

Inclass 19: Dimensional Reduction

[SCS4049] Machine Learning and Data Science

데이터 벡터 $\rightarrow$ 차원 줄이듯.

차원

축소.

주 성분 분석.
principal component
analysis (PCA)

{ PCA $\leftarrow$ SVD.
NMF $\Leftarrow$

nonnegative
factorization.

Seongsik Park (s.park@dgu.ac.kr)

School of AI Convergence & Department of Artificial Intelligence, Dongguk University

Principal component analysis (PCA)

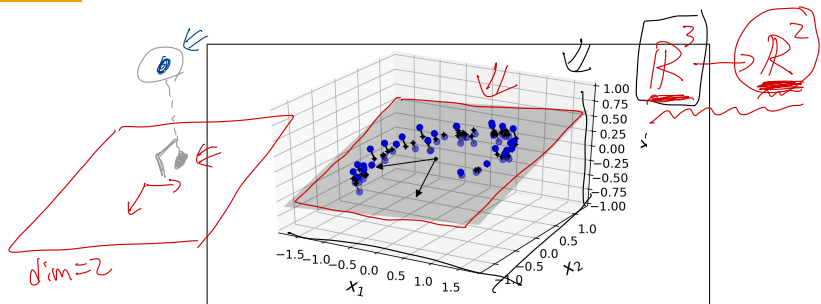


Figure 8-2. A 3D dataset lying close to a 2D subspace

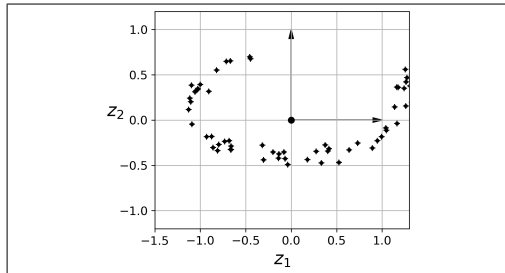


Figure 8-3. The new 2D dataset after projection

Principal component analysis (PCA)

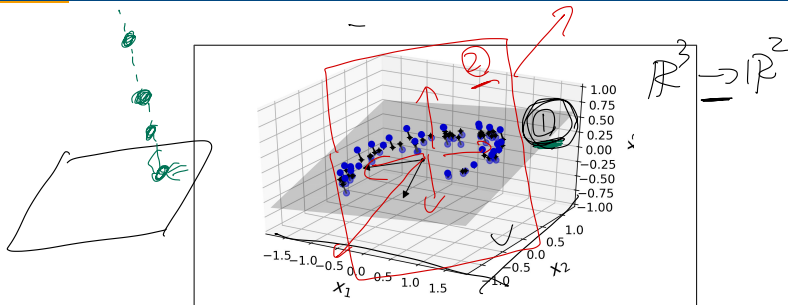


Figure 8-2. A 3D dataset lying close to a 2D subspace

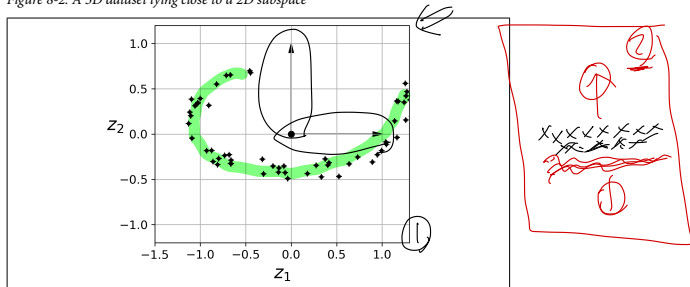


Figure 8-3. The new 2D dataset after projection

Principal component analysis (PCA)

Covariance measures the strength of the linear relationship between two variables

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

Covariance matrix $\mathbf{C}$ for multivariate random variable $\mathbf{x} \in \mathbb{R}^D$

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

$$C_{ij} = \frac{1}{N} \sum_{n=1}^N (x_i - \mu_i)(x_j - \mu_j)$$

Principal component analysis (PCA)

Preserving the variance

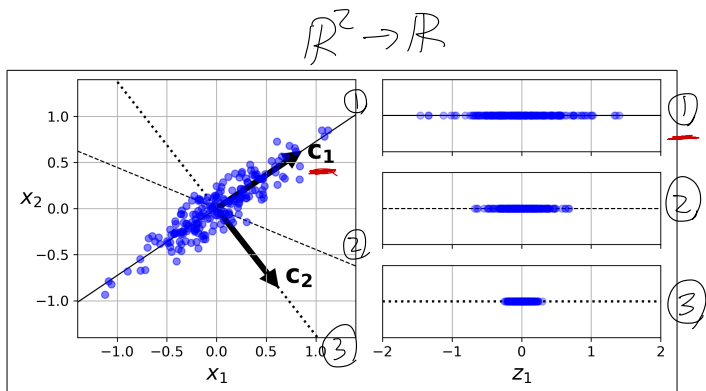


Figure 8-7. Selecting the subspace onto which to project

Principal component analysis (PCA)

For given data $\underline{x_1, x_2, \dots, x_N} \in \mathbb{R}^D$

1. create a matrix $\mathbf{X} \in \mathbb{R}^{D \times N}$ with one column vector per each sample
2. covariance matrix $\mathbf{C} = \mathbb{E}[(X - \mathbb{E}(X))(X - \mathbb{E}(X))^T] \in \mathbb{R}^{D \times D}$
3. find singular vectors and singular values of $\mathbf{C}$
4. principal components = largest singular values and vectors

$\Downarrow$ $\frac{1}{2\sqrt{2}}$

$$\mathbf{C} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T$$

Handwritten annotations: $\mathbf{U}$ is underlined in orange, $\mathbf{\Sigma}$ is circled in red, and $\mathbf{V}$ is underlined in green. Arrows point from the handwritten notes above to these terms.

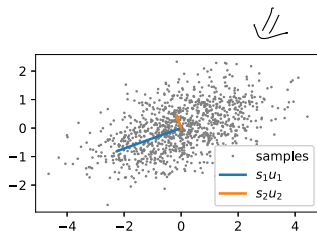
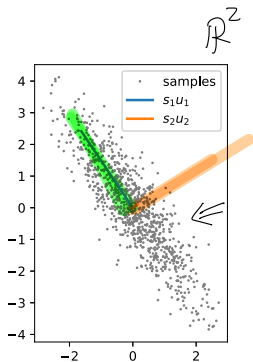
$$\mathbf{C} = \frac{1}{N} \sum (\underline{x - \mu})(\underline{x - \mu})^T$$

$$\mathbf{C} = \mathbf{C}^T$$

$$\mathbf{V} = \mathbf{V}$$

$$(\underline{x - \mu})$$

Principal component analysis (PCA)



$$x_1, \dots, x_{1000} \Rightarrow C \in \mathbb{R}^{2 \times 2}$$

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

Matrix factorization

- Principal component analysis (PCA): orthogonal property
- Vector quantization (VQ): unary property
- Non-negative matrix factorization (NMF): non-negativity

PCA: $\mathbf{X} \approx \mathbf{W}\mathbf{H}$

basis

VQ: $\mathbf{X} \approx \mathbf{W}\mathbf{H}$

NMF: $\mathbf{X} \approx \mathbf{W}\mathbf{H}$

where $\mathbf{W}^T \mathbf{W} = \mathbf{I}$ ← SVD

where $\mathbf{H}$ consists of unary vectors

where $W_{ij} \geq 0, H_{ij} \geq 0$

$\mathbf{X} \in \mathbb{R}^{D \times N}$

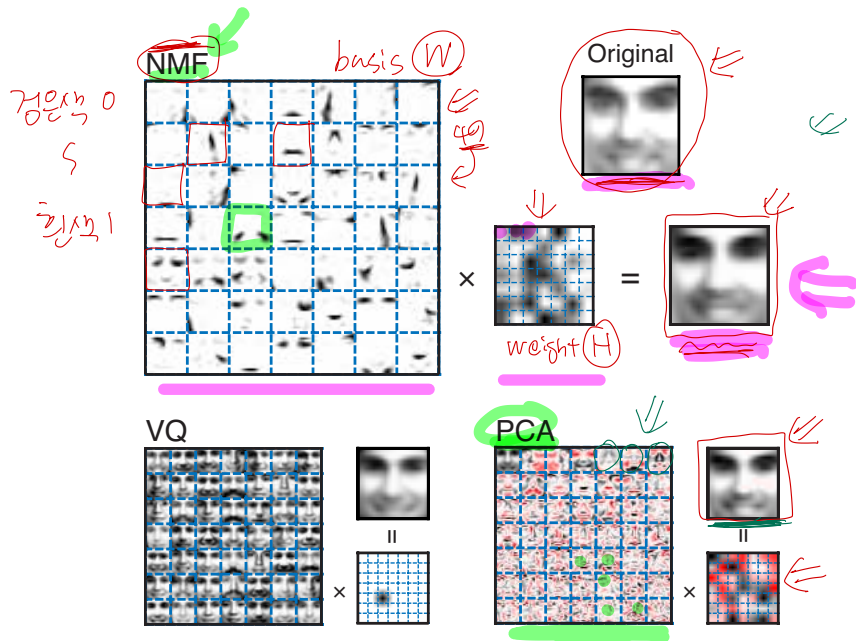
$\approx \mathbf{W}\mathbf{H}$

$\mathbf{X} = [\mathbf{x}_1 \mathbf{x}_2 \dots \mathbf{x}_N]$

$= [\mathbf{w}_1 \mathbf{w}_2 \dots \mathbf{w}_K] [\mathbf{h}_1 \mathbf{h}_2 \dots \mathbf{h}_N]$

DxD' D'xN

Matrix factorization and face image compression



Nonnegative matrix factorization (NMF)

Nonnegative matrix factorization (NMF)

optimization

given $\mathbf{X} \in \mathbb{R}^{n \times m}$ $D \times N$

find $(\mathbf{W}, \mathbf{H}) = \arg \min_{(\mathbf{W}, \mathbf{H})} \|\mathbf{X} - \mathbf{WH}\|_F^2$ Scalar

s.t. $\mathbf{W} \in \mathbb{R}_{+}^{n \times k}$ and $\mathbf{H} \in \mathbb{R}_{+}^{k \times m}$ ($k \leq \text{rank}(\mathbf{A})$)

$$\mathbb{R}_{+} = \text{양수.}$$

$$\underbrace{\mathbf{W}, \mathbf{H}} \stackrel{\text{nmf}}{=} \underbrace{\text{nmf}(\mathbf{X})}$$

PCA vs. NMF for MNIST dataset

