14/15/3
2019

EIGENVALUE DECOMPOSITION (Özdeğer Ayrıştırması)

Denklem Sistemi

$$y_1(t+1) = -y_1(t) - 1.5\, y_2(t)$$
$$y_2(t+1) = 0.5\, y_1(t) + y_2(t)$$

Katsayılar

$$A = \begin{bmatrix} -1 & -1{,}5 \\ 0{,}5 & 1 \end{bmatrix} \qquad \bar{y}(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix}$$

$$\bar{y}(t+1) = A \cdot \bar{y}(t)$$

Çözüm önerisi

$$y_1(t) = x_1 \cdot \lambda^t \Rightarrow y_1(t+1) = x_1 \cdot \lambda^{t+1} \qquad \lambda, x_1, x_2 \in R$$
$$y_2(t) = x_2 \cdot \lambda^t \Rightarrow y_2(t+1) = x_2 \cdot \lambda^{t+1}$$

Denklemde yerine yazarsak

$$\begin{bmatrix} x_1 \cdot \lambda^{t+1} \\ x_2 \cdot \lambda^{t+1} \end{bmatrix} = \lambda^{t+1} \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}} = \underbrace{\begin{bmatrix} -1 & -1{,}5 \\ 0{,}5 & 1 \end{bmatrix}} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \lambda^t$$

$$\lambda \cdot \bar{x} = A \cdot \bar{x}$$

λ = Eigenvalue (Özdeğer)

$\bar{x}$ = Eigen Vector (Özvektör)

A matrisini karakterize eder

Bir vektörü eigenvektörle çarparsak sonuç aynı yönde bir vektör olur

$\propto$ λ ve $\bar{x}$ nasıl bulunur? $\lambda \in R \quad A \in R^{m \times n}, \bar{x} \in R^m$

$$A \cdot \bar{x} = \lambda \cdot \bar{x}$$
$$A \cdot \bar{x} - \lambda \cdot \bar{x} = \bar{0}$$
$$\Rightarrow (A_{m\times n} - \lambda I_{m\times m})\, \bar{x}_{m\times 1} = \bar{0}_{m\times 1}$$

Tersi alınamaz demektir.

- Öyle bir λ değeri bulalım ki
  $A - \lambda . I$ matrisi singüler (tekil) matris olsun.
- Bunun için $\det(A - \lambda I) = \det(\lambda I - A) = 0$
- $\chi_A(\lambda) = \det(\lambda I - A)$ → x'e bağlı polinom

$\boxed{\chi_A(\lambda) = 0}$ Polinom denkleminin çözümü λ'ları verir.

örnek :

$$A = \begin{bmatrix} -1 & -1,5 \\ 0,5 & 1 \end{bmatrix}$$

$$\chi_A(\lambda) = \det(\lambda I - A) = \det\left(\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} - \begin{bmatrix} -1 & -1,5 \\ 0,5 & 1 \end{bmatrix} \right)$$

$$\det \begin{bmatrix} \lambda+1 & 1,5 \\ -0,5 & \lambda-1 \end{bmatrix} = (\lambda+1)(\lambda-1) - 1.5(-0,5)$$

$$\chi_A(\lambda) = \lambda^2 - 1 + 0.75$$

Karakteristik polinom $\Rightarrow$ $\chi_A(\lambda) = \lambda^2 - 0.25 = 0$

$$\lambda^2 = 0.25 \qquad \boxed{\lambda_1 = 0.5} \quad \boxed{\lambda_2 = -0}$$

Her bir eigenvalue karşılık bir eigenvektör vardır.

1) $\lambda_1 = 0.5$ için

$$A.\bar{x}_1 = \lambda_1 . \bar{x}_1$$

$\bar{x}_1 \to \lambda_1$'e karşılık gelen eigen vektör

$$\begin{bmatrix} -1 & -1,5 \\ 0,5 & 1 \end{bmatrix} \bar{x}_1 = 0,5 . \bar{x}_1$$

$$\left(\begin{bmatrix} -1 & -1,5 \\ 0,5 & 1 \end{bmatrix} - 0,5 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \bar{x}_1 = \bar{0}$$

$$\begin{bmatrix} -1,5 & -1,5 \\ 0,5 & 0,5 \end{bmatrix} \bar{x}_1 = \bar{0} \qquad \bar{x}_1 = \begin{bmatrix} x_{11} \\ x_{12} \end{bmatrix} \text{ olsun diyelim.}$$

$$\Rightarrow (-1,5)\, x_{11} - 1,5\, x_{12} = 0 \qquad x_{11} = -x_{12}$$

$$(0,5)\, x_{11} + 0,5\, x_{12} = 0 \qquad \boxed{x_{11} = \alpha} \; \boxed{x_{12} = -\alpha} \text{ olsun diyoruz.}$$

$$\bar{x}_1 = \begin{bmatrix} \alpha \\ -\alpha \end{bmatrix} \quad \alpha \in R \text{ eigenvektör ise}$$

$$\|\bar{x}_1\|^2 = 1 = \bar{x}_1^T . \bar{x}_1$$

$$= [\alpha \;\; -\alpha] \begin{bmatrix} \alpha \\ -\alpha \end{bmatrix} = \alpha^2 + \alpha^2 = 2\alpha^2$$

$$2\alpha^2 = 1$$

$$\alpha^2 = \frac{1}{2} \qquad \alpha = \frac{1}{\sqrt{2}}$$

$$\bar{x}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$$

$\lambda = 0.5$ için $\bar{x}_1$'i bulduk.

$\lambda_2 = -0.5$ için $\quad \bar{X}_2 = \begin{vmatrix} x_{21} \\ x_{22} \end{vmatrix}$ 'yi bulalım

$A.\bar{x}_2 = \lambda_2 . \bar{x}_2$

$$\left(\begin{bmatrix} -1 & -1.5 \\ 0.5 & 1 \end{bmatrix} + \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix} \right) \begin{bmatrix} x_{21} \\ x_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\left. \begin{array}{l} -0.5\, x_{21} - (1.5)\, x_{22} = 0 \\ 0.5\, x_{21} + (1.5)\, x_{22} = 0 \end{array} \right\} \quad \begin{array}{l} x_{21} = -3x_{22} \\ x_{22} = \alpha = \frac{1}{\sqrt{10}} \\ x_{21} = -3\alpha = \frac{-3}{\sqrt{10}} \end{array}$$

$$\|\bar{x}_2\|^2 = 1 \Rightarrow \bar{x}_2^T . \bar{x}_2$$

$$\begin{bmatrix} -3\alpha & \alpha \end{bmatrix} \begin{bmatrix} -3\alpha \\ \alpha \end{bmatrix} = 1$$

$$9\alpha^2 + \alpha^2 = 1$$

$$\alpha^2 = \frac{1}{10} \qquad \alpha = \frac{1}{\sqrt{10}}$$

$$\lambda_2 = -0.5 \Rightarrow \bar{x}_2 = \begin{bmatrix} \frac{-3}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} \end{bmatrix}$$

verilen denklem sisteminin çözümü

$$\bar{y}(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = (0.5)^t \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \text{ ya}$$

$$= (-0.5)^t \begin{bmatrix} \frac{-3}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} \end{bmatrix}$$

toplam çözüm

$$\bar{y}(t) = \lambda_1^t . \bar{x}_1 + \lambda_2^t . \bar{x}_2$$

$$\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = (0.5)^t \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} + (-0.5)^t \begin{bmatrix} \frac{-3}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} \end{bmatrix}$$

Özellik $A \in R^{m\times m}$ matrisi için tüm eigenvalue'lar birbiri farklı ise, $\lambda_1 \neq \lambda_2 \neq \ldots\ldots \neq \lambda_m$ bunlara karşılık gelen eigenvektörler ($\lambda_1 \to \bar{x}_1$, $\lambda_2 \to \bar{x}_2$, ..., $\lambda_m \to \bar{x}_m$) birbirinden bağımsızdır.

$$c_1 . \bar{x}_1 + c_2 . \bar{x}_2 + \ldots\ldots + c_m . \bar{x}_m = \bar{0}$$

$$\boxed{c_1 = c_2 = c_3 \ldots\ldots = c_m = 0}$$

Örnek: $\lambda_1 = 0.5$ $\lambda_2 = -0.5$

$$\overline{x_1} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \qquad \overline{x_2} = \begin{bmatrix} \frac{-3}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} \end{bmatrix}$$

$$c_1 \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{bmatrix} + c_2 \begin{bmatrix} \frac{-3}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$c_1 = c_2 = 0$ için sağlanır

Matrisin Köşegenleştirilmesi $A \in R^{m \times m}$

Eigen Value denklem

$$A \cdot \overline{x_i} = \lambda_i \cdot \overline{x_i} \qquad i = 1, 2, 3, \ldots m$$

$$A \cdot x_1 = \lambda_1 \cdot \overline{x_1}$$

$$A \cdot x_2 = \lambda_2 \cdot \overline{x_2}$$

$$\vdots$$

$$A \cdot x_m = \lambda_m \cdot \overline{x_m}$$

$$\begin{bmatrix} | & | & | \\ A_1 \cdot \overline{x_1} & A_1 \cdot \overline{x_2} & A_1 \cdot \overline{x_m} \\ | & | & | \end{bmatrix}_{m \times n} = \begin{bmatrix} | & | & | \\ \lambda_1 \cdot \overline{x_1} & \lambda_2 \cdot \overline{x_2} & \lambda_m \cdot \overline{x_m} \\ | & | & | \end{bmatrix}_{m}$$

$$\Rightarrow \underbrace{\underset{m \times n}{A} \begin{bmatrix} \overline{x_1} & \overline{x_2} & \ldots & \overline{x_m} \end{bmatrix}_{m \times n}}_{S} = \underbrace{\begin{bmatrix} \overline{x_1} & \overline{x_2} & \ldots & \overline{x_m} \end{bmatrix}_{m \times n}}_{S} \underbrace{\begin{bmatrix} \lambda_1 & & 0 \\ 0 & \lambda_2 & \\ 0 & & \ddots \; \lambda_m \end{bmatrix}_{m \times n}}_{\Lambda}$$

$S = [\overline{x_1} = \overline{x_2} = \ldots \overline{x_m}]$ eigenvalue 'ların oluşturduğu matris

köşegen matrisi $\Lambda = \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots \; \lambda_m \end{bmatrix}$ eigenvalue 'ların oluşturduğu köşegen matris

$$A . S = S . \Lambda$$

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

NOT: Bir matrisin özdeğer ve özvektörlerini bilir isek matrisin kendisini oluşturabiliriz.

Yalnız tüm λ değerleri birbirinden farklı ise

Örnek:

$\lambda = 0.5 \quad \overline{x_1} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{bmatrix}$

$\lambda_2 = -0.5 \quad \overline{x_2} = \begin{bmatrix} \frac{-3}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} \end{bmatrix}$

eigenvalue ve eigenvektörü verilen A matrisini oluşturunuz

$$S = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-3}{\sqrt{10}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{10}} \end{bmatrix} \qquad \Lambda = \begin{bmatrix} 0.5 & 0 \\ 0 & -0.5 \end{bmatrix}$$

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

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

$$A.S = S.\Lambda$$

$$\boxed{S^{-1}.A.S = \Lambda}$$ → Bir matrisin eigenvektörleri biliniyorsa köşegen matrisi Λ bulunabilir

A^n'in Hesaplanması

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

$$A^2 = \underbrace{(S \Lambda S^{-1})(S}_{I} \Lambda S^{-1}) = S \Lambda^2 S^{-1}$$

$$\Lambda = \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \lambda_m \end{bmatrix} \qquad \Lambda^n = \begin{bmatrix} \lambda_1^m & & 0 \\ & \lambda_2^m & \\ 0 & & \lambda_m^m \end{bmatrix} \Rightarrow \boxed{A^n = S.\Lambda^n S^{-1}}$$

matrisin fonksiyonu

$$f(t) = \sum_{i=1}^{k} f(i)\, t^i = f(1).t + f(2)t^2 + \ldots\ldots f(k).t^k$$

verilen $A \in R^{m \times m}$ için

$$f(A) = \sum_{i=1}^{k} f(i).A^i = \sum f(i)\, S \Lambda^i S^{-1}$$

$$= S\left(\sum f(i) \Lambda^i\right) S^{-1}$$

$$e^t = \sum_{i=0}^{\infty} \frac{t^i}{i!}$$

$$= 1 + t + \frac{t^2}{2!} + \frac{t^3}{3!}$$

$$e^A = S \underbrace{\left(\sum_{i=0}^{\infty} \frac{\Lambda^i}{i!}\right)}_{e^\Lambda} S^{-1} \Rightarrow \boxed{e^A = S e^\Lambda S^{-1}}$$

EIGENDECOMPOSITION UYGULAMALARI

Karhunen – Loeve Expansion:

$\bar{x}$ mx1 sıfır-ortalamalı (zero mean) random vektör olarak verilsin.

kovaryans matrisi $\underset{m\times m}{R} = E\{\bar{x}\bar{x}^H\}$

R'yi eigendecomposition yaptığımızı varsayarsak, $R = U \Lambda U^H$

$\underset{m\times m}{U} = \left[\bar{U}_1 \, \bar{U}_2 \cdots \bar{U}_m\right]$ $\bar{U}_i$: R'nin normalize edilmiş eigenvektörü

$\underset{m\times m}{\Lambda} = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_m \end{bmatrix}$ λ_i : R'nin eigenvalue'sı. $U \cdot U^H = I$

$\Rightarrow \bar{y} = U^H \cdot \bar{x}$ $\rightarrow \bar{y}$ sıfır-ortalamalı random vektör

$E\{\bar{y}\bar{y}^H\} = E\{U^H \bar{x}\bar{x}^H U\} = U^H \underbrace{E\{\bar{x}\bar{x}^H\}}_{R} U$

$= \underbrace{U^H U}_{I} \Lambda \underbrace{U^H U}_{I} = \Lambda$ olur.

$\Rightarrow \bar{x} = U \cdot \bar{y}$ olur.

$\Rightarrow \boxed{\bar{x} = \sum_{i=1}^{m} \bar{U}_i \, y_i}$ $\rightarrow$ random vektör $\bar{x}$'i ortogonal vektörlerin (eigenvektörler) lineer kombinasyonu olarak yazabiliriz.

*

Bu açılıma "Karhunen – Loeve" açılımı denir.

Low-Rank Approximation

Bazı durumlarda $\bar{x}$ vektörünü [(*)'da verilen] daha az sayıda eigenvektör kullanarak göstermek isteyebiliriz.

$\tilde{x} \in C^m$

$\tilde{x} = K \bar{x}$ Burada K mxm boyutlu bir matris ancak rankı $r < m$'dir.

$K = \sum_{i=1}^{r} \mu_i \bar{u}_i \bar{U}_i^H = U M_r U^H$ $M_r = \begin{bmatrix} \mu_1 & & & & & 0 \\ & \mu_2 & & & & \\ & & \ddots & & & \\ & & & \mu_r & & \\ & & & & 0 & \\ 0 & & & & & 0 \end{bmatrix}$

$\tilde{\bar{x}}$ ile $\bar{x}$ arasındaki hatayı minimize etmek için

$$K = \sum_{i=1}^{r} \bar{U}_i \bar{U}_i^H = U I_r U^H \qquad I_r = \begin{bmatrix} 1 & & & & & 0 \\ & 1 & & & & \\ & & \ddots & & & \\ & & & 1 & & \\ & & & & 0 & \\ 0 & & & & & \ddots \; 0 \end{bmatrix}_{m \times m}$$

r tane

PRINCIPAL COMPONENT METHOD

$\bar{X} \in R^m$ zero-mean random vektör ise,

$\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n$ n tane gözlem vektörü ise
(observation)

örneklem kovaryans matrisi : $R = \frac{1}{n-1} \sum_{i=1}^{n} \bar{x}_i \bar{x}_i^T$
(sample coveriance)

R matrisinin eigenvektörleri $\bar{a}_1, \bar{a}_2, \dots, \bar{a}_m$ olsun. Bunlara karşılık gelen eigenvalue'lar $\lambda_1, \lambda_2, \dots, \lambda_m$.

$\langle \bar{a}_i, \bar{a}_j \rangle = 0$ eigenvektörler birbirine ortogonal. $\|\bar{a}_i\| = 1$

$X = [\bar{x}_1 \; \bar{x}_2 \; \dots \; \bar{x}_n]$ data matrisi.

$\boxed{y_{1,i} = \bar{a}_1^T \bar{x}_i}$ → 1st principal component (eigenvektör ile datavektörün iç çarpımı)

sample-varyans: $\sigma_{y_1}^2 = \frac{1}{n-1} \sum_{i=1}^{n} y_{1,i}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (\bar{a}_1^T x_i)^2$

$$= \frac{1}{n-1} \sum_{i=1}^{n} \bar{a}_1^T x_i x_i^T \bar{a}_1 = \bar{a}_1^T \underbrace{\left(\frac{1}{n-1} \sum_{i=1}^{n} \bar{x}_i \bar{x}_i^T\right)}_{R} \bar{a}_1$$

$$\Rightarrow \sigma_{y_1}^2 = \bar{a}_1^T R \bar{a}_1 = \bar{a}_1^T \lambda_1 a_1 = \lambda_1 \underbrace{\bar{a}_1^T a_1}_{=1} = \lambda_1$$

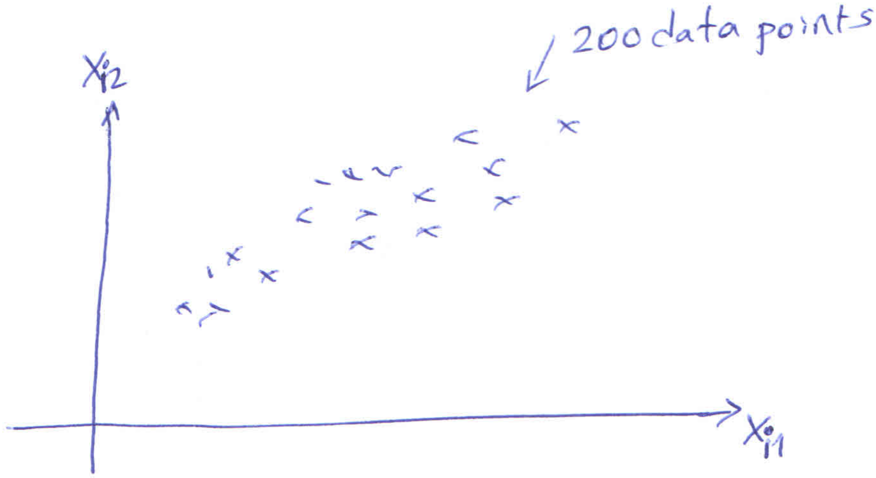
eigenvalue denklemi $\Rightarrow R\bar{a}_1 = \lambda_1 \bar{a}_1$

Benzer şekilde k. principal component

$\boxed{y_{k,i} = \bar{a}_k^T \bar{x}_i}$ → $y_{k,i}$'nin sample-varyansı $= \lambda_k$ olur.

Bu şekilde bir sinyalin varyansının çoğunluğu r tane principal komponent tarafından içerilirse, sinyalin kendisi yerine bu principal componentlar kullanılabilir.

örnek :



$\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_{200}$ → 200 tane datavektörü $\bar{x}_i = \begin{bmatrix} x_{i1} \\ x_{i2} \end{bmatrix}$

Sample-kovaryans matrisi

$$R = \frac{1}{200-1} \sum_{i=1}^{200} \bar{x}_i \bar{x}_i^T = \begin{bmatrix} 24.1893 & 10.6075 \\ 10.6075 & 6.38059 \end{bmatrix}$$

R'nin eigenvalue'ları $\lambda_1 = 29.1343$ $\lambda_2 = 1.4355$

" eigenvektörleri $\bar{a}_1 = [0.9064 \ \ 0.4225]^T$ $\bar{a}_2 = [-0.4225 \ \ 0.9064]^T$

⇒ random vektör $\bar{x}_i = [x_{i1} \ \ x_{i2}]^T$ için

1. principal component = $y_1 = \bar{a}_1^T x_i = 0.9064\, x_{i1} + 0.4225\, x_{i2}$

y_1 → skaler (1. principal component)

y_1 değeri $\bar{x}_i$ vektörünün varyansının 95%'ini barındırır.

Birçok istatistiksel amaç için $\bar{x}_i$ vektörünün iyi bir yaklaşımdır.