Análisis por Componentes principales (PCA)

Contenido

- PCA: Reducción de dimensiónalidad de los datos.

Contenido

- PCA: Reducción de dimensiónalidad de los datos.
- SVD.

Contenido

- PCA: Reducción de dimensiónalidad de los datos.
- SVD.
- Componentes principales y varianza.

Contenido

- PCA: Reducción de dimensiónalidad de los datos.
- SVD.
- Componentes principales y varianza.
- Ejemplo.

Contenido

- PCA: Reducción de dimensiónalidad de los datos.
- SVD.
- Componentes principales y varianza.
- Ejemplo.

Componentes principales

Matriz de m datos de n dimensiones:

$$\mathbf{A} = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_m^T \end{bmatrix}$$

Componentes principales

Matriz de m datos de n dimensiones:

$$\mathbf{A} = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_m^T \end{bmatrix}$$

- Filas de $\mathbf{A}$ son datos
  $\mathbf{x}_i^T = \begin{bmatrix} x_{i1} & x_{i2} & \dots & x_{in} \end{bmatrix}$

Componentes principales

Matriz de m datos de n dimensiones:

$$\mathbf{A} = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_m^T \end{bmatrix}$$

- Filas de $\mathbf{A}$ son datos $\mathbf{x}_i^T = \begin{bmatrix} x_{i1} & x_{i2} & \dots & x_{in} \end{bmatrix}$
- Columnas de $\mathbf{A}$ son descriptores (vectores en $\mathbb{R}^m$).

Componentes principales

Matriz de m datos de n dimensiones:

$$\mathbf{A} = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_m^T \end{bmatrix}$$

- Filas de $\mathbf{A}$ son datos $\mathbf{x}_i^T = \begin{bmatrix} x_{i1} & x_{i2} & \dots & x_{in} \end{bmatrix}$
- Columnas de $\mathbf{A}$ son descriptores (vectores en $\mathbb{R}^m$).

$$\mathbf{x} \in \mathbb{R}^n \longrightarrow \boxed{\mathbf{T} \in \mathbb{R}^{n \times l}} \longrightarrow \mathbf{y} = \mathbf{T}\mathbf{x} \in \mathbb{R}^l$$

$$l \ll n$$

Componentes principales

Matriz de m datos de n dimensiones:

$$\mathbf{A} = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_m^T \end{bmatrix}$$

- Filas de **A** son datos $\mathbf{x}_i^T = \begin{bmatrix} x_{i1} & x_{i2} & \dots & x_{in} \end{bmatrix}$
- Columnas de **A** son descriptores (vectores en $\mathbb{R}^m$).

$$\mathbf{x} \in \mathbb{R}^n \longrightarrow \boxed{\mathbf{T} \in \mathbb{R}^{n\times l}} \longrightarrow \mathbf{y} = \mathbf{T}\mathbf{x} \in \mathbb{R}^l$$

$$l \ll n$$

- Preservar mayoría de la información en los datos en vector de dimensión menor.

Componentes principales

Matriz de m datos de n dimensiones:

$$\mathbf{A} = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_m^T \end{bmatrix}$$

- Filas de $\mathbf{A}$ son datos $\mathbf{x}_i^T = \begin{bmatrix} x_{i1} & x_{i2} & \dots & x_{in} \end{bmatrix}$
- Columnas de $\mathbf{A}$ son descriptores (vectores en $\mathbb{R}^m$).

$$\mathbf{x} \in \mathbb{R}^n \longrightarrow \boxed{\mathbf{T} \in \mathbb{R}^{n \times l}} \longrightarrow \mathbf{y} = \mathbf{T}\mathbf{x} \in \mathbb{R}^l$$

$$l \ll n$$

- Preservar mayoría de la información en los datos en vector de dimensión menor.
- Nuevos descriptores?

- Centrar:

$$\mathbf{x}_i' = \mathbf{x}_i - \frac{1}{m}\sum_{j=1}^{m}\mathbf{x}_j$$

- Centrar:

$$\mathbf{x}_i' = \mathbf{x}_i - \frac{1}{m}\sum_{j=1}^{m} \mathbf{x}_j$$

- Matriz de covarianza:

$$\frac{1}{m}\mathbf{A}^T\mathbf{A} = \begin{bmatrix} \frac{1}{m}\sum_{j=1}^m x_{j1}^2 & \frac{1}{m}\sum_{j=1}^m x_{j1}x_{j2} & \cdots & \frac{1}{m}\sum_{j=1}^m x_{j1}x_{jn} \\ \frac{1}{m}\sum_{j=1}^m x_{j2}x_{j1} & \frac{1}{m}\sum_{j=1}^m x_{j2}^2 & \vdots & \\ \vdots & \vdots & \ddots & \\ \frac{1}{m}\sum_{j=1}^m x_{jn}x_{j1} & \cdots & & \frac{1}{m}\sum_{j=1} x_{jn}^2 \end{bmatrix}$$

Descomposición en valores singulares de $\mathbf{A} \in \mathbb{R}^{m \times n}$

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$$

Descomposición en valores singulares de $\mathbf{A} \in \mathbb{R}^{m\times n}$

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$$

$$= \underbrace{\begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_r \end{bmatrix}}_{\text{base span de columnas}} \underbrace{\begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & \vdots \\ \vdots & 0 & \ddots & \\ 0 & \cdots & & \sigma_r \end{bmatrix}}_{\text{valores singulares}} \left.\begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_r^T \end{bmatrix}\right\} \text{base span de filas}$$

Descomposición en valores singulares de $\mathbf{A} \in \mathbb{R}^{m\times n}$

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$$

$$= \underbrace{\begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_r \end{bmatrix}}_{\text{base span de columnas}} \underbrace{\begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & \vdots \\ \vdots & 0 & \ddots & \\ 0 & \cdots & & \sigma_r \end{bmatrix}}_{\text{valores singulares}} \left.\begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_r^T \end{bmatrix}\right\} \text{base span de filas}$$

- $\sigma_i = \sqrt{\lambda_i}$, donde λ_i son los valores propios de $\mathbf{A}^T\mathbf{A}$

Descomposición en valores singulares de $\mathbf{A} \in \mathbb{R}^{m\times n}$

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$$

$$= \underbrace{\begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_r \\ & & & \end{bmatrix}}_{\text{base span de columnas}} \underbrace{\begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & \vdots \\ \vdots & 0 & \ddots & \\ 0 & \cdots & & \sigma_r \end{bmatrix}}_{\text{valores singulares}} \left.\begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_r^T \end{bmatrix}\right\} \text{base span de filas}$$

- $\sigma_i = \sqrt{\lambda_i}$, donde λ_i son los valores propios de $\mathbf{A}^T\mathbf{A}$
- $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$

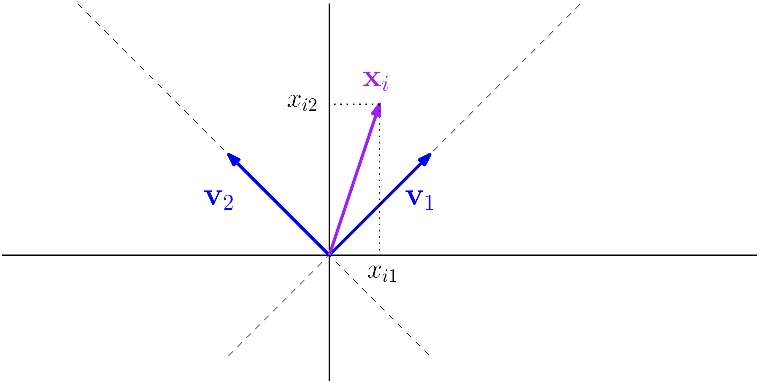
- $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$: vectores propios de $\mathbf{A}^T\mathbf{A}$.

- $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$: vectores propios de $\mathbf{A}^T\mathbf{A}$.
- Un dato:

$$\mathbf{x}_i = \underbrace{(\mathbf{x}_i^T \mathbf{v}_1)}_{c_1} \mathbf{v}_1 + \underbrace{(\mathbf{x}_i^T \mathbf{v}_2)}_{c_2} \mathbf{v}_2 + \cdots + \underbrace{(\mathbf{x}_i^T \mathbf{v}_r)}_{c_r} \mathbf{v}_r$$

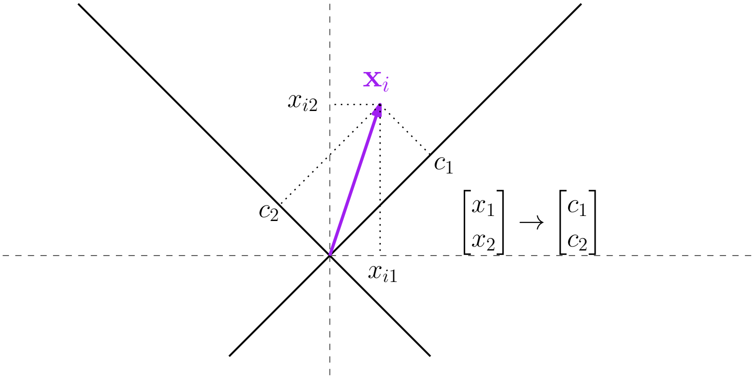
- $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$: vectores propios de $\mathbf{A}^T\mathbf{A}$.
- Un dato:

$$\mathbf{x}_i = \underbrace{(\mathbf{x}_i^T\mathbf{v}_1)}_{c_1}\mathbf{v}_1 + \underbrace{(\mathbf{x}_i^T\mathbf{v}_2)}_{c_2}\mathbf{v}_2 + \cdots + \underbrace{(\mathbf{x}_i^T\mathbf{v}_r)}_{c_r}\mathbf{v}_r$$



- $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$: vectores propios de $\mathbf{A}^T\mathbf{A}$.
- Un dato:

$$\mathbf{x}_i = \underbrace{(\mathbf{x}_i^T \mathbf{v}_1)}_{c_1} \mathbf{v}_1 + \underbrace{(\mathbf{x}_i^T \mathbf{v}_2)}_{c_2} \mathbf{v}_2 + \cdots + \underbrace{(\mathbf{x}_i^T \mathbf{v}_r)}_{c_r} \mathbf{v}_r$$



- Primer componente principal:

$$\mathbf{A}\mathbf{v}_1$$

- Primer componente principal:

$$\mathbf{A}\mathbf{v}_1 = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_m^T \end{bmatrix} \mathbf{v}_1$$

- Primer componente principal:

$$\mathbf{A}\mathbf{v}_1 = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_m^T \end{bmatrix} \mathbf{v}_1 = \begin{bmatrix} \mathbf{x}_1^T \mathbf{v}_1 \\ \mathbf{x}_2^T \mathbf{v}_1 \\ \vdots \\ \mathbf{x}_m^T \mathbf{v}_1 \end{bmatrix}$$

- Primer componente principal:

$$\mathbf{A}\mathbf{v}_1 = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_m^T \end{bmatrix} \mathbf{v}_1 = \begin{bmatrix} \mathbf{x}_1^T \mathbf{v}_1 \\ \mathbf{x}_2^T \mathbf{v}_1 \\ \vdots \\ \mathbf{x}_m^T \mathbf{v}_1 \end{bmatrix} = \sigma_1 \mathbf{u}_1$$

- Primer componente principal:

$$\mathbf{A}\mathbf{v}_1 = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_m^T \end{bmatrix} \mathbf{v}_1 = \begin{bmatrix} \mathbf{x}_1^T\mathbf{v}_1 \\ \mathbf{x}_2^T\mathbf{v}_1 \\ \vdots \\ \mathbf{x}_m^T\mathbf{v}_1 \end{bmatrix} = \sigma_1\mathbf{u}_1 \longrightarrow \text{Varianza } \frac{1}{m}\|\sigma_1\mathbf{u}_1\|^2$$

- Primer componente principal:

$$\mathbf{A}\mathbf{v}_1 = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_m^T \end{bmatrix} \mathbf{v}_1 = \begin{bmatrix} \mathbf{x}_1^T \mathbf{v}_1 \\ \mathbf{x}_2^T \mathbf{v}_1 \\ \vdots \\ \mathbf{x}_m^T \mathbf{v}_1 \end{bmatrix} = \sigma_1 \mathbf{u}_1 \longrightarrow \text{Varianza } \frac{1}{m} \|\sigma_1 \mathbf{u}_1\|^2 = \frac{\sigma_1^2}{m}$$

- Primer componente principal:

$$\mathbf{A}\mathbf{v}_1 = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_m^T \end{bmatrix} \mathbf{v}_1 = \begin{bmatrix} \mathbf{x}_1^T \mathbf{v}_1 \\ \mathbf{x}_2^T \mathbf{v}_1 \\ \vdots \\ \mathbf{x}_m^T \mathbf{v}_1 \end{bmatrix} = \sigma_1 \mathbf{u}_1 \longrightarrow \text{Varianza } \frac{1}{m} \|\sigma_1 \mathbf{u}_1\|^2 = \frac{\sigma_1^2}{m}$$

- Dos componentes principales:

- Primer componente principal:

$$\mathbf{A}\mathbf{v}_1 = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_m^T \end{bmatrix} \mathbf{v}_1 = \begin{bmatrix} \mathbf{x}_1^T \mathbf{v}_1 \\ \mathbf{x}_2^T \mathbf{v}_1 \\ \vdots \\ \mathbf{x}_m^T \mathbf{v}_1 \end{bmatrix} = \sigma_1 \mathbf{u}_1 \longrightarrow \text{Varianza } \frac{1}{m} \|\sigma_1 \mathbf{u}_1\|^2 = \frac{\sigma_1^2}{m}$$

- Dos componentes principales:

$$\mathbf{A}\mathbf{v}_1 + \mathbf{A}\mathbf{v}_2$$

- Primer componente principal:

$$\mathbf{A}\mathbf{v}_1 = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_m^T \end{bmatrix} \mathbf{v}_1 = \begin{bmatrix} \mathbf{x}_1^T \mathbf{v}_1 \\ \mathbf{x}_2^T \mathbf{v}_1 \\ \vdots \\ \mathbf{x}_m^T \mathbf{v}_1 \end{bmatrix} = \sigma_1 \mathbf{u}_1 \longrightarrow \text{Varianza } \frac{1}{m} \|\sigma_1 \mathbf{u}_1\|^2 = \frac{\sigma_1^2}{m}$$

- Dos componentes principales:

$$\mathbf{A}\mathbf{v}_1 + \mathbf{A}\mathbf{v}_2 = \sigma_1 \mathbf{u}_1 + \sigma_2 \mathbf{u}_2$$

- Primer componente principal:

$$\mathbf{A}\mathbf{v}_1 = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_m^T \end{bmatrix} \mathbf{v}_1 = \begin{bmatrix} \mathbf{x}_1^T \mathbf{v}_1 \\ \mathbf{x}_2^T \mathbf{v}_1 \\ \vdots \\ \mathbf{x}_m^T \mathbf{v}_1 \end{bmatrix} = \sigma_1 \mathbf{u}_1 \longrightarrow \text{Varianza } \frac{1}{m} \|\sigma_1 \mathbf{u}_1\|^2 = \frac{\sigma_1^2}{m}$$

- Dos componentes principales:

$$\mathbf{A}\mathbf{v}_1 + \mathbf{A}\mathbf{v}_2 = \sigma_1 \mathbf{u}_1 + \sigma_2 \mathbf{u}_2 \longrightarrow \text{Varianza } \frac{1}{m} \|\sigma_1 \mathbf{u}_1 + \sigma_2 \mathbf{u}_2\|^2$$

- Primer componente principal:

$$\mathbf{A}\mathbf{v}_1 = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_m^T \end{bmatrix} \mathbf{v}_1 = \begin{bmatrix} \mathbf{x}_1^T\mathbf{v}_1 \\ \mathbf{x}_2^T\mathbf{v}_1 \\ \vdots \\ \mathbf{x}_m^T\mathbf{v}_1 \end{bmatrix} = \sigma_1\mathbf{u}_1 \longrightarrow \text{Varianza } \frac{1}{m}\|\sigma_1\mathbf{u}_1\|^2 = \frac{\sigma_1^2}{m}$$

- Dos componentes principales:

$$\mathbf{A}\mathbf{v}_1 + \mathbf{A}\mathbf{v}_2 = \sigma_1\mathbf{u}_1 + \sigma_2\mathbf{u}_2 \longrightarrow \text{Varianza } \frac{1}{m}\|\sigma_1\mathbf{u}_1 + \sigma_2\mathbf{u}_2\|^2 = \frac{\sigma_1^2 + \sigma_2^2}{m}$$

- Primer componente principal:

$$\mathbf{A}\mathbf{v}_1 = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_m^T \end{bmatrix} \mathbf{v}_1 = \begin{bmatrix} \mathbf{x}_1^T \mathbf{v}_1 \\ \mathbf{x}_2^T \mathbf{v}_1 \\ \vdots \\ \mathbf{x}_m^T \mathbf{v}_1 \end{bmatrix} = \sigma_1 \mathbf{u}_1 \longrightarrow \text{Varianza } \frac{1}{m} \|\sigma_1 \mathbf{u}_1\|^2 = \frac{\sigma_1^2}{m}$$

- Dos componentes principales:

$$\mathbf{A}\mathbf{v}_1 + \mathbf{A}\mathbf{v}_2 = \sigma_1 \mathbf{u}_1 + \sigma_2 \mathbf{u}_2 \longrightarrow \text{Varianza } \frac{1}{m} \|\sigma_1 \mathbf{u}_1 + \sigma_2 \mathbf{u}_2\|^2 = \frac{\sigma_1^2 + \sigma_2^2}{m}$$

- $k \leq n$ componentes principales:

$$\mathbf{A}\mathbf{v}_1 + \cdots + \mathbf{A}\mathbf{v}_k = \sigma_1 \mathbf{u}_1 + \sigma_2 \mathbf{u}_2 + \cdots + \sigma_k \mathbf{u}_k$$

- Primer componente principal:

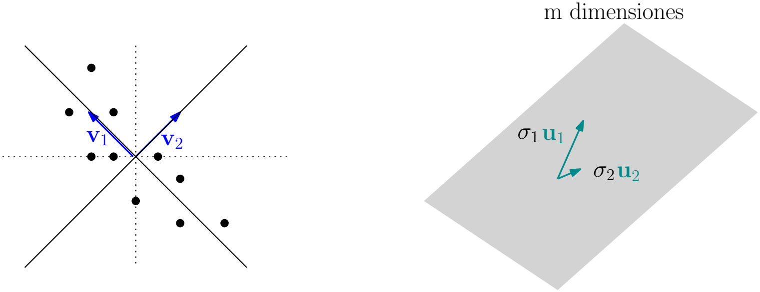
$$\mathbf{A}\mathbf{v}_1 = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_m^T \end{bmatrix} \mathbf{v}_1 = \begin{bmatrix} \mathbf{x}_1^T\mathbf{v}_1 \\ \mathbf{x}_2^T\mathbf{v}_1 \\ \vdots \\ \mathbf{x}_m^T\mathbf{v}_1 \end{bmatrix} = \sigma_1\mathbf{u}_1 \longrightarrow \text{Varianza } \frac{1}{m}\|\sigma_1\mathbf{u}_1\|^2 = \frac{\sigma_1^2}{m}$$

- Dos componentes principales:

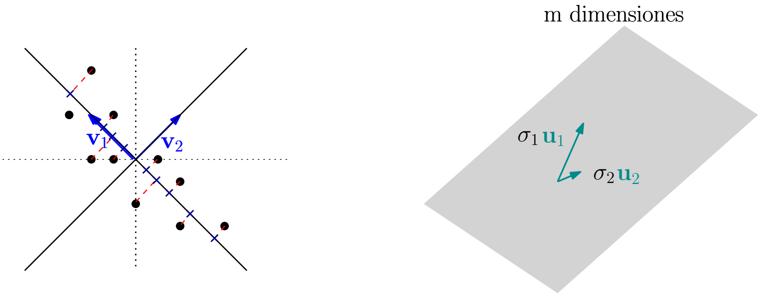
$$\mathbf{A}\mathbf{v}_1 + \mathbf{A}\mathbf{v}_2 = \sigma_1\mathbf{u}_1 + \sigma_2\mathbf{u}_2 \longrightarrow \text{Varianza } \frac{1}{m}\|\sigma_1\mathbf{u}_1 + \sigma_2\mathbf{u}_2\|^2 = \frac{\sigma_1^2 + \sigma_2^2}{m}$$

- $k \leq n$ componentes principales:

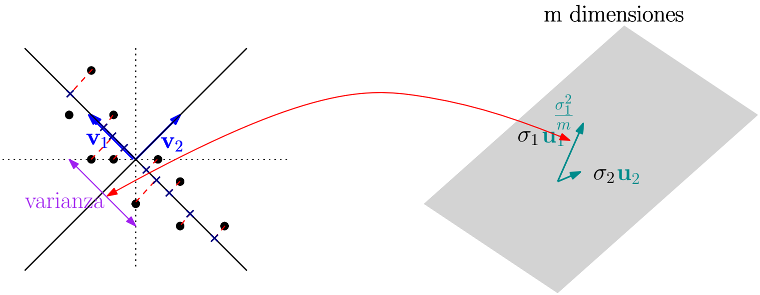
$$\mathbf{A}\mathbf{v}_1 + \cdots + \mathbf{A}\mathbf{v}_k = \sigma_1\mathbf{u}_1 + \sigma_2\mathbf{u}_2 + \cdots + \sigma_k\mathbf{u}_k \longrightarrow \text{Varianza } \frac{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_k^2}{m}$$



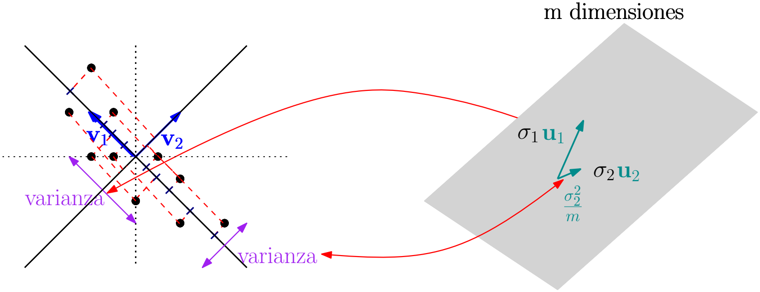
$\mathbf{v}_1$
$\mathbf{v}_2$
m dimensiones
$\sigma_1\,\mathbf{u}_1$
$\sigma_2\,\mathbf{u}_2$



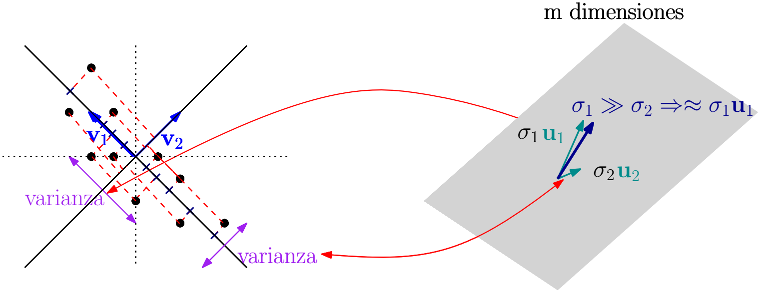
$\mathbf{v}_1$
$\mathbf{v}_2$
m dimensiones
$\sigma_1\,\mathbf{u}_1$
$\sigma_2\,\mathbf{u}_2$



m dimensiones
$\frac{\sigma_1^2}{m}$
$\sigma_1\,\mathbf{u}_1$
$\sigma_2\,\mathbf{u}_2$
$\mathbf{v}_1$
$\mathbf{v}_2$
varianza



m dimensiones
$\mathbf{v}_1$
$\mathbf{v}_2$
varianza
varianza
$\sigma_1\,\mathbf{u}_1$
$\sigma_2\,\mathbf{u}_2$
$\frac{\sigma_2^2}{m}$


m dimensiones
$\sigma_1 \gg \sigma_2 \Rightarrow \approx \sigma_1 \mathbf{u}_1$
$\sigma_1 \mathbf{u}_1$
$\sigma_2 \mathbf{u}_2$
$\mathbf{v}_1$
$\mathbf{v}_2$
varianza
varianza

- En general,

$$\sigma_1, \sigma_2, \ldots, \sigma_k \gg \sigma_{k+1}, \sigma_{k+2}, \ldots, \sigma_r$$

- En general,

$$\sigma_1, \sigma_2, \ldots, \sigma_k \gg \sigma_{k+1}, \sigma_{k+2}, \ldots, \sigma_r$$

$$\Downarrow$$

$$\sigma_1 \mathbf{u}_1 + \sigma_2 \mathbf{u}_2 + \ldots \sigma_k \mathbf{u}_k \approx \sigma_1 \mathbf{u}_1 + \sigma_2 \mathbf{u}_2 + \cdots + \sigma \mathbf{u}_k + \sigma_{k+1} \mathbf{u}_{k+1} + \cdots + \sigma_r \mathbf{u}_r$$

- En general,

$$\sigma_1, \sigma_2, \ldots, \sigma_k \gg \sigma_{k+1}, \sigma_{k+2}, \ldots, \sigma_r$$
$$\Downarrow$$
$$\sigma_1 \mathbf{u}_1 + \sigma_2 \mathbf{u}_2 + \ldots \sigma_k \mathbf{u}_k \approx \sigma_1 \mathbf{u}_1 + \sigma_2 \mathbf{u}_2 + \cdots + \sigma \mathbf{u}_k + \sigma_{k+1} \mathbf{u}_{k+1} + \cdots + \sigma_r \mathbf{u}_r$$

- Nueva matriz de datos:

$$\mathbf{Y} = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_m^T \end{bmatrix} \underbrace{\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \ldots & \mathbf{v}_k \end{bmatrix}}_{\mathbf{V}_k}$$

- En general,

$$\sigma_1, \sigma_2, \ldots, \sigma_k \gg \sigma_{k+1}, \sigma_{k+2}, \ldots, \sigma_r$$
$$\Downarrow$$
$$\sigma_1 \mathbf{u}_1 + \sigma_2 \mathbf{u}_2 + \ldots \sigma_k \mathbf{u}_k \approx \sigma_1 \mathbf{u}_1 + \sigma_2 \mathbf{u}_2 + \cdots + \sigma \mathbf{u}_k + \sigma_{k+1} \mathbf{u}_{k+1} + \cdots + \sigma_r \mathbf{u}_r$$

- Nueva matriz de datos:

$$\mathbf{Y} = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_m^T \end{bmatrix} \underbrace{\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \ldots & \mathbf{v}_k \end{bmatrix}}_{\mathbf{V}_k} = \begin{bmatrix} \mathbf{x}_1^T \mathbf{v}_1 & \mathbf{x}_1^T \mathbf{v}_2 & \ldots & \mathbf{x}_1^T \mathbf{v}_k \\ \mathbf{x}_2^T \mathbf{v}_1 & \mathbf{x}_2^T \mathbf{v}_2 & \ldots & \mathbf{x}_2^T \mathbf{v}_k \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{x}_m^T \mathbf{v}_1 & \mathbf{x}_m^T \mathbf{v}_2 & \ldots & \mathbf{x}_m^T \mathbf{v}_k \end{bmatrix}$$

PCA

1. Centrar datos.

PCA

1. Centrar datos.
2. Hallar valores propios de matriz $\mathbf{A}^T\mathbf{A}$: $\sigma_1^2 \geq \sigma_2^2 \geq \cdots \geq \sigma_n^2 > 0$.

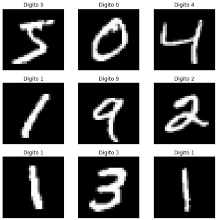
PCA

1. Centrar datos.
2. Hallar valores propios de matriz $\mathbf{A}^T\mathbf{A}$: $\sigma_1^2 \geq \sigma_2^2 \geq \cdots \geq \sigma_n^2 > 0$.
3. Escoger k que contenga porcentaje deseado de la varianza total.

PCA

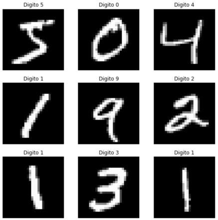
1. Centrar datos.
2. Hallar valores propios de matriz $\mathbf{A}^T\mathbf{A}$: $\sigma_1^2 \geq \sigma_2^2 \geq \cdots \geq \sigma_n^2 > 0$.
3. Escoger k que contenga porcentaje deseado de la varianza total.
4. Calcular proyección de los datos $\mathbf{y} = \mathbf{V}_k^T\mathbf{x}$

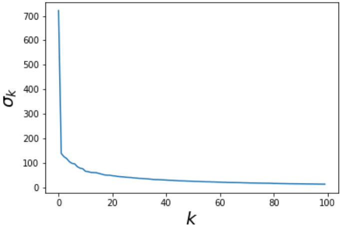
Ejemplo MNIST



- Imágenes $28 \times 28 = 784$ descriptores.
- 10000 imágenes.

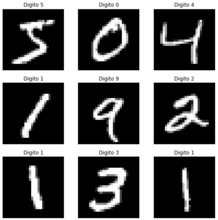
Ejemplo MNIST



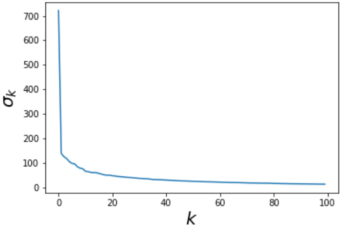


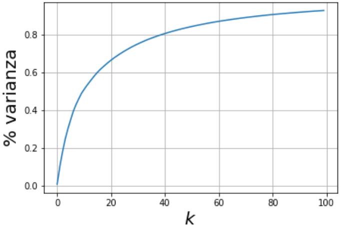
- Imágenes $28 \times 28 = 784$ descriptores.
- 10000 imágenes.

Ejemplo MNIST



- Imágenes $28 \times 28 = 784$ descriptores.
- 10000 imágenes.





Resumen

- PCA: Datos $\mathbf{A} \Rightarrow$ datos en menor dimensión $\mathbf{Y} = \mathbf{A}\mathbf{V}_k$

Resumen

- PCA: Datos $\mathbf{A} \Rightarrow$ datos en menor dimensión $\mathbf{Y} = \mathbf{A}\mathbf{V}_k$
- Porcentaje de varianza explicada.

Resumen

- PCA: Datos $\mathbf{A} \Rightarrow$ datos en menor dimensión $\mathbf{Y} = \mathbf{A}\mathbf{V}_k$
- Porcentaje de varianza explicada.
- Típicamente podemos usar $k \ll m$ componentes.