Skoltech

4.

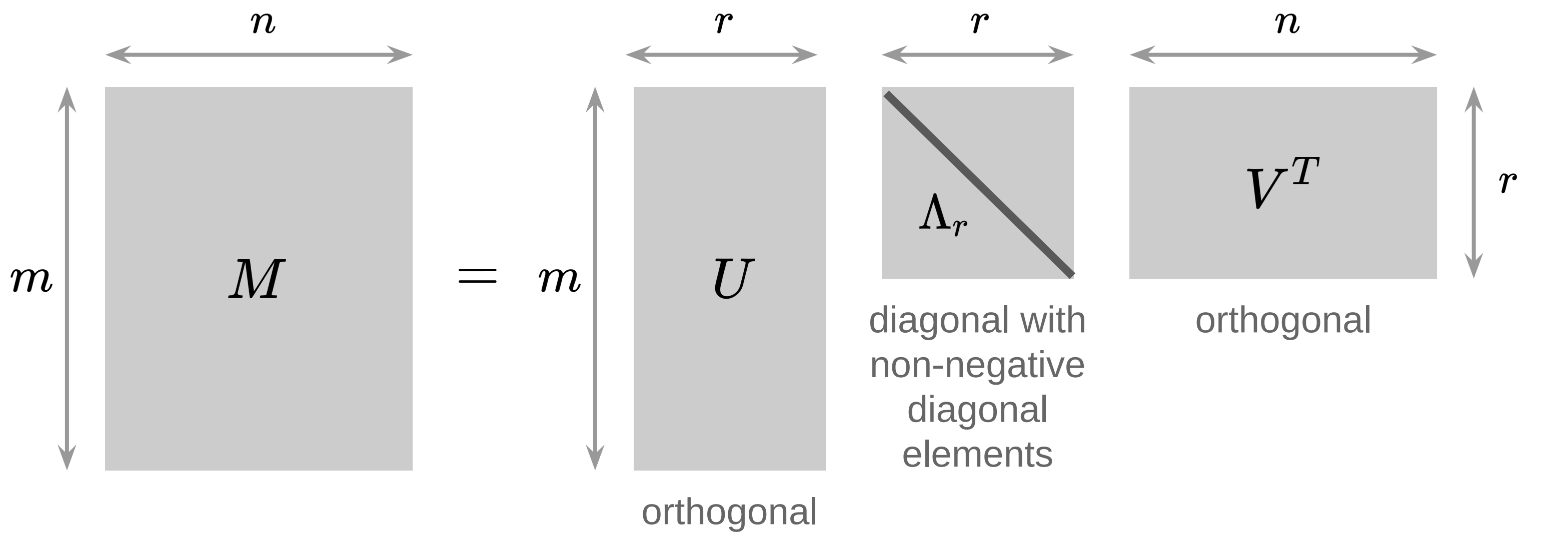
Principal Component Analysis from Singular Value Decomposition technique

May, 2023

Table of Content

Skoltech

Singular Value Decomposition (SVD)



$$Rank(M) = r \leq \min(m, n)$$

Skoltech

Derive SVD Form

Let $\overline{X} = \begin{pmatrix} \overline{X}_1 & \overline{X}_2 & \dots & \overline{X}_n \end{pmatrix}$ is p × n centered data matrix and p < n.

Apply SVD: $\overline{X} = U_p \times \Lambda_p \times V_p^T$

$$\begin{array}{c} \overline{X} \\ p \begin{pmatrix} \overline{X}_1 & \overline{X}_2 & & \dots & & \overline{X}_n \end{pmatrix} \\ n \end{array} = \begin{array}{c} U_p \\ \begin{pmatrix} e_1 & e_2 & \dots & e_p \end{pmatrix} \end{array} \times \begin{array}{c} \Lambda_p \\ \begin{pmatrix} \lambda_1 & \dots & 0 & & & \\ \dots & \dots & \dots & & 0 & \\ 0 & \dots & \lambda_q & & & \\ & & & \lambda_{q+1} & \dots & 0 \\ & 0 & & \dots & \dots & \dots \\ & & & 0 & \dots & \lambda_p \end{pmatrix} \end{array} \times \begin{array}{c} V_p^T \\ \begin{pmatrix} V_1^{(p)} & V_2^{(p)} & & \dots & & V_n^{(p)} \end{pmatrix} \end{array}$$

Orthogonal matrix with orthonormal vectors-columns $e_1, e_2, \dots, e_p \in R^p$

Diagonal matrix

Orthogonal matrix

Derive SVD Form

Covariance matrix : $$\Sigma = \frac{1}{n}\overline{X} \times \overline{X}^T = \frac{1}{n}\sum_{i=1}^{n}\left(\left(X_i - \overline{X}\right) \times \left(X_i - \overline{X}\right)^T\right)$$

From : $$\overline{X} = U_p \times \Lambda_p \times V_p^T$$

$$U_p^T \overline{X} = \Lambda_p \times V_p^T$$

transformed centered vectors

Write in column form : $$\Lambda_p \times V_i^{(p)} = U_p^T\left(X_i - \overline{X}\right) = Y_{p,i} = \begin{pmatrix} y_{i1} \\ \dots \\ y_{ip} \end{pmatrix}$$

Then, consider about Λ_p :

$$\Lambda_p = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \lambda_p \end{pmatrix} = \begin{pmatrix} \Lambda_q & 0 \\ 0 & \Lambda_{p-q} \end{pmatrix}$$

Derive SVD Form

We have : $$Y_{i,p} = \begin{pmatrix} y_{i,1} \\ \dots \\ y_{i,p} \end{pmatrix} = \begin{pmatrix} Y_{i,q} \\ Y_{i,p-q} \end{pmatrix}$$

where : $$Y_{i,q} = \begin{pmatrix} y_{i,1} \\ \dots \\ y_{i,q} \end{pmatrix} = (\Lambda_q \quad 0) \times V_i^{(p)} \in R^q$$

$$Y_{i,p-q} = \begin{pmatrix} y_{i,q+1} \\ \dots \\ y_{i,p} \end{pmatrix} = (0 \quad \Lambda_{p-q}) \times V_i^{(p)} \in R^{p-q}$$

Derive SVD Form

Let the eigenvalues are small : $\lambda_{q+1}, \lambda_{q+2}, \dots, \lambda_p \approx 0$

$$\Lambda_{p-q} = \begin{pmatrix} \lambda_{q+1} & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \lambda_p \end{pmatrix} \approx 0$$

So,
$$\begin{aligned} Y_{i,p-q} &= (0 \quad \Lambda_{p-q}) \times V_i^{(p)} \\ &\approx (0 \quad 0) \times V_i^{(p)} = 0 \end{aligned}$$

Transformed features:

$$Y_{i,p} = \begin{pmatrix} Y_{i,q} \\ Y_{i,p-q} \end{pmatrix} \approx \begin{pmatrix} Y_{i,q} \\ 0 \end{pmatrix}, \qquad i = 1, 2, \dots, n$$

Derive SVD Form

$$\overline{X} \approx U_q \times \Lambda_q \times V_q^T$$

$$\underset{p \times n}{\overline{X}} = \begin{bmatrix} \overline{X}_1 & \overline{X}_2 & \cdots & \overline{X}_n \end{bmatrix} \approx \underset{U_q}{\begin{bmatrix} e_1 & e_2 & \cdots & e_q \end{bmatrix}} \times \underset{\Lambda_q}{\begin{pmatrix} \lambda_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \lambda_q \end{pmatrix}} \times \underset{V_q^T}{\begin{bmatrix} V_1^{(p)} & V_2^{(p)} & \cdots & V_n^{(q)} \end{bmatrix}}$$

$p \times q$ orthogonal matrix spanned by $e_1, e_2, \ldots, e_q \in R^q$

Diagonal matrix

Orthogonal matrix

$$Y_{i,q} = \begin{pmatrix} y_{i1} \\ \dots \\ y_{iq} \end{pmatrix} = \Lambda_q \times V_i^{(q)} = U_q^T \times \left(X_i - \overline{X} \right) \in R^q, \quad i = 1,\, 2,\, \ldots,\, n$$

Skoltech

Derive SVD Form

Then : orthonormal vectors $e_1, e_2, \dots, e_q \in R^q$ are p-dimensional eigenvectors of p × p matrix and matrix Σ corresponding to q largest eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_q$ of the matrix Σ

$e_1, e_2, \dots, e_q \in R^q$ is the first q PCA components and $U_q = p \underbrace{\begin{bmatrix} e_1 & \dots & e_q \end{bmatrix}}_{q}$

$$\overline{X} \approx U_q \times Y_{i,q} \quad \Rightarrow \quad Y_{i,q} = U_q^T \times \left(X_i - \overline{X} \right) \in R^q$$

Reduced PCA - features

$$\hat{X} \approx \hat{X}_{PCA,i} = \overline{X} + U_q \times Y_{i,q} \in R^p , \quad i = 1,\, 2,\, \dots,\, n$$

PCA Steps

- Step 1 : based on the training dataset, PCA transform the original p-dimensional feature vector $X = \{x_1, x_2, \ldots, x_p\}$ to the vector consisting of new features $Y = \{y_1, y_2, \ldots, y_p\}$
- Step 2 : Removing the number of last transformed features which are non-informative respected to the chosen criteria

Skoltech

Rules of Thumb

Consider how many features should be retained:

- Base on at least the faction of P, for example, $P \sim 0.9,\ 0.95$

$$q(P) = \text{minimal}\, q : \frac{\sum_{k=1}^{q} \lambda_k}{\sum_{k=1}^{p} \lambda_k} \geq P$$

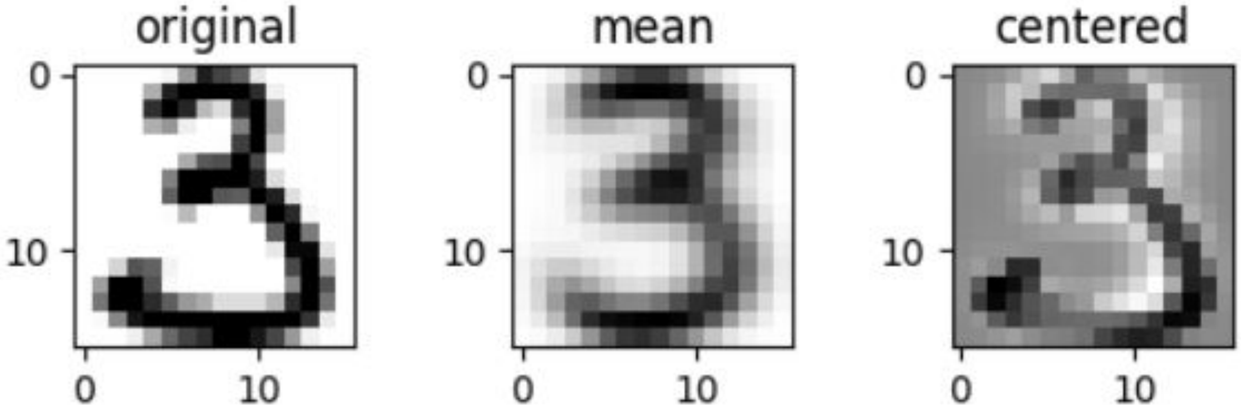
- Base on quantities

$$\left\{ \frac{\lambda_q}{\sum_{k=q+1}^{p} \lambda_k} \right\},\ \{\lambda_q - \lambda_{q+1}\},$$

Skoltech

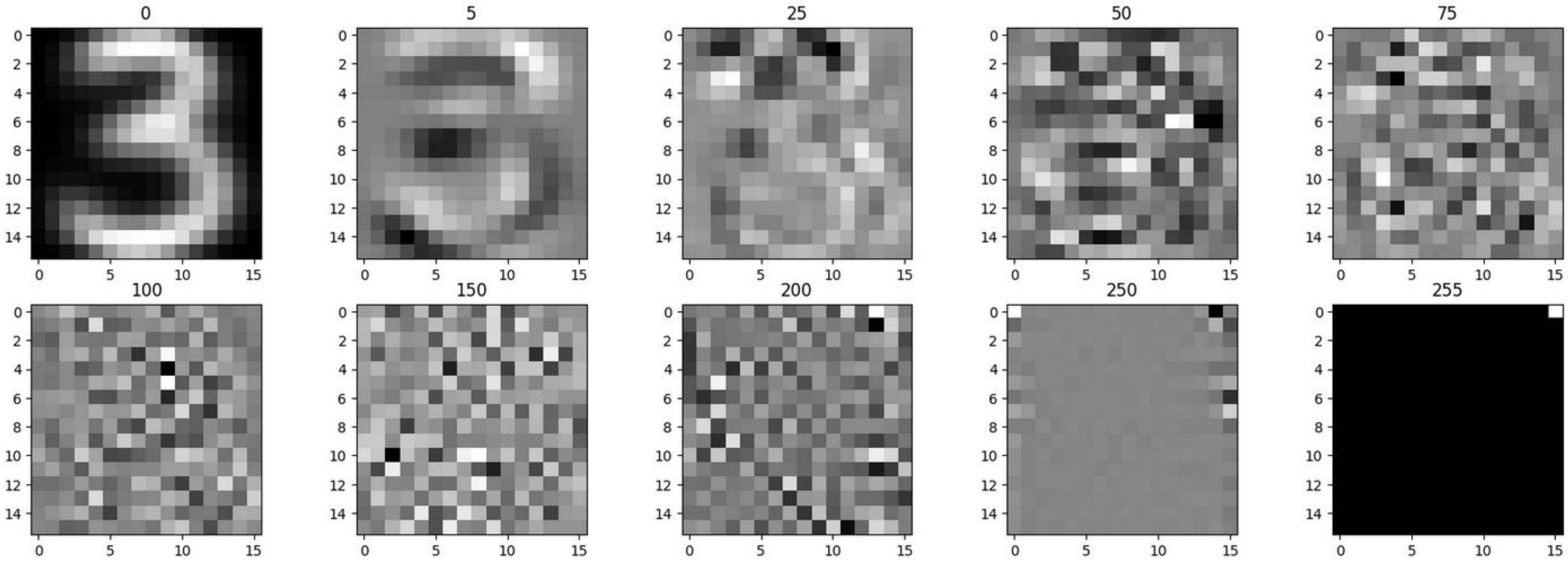
Example : Handwritten Digits USPS dataset

Example of original, mean, and centered data of digit 3



Example : Handwritten Digits USPS dataset

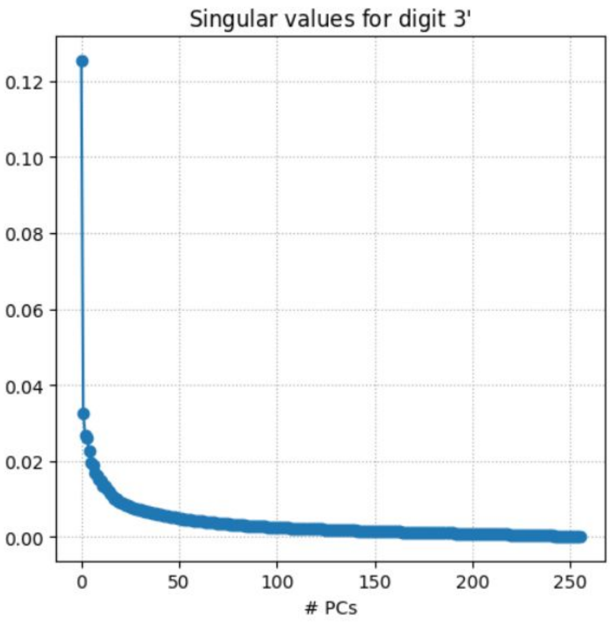
Principal digit 3 of components: 0, 5, 25, 50, 75, 100, 150, 200, 250, 255

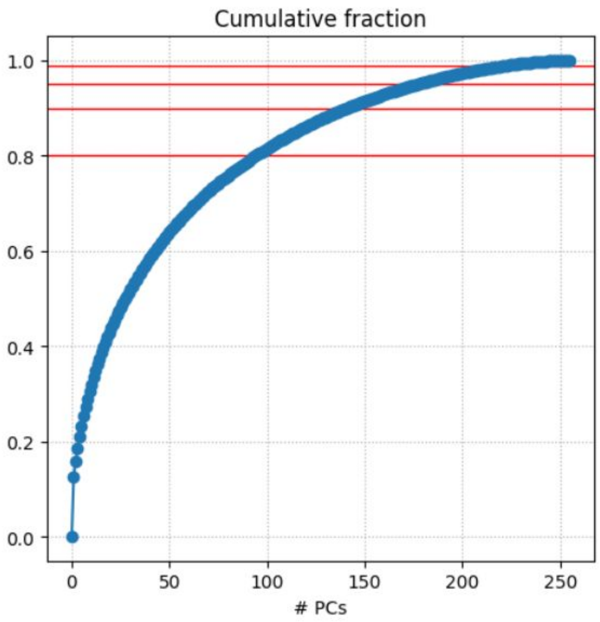


Skoltech

Example : Handwritten Digits USPS dataset

Singular values plot and its cumulative fraction

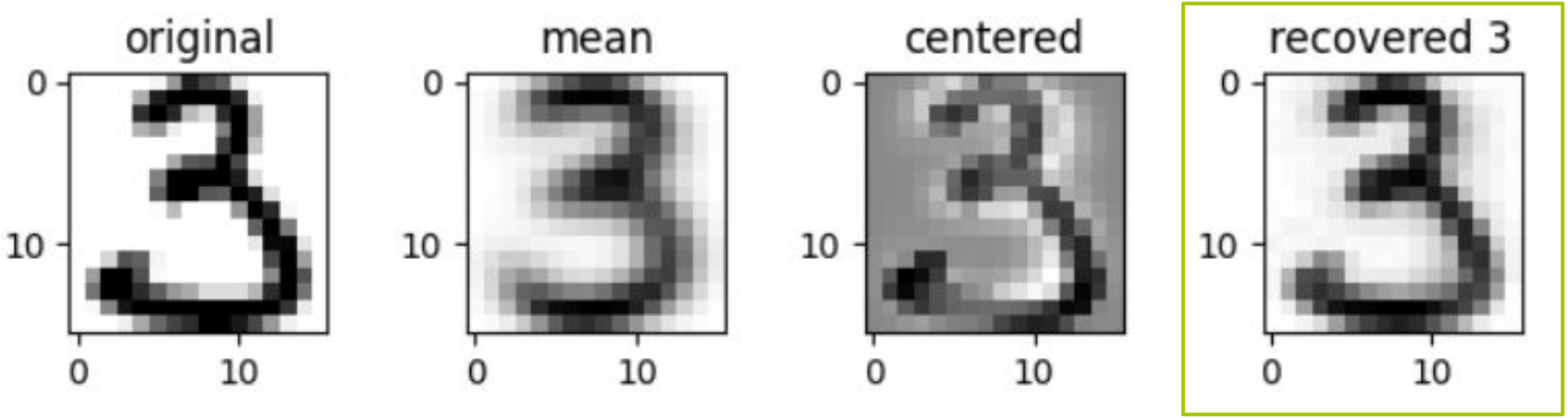




Cumulative fraction reach to 0.9501 at 177th-component

Example : Handwritten Digits USPS dataset

Example of original digit 3 and the recovered digit with
79-dimensional reduced feature vector



Skoltech