사전강의: Singular Value Decomposition. ← 이론은 이게 전부.

본강의: SVD 보충, PCA 본론.
- dimensional reduction 차원 축소.
- reduction / ~~reconstruction~~
  projection /

다음강의: PCA 실습

Inclass 20: Principal Component Analysis

[SCS4049] Machine Learning and Data Science

Seongsik Park (s.park@dgu.edu)

Department of Artificial Intelligence, Dongguk University

SVD : 행렬을 3단계로 분해.

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

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

$$\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \cdots)$$

$$x \longrightarrow \boxed{V^T} \xrightarrow{V^T x} \boxed{\Sigma} \xrightarrow{\Sigma V^T x} \boxed{U} \xrightarrow{U \Sigma V^T x}$$

$$y = Ax$$

① $\underset{D\times D}{V}^T \underset{D\times 1}{x} = \begin{bmatrix} - & v_1^T & - \\ - & v_2^T & - \\ - & v_3^T & - \\ & \vdots & \end{bmatrix} x = \begin{bmatrix} v_1^T x \\ v_2^T x \\ \vdots \end{bmatrix}$ $D \times 1$ vector.

$V = \begin{bmatrix} | & | & \\ v_1 & v_2 & \cdots \\ | & | & \end{bmatrix}$ $\quad V^T = \begin{bmatrix} - & v_1^T & - \\ - & v_2^T & - \\ & \vdots & \end{bmatrix}$

② $\underbrace{\Sigma V^T x}_{\not A} = \underbrace{\begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_D \end{bmatrix}}_{D\times D} \underbrace{\begin{bmatrix} v_1^T x \\ v_2^T x \\ \vdots \\ v_D^T x \end{bmatrix}}_{D\times 1} = \begin{bmatrix} \sigma_1 v_1^T x \\ \sigma_2 v_2^T x \\ \vdots \\ \sigma_D v_D^T x \end{bmatrix}$

Scalar 곱.

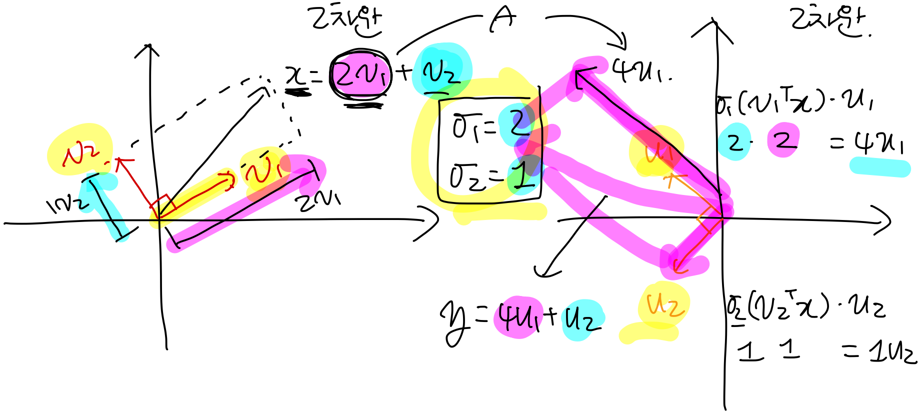
③

$$\underline{U}\,\underbrace{V\Sigma^T x} = \begin{bmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_D \\ | & | & & | \end{bmatrix} \begin{bmatrix} \boxed{\sigma_1 v_1^T x} \\ \boxed{\sigma_2 v_2^T x} \\ \vdots \\ \sigma_D v_D^T x \end{bmatrix}$$

D×D 행렬 D×1 벡터.

$$= \begin{matrix} | \\ u_1 \\ | \end{matrix} \cdot \sigma_1 v_1^T x + \begin{matrix} | \\ u_2 \\ | \end{matrix} \cdot \sigma_2 v_2^T x + \cdots + \begin{matrix} | \\ u_D \\ | \end{matrix} \cdot \sigma_D v_D^T x$$

$$= \overbrace{\sum_{d=1}^{D} \overbrace{\left(\sigma_d \underbrace{v_d^T x}_{①}\right)}^{②} \cdot \begin{matrix} | \\ u_d \\ | \end{matrix}}^{③} = \underline{A}\,\underline{x}$$



x가 사는 동네 ⟶ y가 사는 동네

$$\boxed{y = Ax}$$

2x1 2x2 2x1

① 주어진 벡터 x를 $v_1, v_2, \cdots, v_D$ 방향의

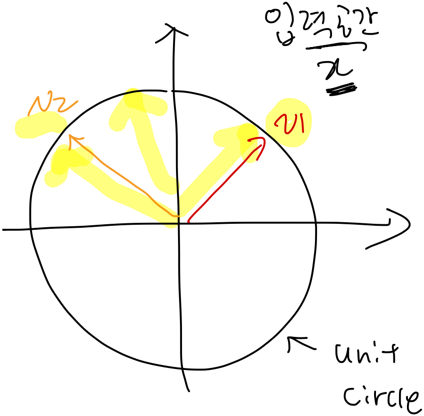
$V^T x$ 성분을 검사, 계산.

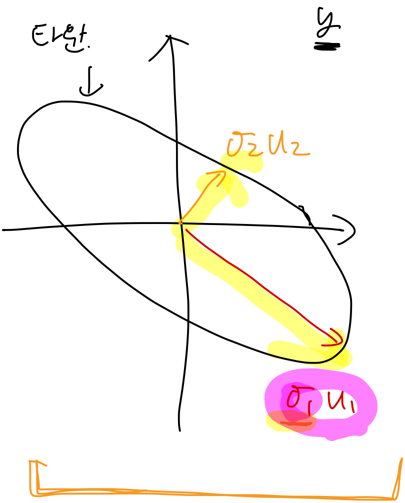
② 1번에서 구한 성분값 $(v_d^T x)$ 에다가

singular value σ_d를 곱함 $\sigma_d (v_d^T x)$.

③ 2번에서 구한 성분 x 배수에 u_d로 span.

$$\sum_{d=1}^{D} \sigma_d (v_d^T x) \cdot u_d$$


입력공간
x
v2
v1
unit
circle


타원
y
σ2u2
σ1u1

차원 축소.

- PCA
- non-negative matrix factorization (NNMF).

차원을 줄여야 할 일이 생김.

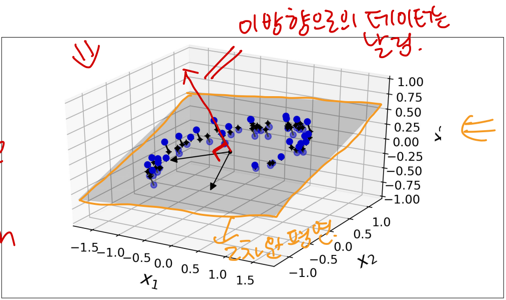
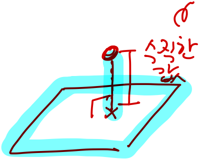


Figure 8-2. A 3D dataset lying close to a 2D subspace

3차원의 데이터.

2차원으로 축소



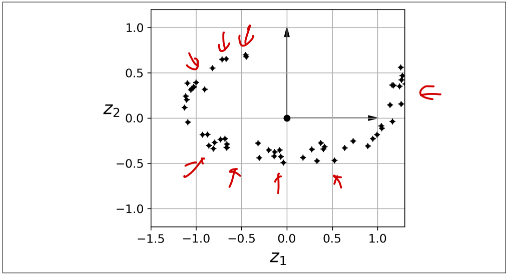
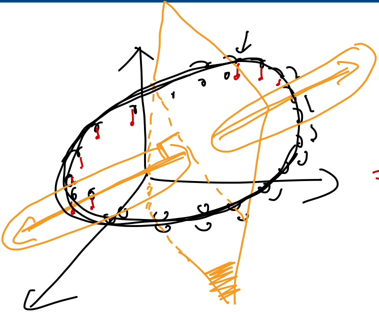
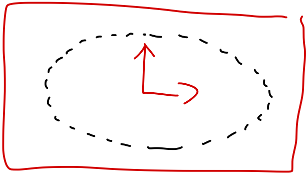


Figure 8-3. The new 2D dataset after projection

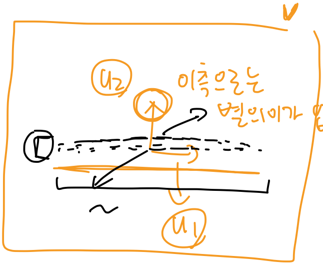


손실된정보,
손실된분포
= 평면에
수직한성분.

②.



①



②

Covariance matrix

Covariance measures the strength of the linear relationship between two variables

$$\sigma_{xy} = \mathrm{E}[(x - \mu_x)(y - \mu_y)] \tag{1}$$

Covariance matrix C for multivariate random variable X

$$C_{ij} = \mathrm{E}[(x_i - \mu_i)(x_j - \mu_j)] \tag{2}$$

Preserving the variance

Principal component analysis

주성분 = variance를 최대화 하는 축방향, 성분.

2차원 데이터. → 1차원 축소.

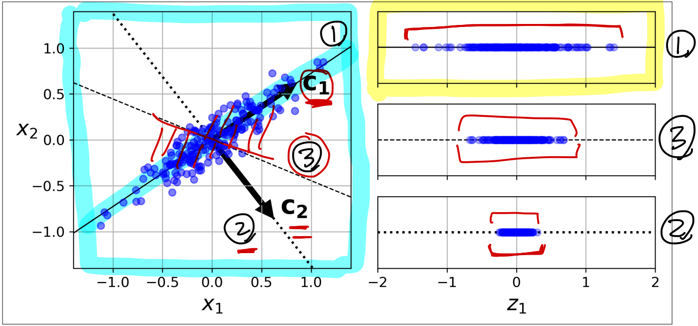


Figure 8-7. Selecting the subspace onto which to project

Principal component analysis (PCA)

= 분산이 큰 축방향이 어딘가요?

For given data $x_1, x_2, \ldots, x_N \in \Re^D$

1. create a matrix $\boxed{X \in \Re^{D \times N}}$ with one column vector per each sample

   D차원, N개의 sample.

2. covariance matrix $\boxed{C = \mathrm{E}\left[(X - \mathrm{E}(X))(X - \mathrm{E}(X))^T\right]} \in \Re^{D \times D}$
3. find singular vectors and singular values of C

   SVD 적용.

4. principal components = largest singular values and vectors

D차원 입력 → 1차원 축소: v_1

→ 2차원: v_1, v_2

→ 3차원: v_1, v_2, v_3

⋮

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

$$\begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_D \end{bmatrix}$$

$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_D$ 정렬되어서 나온다.

PCA = variance를 유지하는 축방향 찾기.

principal component. 주성분.

① covariance matrix를 구한다.

② cov. matrix에 SVD를 수행한다.

③ singular value가 큰 순서대로 V의 column vector를 골라간다.

= principal component.

cov.

$$C = U \Sigma V^T \qquad U = V$$

$$= \begin{bmatrix} \frac{\sqrt{3}}{2} & 0 & -\frac{1}{2} \\ \frac{1}{2} & 0 & \frac{\sqrt{3}}{2} \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 100 \end{bmatrix} \begin{array}{ccc} \begin{bmatrix} \frac{\sqrt{3}}{2} & 0 & -\frac{1}{2} \\ \frac{1}{2} & 0 & \frac{\sqrt{3}}{2} \\ 0 & 1 & 0 \end{bmatrix}^T \\ \begin{array}{ccc} v_1 & v_2 & v_3 \end{array} \end{array}$$

1차원 축소: $\begin{bmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \\ 0 \end{bmatrix}$

v_3

2개: $v_3 + v_2$.

PCA: "큰 순서로 골라라"

singular value.

V의 column vector.

SVD:

$$\begin{bmatrix} \frac{\sqrt{3}}{2} & 0 & -\frac{1}{2} \\ \frac{1}{2} & 0 & \frac{\sqrt{3}}{2} \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & & \\ & 10 & \\ & & 100 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{3}}{2} & 0 & -\frac{1}{2} \\ \frac{1}{2} & 0 & \frac{\sqrt{3}}{2} \\ 0 & 1 & 0 \end{bmatrix}^T$$

$$= \begin{bmatrix} -\frac{1}{2} & 0 & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 100 & & \\ & 10 & \\ & & 1 \end{bmatrix} \begin{bmatrix} -\frac{1}{2} & 0 & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix}^T$$

내림차순 정렬.

선택

① reduction, projection

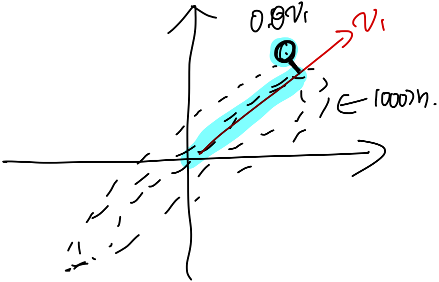
정렬됨.

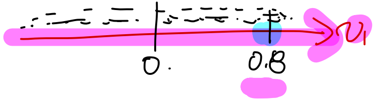
$\mathbb{X} \in \mathbb{R}^{2\times1000}$ $\quad$ $\mathbb{C} \in \mathbb{R}^{2\times2}$ $\quad$ $\mathbb{C} = U\Sigma \underline{V}^T$

$[v_1 \ v_2]^T$

1차원 : v_1

$$v_1^T \mathbb{X} = v_1^T \begin{bmatrix} x_1 & x_2 & \cdots & x_{1000} \end{bmatrix} = \begin{array}{ccc} 0.8 & 1.2 & \\ \begin{bmatrix} v_1^T x_1 & v_1^T x_2 \cdots & v_1^T x_{1000} \end{bmatrix} \end{array} \in \mathbb{R}^{1\times1000}.$$





$$\mathbb{X} \in \mathbb{R}^{3\times 1000} \qquad \mathbb{C} \in \mathbb{R}^{3\times 3} \qquad \mathbb{C} = U \Sigma V^T$$

정렬.

2차원 축소.

$$\boxed{\begin{bmatrix} | & | \\ v_1 & v_2 \\ | & | \end{bmatrix}^T} \mathbb{X} = \begin{bmatrix} \underline{v_1^T x_1} & v_1^T x_2 & v_1^T x_3 & \cdots \\ \underline{v_2^T x_2} & v_1^T x_2 & v_2^T x_3 & \end{bmatrix}$$

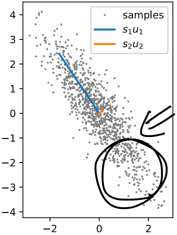
$$\in \mathbb{R}^{2\times 1000}$$

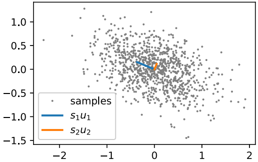
dimension reduction

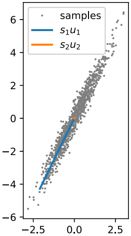
Principal component analysis (PCA)

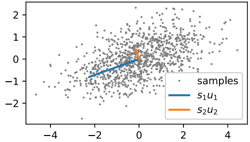
① 생체신호

② 이미지.









회색점들 = 데이터

⇓

covariance

⇓

SVD

⇓

V 벡터들

×

Σ 값들

(장축, 단축).

⇓

다시그림.

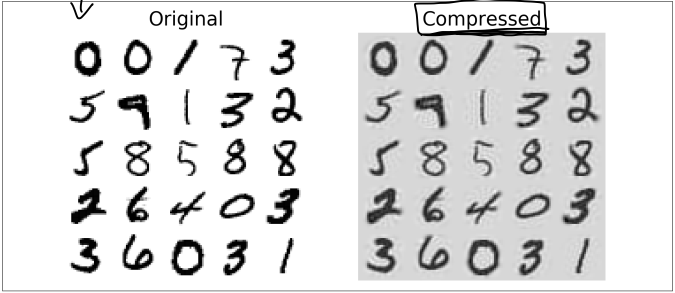


Figure 8-9. MNIST compression preserving 95% of the variance