آموزش یادگیری ماشین (Machine Learning)
(تئوری – عملی) – بخش دوم

درس پنجم: تقلیل ابعاد

مدرس:
فرشید شیرافکن
دانشجوی دکترای بیو انفورماتیک
دانشگاه تهران

مقدمه

ماتریس قطری:

$$A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & a_{mm} \end{bmatrix}$$

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}_{2\times 2}$$

ماتریس‌متعامد (Orthogonal) :

$$AA^T = A^T A = I$$

$$0^2 + \left(\frac{3}{5}\right)^2 + \left(\frac{4}{5}\right)^2 = 1$$

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{3}{5} & -\frac{4}{5} \\ 0 & \frac{4}{5} & \frac{3}{5} \end{bmatrix} \times A^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{3}{5} & \frac{4}{5} \\ 0 & -\frac{4}{5} & \frac{3}{5} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

دترمینان ماتریس

$$|A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

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

$$|A| = 4 \times 2 - 1 \times 1 = 8 - 1 = 7$$

$$\begin{vmatrix} -1 & 4 & 3 \\ 2 & 6 & 4 \\ 3 & -2 & 8 \end{vmatrix} = (-1)\begin{vmatrix} 6 & 4 \\ -2 & 8 \end{vmatrix} - 4\begin{vmatrix} 2 & 4 \\ 3 & 8 \end{vmatrix} + 3\begin{vmatrix} 2 & 6 \\ 3 & -2 \end{vmatrix}$$

$$= (-1)(48 + 8) - 4(16 - 12) + 3(-4 - 18) =$$

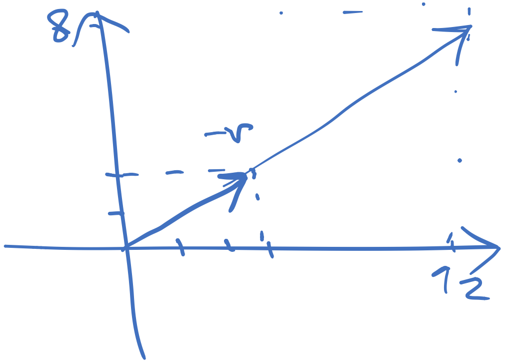
بردار ویژه و مقدار ویژه یک ماتریس

بردار v، بردار ویژه ماتریس مربعی A است، اگر ثابت لاندا وجود داشته باشد به‌طوری که:

$$\boxed{\mathbf{A}\,.\,\mathbf{v} = \lambda\,.\,\mathbf{v}}$$

$$A = \begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix}_{2\times2} \qquad v = \begin{bmatrix} 3 \\ 2 \end{bmatrix}_{2\times1}$$

$$\underbrace{\begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} 3 \\ 2 \end{bmatrix}}_{v} = \begin{bmatrix} 12 \\ 8 \end{bmatrix}_{2\times1} = \underbrace{4}_{\lambda} \times \underbrace{\begin{bmatrix} 3 \\ 2 \end{bmatrix}}_{v}$$



محاسبه مقدار ویژه

$$\boxed{A \cdot v = \lambda \cdot v}$$

$$A \cdot v = \lambda \cdot v$$

$$A \cdot v - \lambda \cdot I \cdot v = 0$$

$$\boxed{(A - \lambda \cdot I) \cdot v = 0}$$

مقادیر ویژه: ریشه‌های $|A - \lambda \cdot I|$

مثال

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

$v \,,\, \lambda$

$$A \cdot v = \lambda \cdot v$$

$$\boxed{|A - \lambda I| = 0}$$

$$\begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \underbrace{\begin{bmatrix} -\lambda & 1 \\ -2 & -3-\lambda \end{bmatrix}}_{A - \lambda I}$$

$$|A - \lambda I| = (-\lambda)(-3-\lambda) - (-2)(1) = \lambda^2 + 3\lambda + 2 = 0$$

مقادیر ویژه

$$\begin{cases} \lambda_1 = -1 \longrightarrow v_1 = \begin{bmatrix} ? \\ ? \end{bmatrix} \\ \lambda_2 = -2 \longrightarrow v_2 = \begin{bmatrix} ? \\ ? \end{bmatrix} \end{cases}$$

$$\boxed{A . v = \lambda . v}$$

$$v_1 = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ -x \end{bmatrix} = x \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}$$

$$A . v_1 = \lambda_1 . v_1$$

$$\begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = (-1) \begin{bmatrix} x \\ y \end{bmatrix} \Rightarrow \begin{bmatrix} y \\ -2x-3y \end{bmatrix} = \begin{bmatrix} -x \\ -y \end{bmatrix}$$

$$\begin{bmatrix} y \\ -2x-3y \end{bmatrix} - \begin{bmatrix} -x \\ -y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{cases} y + x = 0 \\ -2x-3y+y = 0 \end{cases} \Rightarrow y = -x$$

$$\boxed{A \cdot v = \lambda \cdot v}$$

$$A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \qquad \lambda_2 = -2$$

$$A \cdot v_2 = \lambda_2 \cdot v_2$$

$$\begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \begin{bmatrix} m \\ n \end{bmatrix} = -2 \begin{bmatrix} m \\ n \end{bmatrix}$$

$$v_2 = \begin{bmatrix} m \\ n \end{bmatrix} = \begin{bmatrix} m \\ -2m \end{bmatrix} = m \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

$$\begin{bmatrix} n \\ -2m-3n \end{bmatrix} + 2 \begin{bmatrix} m \\ n \end{bmatrix} = 0 \Rightarrow \begin{bmatrix} n+2m \\ -2m-n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{cases} n + 2m = 0 \\ -2m - n = 0 \end{cases} \Rightarrow n = -2m$$

مقادیر منفرد ماتریس

مقادیر منفرد(تکین) ماتریس $A_{m\times n}$ برابر است با: جذر مقادیر ویژه ماتریس A^TA یا AA^T

$\sqrt{10}$

$\sqrt{12}$

$$A = \begin{bmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{bmatrix}_{2\times 3}$$

$$AA^T = \begin{bmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ 1 & 3 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 11 & 1 \\ 1 & 11 \end{bmatrix}$$

$$\begin{vmatrix} 11-\lambda & 1 \\ 1 & 11-\lambda \end{vmatrix} = 0 \Rightarrow (11-\lambda)^2 - 1 = 0 \Rightarrow 11-\lambda = \pm 1$$

$$\lambda_1 = 10 \;,\; \lambda_2 = 12$$

تجزیه مقدارهای منفرد

Singular Value Decomposition (SVD)

قطری

$$X = USV^T$$

$$\begin{bmatrix} - & - \\ - & - \end{bmatrix}_{2\times2} = \begin{bmatrix} m & p \\ n & q \end{bmatrix}_{2\times2} \times \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}_{2\times2} \times \begin{bmatrix} a & b \\ c & d \end{bmatrix}_{2\times2}$$

$$U \times U^T = I \qquad V \times V^T = I$$

$$X^T X = \begin{bmatrix} - & - \\ - & - \end{bmatrix}$$

$$X^T X = V S U^T \, U S V^T = V S^2 V^T$$

$\sigma_1 \quad \lambda_1 \rightarrow V_1 = \begin{pmatrix} a \\ b \end{pmatrix}$

$\sigma_2 \quad \lambda_2 \rightarrow V_2 = \begin{pmatrix} c \\ d \end{pmatrix}$

$$\boxed{\begin{array}{l} X V_2 = \sigma_2 u_2 \\ u_2 = \begin{pmatrix} p \\ q \end{pmatrix} \end{array}} \qquad \boxed{\begin{array}{l} X V_1 = \sigma_1 u_1 \\ u_1 = \begin{pmatrix} m \\ n \end{pmatrix} \end{array}}$$

مثال

$$X = \begin{pmatrix} 2 & 2 \\ -1 & 1 \end{pmatrix}$$

$$X = U S V^T$$

$$X^T X = \begin{bmatrix} 2 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix} \qquad \begin{vmatrix} 5-\lambda & 3 \\ 3 & 5-\lambda \end{vmatrix} = 0$$

$$(5-\lambda)^2 - 9 = 0 \Rightarrow 5-\lambda = \pm 3 \Rightarrow \lambda_1 = 2 \,,\, \lambda_2 = 8$$

$$\sigma_1 = \sqrt{2} \,,\, \sigma_2 = \sqrt{8} = 2\sqrt{2}$$

$$S = \begin{bmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{8} \end{bmatrix} \qquad \lambda_1 = 2 \Rightarrow v_1 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

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

$$\lambda_2 = 8 \Rightarrow v_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

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

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

$$X V_1 = \sigma_1 u_1$$

$$\begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \sqrt{2} \begin{bmatrix} a \\ b \end{bmatrix} \longrightarrow \begin{matrix} a = 0 \\ b = 1 \end{matrix}$$

$$u_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$X V_2 = \sigma_2 u_2$$

$$\begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \sqrt{8} \begin{bmatrix} c \\ d \end{bmatrix} \Rightarrow \begin{matrix} c = 1 \\ d = 0 \end{matrix} \Rightarrow u_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

ماتریس کواریانس

X_1	X_2
1	2
3	5
5	8

$$X = \begin{bmatrix} 1 & 2 \\ 3 & 5 \\ 5 & 8 \end{bmatrix} \qquad \overline{X_1} = 3 \qquad \overline{X_2} = 5$$

$$cov(x,y) = \frac{1}{m-1}\sum_{i=1}^{m}(x_i - \overline{x})(y - \overline{y})$$

$$v(x_1) = \frac{1}{3-1}\left[(1-3)^2 + (3-3)^2 + (5-3)^2\right] = 4$$

$$v(x_2) = \frac{1}{3-1}\left[(2-5)^2 + (5-5)^2 + (8-5)^2\right] = 9$$

$$cov(x_1, x_2) = \frac{1}{3-1}\left[(1-3)(2-5) + (3-3)(5-5) + (5-3)(8-5)\right] = 6$$

$$cov(x_1, x_2) = \begin{bmatrix} v(x_1) & cov(x_1, x_2) \\ cov(x_2, x_1) & v(x_2) \end{bmatrix} = \begin{bmatrix} 4 & 6 \\ 6 & 9 \end{bmatrix}$$

ماتریس کواریانس(روش دوم)

X_1	X_2
1	2
3	5
5	8

$$X = \begin{bmatrix} 1 & 2 \\ 3 & 5 \\ 5 & 8 \end{bmatrix} \Longrightarrow \begin{bmatrix} -2 & -3 \\ 0 & 0 \\ 2 & 3 \end{bmatrix}$$

$$\bar{x}_1 = 3 \qquad \bar{x}_2 = 5$$

$$C = \frac{1}{2} X^T X = \frac{1}{2} \begin{bmatrix} -2 & 0 & 2 \\ -3 & 0 & 3 \end{bmatrix} \begin{bmatrix} -2 & -3 \\ 0 & 0 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 4 & 6 \\ 6 & 9 \end{bmatrix}$$

تقلیل ابعاد

الگوریتم تقلیل ابعاد، یکی از الگوریتم های یادگیری بدون ناظر است.

کاربردهای تقلیل ابعاد:

۱– به تصویر کشیدن داده‌ها (Data Visualization)

هدف از تجسم داده‌ها، تقلیل داده‌ها به K=2 یا K=3 است طوری که بتوان داده‌ها را به‌صورت نمودار ترسیم کرد و تجسم آن‌ها ساده شود.

۲– فشرده کردن داده‌ها (Data Compression)

فشرده‌سازی به‌منظور کاهش حجم حافظه کامپیوتر و تسریع الگوریتم‌های یادگیری انجام می‌شود.

هدف فشرده‌سازی، تقلیل ابعاد داده‌ها از n به K است طوری که درصد بالایی از واریانس داده‌ها حفظ شود.

۱۰۰ ۱۰۰۰

تجسم داده‌ها

تجسم داده‌ها و مشاهده ساختار آن‌ها می‌تواند باعث بهتر شدن طراحی الگوریتم‌های یادگیری شود.

$x^{(i)} \in R^{40}$

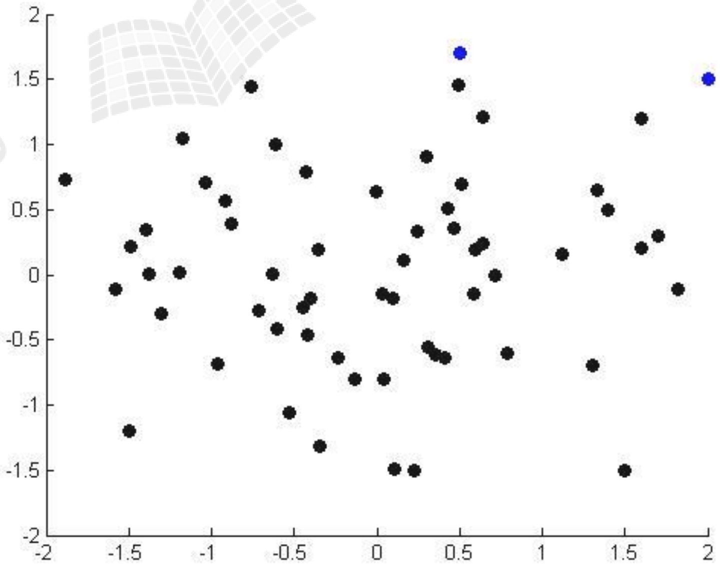
$40 \rightarrow 2$

Country	X1	X2	…	X40
Canada	1.577	80.7	…	67.293
China	5.878	73	…	10.22
India	1.632	64.7	…	0.735
Russia	1.48	65.5	…	0.72
…	…	…		…

Country	Z1	Z2
Canada	1.6	1.2
China	1.7	0.3
India	1.6	0.2
Russia	1.4	0.5
…	…	…

$z^{(i)} \in R^{2}$

Z2



Z1

استخراج ویژگی

$$\begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix} \rightarrow \begin{bmatrix} y_1 \\ \vdots \\ y_{d'} \end{bmatrix} = f\left(\begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix}\right)$$

استخراج ویژگی **(Feature Extraction)**:

- تبدیل خطی یا غیر خطی بر روی فضای ویژگی اصلی (original)
- ایجاد ویژگی‌های جدید از ویژگی‌های اصلی

$$\begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix} \rightarrow \begin{bmatrix} x_{i_1} \\ \vdots \\ x_{i_{d'}} \end{bmatrix}$$

انتخاب ویژگی **(Feature Selection)**:

- انتخاب یک زیرمجموعه از مجموعه ویژگی‌های داده شده $(d' < d)$

$$\begin{bmatrix} x_1 \\ \vdots \\ x_{100} \end{bmatrix} \xrightarrow{FS} \begin{bmatrix} x_2 \\ x_5 \\ x_{90} \end{bmatrix}$$

$d = 100$

$d' = 3$

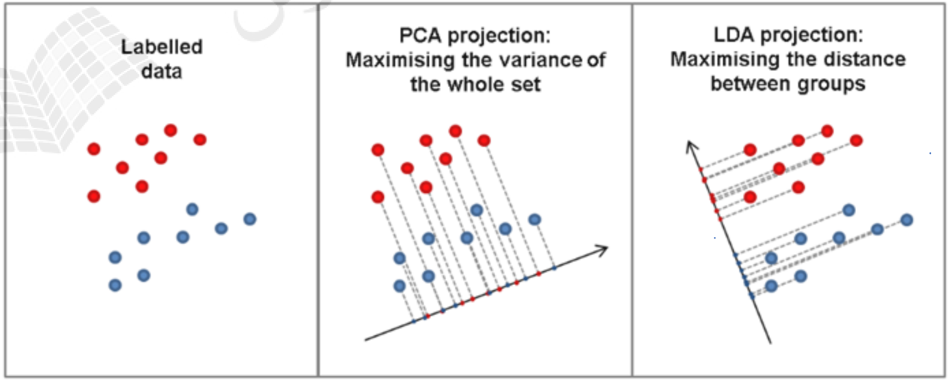
$$\begin{bmatrix} x_1 \\ \vdots \\ x_{100} \end{bmatrix} \xrightarrow{FE} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

$y_1 = x_2 + 6x_{70}$

الگوریتم‌های خطی Feature extraction

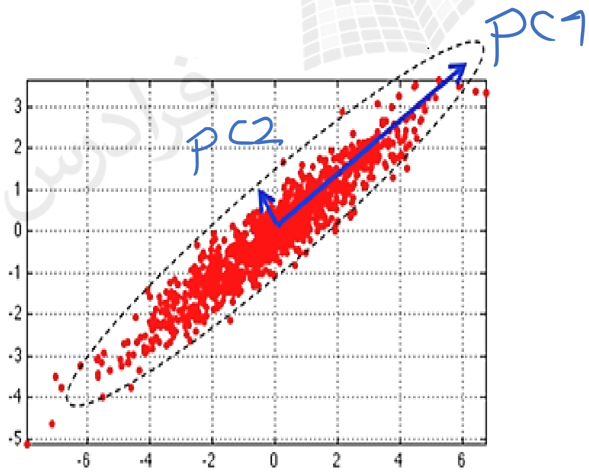
- بدون نظارت : Principal Component Analysis (PCA)
- با نظارت : Linear Discriminant Analysis (LDA)

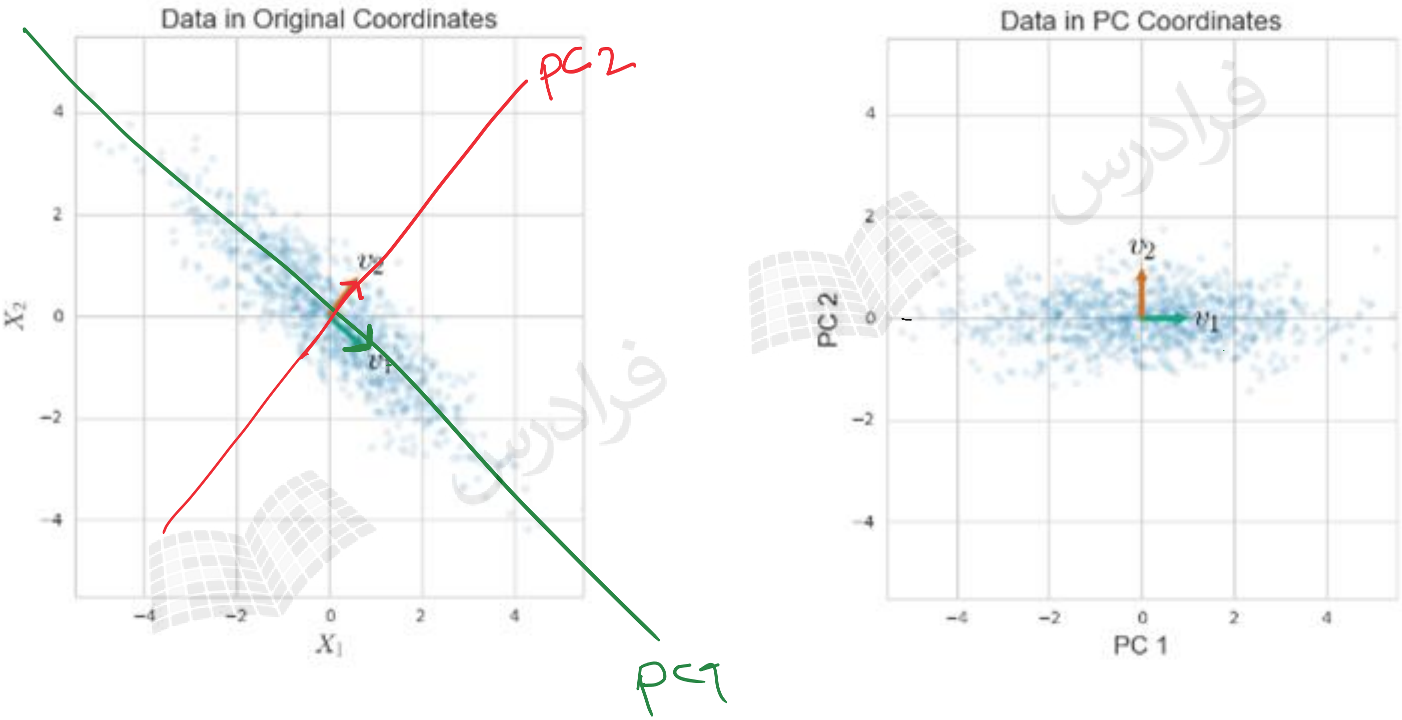
در LDA از برچسب‌ها استفاده می‌کند و راستایی را پیدا می‌کند که اگر داده‌ها بر روی آن پروجکت شوند، داده‌های دو کلاس از هم دیگر فاصله دارند و داده‌های یک کلاس به هم نزدیکتر باشند.



تحلیل عناصر اصلی

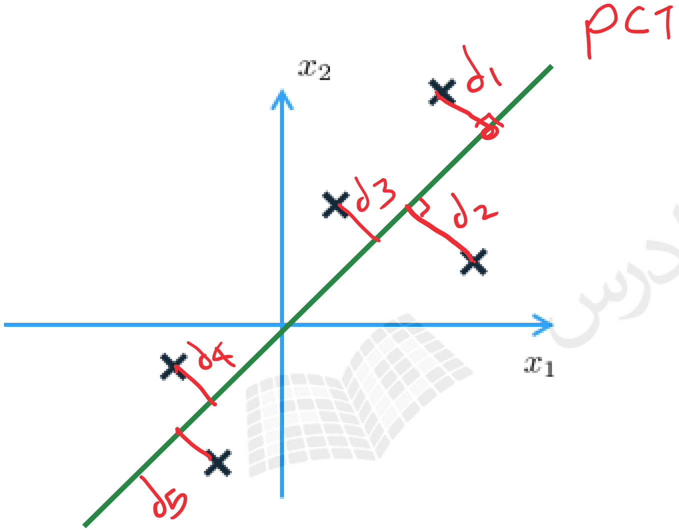
Principal Component Analysis (PCA) is a dimension-reduction tool that can be used to reduce a large set of variables to a small set that still contains most of the information in the large set.




Data in Original Coordinates
X_1
X_2
v_1
v_2
PC 2
PC1
Data in PC Coordinates
PC 1
PC 2
v_1
v_2

PCA

یافتن خطی که اگر از نقاط بر آن خط عمود کنیم، مجموع مربعات فاصله‌های نقاط تا تصاویر آن‌ها بر آن خط کمینه شود.



$$\underbrace{d_1^2 + d_2^2 + \cdots + d_5^2}_{\text{حداقل}}$$

مثال

X_1	X_2
1	2
3	4
5	6

$$X = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$$

$\bar{X}_1 = 3$

$\bar{X}_2 = 4$

مرحله ۱

$$X = \begin{bmatrix} -2 & -2 \\ 0 & 0 \\ 2 & 2 \end{bmatrix}$$

مرحله ۲

$$C = \frac{1}{3-1} X^T X = \frac{1}{2} \begin{bmatrix} -2 & 0 & 2 \\ -2 & 0 & 2 \end{bmatrix} \times \begin{bmatrix} -2 & -2 \\ 0 & 0 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 4 \\ 4 & 4 \end{bmatrix}$$

مرحله ۳

$$\begin{vmatrix} 4-\lambda & 4 \\ 4 & 4-\lambda \end{vmatrix} = 0 \Rightarrow (4-\lambda)^2 = 16 \Rightarrow \begin{cases} \lambda_1 = 8 \Rightarrow \begin{bmatrix} 1 \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} 0.7 \\ 0.7 \end{bmatrix} \\ \lambda_2 = 0 \Rightarrow \begin{bmatrix} -1 \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} -0.7 \\ 0.7 \end{bmatrix} \end{cases}$$

$$V = \begin{array}{cc} v_1 & v_2 \\ \end{array}$$
$$V = \begin{bmatrix} 0.7 & -0.7 \\ 0.7 & 0.7 \end{bmatrix}$$

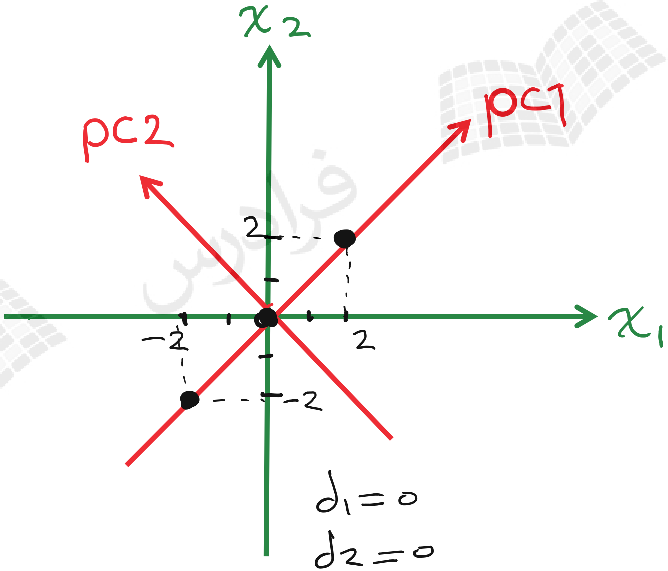
$pc1: y = x$

$pc2: y = -x$

$$\boxed{P = X \cdot V}$$

$$P = \begin{bmatrix} -2 & -2 \\ 0 & 0 \\ 2 & 2 \end{bmatrix} \times V$$

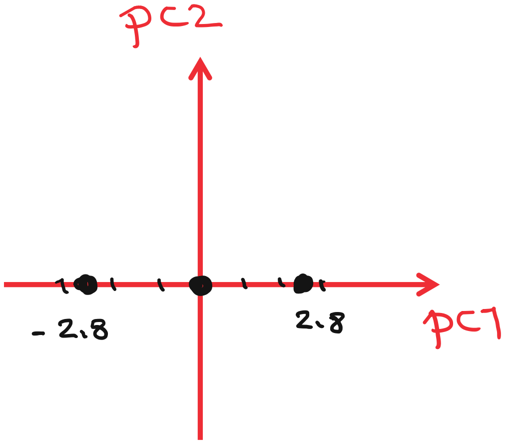
$$P = \begin{array}{cc} pc1 & pc2 \\ \end{array}$$
$$P = \begin{bmatrix} -2.8 & 0 \\ 0 & 0 \\ 2.8 & 0 \end{bmatrix}$$



$d_1 = 0$

$d_2 = 0$

$d_3 = 0$



$2D \rightarrow 1D$

$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \begin{bmatrix} pc1 \end{bmatrix}$

اگر بخواهیم یک بردار روی فضا پیدا کنیم که داده‌ها روی آن پروجکت شوند و دارای بیشترین واریانس باشند، آن بردار، بردار ویژه متناظر با بزرگترین مقدار ویژه ماتریس کواریانس است.

بردارهای متناظر با مقادیر ویژه کوچک را می‌توان دور ریخت تا ابعاد کاهش پیدا کنند تا در فضای کوچکتری عملیات را انجام داد. بردارهای متناظر با مقادیر ویژه کوچک، در بازسازی تاثیر زیادی ندارند.

$$X_{m\times d} \longrightarrow X_{center} \longrightarrow C = \frac{1}{2} X^T X \Longrightarrow \begin{matrix} \lambda_1 \to v_1 \\ \lambda_2 \to v_2 \end{matrix}$$

$$V = \begin{bmatrix} | & | \\ v_1 & v_2 \\ | & | \end{bmatrix} \qquad \boxed{P = X \times V}$$

مثال

X_1	X_2
1	2
3	5
5	8

$$C = \begin{bmatrix} 4 & 6 \\ 6 & 9 \end{bmatrix}$$

$$\begin{vmatrix} 4-\lambda & 6 \\ 6 & 9-\lambda \end{vmatrix} = 0$$

$$\Rightarrow \begin{cases} \lambda_1 = 13 \Rightarrow V_1 = \begin{bmatrix} 1 \\ 1.5 \end{bmatrix} \Longrightarrow \begin{bmatrix} 0.5 \\ 0.8 \end{bmatrix} \\ \lambda_2 = 0 \Rightarrow V_2 = \begin{bmatrix} 1 \\ -0.6 \end{bmatrix} \Longrightarrow \begin{bmatrix} 0.8 \\ -0.5 \end{bmatrix} \end{cases}$$

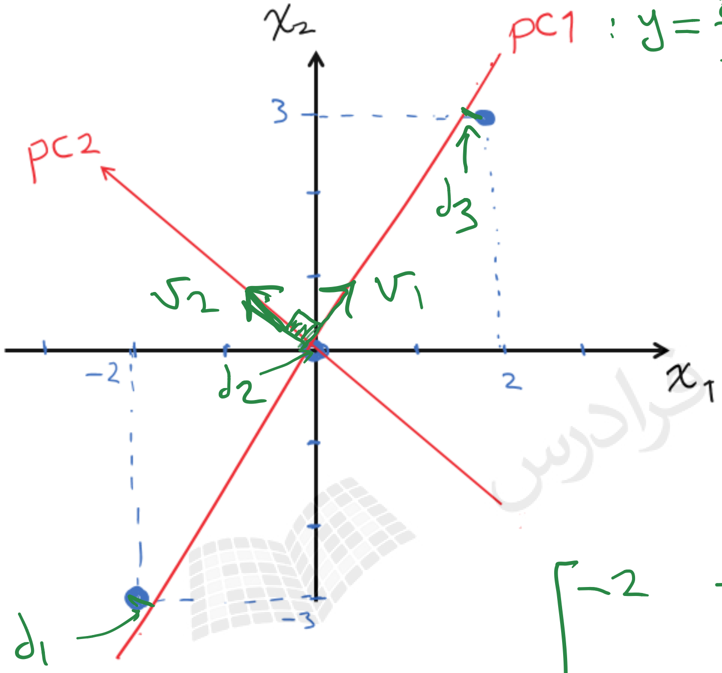
$PC1$

$y = mx$

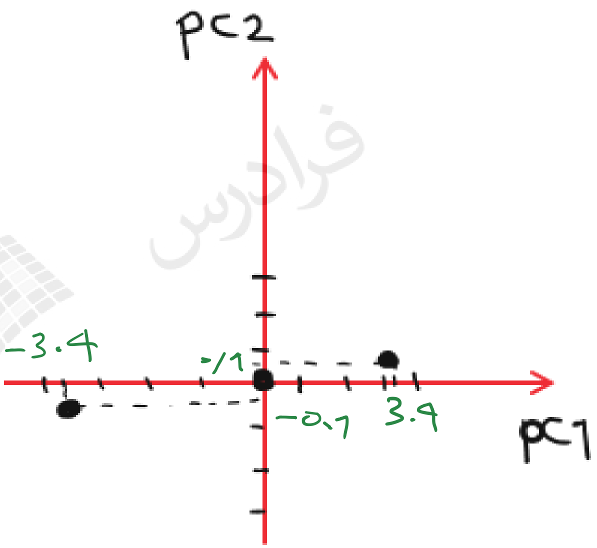
$$\boxed{y = \frac{8}{5}x}$$

$PC2$

$$\boxed{y = -\frac{5}{8}x}$$



$pc1 : y = \frac{8}{5}x$



$$SS = d_1^2 + d_2^2 + d_3^2 = 0.02$$

$$\begin{bmatrix} -2 & -3 \\ 0 & 0 \\ 2 & 3 \end{bmatrix} \times \begin{bmatrix} 0.5 & 0.8 \\ 0.8 & -0.5 \end{bmatrix} = \begin{bmatrix} -3.4 & -0.1 \\ 0 & 0 \\ 3.4 & 0.1 \end{bmatrix}$$

الگوریتم PCA

$d \longrightarrow d'$

1. $\mu = \frac{1}{m}\sum_{i=1}^{m} x^{(i)}$

2. $X \leftarrow$ centered X

3. $C = \frac{1}{m-1} X^T X$

4. Calculate eigenvalue and eigenvectors of C

5. Pick d' eigenvectors corresponding to the largest eigenvalues and put them in the columns of $V = [v_1, \dots, v_{d'}]$

   First PC   d'-th PC

6. $p = X V$

به کمک SVD

$$X = \begin{bmatrix} x_1^{(1)} & \cdots & x_d^{(1)} \\ \vdots & & \vdots \\ x_1^{(m)} & \cdots & x_d^{(m)} \end{bmatrix}$$

$$\begin{array}{ccc} \downarrow & & \downarrow \\ \mu_1 & & \mu_d \end{array}$$

۱- محاسبه میانگین هر ویژگی:

$$\mu_j = \frac{1}{m}\sum_{i=1}^{m} x_j^{(i)}$$

۲- نرمالیزه کردن ویژگی‌ها:

$$x_j^{(i)} \leftarrow x_j - \mu_j$$

۳- اگر مقیاس اندازه‌گیری ویژگی‌ها از نظر عددی بسیار متفاوت باشند، در ضرایبی متناسب با دامنه آن‌ها ضرب می‌شوند تا میزان عددی آن‌ها قابل قیاس باشد.

۴- تهیه ماتریس کواریانس:

$$\Sigma = \frac{1}{m}\sum_{i=1}^{n} (x^{(i)})(x^{(i)})^T$$

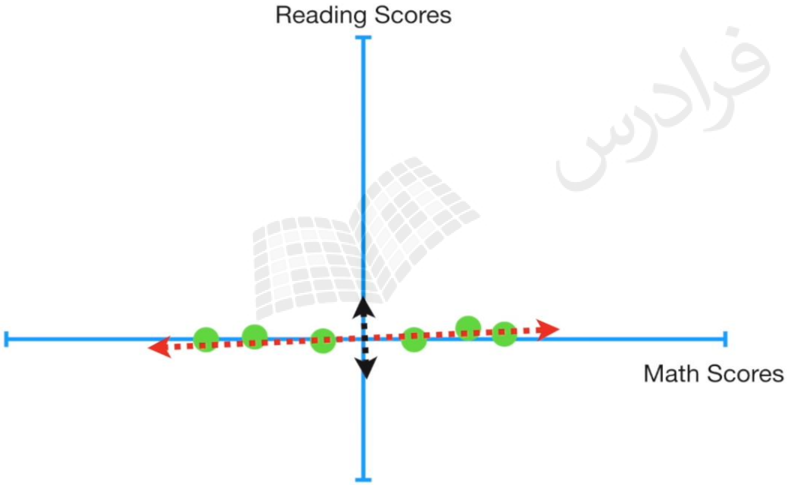
۵- محاسبه بردارهای ویژه ماتریس سیگما با استفاده از روش تجزیه SVD

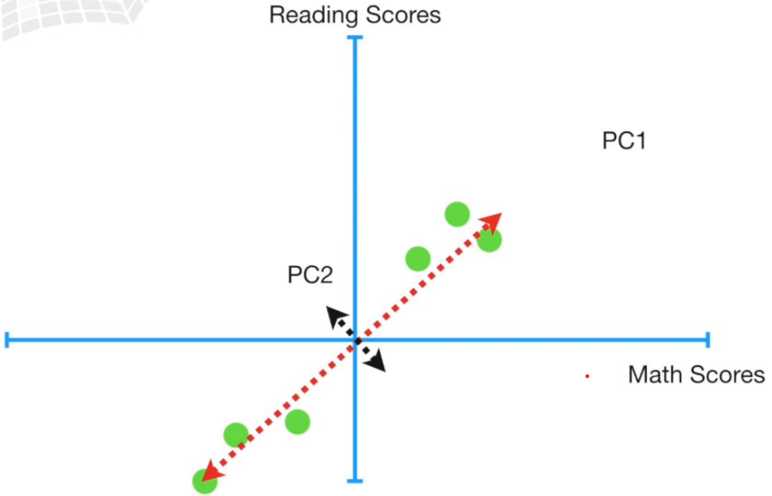
```
[ U , S , V ] = svd(Sigma);
```

مثال

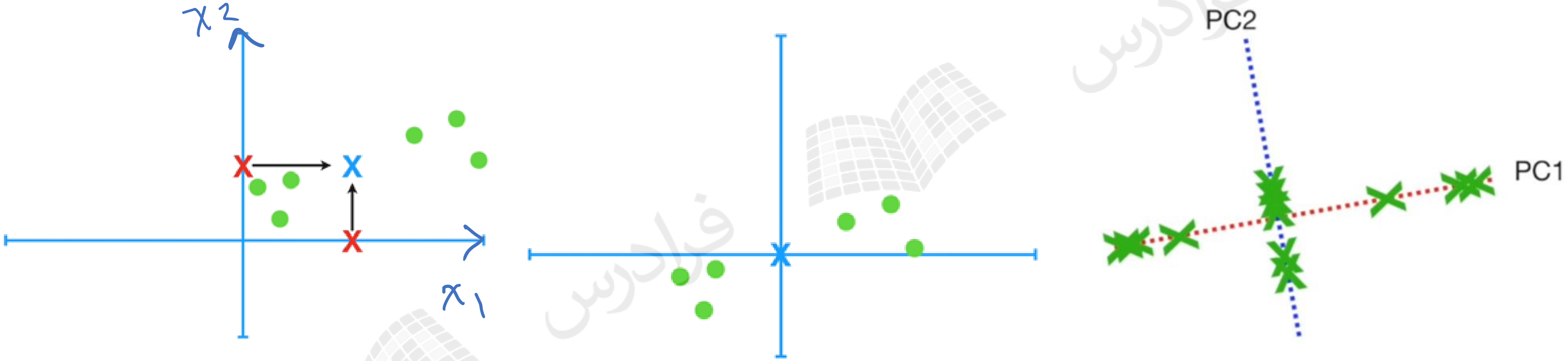
مقیاس

	student1	student2	student3	student4	…	
Math	95 9.5	88 8.8	93 9.3	75 7.5		→ ۰-۱۰۰
Reading	9	8	10	7		→ ۰-۱۰





مثال



دو دیدگاه معادل

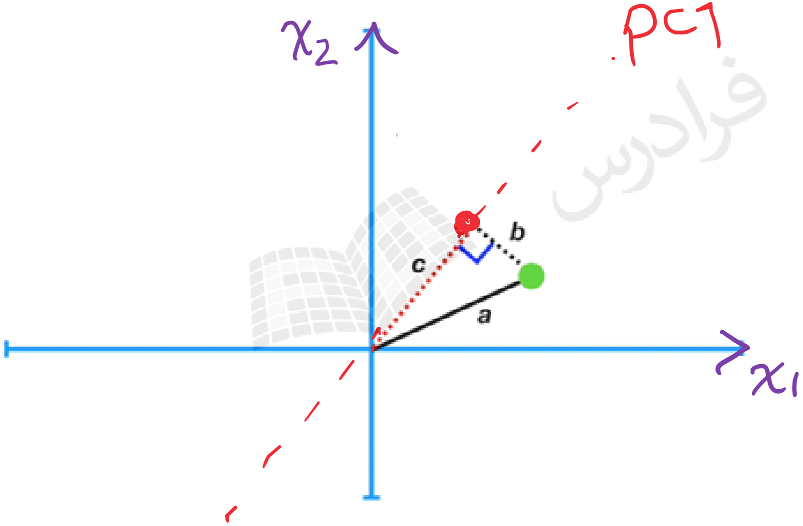
$$a^2 = b^2 + c^2$$

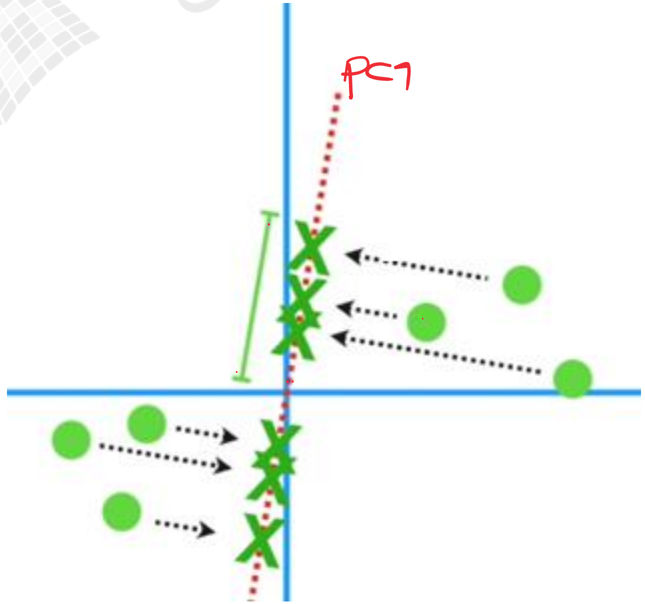
ثابت کم زیاد

PCA can either:

minimize the distance to the line or

maximize the distance from the projected point to the origin.

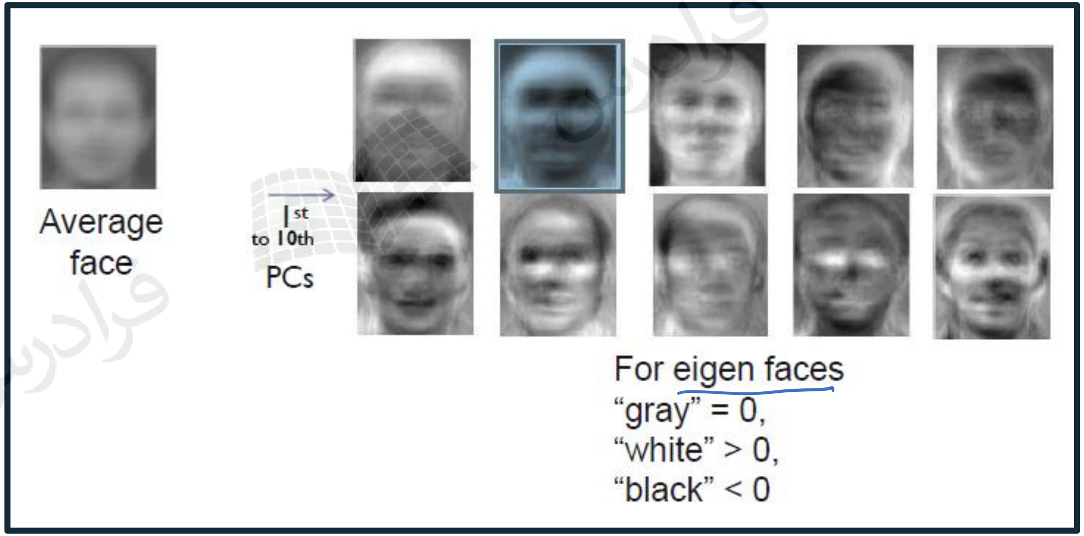




PCA روی تصویر

۱۱۲ × ۹۲ = ۱۰۳۰۴

ORL Database

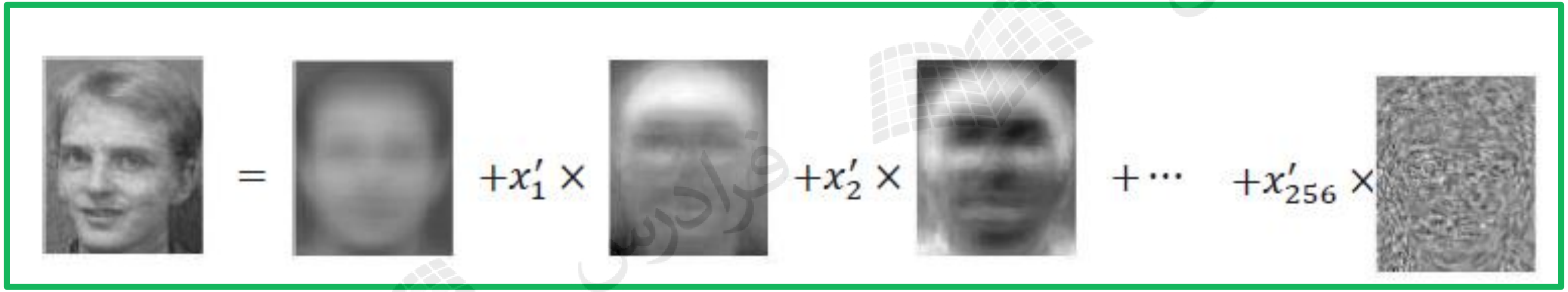


۱) ۲) ۳) c ۴) ۷

۱۰۳۰۴

بازسازی تصویر

با ترکیب ۲۵۶ تصویر می‌توان تصویر را با دقت خوبی بازسازی کرد.



Feature vector=$[x'_1, x'_2, \dots, x'_{d'}]$

$x'_i = PC_i^T \boldsymbol{x} \longrightarrow$ The projection of $\boldsymbol{x}$ on the i-th PC

انتخاب k مناسب

$$\frac{\sum_{i=1}^{k} S_{ii}}{\sum_{i=1}^{n} S_{ii}}$$

$$\underset{1000}{n} \rightarrow \boxed{PCA} \rightarrow \underset{100}{K}$$

هدف: حداقل ٪95 واریانس داده‌ها حفظ شود.

SVD

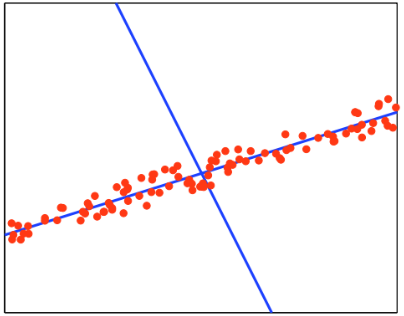
$$S = \begin{bmatrix} S_{11} & 0 & 0 \\ 0 & S_{22} & 0 \\ 0 & 0 & S_{33} \end{bmatrix}$$

$n = 3$

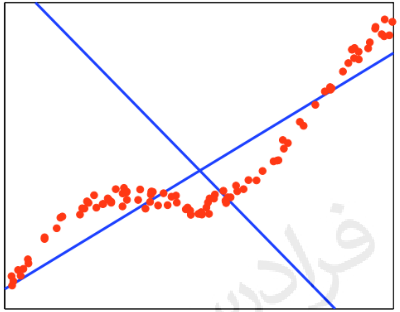
$\boxed{K = 2}$

$$\frac{S_{11} + S_{22}}{S_{11} + S_{22} + S_{33}} > 0.95$$

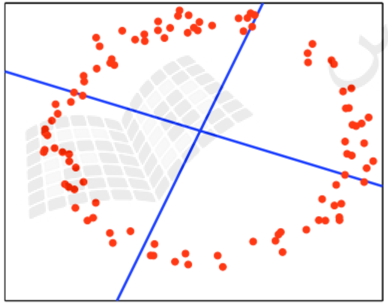
مثال



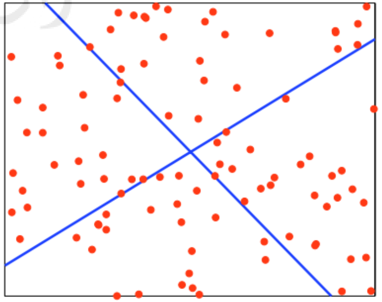
$\lambda_1 = 0.0938$
$\lambda_2 = 0.0007$



$\lambda_1 = 0.1260$
$\lambda_2 = 0.0054$



$\lambda_1 = 0.0884$
$\lambda_2 = 0.0725$



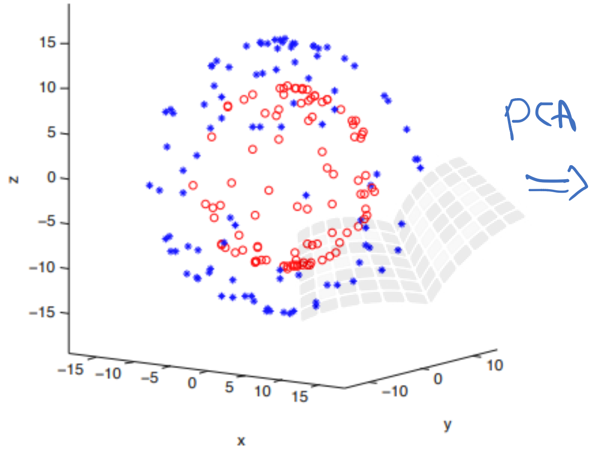
$\lambda_1 = 0.0881$
$\lambda_2 = 0.0769$

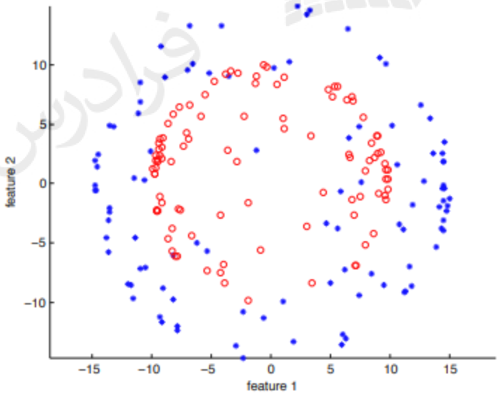
Kernel PCA

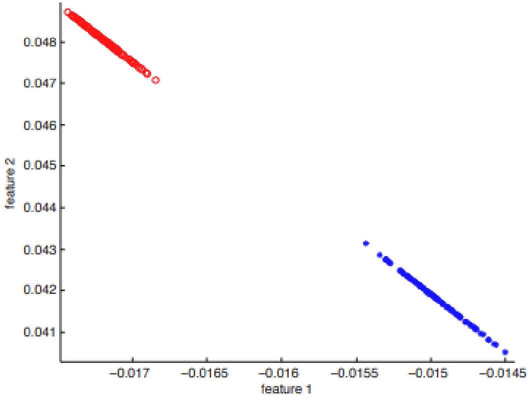
تجزیه و تحلیل مؤلفه اصلی می‌تواند با استفاده از ترفند هسته در یک روش غیر خطی استفاده شود.

تکنیک حاصل قادر به ساخت نقشه‌های غیر خطی است که واریانس را در داده‌ها به حداکثر می‌رساند.

3D → 2D







Gaussian Kernel
(σ : 20)

این اسلایدها بر مبنای نکات مطرح شده در فرادرس

«آموزش یادگیری ماشین (Machine Learning) (تئوری - عملی) – بخش دوم»

تهیه شده است.

برای کسب اطلاعات بیشتر در مورد این آموزش به لینک زیر مراجعه نمایید.

faradars.org/fvdm94062