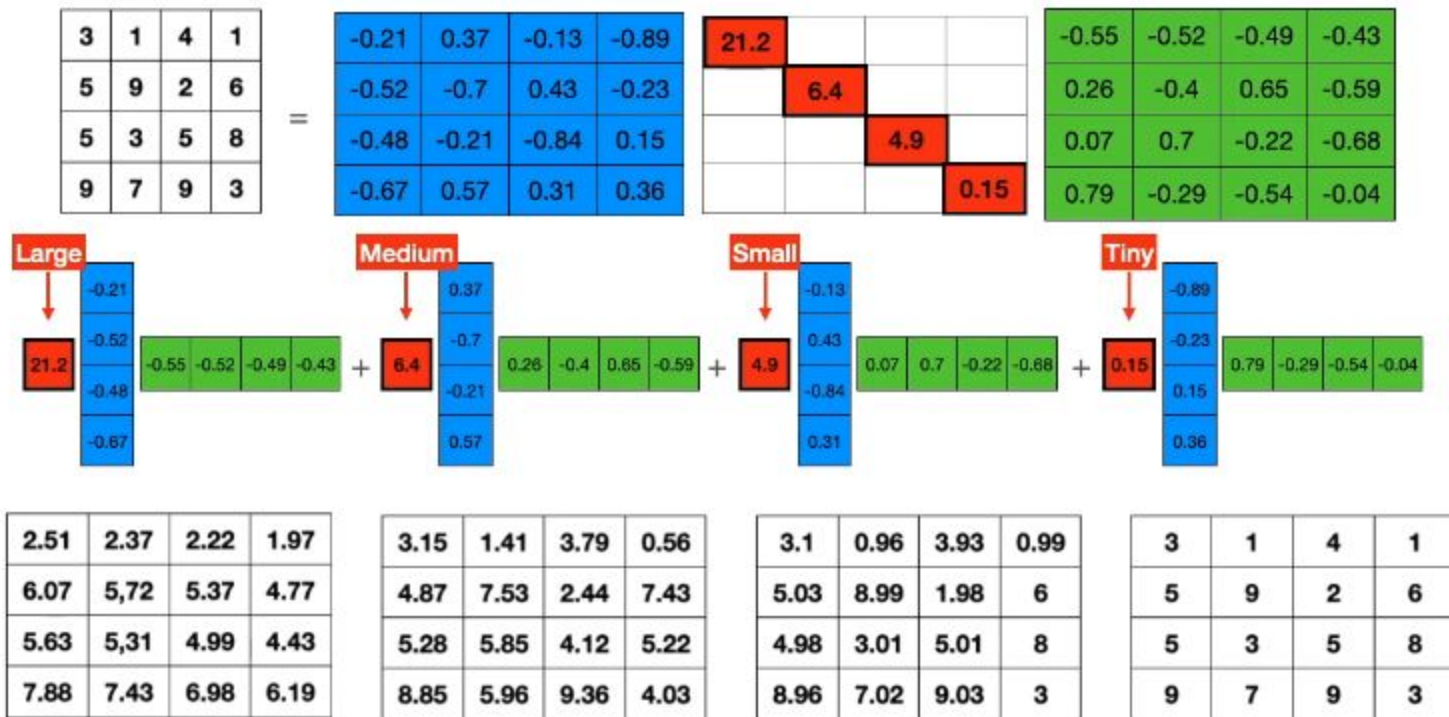
Rank 1

Rank 1





# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)



# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)



3	1	4	1
5	9	2	6
5	3	5	8
9	7	9	3

=

-0.21	0.37	-0.13	-0.89
-0.52	-0.7	0.43	-0.23
-0.48	-0.21	-0.84	0.15
-0.67	0.57	0.31	0.36

21.2			
	6.4		
		4.9	
			0.15

-0.55	-0.52	-0.49	-0.43
0.26	-0.4	0.65	-0.59
0.07	0.7	-0.22	-0.68
0.79	-0.29	-0.54	-0.04

Large

21.2
-0.21
-0.52
-0.48
-0.67

Medium

6.4
0.37
-0.7
-0.21
0.57

-0.55	-0.52	-0.49	-0.43
-------	-------	-------	-------

+

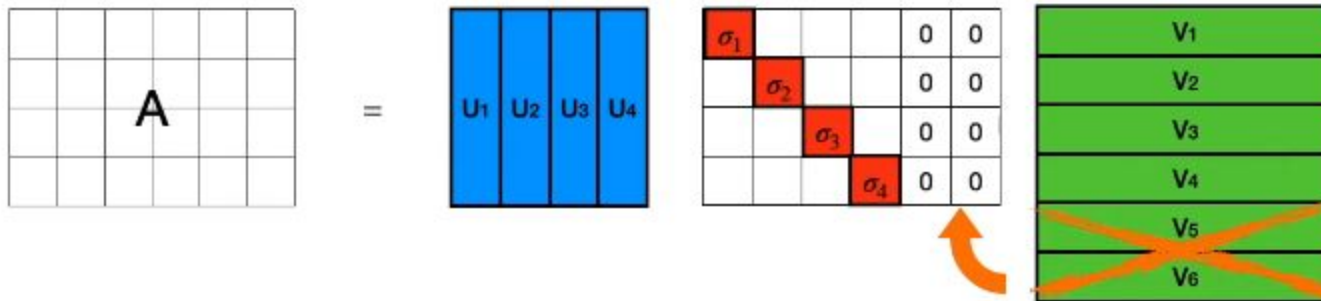
0.26	-0.4	0.65	-0.59
------	------	------	-------

2.51	2.37	2.22	1.97
6.07	5.72	5.37	4.77
5.63	5.31	4.99	4.43
7.88	7.43	6.98	6.19

3.15	1.41	3.79	0.56
4.87	7.53	2.44	7.43
5.28	5.85	4.12	5.22
8.85	5.96	9.36	4.03

# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)

Μη τετραγωνικός πίνακας;





## Ανάλυση διανυσμάτων πάνω σε ορθογώνιους άξονες

- Αν ένας πίνακας δεν είναι τετραγωνικός, η ανάλυση ιδιοτιμών δεν ορίζεται.
- Αντ' αυτού, χρησιμοποιούμε την ανάλυση σε ιδιάζουσες τιμές

Κάθε πίνακας  $A$  ( $m \times n$ ) μπορεί να γραφεί ως  $A = U \Sigma V^T$

Σημαντική εφαρμογή: Μπορούμε να γενικεύσουμε, εν μέρει, την αντιστροφή ενός μη τετραγωνικού πίνακα

# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)



$$A = U\Sigma V^T$$

Όπου U: Ορθογώνιος πίνακας  $m \times m$

→ Περιστροφή (Rotation)

Σ: Διαγώνιος πίνακας διάστασης  $m \times n$   
(όχι απαραίτητα τετραγωνικός)

→ Έκταση (Stretch)

V: Ορθογώνιος πίνακας  $n \times n$

→ Περιστροφή (Rotation)

$$A = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \end{bmatrix}$$

Left singular vector      singular values      Right singular vector

# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)



Τι είναι τα  $U, \Sigma, V^T$

Ίδιος πίνακας

$$A = \underline{V} \underline{\Lambda} V^{-1}$$

Διαφορετικοί Πίνακες

$$A = \underline{U} \underline{\Sigma} V^T$$

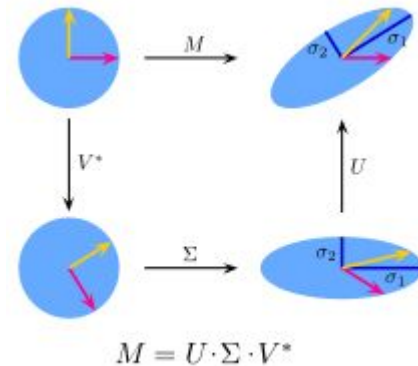
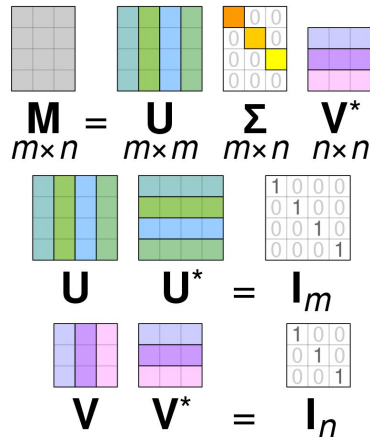
# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)

λ, για A<sup>T</sup>A  
σ<sup>2</sup>, για A

$$A = U\Sigma V^T$$

$$\boxed{A^T A} = (V\Sigma^T U^T)U\Sigma V^T = V\Sigma^T (U^T U)\Sigma V^T = V(\Sigma^T \Sigma)V^T$$

$$\overline{AA^T} = U\Sigma V^T (V\Sigma^T U^T) = U\Sigma(V^T V)\Sigma^T U^T = U(\Sigma\Sigma^T)U^T$$





# Βήμα προς βήμα με python

## Deep Learning Book Series - 2.8 Singular Value Decomposition

A : πίνακας που μπορούμε να τον “δούμε” ως γραμμικό μετασχηματισμό που μπορεί να αναλυθεί σε 3 υπό- μετασχηματισμούς (αντιστοιχούν σε 3 πίνακες) :

1. Περιστροφή,
2. Έκταση,
3. Περιστροφή.



# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)



[www.github.com/luisguiserrano/singular-value-decomposition](https://www.github.com/luisguiserrano/singular-value-decomposition)

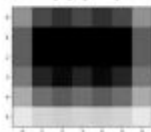
0	1	1	0	1	1	0
1	1	1	1	1	1	1
1	1	1	1	1	1	1
0	1	1	1	1	1	0
0	0	1	1	1	0	0
0	0	0	1	0	0	0

```
[[-0.36 -0.   -0.73 -0.05  0.56  0.13]
 [-0.54  0.35  0.27 -0.08 -0.16  0.69]
 [-0.54  0.35  0.27 -0.08  0.16 -0.69]
 [-0.45 -0.35 -0.27  0.52 -0.56 -0.13]
 [-0.28 -0.71  0.18 -0.62 -0.   -0. ]
 [-0.08 -0.35  0.46  0.57  0.56  0.13]]
```

```
[4.74 0.  0.  0.  0.  0.  0. ]
[0.  1.41 0.  0.  0.  0.  0. ]
[0.  0.  1.41 0.  0.  0.  0. ]
[0.  0.  0.  0.73 0.  0.  0. ]
[0.  0.  0.  0.  0.  0.  0. ]
[0.  0.  0.  0.  0.  0.  0. ]
```

```
[[-0.23 -0.4 -0.46 -0.4 -0.46 -0.4 -0.23]
 [0.5  0.25 -0.25 -0.5 -0.25  0.25  0.5]
 [0.39 -0.32 -0.19  0.65 -0.19 -0.32  0.39]
 [-0.22  0.42 -0.44  0.42 -0.44  0.42 -0.22]
 [0.56 -0.43  0.03  0.  -0.03  0.43 -0.56]
 [-0.42 -0.55 -0.16  0.  0.16  0.55  0.42]
 [-0.12 -0.11  0.69 -0.  -0.69  0.11  0.12]]
```

4.74



# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)



[www.github.com/luisguiserrano/singular\\_value\\_decomposition](https://www.github.com/luisguiserrano/singular_value_decomposition)

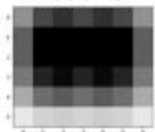
0	1	1	0	1	1	0
1	1	1	1	1	1	1
1	1	1	1	1	1	1
0	1	1	1	1	1	0
0	0	1	1	1	0	0
0	0	0	1	0	0	0

```
[[-0.36  -0.   -0.73 -0.05  0.56  0.13]
 [-0.54  0.35  0.27 -0.08 -0.16  0.69]
 [-0.54  0.35  0.27 -0.08  0.16 -0.69]
 [-0.45 -0.35 -0.27  0.52 -0.56 -0.13]
 [-0.28 -0.71  0.18 -0.62 -0.   -0. ]
 [-0.08 -0.35  0.46  0.57  0.56  0.13]]
```

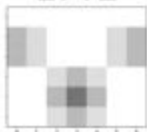
```
[[4.74 0.   0.   0.   0.   0.   0. ]
 [0.  1.41 0.   0.   0.   0.   0. ]
 [0.  0.  1.41 0.   0.   0.   0. ]
 [0.  0.  0.  0.73 0.   0.   0. ]
 [0.  0.  0.  0.   0.   0.   0. ]
 [0.  0.  0.  0.   0.   0.   0. ]
```

```
[[-0.23 -0.4  -0.46 -0.4  -0.46 -0.4  -0.23]
 [ 0.5  0.25 -0.25 -0.5  -0.25  0.25  0.5 ]
 [ 0.39 -0.32 -0.19  0.65 -0.19 -0.32  0.39]
 [-0.22  0.42 -0.44  0.42 -0.44  0.42 -0.22]
 [ 0.56 -0.43  0.03  0.  -0.03  0.43 -0.56]
 [-0.42 -0.55 -0.16  0.  0.16  0.55  0.42]
 [-0.12 -0.11  0.69 -0.  -0.69  0.11  0.12]]
```

4.74



1.41



# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)



[www.github.com/luisguiserrano/singular\\_value\\_decomposition](https://www.github.com/luisguiserrano/singular_value_decomposition)

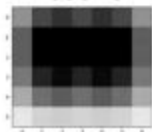
0	1	1	0	1	1	0
1	1	1	1	1	1	1
1	1	1	1	1	1	1
0	1	1	1	1	1	0
0	0	1	1	1	0	0
0	0	0	1	0	0	0

```
[[-0.36 -0. -0.73 -0.05 0.56 0.13]
 [-0.54 0.35 0.27 -0.08 -0.16 0.69]
 [-0.54 0.35 0.27 -0.08 0.16 -0.69]
 [-0.45 -0.35 -0.27 0.52 -0.56 -0.13]
 [-0.28 -0.71 0.18 -0.62 -0. -0. ]
 [-0.08 -0.35 0.46 0.57 0.56 0.13]]
```

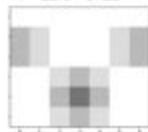
```
[[4.74 0. 0. 0. 0. 0. ]
 [0. 1.41 0. 0. 0. 0. ]
 [0. 0. 1.41 0. 0. 0. ]
 [0. 0. 0. 0.73 0. 0. ]
 [0. 0. 0. 0. 0. 0. ]
 [0. 0. 0. 0. 0. 0. ]]
```

```
[[-0.23 -0.4 -0.46 -0.4 -0.46 -0.4 -0.23]
 [0.5 0.25 -0.25 -0.5 -0.25 0.25 0.5 ]
 [0.39 -0.32 -0.19 0.65 -0.19 -0.32 0.39]
 [-0.22 0.42 -0.44 0.42 -0.44 0.42 -0.22]
 [0.56 -0.43 0.03 0. -0.03 0.43 -0.56]
 [-0.42 -0.55 -0.16 0. 0.16 0.55 0.42]
 [-0.12 -0.11 0.69 -0. -0.69 0.11 0.12]]
```

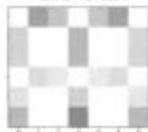
4.74



1.41



1.41



# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)



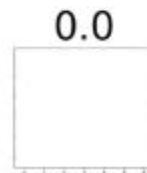
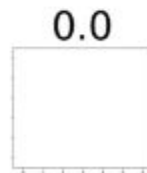
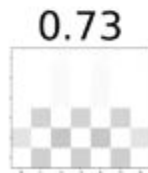
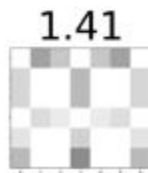
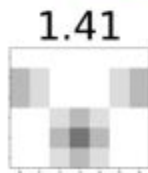
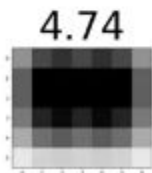
[www.github.com/luisguiserrano/singular\\_value\\_decomposition](https://www.github.com/luisguiserrano/singular_value_decomposition)

0	1	1	0	1	1	0
1	1	1	1	1	1	1
1	1	1	1	1	1	1
0	1	1	1	1	1	0
0	0	1	1	1	0	0
0	0	0	1	0	0	0

```
[[-0.36 -0. -0.73 -0.05 0.56 0.13]
 [-0.54 0.35 0.27 -0.08 -0.16 0.69]
 [-0.54 0.35 0.27 -0.08 0.16 -0.69]
 [-0.45 -0.35 -0.27 0.52 -0.56 -0.13]
 [-0.28 -0.71 0.18 -0.62 -0. -0.]
 [-0.08 -0.35 0.46 0.57 0.56 0.13]]
```

```
[[4.74 0. 0. 0. 0. 0.]
 [0. 1.41 0. 0. 0. 0.]
 [0. 0. 1.41 0. 0. 0.]
 [0. 0. 0. 0.73 0. 0.]
 [0. 0. 0. 0. 0. 0.]
 [0. 0. 0. 0. 0. 0.]]
```

```
[[-0.23 -0.4 -0.46 -0.4 -0.46 -0.4 -0.23]
 [ 0.5 0.25 -0.25 -0.5 -0.25 0.25 0.5 ]
 [ 0.39 -0.32 -0.19 0.65 -0.19 -0.32 0.39]
 [-0.22 0.42 -0.44 0.42 -0.44 0.42 -0.22]
 [ 0.56 -0.43 0.03 0. -0.03 0.43 -0.56]
 [-0.42 -0.55 -0.16 0. 0.16 0.55 0.42]
 [-0.12 -0.11 0.69 -0. -0.69 0.11 0.12]]
```



# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)



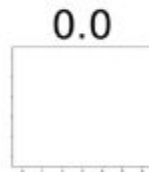
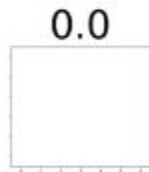
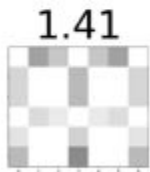
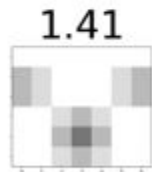
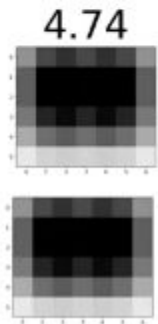
[www.github.com/luisguiserrano/singular\\_value\\_decomposition](https://www.github.com/luisguiserrano/singular_value_decomposition)

0	1	1	0	1	1	0
1	1	1	1	1	1	1
1	1	1	1	1	1	1
0	1	1	1	1	1	0
0	0	1	1	1	0	0
0	0	0	1	0	0	0

```
[[-0.36  -0.   -0.73 -0.05  0.56  0.13]
 [-0.54  0.35  0.27 -0.08 -0.16  0.69]
 [-0.54  0.35  0.27 -0.08  0.16 -0.69]
 [-0.45 -0.35 -0.27  0.52 -0.56 -0.13]
 [-0.28 -0.71  0.18 -0.62 -0.   -0. ]
 [-0.08 -0.35  0.46  0.57  0.56  0.13]]
```

```
[[4.74 0.   0.   0.   0.   0.   0. ]
 [0.  1.41 0.   0.   0.   0.   0. ]
 [0.   0.  1.41 0.   0.   0.   0. ]
 [0.   0.   0.  0.73 0.   0.   0. ]
 [0.   0.   0.   0.   0.   0.   0. ]
 [0.   0.   0.   0.   0.   0.   0. ]]
```

```
[[-0.23 -0.4 -0.46 -0.4 -0.46 -0.4 -0.23]
 [ 0.5  0.25 -0.25 -0.5 -0.25  0.25  0.5 ]
 [ 0.39 -0.32 -0.19  0.65 -0.19 -0.32  0.39]
 [-0.22  0.42 -0.44  0.42 -0.44  0.42 -0.22]
 [ 0.56 -0.43  0.03  0.  -0.03  0.43 -0.56]
 [-0.42 -0.55 -0.16  0.  0.16  0.55  0.42]
 [-0.12 -0.11  0.69 -0.  -0.69  0.11  0.12]]
```



# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)



[www.github.com/luiguiserrano/singular\\_value\\_decomposition](https://www.github.com/luiguiserrano/singular_value_decomposition)

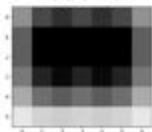
0	1	1	0	1	1	0
1	1	1	1	1	1	1
1	1	1	1	1	1	1
0	1	1	1	1	1	0
0	0	1	1	1	0	0
0	0	0	1	0	0	0

```
[[-0.36  -0.    -0.73 -0.05  0.56  0.13]
 [-0.54  0.35  0.27 -0.08 -0.16  0.69]
 [-0.54  0.35  0.27 -0.08  0.16 -0.69]
 [-0.45 -0.35 -0.27  0.52 -0.56 -0.13]
 [-0.28 -0.71  0.18 -0.62 -0.    -0. ]
 [-0.08 -0.35  0.46  0.57  0.56  0.13]]
```

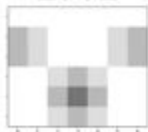
```
[[4.74 0.    0.    0.    0.    0. ]
 [0.  1.41 0.    0.    0.    0. ]
 [0.    0.  1.41 0.    0.    0. ]
 [0.    0.    0.  0.73 0.    0. ]
 [0.    0.    0.    0.    0.    0. ]
 [0.    0.    0.    0.    0.    0. ]]
```

```
[[-0.23 -0.4  -0.46 -0.4  -0.46 -0.4  -0.23]
 [ 0.5  0.25 -0.25 -0.5  -0.25  0.25  0.5 ]
 [ 0.39 -0.32 -0.19  0.65 -0.19 -0.32  0.39]
 [-0.22  0.42 -0.44  0.42 -0.44  0.42 -0.22]
 [ 0.56 -0.43  0.03  0.   -0.03  0.43 -0.56]
 [-0.42 -0.55 -0.16  0.   0.16  0.55  0.42]
 [-0.12 -0.11  0.69 -0.   -0.69  0.11  0.12]]
```

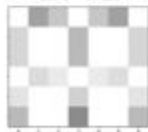
4.74



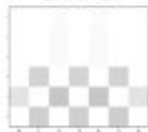
1.41



1.41



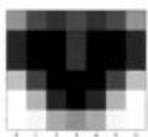
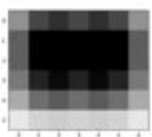
0.73



0.0



0.0



# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)



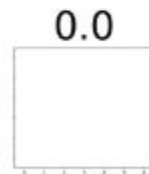
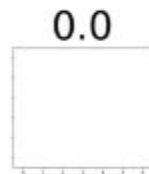
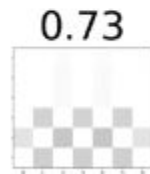
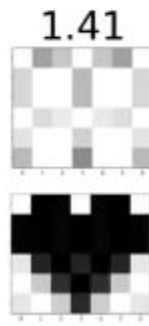
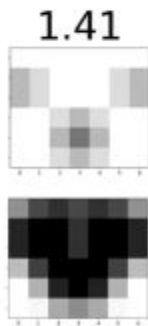
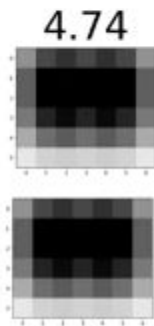
[www.github.com/luisguiserrano/singular\\_value\\_decomposition](https://www.github.com/luisguiserrano/singular_value_decomposition)

0	1	1	0	1	1	0
1	1	1	1	1	1	1
1	1	1	1	1	1	1
0	1	1	1	1	1	0
0	0	1	1	1	0	0
0	0	0	1	0	0	0

```
[[-0.36 -0.    -0.73 -0.05  0.56  0.13]
 [-0.54  0.35  0.27 -0.08 -0.16  0.69]
 [-0.54  0.35  0.27 -0.08  0.16 -0.69]
 [-0.45 -0.35 -0.27  0.52 -0.56 -0.13]
 [-0.28 -0.71  0.18 -0.62 -0.    -0.   ]
 [-0.08 -0.35  0.46  0.57  0.56  0.13]]
```

```
[[4.74 0.    0.    0.    0.    0.   ]
 [0.  1.41 0.    0.    0.    0.   ]
 [0.    0.  1.41 0.    0.    0.   ]
 [0.    0.    0.  0.73 0.    0.   ]
 [0.    0.    0.    0.    0.    0.   ]
 [0.    0.    0.    0.    0.    0.   ]]
```

```
[[-0.23 -0.4  -0.46 -0.4  -0.46 -0.4  -0.23]
 [ 0.5  0.25 -0.25 -0.5  -0.25  0.25  0.5 ]
 [ 0.39 -0.32 -0.19  0.65 -0.19 -0.32  0.39]
 [-0.22  0.42 -0.44  0.42 -0.44  0.42 -0.22]
 [ 0.56 -0.43  0.03  0.   -0.03  0.43 -0.56]
 [-0.42 -0.55 -0.16  0.   0.16  0.55  0.42]
 [-0.12 -0.11  0.69 -0.   -0.69  0.11  0.12]]
```



# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)



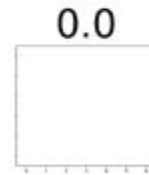
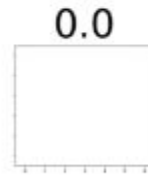
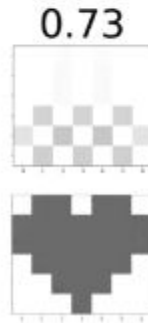
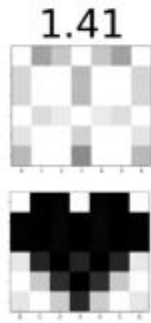
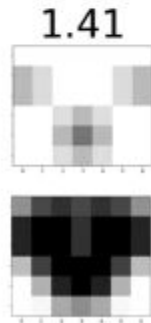
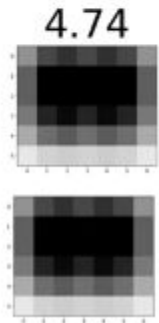
[www.github.com/luisguiserrano/singular\\_value\\_decomposition](https://www.github.com/luisguiserrano/singular_value_decomposition)

0	1	1	0	1	1	0
1	1	1	1	1	1	1
1	1	1	1	1	1	1
0	1	1	1	1	1	0
0	0	1	1	1	0	0
0	0	0	1	0	0	0

```
[[-0.36 -0.   -0.73 -0.05  0.56  0.13]  
 [-0.54  0.35  0.27 -0.08 -0.16  0.69]  
 [-0.54  0.35  0.27 -0.08  0.16 -0.69]  
 [-0.45 -0.35 -0.27  0.52 -0.56 -0.13]  
 [-0.28 -0.71  0.18 -0.62 -0.   -0. ]  
 [-0.08 -0.35  0.46  0.57  0.56  0.13]]
```

```
[[4.74 0.   0.   0.   0.   0.   0. ]  
 [0.   1.41 0.   0.   0.   0.   0. ]  
 [0.   0.   1.41 0.   0.   0.   0. ]  
 [0.   0.   0.   0.73 0.   0.   0. ]  
 [0.   0.   0.   0.   0.   0.   0. ]  
 [0.   0.   0.   0.   0.   0.   0. ]
```

```
[[-0.23 -0.4  -0.46 -0.4  -0.46 -0.4  -0.23]  
 [ 0.5  0.25 -0.25 -0.5  -0.25  0.25  0.5 ]  
 [ 0.39 -0.32 -0.19  0.65 -0.19 -0.32  0.39]  
 [-0.22  0.42 -0.44  0.42 -0.44  0.42 -0.22]  
 [ 0.56 -0.43  0.03  0.   -0.03  0.43 -0.56]  
 [-0.42 -0.55 -0.16  0.   0.16  0.55  0.42]  
 [-0.12 -0.11  0.69 -0.   -0.69  0.11  0.12]]
```





# Ανάλυση σε Ιδιάζουσες Τιμές (Singular Value Decomposition)



[www.github.com/luisguiserrano/singular\\_value\\_decomposition](https://www.github.com/luisguiserrano/singular_value_decomposition)

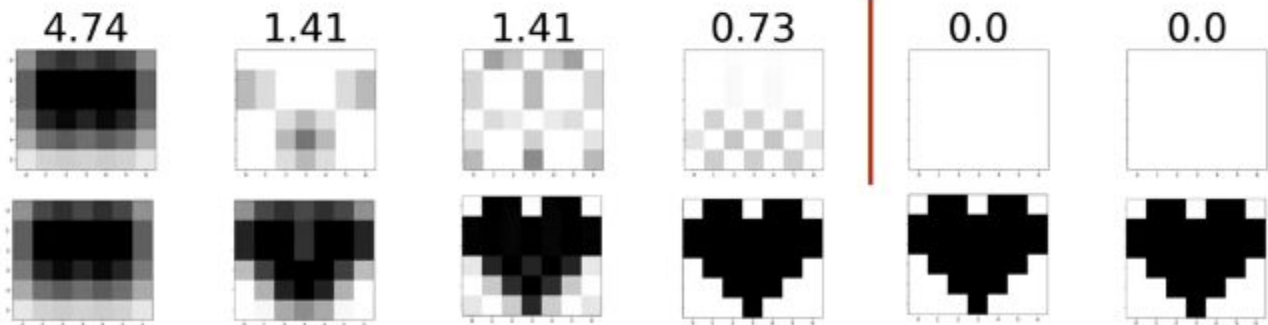
0	1	1	0	1	1	0
1	1	1	1	1	1	1
1	1	1	1	1	1	1
0	1	1	1	1	1	0
0	0	1	1	1	0	0
0	0	0	1	0	0	0

Rank 4

```
[[-0.36 -0.    -0.73 -0.05  0.56  0.13]
 [-0.54  0.35  0.27 -0.08 -0.16  0.69]
 [-0.54  0.35  0.27 -0.08  0.16 -0.69]
 [-0.45 -0.35 -0.27  0.52 -0.56 -0.13]
 [-0.28 -0.71  0.18 -0.62 -0.    -0. ]
 [-0.08 -0.35  0.46  0.57  0.56  0.13]]
```

```
[[4.74 0.    0.    0.    0.    0.    0. ]
 [0.   1.41 0.    0.    0.    0.    0. ]
 [0.   0.   1.41 0.    0.    0.    0. ]
 [0.   0.   0.   0.73 0.    0.    0. ]
 [0.   0.   0.   0.   0.    0.    0. ]
 [0.   0.   0.   0.   0.    0.    0. ]]
```

```
[[-0.23 -0.4  -0.46 -0.4  -0.46 -0.4  -0.23]
 [ 0.5  0.25 -0.25 -0.5  -0.25  0.25  0.5 ]
 [ 0.39 -0.32 -0.19  0.65 -0.19 -0.32  0.39]
 [-0.22  0.42 -0.44  0.42 -0.44  0.42 -0.22]
 [ 0.56 -0.43  0.03  0.   -0.03  0.43 -0.56]
 [-0.42 -0.55 -0.16  0.   0.16  0.55  0.42]
 [-0.12 -0.11  0.69 -0.   -0.69  0.11  0.12]]
```





# Moore-Penrose Pseudoinverse

Η αντιστροφή πίνακα δεν μπορεί να ορισθεί για μη τετραγωνικούς πίνακες.

Λύση: ψευδοαντίστροφος Moore-Penrose

U,D,V: SVD του A

$$A_+ = VD^+U$$

D<sup>+</sup>: Ο ψευδο-αντίστροφος ενός διαγώνιου πίνακα D

→ λαμβάνει το αντίστροφο των μη μηδενικών στοιχείων του, και έπειτα παίρνει τον ανάστροφο αυτού.

A<sub>m×n</sub> - n>m : Μία ή περισσότερες πιθανές λύσεις με χρήση του A<sup>+</sup>

(x=A<sup>+</sup>y με την μικρότερη ||x||<sup>2</sup> ανάμεσα σε όλες τις πιθανές λύσεις)

m>n : Μπορεί να μην υπάρχει λύση

Ο A<sup>+</sup> δίνει τα x για τα οποία το Ax→ y σύμφωνα με τη νόρμα ||Ax - y||<sup>2</sup>

# Βιβλιογραφία

Strang Gilbert

- Linear Algebra and Learning from Data (math.mit.edu/learningfromdata)
- Introduction to Linear Algebra - Fifth Edition
- [MIT 18.065 Matrix Methods in Data Analysis, Signal Processing, and Machine Learning](#)

Marc Peter Deisenroth, A. Aldo Faisal, and Cheng Soon Ong

- [Mathematics for Machine Learning - Book](#)
- [Mathematics for Machine Learning -github](#)

Ian Goodfellow, Yoshua Bengio, Aaron Courville

- Deep Learning, [Chapter 2: Linear Algebra](#)

[2.3. Linear Algebra — Dive into Deep Learning 0.16.1 documentation](#)