DIAMOND SPONSORS

PLATINUM SPONSORS

GOLD SPONSORS

SILVER SPONSORS

ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ & ΗΛΕΚΤΡΟΝΙΚΩΝ ΜΗΧΑΝΙΚΩΝ

ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ

Διαφάνειες διαλέξεων

Υπολογιστική Νοημοσύνη

Βαθιά μάθηση

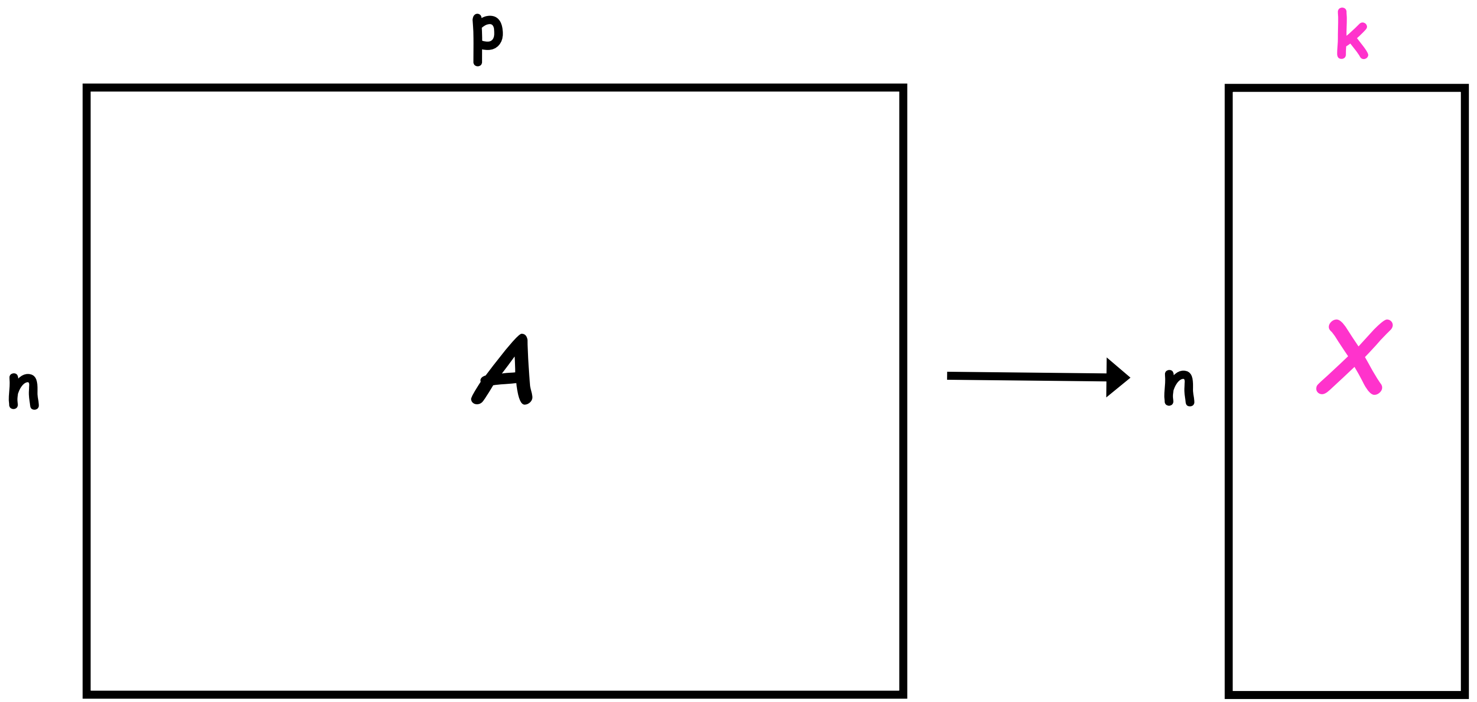
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ & ΗΛΕΚΤΡΟΝΙΚΩΝ ΜΗΧΑΝΙΚΩΝ
ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ

Στοιχεία Μηχανικής Μάθησης

Ανάλυση κυρίων συνιστωσών
(Principal Component Analysis)

Ανάλυση κυρίων συνιστωσών
(Principal Component Analysis)

- Περιγραφή φυσικών δεδομένων (data) (p-διαστάσεων) από ένα υποσύνολο (k) νέων (συνθετικών) διαστάσεων.



Εισαγωγή

- Ανακαλύφθηκε από τον Karl Pearson (1857-1936) το 1901.

- Αναπτύχθηκε ανεξάρτητα από τον Harold Hotelling (1895-1973) το 1930.

Εισαγωγή

- Στην επεξεργασία σήματος ονομάζεται διακριτός Karhunen–Loève μετασχηματισμός (KLT).
  - **Kari Karhunen** (1915–1992)
  - **Michel Loève** (1907–1979)

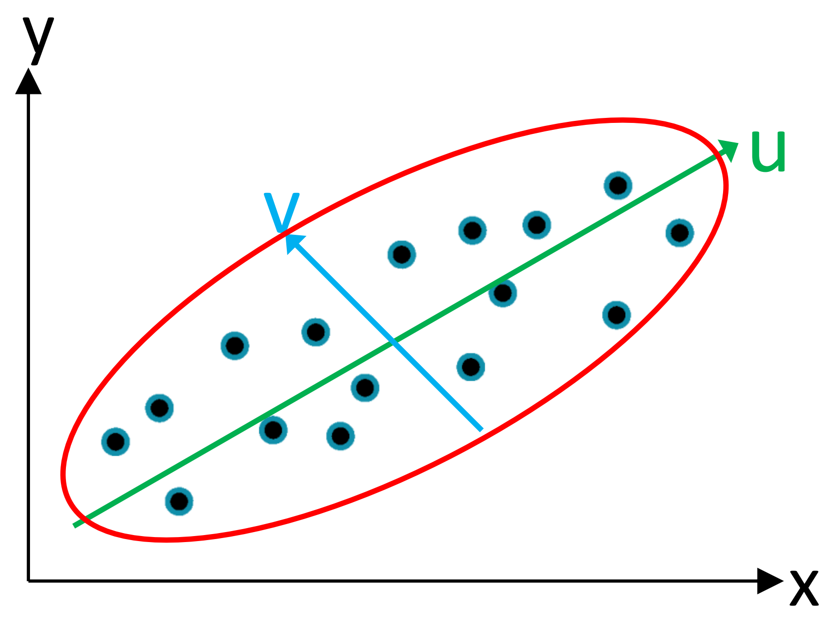
- Άλλα ονόματα:
  - Singular value decomposition (SVD)
  - Eigendecomposition of a matrix
  - Factor Analysis
  - Και άλλα πολλά…

Τι είναι η PCA;

- Η ανάλυση κυρίων συνιστωσών Principal Component Analysis (PCA) είναι μια στατιστική διαδικασία με την οποία αναπαριστούμε ένα πίνακα συνδιακύμανσης ενός συνόλου «αρχικών» μεταβλητών μέσα από ένα διαφορετικό (και συνήθως μικρότερο) σύνολο «νέων» μεταβλητών οι οποίες προκύπτουν από τον γραμμικό συνδυασμό των «αρχικών» μεταβλητών.

Η PCA στο απλό…

- Για ένα γνωστό σύνολο δεδομένων {x,y}:
- Καθορίζουμε την διεύθυνση στην οποία έχω την μεγαλύτερη διασπορά (u).
- Μετά βρίσκω την επόμενη διεύθυνση στην οποία έχω την αμέσως μεγαλύτερη διασπορά (v).
- PCA = {u,v}
- Έτσι βρίσκουμε το νέο σύστημα συντεταγμένων που αναπαριστά με πιο συμπαγή τρόπο τα «αρχικά» δεδομένα



Η PCA στο απλό…

- **Αλγεβρικά**: η PCA αποτελείται από ιδιαίτερους γραμμικούς συνδυασμούς των «αρχικών» μεταβλητών
  - «Ιδιαίτερους» Σχηματίζουν μια προβολή των «αρχικών» μεταβλητών που αντιπροσωπεύει καλύτερα τα δεδομένα ως προς το μέσο τετραγωνικό σφάλμα.

- **Γεωμετρικά**: η PCA μέσα από τους γραμμικούς συνδυασμούς των «αρχικών» μεταβλητών δημιουργούν ένα νέο σύστημα συντεταγμένων που δημιουργείται μετατοπίζοντας και περιστρέφοντας το παλαιό σύστημα. Οι νέοι άξονες καθορίζουν τις κατευθύνσεις που παρουσιάζουν την/τις μέγιστη/ες μεταβολές των δεδομένων.

Η PCA στο απλό…

- Υπολογιστικά: Οι κύριες συνιστώσες της PCA βρίσκονται από τον υπολογισμό των ιδιοδιανυσμάτων και των αντιστοίχων ιδιοτιμών του πίνακα συνδιακύμανσης, όπως αυτός υπολογίζεται από τα δεδομένα που υπάρχουν.
  - Το ιδιοδιάνυσμα με την μεγαλύτερη ιδιοτιμή είναι η κατεύθυνση της μεγαλύτερης διασποράς.

Ιδιοδιανύσματα – Ιδιοτιμές (Eigenvectors, Eigenvalues)

- Δεδομένου ενός πίνακα: **M** $\in \mathbb{R}^{d\times d}$ μία πολύ σημαντική κατηγορία γραμμικών εξισώσεων είναι της μορφής:

$$\mathbf{M}\cdot\mathbf{x}=\lambda\mathbf{x}, \text{ με } \lambda\in\mathbb{R}^{1}, \boldsymbol{x}\in\mathbb{R}^{d\times 1}$$

- Η παραπάνω εξίσωση γράφεται:

$$(\mathbf{M}-\lambda\mathbf{I})\mathbf{x}=\mathbf{0}, \text{ όπου } \{\mathbf{I},\mathbf{0}\}\in\mathbb{R}^{d\times d}$$

- $\mathbf{I}$ ο μοναδιαίος πίνακας
- $\mathbf{0}$ ο μηδενικός πίνακας
- $(\mathbf{M}-\lambda\mathbf{I})\mathbf{x}=\mathbf{0}$, χαρακτηριστικό πολυώνυμο

Ιδιοδιανύσματα – Ιδιοτιμές (Eigenvectors, Eigenvalues)

- Μια λύση $\boldsymbol{x} = \boldsymbol{e}_i \in \mathbb{R}^{d\times 1}$ της $(\mathbf{M} - \lambda_\mathrm{i}\mathbf{I})\boldsymbol{e}_i = \mathbf{0}$ και η αντίστοιχη τιμή λ_i καλούνται ιδιοδιάνυσμα και ιδιοτιμή αντίστοιχα.
- Συνολικά υπάρχουν d-(πιθανώς μη διαφορετικά) ιδιοδιανύσματα λύσης $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots \boldsymbol{e}_d\}$.
- Κάθε $\boldsymbol{e}_i$ συνοδεύεται από την αντίστοιχη ιδιοτιμή τους $\{\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \ldots \boldsymbol{\lambda}_d\}$
- Συνεπώς ισχύει: $\mathbf{M} \cdot \boldsymbol{e}_j = \lambda_\mathrm{j}\boldsymbol{e}_j$
- Ιδιότητα: det(M)=$|M| = \prod_{i=1}^{d} \boldsymbol{\lambda_i}$
- Ιδιότητα: $\left\|\boldsymbol{e}_j\right\|^2 = 1, (\boldsymbol{e}_j)^T(\boldsymbol{e}_j)$=1. Τα $\boldsymbol{e}_j$ είναι μοναδιαία. Αν δεν προκύψουν τα κανονικοποιώ.

Ιδιοδιανύσματα – Ιδιοτιμές

- Παράδειγμα:

$$M = \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix}$$

$$e_1 = \begin{bmatrix} 1/\sqrt{5} \\ 2\sqrt{5} \end{bmatrix} \quad \text{και} \quad \lambda_1 = 7$$

όντως: αν $M \cdot e_1 = \lambda_1 \cdot e_1$

$$\begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix} \cdot \begin{bmatrix} 1\sqrt{5} \\ 2\sqrt{5} \end{bmatrix} = 7 \begin{bmatrix} 1/\sqrt{5} \\ 2 \cdot \sqrt{5} \end{bmatrix}$$

επίσης: $(1/\sqrt{5})^2 + (2\sqrt{5})^2 = 1$ (ΜΟΝΑΔΙΑΙΑ).

Εύρεση ιδιοτιμών και ιδιοδιανυσμάτων

- Υπάρχουν 2 κύριες μέθοδοι:
  - $O(n^3)$ running time method (with pivotal condensation)
    - Υποστηρίζει την εύρεση λύσης για συμμετρικούς πίνακες, όπως ο πίνακας της συνδιακύμανσης Σ.
    - Η λύση είναι ακριβής χωρίς σφάλματα.
  - *Power iteration method*

$O(n^3)$ running time method

- $(\mathbf{M} - \lambda\mathbf{I})\mathbf{e} = \mathbf{0}$
  - Θα πρέπει να ισχύει: det $(\mathbf{M} - \lambda\mathbf{I})$=0
  - Η ανωτέρω ορίζουσα είναι ένα πολυώνυμο d- τάξης ως προς λ και έχει υπολογιστική πολυπλοκότητα $O(n^3)$.
  - Λαμβάνουμε d- τιμές για το λ.
  - Για κάθε τιμή λύνουμε την $\mathbf{M}\mathbf{e} = \lambda\mathbf{e}$ και υπολογίζουμε το $\mathbf{e}$.
  - Κανονικοποιούμε το $\mathbf{e}$ ώστε να έχει μέτρο 1.

$O(n^3)$ running time method

- Παράδειγμα:

Έστω ο πίνακας: $M \in \mathbb{R}^{2\times 2}$, $M = \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix}$

(ο πίνακας είναι συμμετρικός)

Τότε $M-\lambda I$: $\begin{bmatrix} 3-\lambda & 2 \\ 2 & 6-\lambda \end{bmatrix}$ και η ορίζουσα: $(3-\lambda)\cdot(6-\lambda)-4$

έχω $\det(M-\lambda I)=0 \Rightarrow \lambda^2-9\lambda+14=0$

$\Delta = 81-4\cdot 14 = 81-56 = 25$, $\sqrt{\Delta}=\sqrt{25}=5$.

$\lambda_{1,2} = \frac{-(-9)\pm 5}{2}$

$\lambda_1 = 7 \leftarrow$ 1η κύρια

$\lambda_2 = 2 \leftarrow$ 2η κύρια

Θεωρώ $e=\begin{bmatrix} x \\ y \end{bmatrix}$, έτσι πρέπει να λύσω:

$$\begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 7 \begin{bmatrix} x \\ y \end{bmatrix} \Rightarrow$$

$$\begin{matrix} 3x+2y=7x & (1) & \Rightarrow & 2y=4x \Rightarrow y=2x \\ 2x+6y=7y & (2) & & 2x=y \Rightarrow y=2x \end{matrix} \quad !$$

πιθανό ιδιοδιάνυσμα $e=\begin{bmatrix} x \\ y \end{bmatrix}=\begin{bmatrix} x \\ 2x \end{bmatrix} \sim \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

κανονικοποιώ το $e \rightsquigarrow \hat{e}$ ή (e)

$$\|e\|=\sqrt{1^2+2^2}=\sqrt{5} \quad \text{και τελικά:}$$

$$\boxed{e_1=\begin{bmatrix} 1/\sqrt{5} \\ 2\sqrt{5} \end{bmatrix} \text{ και } \lambda_1=7}$$

ΥΠΟΛΟΓΙΣΜΟΣ $2^{ης}$ ιδιοτιμής / ιδιοδιανύσματος.

$\lambda_2 = 2$, $e = \begin{bmatrix} x \\ y \end{bmatrix}$ άρα: $\begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = 2 \begin{bmatrix} x \\ y \end{bmatrix}$

$$\Rightarrow \left.\begin{array}{l} 3x + 2y = 2x \Rightarrow 2y = -x \\ 2x + 6y = 2y \Rightarrow 4y = -2x \end{array}\right\} \boxed{x = -2y}$$

Άρα: $e = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -2y \\ y \end{bmatrix}$ ή $\begin{bmatrix} x \\ -x/2 \end{bmatrix}$ έτσι ώστε να είναι θετικό

άρα: $e = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ και $\|e\| = \sqrt{2^2 + 1^2} = \sqrt{5}$

τελικά $\boxed{e_2 = \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix} \text{ με } \lambda_2 = 2}$

Power iteration method – 1ο ζεύγος τιμών

- Έστω M ο πίνακας του οποίου τις ιδιοτιμές και τα ιδιοδιανύσματα αναζητάμε:
- Υπολογισμός της πρώτης ιδιοτιμής και του αντιστοίχου ιδιοδιανύσματος:
  - Για ένα οποιοδήποτε $\mathbf{x}_0$ υπολογίσατε επαναληπτικά: $\boldsymbol{x}_{k+1} = \frac{M\boldsymbol{x}_k}{\|M\boldsymbol{x}_k\|}$, με $\|M\boldsymbol{x}_k\| = \|\mathrm{N}\|$ η νόρμα Frobenius:
  - $$\|\mathrm{N}\| = \sqrt{\sum_{i=1}^{\#\sigma\tau o\iota\chi\varepsilon\iota\omega\nu \mathrm{N}} (N_i)^2}$$
  - Τερματισμός όταν: $\|\boldsymbol{x}_k - \boldsymbol{x}_{k+1}\| < \varepsilon$.
  - Ανέθεσε στο αναζητούμενο ιδιοδιάνυσμα το $\boldsymbol{x} = \boldsymbol{x}_k$
  - Εφόσον ισχύει $\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{x} = \mathbf{1}$, έχουμε: $\mathbf{Mx} = \lambda\mathbf{x}$ και:
    $$\boldsymbol{x}^{\boldsymbol{T}}\mathbf{Mx} = \boldsymbol{x}^{\boldsymbol{T}}\lambda\mathbf{x} = \lambda\boldsymbol{x}^{\boldsymbol{T}}\mathbf{x} = \boldsymbol{\lambda}$$

Power iteration method – 1ο ζεύγος τιμών

- Παράδειγμα:

έστω $M = \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix}$ και επιλέγω για $x_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

τότε: $x_1 = M \cdot x = \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 8 \end{bmatrix}$

η νόρμα Frobenius: $\sqrt{5^2+8^2} = \sqrt{89} = 9.434$. και x_1:

$x_1 = \begin{bmatrix} 5/\sqrt{89} \\ 8/\sqrt{89} \end{bmatrix} = \begin{bmatrix} 0.530 \\ 0.848 \end{bmatrix}$. Ομοίως $x_2 = M \cdot x_1 =$

$\begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix} \cdot \begin{bmatrix} 0,530 \\ 0.848 \end{bmatrix} = \begin{bmatrix} 3.286 \\ 6.148 \end{bmatrix}$. η νόρμα $F = 6.971$

$\Rightarrow x_2 = \begin{bmatrix} 0.471 \\ 0.882 \end{bmatrix}$ συγκλίνουμε πρός ένα διάνυσμα του οποίου ο 2ος συντελεστής είναι ο διπλάσιος του πρώτου.

και μετά από μερικές πράξεις: $x = \begin{bmatrix} 0.447 \\ 0.894 \end{bmatrix}$

και η αντίστοιχη λ:

$$\lambda = x^T M x = \begin{bmatrix} 0.447 & 0.894 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} 0.447 \\ 0.894 \end{bmatrix}$$

$$= \underline{6.993}$$

κανονικά είναι 7.
η μέθοδος βάζει ένα
σφάλμα...

Power iteration method
Επόμενο ζεύγος τιμών

- Δημιουργία του νέου πίνακα:

$$\mathrm{M}^* = \mathrm{M} - \lambda \mathbf{x}\mathbf{x}^{\boldsymbol{T}}$$

- Επανάληψη της προηγούμενης διαδικασίας για τον υπολογισμό του επόμενου ζεύγους ιδιοδιανύσματος – ιδιοτιμής με βάση τον πίνακα M^*.
- Ιδιότητες: το ιδιοδιάνυσμα $\mathbf{x}$ είναι ιδιοδιάνυσμα του M^* με ιδιοτιμή 0.

$$M^*\mathbf{x} = (M - \lambda\mathbf{x}\mathbf{x}^{\mathrm{T}})\mathbf{x} = M\mathbf{x} - \lambda\mathbf{x}\mathbf{x}^{\mathrm{T}}\mathbf{x} = M\mathbf{x} - \lambda\mathbf{x} = 0$$

Power iteration method
Επόμενο ζεύγος τιμών

- Ιδιότητες: Ομοίως αν **v**, λ_v είναι ένα eigenpair του συμμετρικού πίνακα M (εκτός του πρώτου ζεύγους), τότε είναι eigenpairs του M^*.

$$
\begin{aligned}
M^*\mathbf{v} = (M^*)^{\mathrm{T}}\mathbf{v} &= (M - \lambda \mathbf{x}\mathbf{x}^{\mathrm{T}})^{\mathrm{T}}\mathbf{v} \\
&= M^{\mathrm{T}}\mathbf{v} - \lambda \mathbf{x}(\mathbf{x}^{\mathrm{T}}\mathbf{v}) = M^{\mathrm{T}}\mathbf{v} = \lambda_v \mathbf{v}
\end{aligned}
$$

Power iteration method
Επόμενο ζεύγος τιμών

- Συνέχεια:

$$M^* = M - \lambda \mathbf{x}\mathbf{x}^T =$$

$$\begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix} - 6.993 \begin{bmatrix} 0.447 \\ 0.894 \end{bmatrix} \begin{bmatrix} 0.447 & 0.894 \end{bmatrix} =$$

$$= \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix} - \begin{bmatrix} 1.397 & 2.795 \\ 2.795 & 5.589 \end{bmatrix} = \begin{bmatrix} 1.603 & -0.795 \\ -0.795 & 0.411 \end{bmatrix}$$

Επανάληψη της προηγούμενης διαδικασίας για τον υπολογισμό του επόμενου ζεύγους ιδιοδιανύσματος – ιδιοτιμής με βάση τον πίνακα M^*.

Ο πίνακας των ιδιοδιανυσμάτων

- Δεδομένου ενός πίνακα: $\mathbf{M} \in \mathbb{R}^{d\times d}$ υπολογίζουμε τα ιδιοδιανύσματα του:

$$\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots \boldsymbol{e}_d$$

- Έστω **E** ο πίνακας, η i-στήλη του οποίου είναι το $\boldsymbol{e}_i$ ιδιοδιάνυσμα:

$$\mathbf{E} = \begin{pmatrix} \cdots & \uparrow & \cdots \\ \vdots & \boldsymbol{e}_i & \vdots \\ \cdots & \downarrow & \cdots \end{pmatrix}$$

- Τότε: $\mathbf{E}\mathbf{E}^{\mathrm{T}} = \mathbf{E}^{\mathrm{T}}\mathbf{E} = \mathbf{I}$ (τα ιδιοδιανύσματα είναι ορθοκανονικά)

Παράδειγμα

$$\mathbf{E}= \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$$

$$\mathbf{E}^T= \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$$

$$\mathbf{E}\mathbf{E}^T= \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$$

$$\begin{bmatrix} 4/5+1/5 & -2/5+2/5 \\ -2/5+2/5 & 1/5+4/5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Ο πίνακας των ιδιοδιανυσμάτων

- Δεδομένου του πίνακα των ιδιοδιανυσμάτων **E**, ενός συμμετρικού πίνακα **A**$\in \mathbb{R}^{3\times 3}$:

$$\mathbf{A}\mathbf{e}_1 = \lambda_1\mathbf{e}_1, \mathbf{A}\mathbf{e}_2 = \lambda_2\mathbf{e}_2, \mathbf{A}\mathbf{e}_3 = \lambda_3\mathbf{e}_3$$

$$\mathbf{A}\cdot[\mathbf{e}_1\ \mathbf{e}_2\ \mathbf{e}_3] = [\mathbf{e}_1\ \mathbf{e}_2\ \mathbf{e}_3]\cdot\begin{bmatrix}\lambda_1 & 0 & 0\\ 0 & \lambda_2 & 0\\ 0 & 0 & \lambda_3\end{bmatrix}$$

ή $$\mathbf{A}\cdot\mathbf{E} = \mathbf{E}\cdot\mathbf{\Lambda}$$

και: $$\mathbf{A} = \mathbf{E}\cdot\mathbf{\Lambda}\cdot\mathbf{E}^{\mathbf{T}}$$

Ο πίνακας των ιδιοδιανυσμάτων

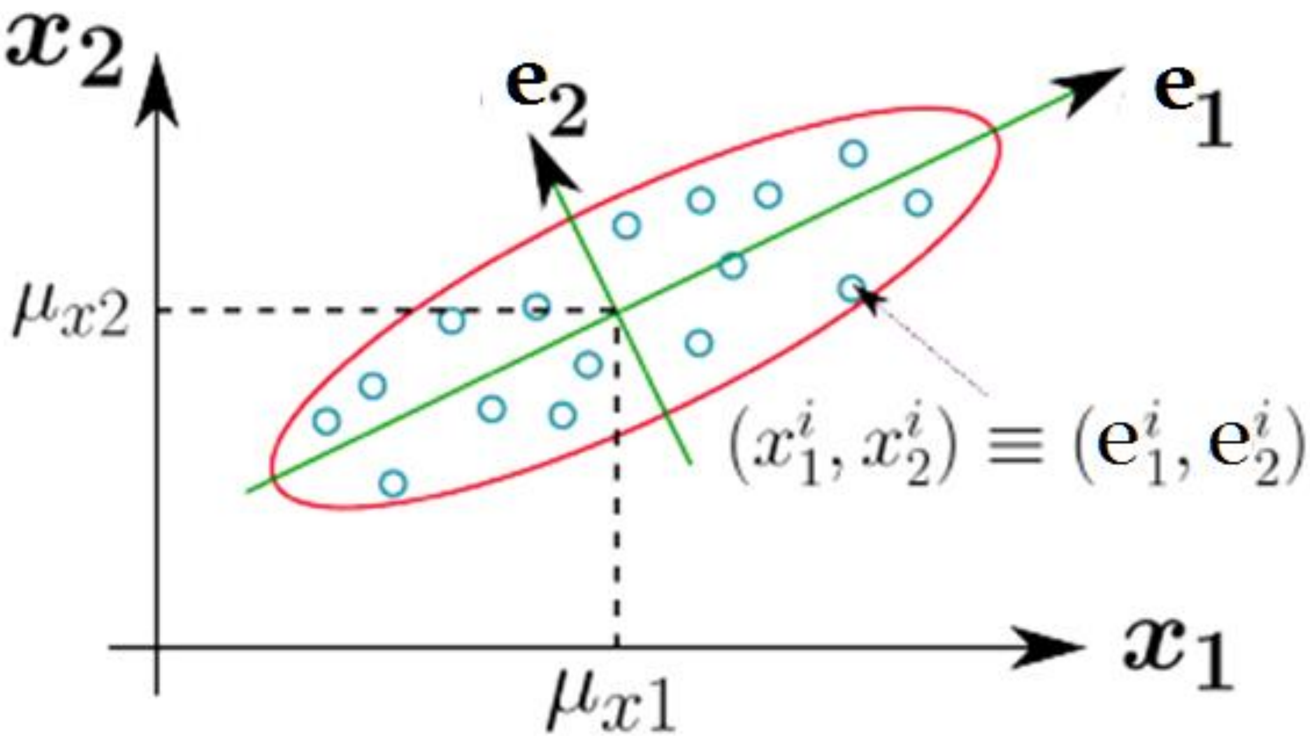
- Αντί για οποιοδήποτε **Α**, ας θεωρήσουμε τον συμμετρικό πίνακα συνδιακύμανσης **Σ**$\in \mathbb{R}^{d\times \boldsymbol{d}}$ ενός πληθυσμού. Τότε υπάρχει α) ένας πίνακας **Ε** (των ιδιοδιανυσμάτων του) και β) ένας διαγώνιος πίνακας Λ των ιδιοτιμών του έτσι ώστε:
$$\mathbf{E}^{\mathbf{T}} \cdot \mathbf{\Sigma} \cdot \mathbf{E} = \mathbf{\Lambda}$$
- Πρακτικά ο πίνακας **Ε** μετασχηματίζει γραμμικά (περιστρέφει και μετατοπίζει) τον πίνακα **Σ**.
- Οι νέες συντεταγμένες είναι ασυσχέτιστες (decorrelated).

Η PCA στην πράξη

- Έστω ένα συνόλου δεδομένων στο αρχικό σύστημα συντεταγμένων: $X = [x_1, x_2, \ldots, x_d]$.
  - Υπολογίσατε τις μέσες τιμές: $\boldsymbol{\mu}_{\mathbf{x}}=[\mu_{x1}, \mu_{x2}, \ldots, \mu_{xd}]$
  - Υπολογίσατε τον πίνακα συνδιακύμανσης $\boldsymbol{\Sigma}$.
  - Υπολογίσατε τα ιδιοδιανύσματα $\mathbf{E} = [\mathbf{e}_1, \ldots, \mathbf{e}_d]$
- Κάθε σημείο $\mathbf{p}_x$ στο αρχικό σύστημα συντεταγμένων μετασχηματίζεται στο νέο σύστημα συντεταγμένων $[\mathbf{e}_1, \ldots, \mathbf{e}_d]$ ως:

$$\mathbf{p}_{\mathbf{e}} = (\mathbf{p}_x - \boldsymbol{\mu}_{\mathbf{x}})\mathbf{E}$$

Η PCA στην πράξη



Η PCA πιο μαθηματικά…

- Έστω ότι θέλουμε να αναπαραστήσουμε ένα πληθυσμό από n το πλήθος δεδομένα – διανύσματα που βρίσκονται σε ένα χώρο d-διαστάσεων με ένα μόνο διάνυσμα.
- Δλδ, ψάχνουμε την καλύτερη αναπαράσταση των $\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_n \in \mathbb{R}^d$ με το διάνυσμα $\mathbf{x}_0 \in \mathbb{R}^d$.
- Ένας κλασσικός τρόπος είναι να ελαχιστοποιήσουμε το άθροισμα των τετραγωνικών αποστάσεων μεταξύ του $\mathbf{x}_0$ και των $\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_n$.
- Ορίζω την συνάρτηση: $J_0(\mathbf{x}_0) = \sum_{k=1}^{n} \|\mathbf{x}_0 - \mathbf{x}_k\|^2,$

Η PCA πιο μαθηματικά…

- Ισχυριζόμαστε ότι αυτό γίνεται αν $\mathbf{x}_0$=**m,** όπου **m** είναι το διάνυσμα της μέσης τιμής του πληθυσμού:

$$\mathbf{m} = \frac{1}{n} \sum_{k=1}^{n} \mathbf{x}_k .$$

- Ακολουθεί απόδειξη:

$$J_0(\mathbf{x}_0) = \sum_{k=1}^{n} ||(\mathbf{x}_0 - \mathbf{m}) - (\mathbf{x}_k - \mathbf{m})||^2$$

$$= \sum_{k=1}^{n} ||\mathbf{x}_0 - \mathbf{m}||^2 - 2\sum_{k=1}^{n} (\mathbf{x}_0 - \mathbf{m})^t(\mathbf{x}_k - \mathbf{m}) + \sum_{k=1}^{n} ||\mathbf{x}_k - \mathbf{m}||^2$$

$$= \sum_{k=1}^{n} ||\mathbf{x}_0 - \mathbf{m}||^2 - 2(\mathbf{x}_0 - \mathbf{m})^t \sum_{k=1}^{n} (\mathbf{x}_k - \mathbf{m}) + \sum_{k=1}^{n} ||\mathbf{x}_k - \mathbf{m}||^2$$

Ο όρος: $\sum_{k=1}^{n} (\mathbf{x}_k - \mathbf{m}) = \sum_{k=1}^{n} \mathbf{x}_k - \sum_{k=1}^{n} \mathbf{m} = n\mathbf{m} - n\mathbf{m} = 0$

$$= \sum_{k=1}^{n} ||\mathbf{x}_0 - \mathbf{m}||^2 + \underbrace{\sum_{k=1}^{n} ||\mathbf{x}_k - \mathbf{m}||^2}_{\text{Ανεξάρτητο του } \mathbf{x}_0}.$$

Η παραπάνω έκφραση ελαχιστοποιείται αν $\mathbf{x}_0 = \mathbf{m}$

Η PCA πιο μαθηματικά…

- Ο μέσος όρος είναι μια αναπαράσταση μηδενικής διάστασης των δεδομένων μας.
- Μπορούμε να λάβουμε μια πιο ενδιαφέρουσα μονοδιάστατη αναπαράσταση αν προβάλουμε τα δεδομένα σε μια γραμμή που περνάει από την μέση τιμή.
  - Έστω **e** ένα μοναδιαίο διάνυσμα στην επιθυμητή κατεύθυνση.
  - Τότε η εξίσωση της ευθείας θα είναι:

$$\mathbf{x} = \mathbf{m} + a\mathbf{e},$$

- Η τιμή *a* αντιστοιχεί στην απόσταση κάθε σημείου **x** από το κέντρο των δεδομένων **m**.
- Συνεπώς αναπαριστούμε το σημείο $\mathbf{x}_k$ με το $\mathbf{m}+a_k\mathbf{e}$
  - Η διαφορά τους: $(\mathbf{m}+a_k\mathbf{e}) - \mathbf{x}_k$
  - Το τετράγωνο του μέτρου της διαφοράς τους: $||(\mathbf{m}+a_k\mathbf{e}) - \mathbf{x}k||^2$
- Για να βρούμε την νέα αναπαράσταση ελαχιστοποιούμε την:

$$J_1(a_1, \ldots, a_n, \mathbf{e}) = \sum_{k=1}^{n} ||(\mathbf{m} + a_k\mathbf{e}) - \mathbf{x}_k||^2 = \sum_{k=1}^{n} ||a_k\mathbf{e} - (\mathbf{x}_k - \mathbf{m})||^2$$

$$= \sum_{k=1}^{n} a_k^2||\mathbf{e}||^2 - 2\sum_{k=1}^{n} a_k\mathbf{e}^t(\mathbf{x}_k - \mathbf{m}) + \sum_{k=1}^{n} ||\mathbf{x}_k - \mathbf{m}||^2.$$

- Αν θεωρήσουμε ότι $\|\boldsymbol{e}\| = 1$ τότε λαμβάνοντας την μερική παράγωγο σε σχέση με το a_k και μηδενίζοντας, έχω:

$$a_k = \mathbf{e}^t(\mathbf{x}_k - \mathbf{m}).$$

- Γεωμετρική ερμηνεία: λαμβάνουμε μια λύση τύπου ελαχίστων τετραγώνων αν προβάλουμε το διάνυσμα $\mathbf{x}_k$ σε μια γραμμή με κατεύθυνση το $\boldsymbol{e}$ και η οποία περνάει από το σημείο της μέσης τιμής **m.**

Ένα πολύ πιο ενδιαφέρον ερώτημα

- Ποια είναι η βέλτιστη κατεύθυνση **e**;
  - Θεωρώ τον πίνακα διασποράς (scatter matrix):

$$\mathbf{S} = \sum_{k=1}^{n} (\mathbf{x}_k - \mathbf{m})(\mathbf{x}_k - \mathbf{m})^t$$

  - Ισχύει: **S**=(n-1)**Σ** όπου **Σ** ο πίνακας συνδιακύμανσης

$$J_1(a_1, \ldots, a_n, \mathbf{e}) = \sum_{k=1}^{n} ||(\mathbf{m} + a_k\mathbf{e}) - \mathbf{x}_k||^2 = \sum_{k=1}^{n} ||a_k\mathbf{e} - (\mathbf{x}_k - \mathbf{m})||^2$$

$$= \sum_{k=1}^{n} a_k^2 ||\mathbf{e}||^2 - 2\sum_{k=1}^{n} a_k \mathbf{e}^t(\mathbf{x}_k - \mathbf{m}) + \sum_{k=1}^{n} ||\mathbf{x}_k - \mathbf{m}||^2.$$

$$a_k = \mathbf{e}^t(\mathbf{x}_k - \mathbf{m})$$

Ένα πολύ πιο ενδιαφέρον ερώτημα

$$J_1(\mathbf{e}) = \sum_{k=1}^{n} a_k^2 - 2\sum_{k=1}^{n} a_k^2 + \sum_{k=1}^{n} ||\mathbf{x}_k - \mathbf{m}||^2$$

$$a_k = \mathbf{e}^t(\mathbf{x}_k - \mathbf{m})$$

$$= -\sum_{k=1}^{n} [\mathbf{e}^t(\mathbf{x}_k - \mathbf{m})]^2 + \sum_{k=1}^{n} ||\mathbf{x}_k - \mathbf{m}||^2$$

$$= -\sum_{k=1}^{n} \mathbf{e}^t(\mathbf{x}_k - \mathbf{m})(\mathbf{x}_k - \mathbf{m})^t\mathbf{e} + \sum_{k=1}^{n} ||\mathbf{x}_k - \mathbf{m}||^2$$

$$= -\mathbf{e}^t\mathbf{S}\mathbf{e} + \sum_{k=1}^{n} ||\mathbf{x}_k - \mathbf{m}||^2.$$

- Φαίνεται ότι το διάνυσμα **e** το οποίο ελαχιστοποιεί το J_1(**e**) μεγιστοποιεί τον όρο:

$$\mathbf{e}^t\mathbf{S}\mathbf{e}$$

Μεγιστοποίηση του $\mathbf{e}^t\mathbf{S}\mathbf{e}$

- Υπό τον περιορισμό ότι: $\|\mathbf{e}\| = 1$
- Χρήση πολλαπλασιαστών Lagrange,
  - Έστω λ ο πολλαπλασιαστής
  - Θέτουμε: $u = \mathbf{e}^t\mathbf{S}\mathbf{e} - \lambda(\mathbf{e}^t\mathbf{e} - 1)$
  - Παραγωγίζουμε ως προς $\mathbf{e}$:

$$\frac{\partial u}{\partial \mathbf{e}} = 2\mathbf{S}\mathbf{e} - 2\lambda\mathbf{e}$$

  - Και μηδενίζουμε: $\mathbf{S}\mathbf{e} = \lambda\mathbf{e}.$
  - Από ότι φαίνεται για να μεγιστοποιήσουμε το $\mathbf{e}^t\mathbf{S}\mathbf{e}$ **πρέπει να επιλέξουμε το ιδιοδιάνυσμα που αντιστοιχεί στην μεγαλύτερη ιδιοτιμή του S.**

Μεγιστοποίηση του $\mathbf{e}^t\mathbf{S}\mathbf{e}$

- Επειδή: $\mathbf{S}\mathbf{e} = \lambda\mathbf{e}.$ έχω: $\mathbf{e}^t\mathbf{S}\mathbf{e} = \lambda\mathbf{e}^t\mathbf{e} = \lambda,$
- Με άλλα λόγια: για να βρούμε την **καλύτερη** (*καλύτερη ως προς το μέσο τετραγωνικό σφάλμα*) προβολή σε μια διάσταση των δεδομένων μας $\mathbf{x}_k$ πρέπει να προβάλουμε σε μια γραμμή που περνάει από το μέσο **m** των δεδομένων και **προς την κατεύθυνση του ιδιοδιανύσματος του πίνακα διασποράς (scatter) με την μεγαλύτερη ιδιοτιμή**.

Ανάλυση κύριων συνιστωσών

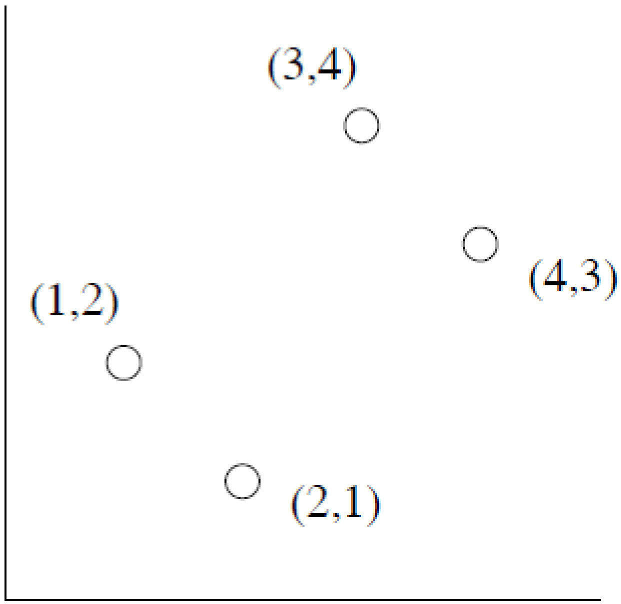
- Αν πάμε σε περισσότερες προβολές:

$$\boldsymbol{x} = \boldsymbol{m} + \sum_{i=1}^{d'} a_i \boldsymbol{e_i}$$

Με: d'<=d και $\boldsymbol{e_i}$ τα ιδιοδιανύσματα που αντιστοιχούν σε φθίνουσα σειρά των ιδιοτιμών λ_i.

- Οι συντελεστές a_i που αντιστοιχούν στις προβολές στους νέους άξονες $\boldsymbol{e_i}$ ονομάζονται κύριοι συντελεστές (**principal component**).

Παράδειγμα



Four points in a two-dimensional space

$$M = \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 3 & 4 \\ 4 & 3 \end{bmatrix}$$

$$M^{\mathrm{T}}M = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 3 & 4 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 30 & 28 \\ 28 & 30 \end{bmatrix}$$

$$(30 - \lambda)(30 - \lambda) - 28 \times 28 = 0$$

Παράδειγμα

$$\lambda = 58 \longrightarrow \begin{bmatrix} 30 & 28 \\ 28 & 30 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 58 \begin{bmatrix} x \\ y \end{bmatrix} \longrightarrow \begin{array}{rcl} \text{30x+28y} & = & \text{58x} \\ \text{28x+30y} & = & \text{58y} \end{array} \quad x = y.$$

$$\lambda = 58 \longrightarrow \mathbf{e}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

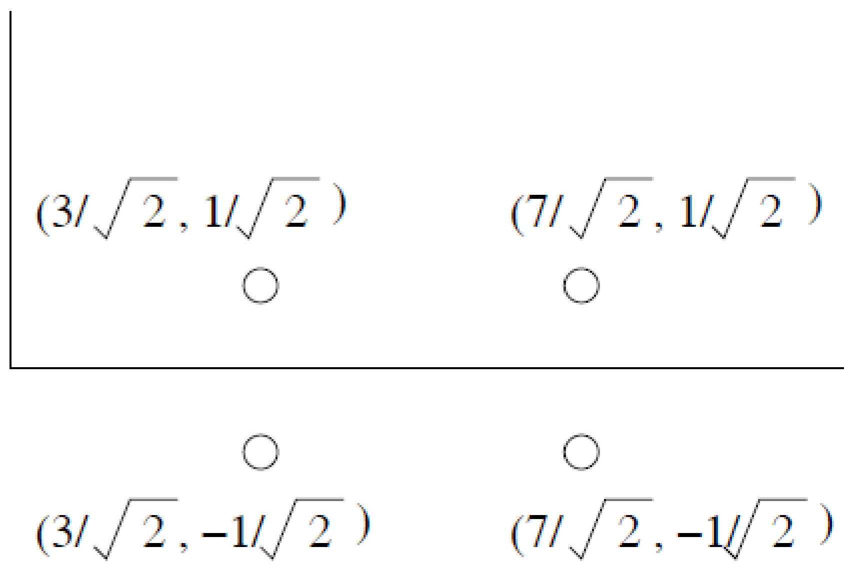
$$\lambda = 2. \longrightarrow \begin{bmatrix} 30 & 28 \\ 28 & 30 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 2 \begin{bmatrix} x \\ y \end{bmatrix} \longrightarrow \begin{array}{rcl} \text{30x+28y} & = & \text{2x} \\ \text{28x+30y} & = & \text{2y} \end{array} \quad x = -y.$$

$$\lambda = 2. \longrightarrow \mathbf{e}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

$$E = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

Παράδειγμα

$$ME = \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 3 & 4 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 3/\sqrt{2} & 1/\sqrt{2} \\ 3/\sqrt{2} & -1/\sqrt{2} \\ 7/\sqrt{2} & 1/\sqrt{2} \\ 7/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$$



The points of Fig. 11.1 in the new coordinate system

Εφαρμογή: Αναγνώριση προσώπου (face recognition)

- Οι στόχοι της χρήσης PCA στην αναγνώριση προσώπου είναι οι εξής:
  - Ελάττωση των δεδομένων (Data Reduction)
  - Επιλογή Χαρακτηριστικών (Feature Selection)

Εφαρμογή: Αναγνώριση προσώπου (face recognition)

- Στην αναγνώριση προσώπου μία εικόνα εισόδου με n pixels μπορεί να θεωρηθεί ως **ένα** σημείο σε ένα n-διάστατο χώρο που ονομάζεται χώρος εικόνας.
  - Οι συντεταγμένες εκείνου του σημείου είναι οι τιμές κάθε pixel της εικόνας:

$$\mathbf{p}_x=(i_1,i_2,\ldots,i_n)$$

  - Οι τιμές αυτές λαμβάνονται διατρέχοντας την εικόνα είτε γραμμή-γραμμή είτε στήλη-στήλη.

Εφαρμογή: Αναγνώριση προσώπου (face recognition)

- Π.χ: Εικόνα 128×128 pixels.

$$= \begin{bmatrix} 150 & 152 & \cdot & 151 \\ 131 & 133 & \cdot & 72 \\ \cdot & \cdot & \cdot & \cdot \\ 144 & 171 & \cdot & 67 \end{bmatrix}_{128 \times 128}$$

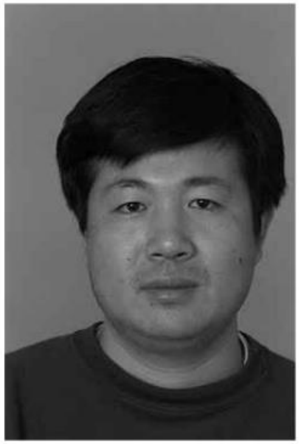
$$16384 \text{ vector}$$

$$[150, 152, \cdots 151, 131, 133, \cdots 72, \cdots 144, 171, \cdots 67]_{16K}$$

Εφαρμογή: Αναγνώριση προσώπου (face recognition)

- Οι εικόνες προσώπου γενικά παρουσιάζουν μια αρκετά πλεονάζουσα (redundant) πληροφορία:
  - Όλα τα pixel παρασκηνίου είναι ίδια
  - Κάθε υποκείμενο (άνθρωπος) έχει τα ίδια χαρακτηριστικά προσώπου.



- Στον χώρο των δεδομένων κάθε pixel είναι μια μεταβλητή
  - Έτσι η διάσταση του χώρου είναι πολύ υψηλή.
  - Η ελάττωση της διάστασης γίνεται με την PCA.
- Έστω $D \in \mathbb{R}^{N \times n}$ ο πίνακας που αποτελείται από N πρόσωπα των n-pixels.
  - Κάθε γραμμή είναι και μια εικόνα.

$$\boldsymbol{D} = \begin{bmatrix} 150 & 152 & \cdots & 254 & 255 & \cdots & 252 \\ 131 & 133 & \cdots & 221 & 223 & \cdots & 241 \\ \cdot & \cdot & & \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & \cdot & & \cdot \\ 144 & 171 & \cdots & 244 & 245 & \cdots & 223 \end{bmatrix}_{N \times n}$$

- Έστω ότι η μέση τιμή των στηλών είναι:

$$[120\ 140\ \cdots\ 230\ 230\ \cdots\ 240]$$

- Δημιουργούμε τον κεντραρισμένο (centered) πίνακα **U** αφαιρώντας την μέση τιμή από τον πίνακα **D**.

$$\boldsymbol{U} = \begin{bmatrix} 30 & 12 & \cdots & 24 & 25 & \cdots & 12 \\ 11 & -7 & \cdots & -9 & -7 & \cdots & 1 \\ . & . & & . & . & & . \\ 24 & 31 & \cdots & 14 & 15 & \cdots & -17 \end{bmatrix}_{N \times n}$$

- Υπολογίζω τον πίνακα συνδιακύμανσης $\boldsymbol{\Sigma} \in \mathbb{R}^{n \times n}$ από τον **U**.

$$\Sigma = \boldsymbol{U}^T \boldsymbol{U} / (N-1)$$

- Βρίσκω τον πίνακα ιδιοδιανυσμάτων **E** και τον αντίστοιχο πίνακα ιδιοτιμών **Λ** του πίνακα **Σ** που ικανοποιούν την:

$$\Sigma = \mathbf{E} \cdot \mathbf{\Lambda} \cdot \mathbf{E}^{\mathrm{T}}$$

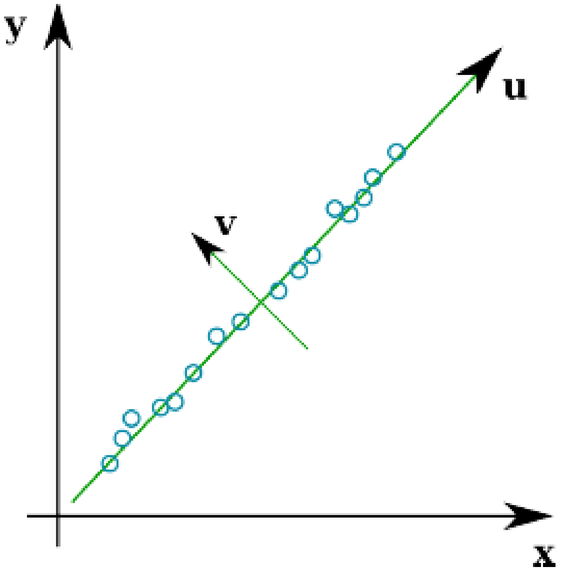
- Κανονικοποιούμε (αν χρειάζεται) τα $\mathbf{e}_i$ ιδιοδιανύσματα έτσι ώστε να αποτελούν μια ορθοκανονική βάση:

$$\forall \boldsymbol{e}_i, \boldsymbol{e}_j \in \boldsymbol{E}, < \boldsymbol{e}_i, \boldsymbol{e}_j > = \delta_{ij} = \begin{cases} 1, \alpha\nu\ i = j \\ 0, \alpha\nu\ i \neq j \end{cases}$$

- *Ιδιοδιανύσματα με μικρές ιδιοτιμές συμβάλλουν με μικρή πληροφορία στην αναπαράσταση των δεδομένων.*
- *Συνεπώς μείωση των διαστάσεων των δεδομένων επιτυγχάνεται αγνοώντας τα ιδιοδιανύσματα με μικρές ιδιοτιμές.*

- Το σύνολο των ιδιοδιανυσμάτων *m* (με *m*<*n*) του **Σ** που έχουν τις m-μεγαλύτερες ιδιοτιμές, ελαχιστοποιεί το μέσο τετραγωνικό σφάλμα ανακατασκευής.

- Αν και τυπικά απαιτούνται *n*-μεταβλητές για την αναπαράσταση ενός αρχικού δείγματος εικόνας **X**, μεγάλο μέρος της μεταβλητότητας των δεδομένων μπορεί να οφείλεται σε μικρότερο αριθμό **m** κυριών συνιστωσών.



EigenFaces

- Έτσι, τα αρχικά δεδομένα που αποτελούνται από N εικόνες των n-μεταβλητών μπορούν να ελαττωθούν σε N εικόνες των m-μεταβλητών (ιδιοδιανύσματα).

- Στην αναγνώριση προσώπου τα ιδιοδιανύσματα καλούνται «**ιδιόφατσες**» - **eigenfaces**.

Δεδομένα εκπαίδευσης

- **D**: Training Set: 14 Εικόνες 384×256 pixels

EigenFaces

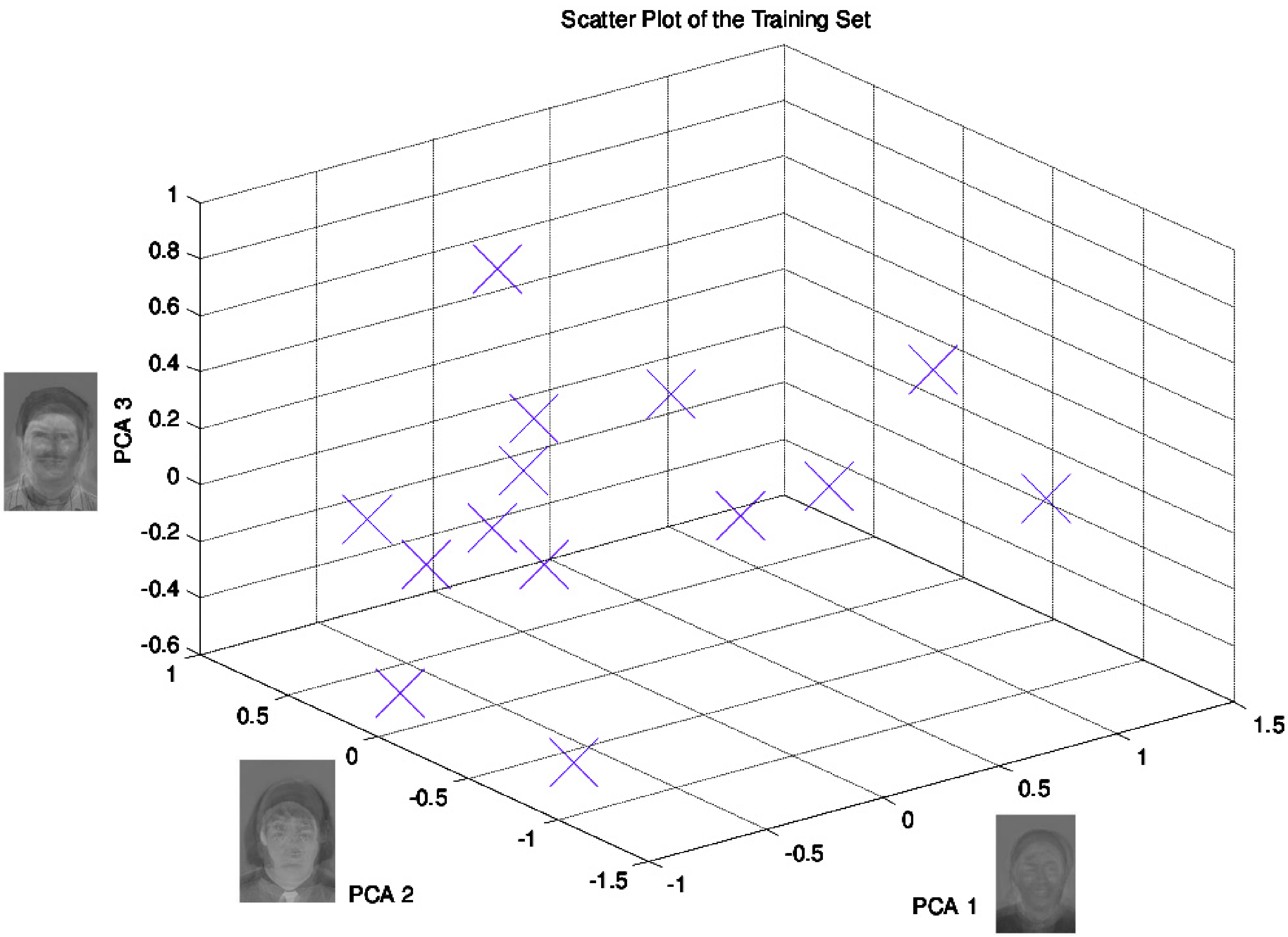
- Μέση τιμή:

EigenFaces

- Τα 4 ιδιοδιανύσματα με τις πιο ισχυρές ιδιοτιμές.

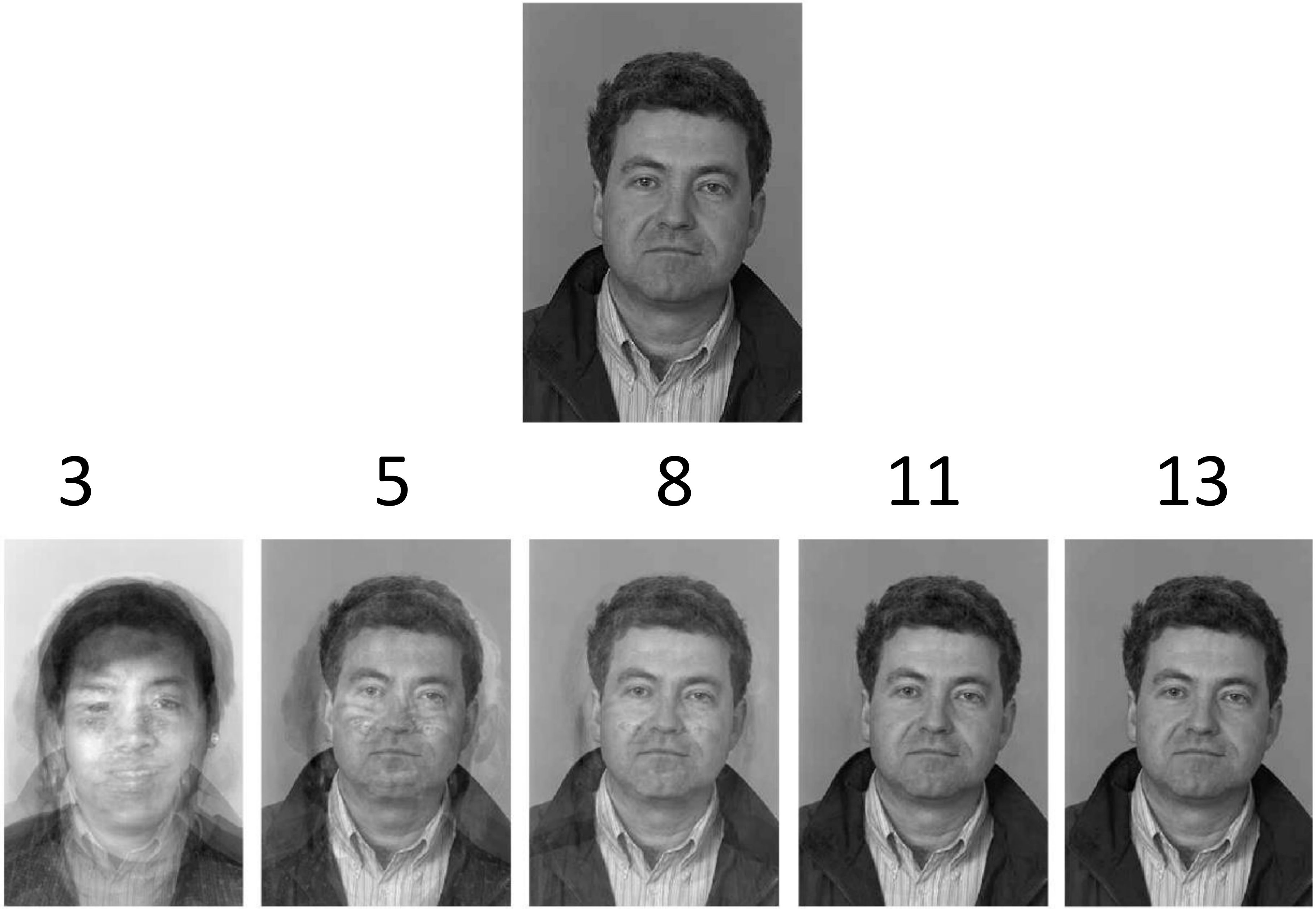
EigenFaces

- Η προβολή τους στα 3 πρώτα ιδιοδιανύσματα:



Ανακατασκευή προσώπων

- Αρχική:


3
5
8
11
13

Ανακατασκευή προσώπων

- Αρχική:

3

5

8

11

13

Πόσες κύριες συνιστώσες χρειαζόμαστε;;

- Δεν υπάρχει απόλυτη απάντηση
- Μια προσέγγιση είναι να δούμε πόση διασπορά αντιπροσωπεύει η κάθε μια κύρια συνιστώσα.
  – Για το i-ιδιοδιάνυσμα:

$$r_i = 100 \times \frac{\lambda_i}{\sum_{j=1}^{n} \lambda_j}$$

- Μια ευριστική (heuristic) μέθοδος δεν λαμβάνει υπόψη μια κύρια συνιστώσα όταν $r_i < 1\%$.

Η κατάρα της διαστατικότητας.

- Η αναγνώριση προσώπου αποτελεί ένα χαρακτηριστικό παράδειγμα της κατάρας της διαστατικότητας.
  - small sample size problem.
- Το πλήθος των μεταβλητών (pixels) n είναι πολύ μεγάλο αλλά το πλήθος των δειγμάτων – εικόνων στο σύνολο εκπαίδευσης N είναι πολύ μικρότερο.
- Καθώς ο πίνακας συνδιακύμανσης **Σ** είναι πολύ μεγάλος (n×n) ο υπολογισμός των ιδιοδιανυσμάτων είναι μια δύσκολη υπόθεση.
  - Δεδομένου ότι υπάρχουν μόνο N εικόνες εκπαίδευσης ο βαθμός (rank) του **Σ** είναι το πολύ N-1 και ως εκ τούτου υπάρχουν το πολύ N-1 ιδιοδιανύσματα με μη μηδενικές ιδιοτιμές.

Μετασχηματισμός KL
Karhunen Lowe Transform

- Οι Karhunen και Lowe βρήκαν μια πολύ ωραία ιδέα για τον αποδοτικό υπολογισμό των κυρίων συνιστωσών.
  - $\mathbf{U}^{\mathrm{T}}\mathbf{U} \in \mathbb{R}^{n\times n}$
  - Ενώ $\mathbf{U}\mathbf{U}^{\mathrm{T}} \in \mathbb{R}^{N\times N}$
- Δεδομένου N<<n υπολογίζουμε τις ιδιοτιμές του $\mathbf{U}\mathbf{U}^{\mathrm{T}}$ με μεγαλύτερη ευκολία.

$$\boldsymbol{U}\boldsymbol{U}^{T}\boldsymbol{E}' = \boldsymbol{E}'\boldsymbol{\Lambda}$$

- Πολλαπλασιάζοντας από δεξιά με $\mathbf{U}^{\mathrm{T}}$ έχουμε:

$$\boldsymbol{U}^{T}\boldsymbol{U}\boldsymbol{U}^{T}\boldsymbol{E}' = \boldsymbol{U}^{T}\boldsymbol{E}'\boldsymbol{\Lambda}$$

$$\boldsymbol{U}^T\boldsymbol{U}\boldsymbol{U}^T\boldsymbol{E}' = \boldsymbol{U}^T\boldsymbol{E}'\boldsymbol{\Lambda}$$

- Με μια αναδιοργάνωση…

$$\boldsymbol{U}^T\boldsymbol{U}(\boldsymbol{U}^T\boldsymbol{E}') = (\boldsymbol{U}^T\boldsymbol{E}')\boldsymbol{\Lambda}$$

- Άρα τα $(\boldsymbol{U}^T\boldsymbol{E}')$ είναι τα ιδιοδιανύσματα των $\boldsymbol{U}^T\boldsymbol{U}$.
- Αρα:
  - Βρίσκουμε τα ιδιοδιανύσματα $\boldsymbol{E}'$ του «μικρού» πίνακα $\mathbf{UU}^T$
  - Στην συνέχεια πολ/μος με $\mathbf{U}^t$ για να βρώ τα ιδιοδιανύσματα του $\boldsymbol{U}^T\boldsymbol{U}$.
- Τα προκύπτοντα ιδιοδιανύσματα δεν είναι ορθοκανονικά. Αυτό επιτυγχάνεται με την διαίρεση κάθε στήλης του πίνακα ιδιοδιανυσμάτων $(\boldsymbol{U}^T\boldsymbol{E}')$ με το μέτρο:
  - $\sqrt{(\mathrm{N}-1)\lambda_i}$ αν έχουμε τον scatter πινακα **S**.
  - $\sqrt{\lambda_i}$ αν έχουμε τον πίνακα συνδιακυμανσης **σ**.

$$\boldsymbol{U^T U U^T E' = U^T E' \Lambda}$$

- Υπάρχουν το πολύ N-1 μη μηδενικές ιδιοτιμές στον πίνακα $\mathbf{U}\mathbf{U}^{T}$
- Αυτό δεν αποτελεί περιορισμό, καθώς όπως γνωρίζουμε ο βαθμός (rank) του $\mathbf{U}^{T}\mathbf{U}$ είναι το πολύ N-1.