Makine Öğrenmesi

Feature selection, PCA, SVD, LDA ve kümeleme

Orijinal not: Ders7.pdf

UYARI Bu belge, el yazısı ders notlarından Claude (Anthropic) yardımıyla dijitale aktarılmıştır. Yardımcı kaynak olarak hazırlanmıştır; hata veya eksik içerebilir. Lütfen orijinal notları ve ders kaynaklarını esas alın.

NOTICE *This document was transcribed from handwritten lecture notes with the help of Claude (Anthropic). It is meant as a supplementary resource and may contain errors or omissions. Always check against the original notes and official course material.*

orj. s. 1

hiperspektral ? *görüntüleme*

220 dalga boyu ? *her* kanal

Verinin boyutunu azaltmak için

① Feature selection

$$\{x^t, r^t\}_{t=1}^N$$

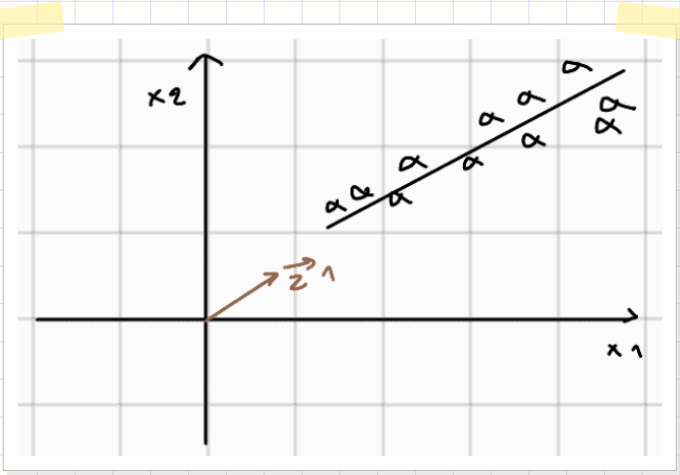
$$x^t_{N_1 \times 1} \rightarrow x^t_{N_2 \times 1} \qquad N_2 < N_1$$

Subset Selection

$$F = \emptyset$$

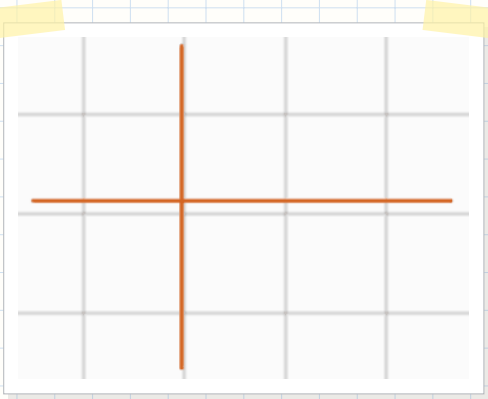
while not converge:
  for $x_i \notin F$
    $F_i \leftarrow F \cup x_i$
    $\mathcal{L}_i(x \mid F_i)$
  $i = \arg\min \mathcal{L}_i(x_i \mid F_i)$
  $F \leftarrow F_i$

Başka bir yöntem var. O da PCA



orijinal çizim, s. 1

regresyonu $\sum_{t=1}^{N} (\hat{r}^t - r^t)^2$



orijinal çizim, s. 1

$(Xz)^T(Xz) = z^T \underbrace{X^T X}_{\Sigma} z$

bunu max yapıyoruz

orj. s. 2

$$z_1 = \arg\max_z \underbrace{\left(z^T \Sigma z - \alpha(z_1^T z_1 - 1)\right)}_{\mathcal{L}(z|x)}$$

$$\frac{\partial \mathcal{L}}{\partial z} = 0$$

① $\boxed{\frac{\vec{x_i} - \vec{\mu}}{\Sigma}}$ ② [min, max]

$$2\Sigma z - 2\alpha z = 0$$
$$2\Sigma z = 2\alpha z$$

$$\boxed{\Sigma z = \alpha z}$$

$$z^T \Sigma z = z^T \alpha z$$
$$= \alpha z^T \cdot z = \alpha$$

$$z_1 \Sigma z_1 = \alpha_1^2 \qquad \lambda_1$$
$$z_2 \Sigma z_2 = \alpha_2^2 \qquad \lambda_2$$

$$x_{N\times 1}$$

$$z_2 = \arg\max_z \left(z^T \Sigma z - \alpha(z^T z - 1) - \beta(z^T \cdot z_1) \right)$$

$$z_1^T \;/\; 2\Sigma z - 2\alpha z - \beta z_1 = 0$$
$$2z_1^T \Sigma z - 2\alpha z_1^T z - \beta z_1^T z_1 = 0$$
$$2z_1^T \Sigma z - 2\alpha z_1^T \cdot z = \beta$$
$$2z_1^T \Sigma z - 2\alpha z_1^T z = \beta$$
$$\underbrace{2z_1^T \Sigma z}$$

orj. s. 3

Principle Component analizi

← Dimension reduction

$$N \rightarrow M$$

$$z_1, z_2, z_3, z_4, \ldots, z_M$$

$$x \cdot z_1 = \begin{bmatrix} y_1 \end{bmatrix}$$
$$x \cdot z_2 = \begin{bmatrix} y_2 \end{bmatrix}$$
$$x \cdot z_n = \begin{bmatrix} y_n \end{bmatrix}$$

$$\underbrace{\begin{bmatrix} x_1 & x_2 & x_3 & \cdots \end{bmatrix}}_{X^T} \qquad \begin{bmatrix} z_1 & z_2 & z_3 & \cdots & z_n \end{bmatrix}$$

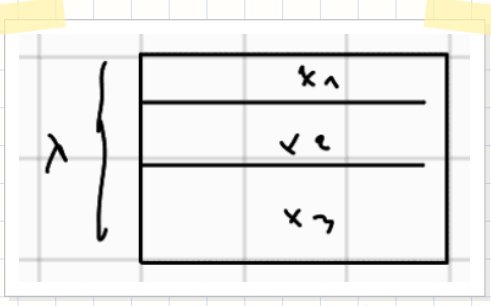
x.shape = (100, 200, 224)

x.view(-1, 224)
↑ kullanın

Haftaya perceptron bakacağız

Feature Embedding

$$WDW^T = X \cdot X^T$$



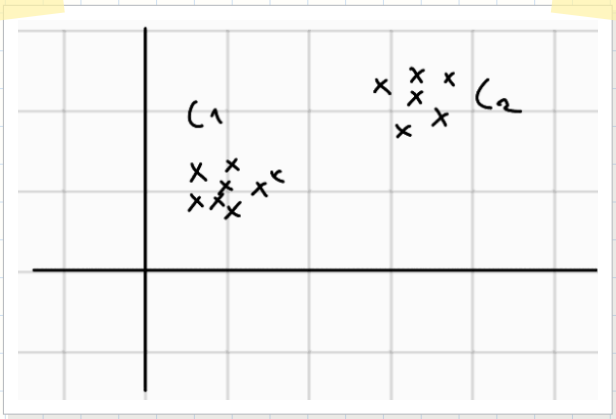
orijinal çizim, s. 3

Eigen faces

Singular value composition

$$\text{SVD}: \quad x = [x_1\ x_2\ \cdots\ x_N]_{N\times d}$$
$$= \underset{N\times f}{V}\ \underset{f\times f}{\Lambda}\ \underset{f\times d}{W^T}$$

orj. s. 4



orijinal çizim, s. 4

→ Linear Discriminant Analysis

$$z = w^T x$$

$$m_1 = \frac{\sum_t w^T x^t r^t}{\sum_t r^t} = w^T \cdot \vec{m}_1$$

$$m_2 = \frac{\sum_t w^T x^t (1-r^t)}{\sum_t (1-r^t)} = w^T \vec{m}_2$$

$$s_1^2 = \sum_t (w^T x^t - m_1)^2\, r^t$$

$$s_2^2 = \sum_t (w^T x^t - m_2)^2\, (1-r^t)$$

$$J(w) = \frac{(m_1 - m_2)^2}{s_1^2 + s_2^2} = \frac{(w^T\vec{m}_1 - w^T\vec{m}_2)^2}{s_1^2 + s_2^2}$$
$$= \frac{w^T(\vec{m}_1 - \vec{m}_2)^T(\vec{m}_1 - \vec{m}_2)\,w}{s_1^2 + s_2^2}$$
$$= \frac{w^T S_B\, w}{s_1^2 + s_2^2}$$

$$s_1^2 = \sum_t (w^T x^t - m_1)^2\, r^t$$
$$= w^T \cdot S_1\, w \qquad S_1 = \sum_t w^T (x^t - \vec{m}_1)(x^t - \vec{m}_1)^T r^t$$

$$s_1^2 + s_2^2 = w^T S_W w = \vec{w}(S_1^2 + S_2^2) w$$

$$J(w) = \frac{w^T S_B w}{w^T \cdot S_W w}$$

$$\frac{\partial J(w)}{\partial w} = 0$$

$$= \frac{w^T(\vec{m}_1 - \vec{m}_2)}{w^T S_W w} \; 2$$

orj. s. 5

Soft k means

$P(x \mid C_i)$ =?? esnek label

belli bir eşiğin altında ?kullanmayız

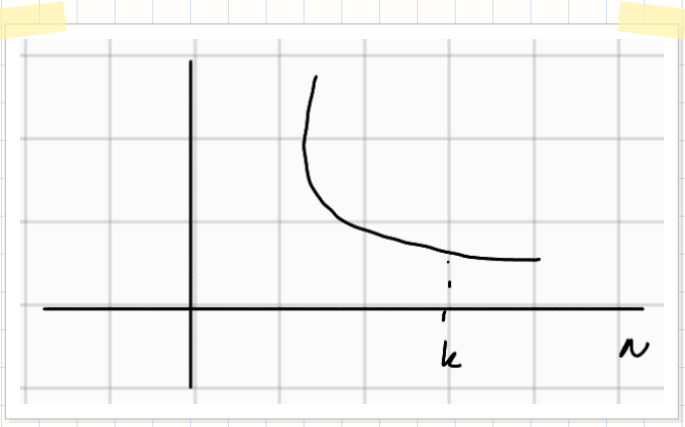
Expectation maximization



orijinal çizim, s. 5

hierarchical clustering

DB scan



orijinal çizim, s. 5