Projections & PCA

$$\vec{x} = a_1\vec{u}_1 + \cdots + a_r\vec{u}_r + \cdots + a_d\vec{u}_d$$
$$\vec{x}' = a_1\vec{u}_1 + \cdots + a_r u_r$$
$$= \begin{pmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_r \\ | & | & & | \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_r \end{pmatrix} = \underbrace{U_r}_{\text{first } r \text{ basis vectors of } U} \underbrace{a_r}_{\text{first } r \text{ coords of } \vec{a}}$$

$$\vec{a} = U^T\vec{x}$$
$$\vec{a}_r = U_r^T\vec{x}$$

$$x' = U_r a_r = \boxed{U_r U_r^T}\vec{x} = P_r\vec{x}$$

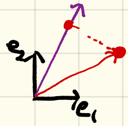
P_r = orogonal projetion matrix for the subspace of the first r basis vector of U

Props

- sym (wts $P_r = P_r^T$)

$$P_r^T = (U_r U_r^T)^T = U_r^{T^T} U_r^T = U_r U_r^T = P_r \Rightarrow P_r^T = P_r \checkmark$$

- show projection matrix (w.t.s. $P_r^2 = P_r$)



$$P_r^2 = (U_r U_r^T)^2 = (U_r U_r^T)(U_r U_r^T)$$
$$= U_r \underbrace{(U_r^T U_r)}_{I_{r \times r}} U_r^T = U_r U_r^T = P_r \Rightarrow P_r^2 = P_r$$

it is a projection matrix

Error vector

$$\varepsilon = \sum_{i=r+1}^{d} a_i \vec{u}_i = \vec{x} - \vec{x}'$$

x' & ε are ortogal

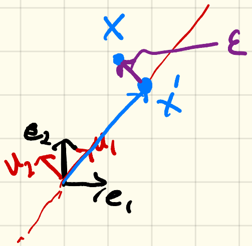
$$x' \cdot \varepsilon = \sum_{i=1}^{r} \sum_{j=r+1}^{d} a_i a_j \, u_i \cdot u_j = 0$$

$S_r = \text{span}(u_1, \ldots, u_r)$

$S_{d-r} = \text{span}(u_{r+1}, \ldots, u_d)$

S_r is orthogaal to S_{d-r}

S_{d-r} is orthogal complement of S_r



$$\vec{x}' \cdot \varepsilon = 0$$

$$\vec{x} = \begin{pmatrix} -.343 \\ -.754 \\ .241 \end{pmatrix}$$

project onto 1^{st} basis vector

$$u_1 = \begin{pmatrix} -.39 \\ .089 \\ -.916 \end{pmatrix}$$

$$\vec{a} = \begin{pmatrix} -.154 \\ .828 \\ -.190 \end{pmatrix}$$

Construct projection matrix

$$P_1 = \vec{u}_1 \vec{u}_1^T = \begin{pmatrix} -.39 \\ .089 \\ -.916 \end{pmatrix} \begin{pmatrix} -.39 & .089 & -.916 \end{pmatrix} =$$

$$\begin{pmatrix} 0.152 & -0.035 & 0.357 \\ -0.035 & 0.008 & -0.082 \\ 0.357 & -0.082 & 0.839 \end{pmatrix}$$

$$\vec{x}' = P_1 \vec{x} = \begin{pmatrix} 0.152 & -0.035 & 0.357 \\ -0.035 & 0.008 & -0.082 \\ 0.357 & -0.082 & 0.839 \end{pmatrix} \begin{pmatrix} -.343 \\ -.754 \\ .241 \end{pmatrix} = \begin{pmatrix} .060 \\ -.014 \\ .141 \end{pmatrix}$$

$$\varepsilon = a_2 u_2 + a_3 u_3 = \vec{x} - \vec{x}' = \begin{pmatrix} -.40 \\ -.74 \\ .1 \end{pmatrix}$$

Check orthogal

$$\mathbf{x}'^T \epsilon = \begin{pmatrix} 0.060 & -0.014 & 0.141 \end{pmatrix} \begin{pmatrix} -0.40 \\ -0.74 \\ 0.10 \end{pmatrix} = 0$$

PCA:

Find r-dim basis that "best captures" the variance of the data

- 1^{st} principal component captures largest projected variance
- 2^{nd} principal component captures 2^{nd} largest projected variance

⋮

directions that maxize projected variance minimizes mean squared error

Best 1D approx

$r = 1$

$\Rightarrow$ find a subsape (a line) ℓ s.t. projecting onto ℓ best approximates D

def data matrix centerd at mean

$$\bar{D} = D - \underbrace{\mathbb{1}}_{\text{vector of all 1's}}\mu^T$$

$\Rightarrow \bar{\mu}$ as the center of $\bar{D}$

$\bar{\mu} = 0$

want to find $\vec{u}$ define subspace

$\|\vec{u}\| = \vec{u} \cdot \vec{u} = \vec{u}^T \vec{u} = 1$

Let $\vec{u}$ be our vector defining subspace

Projectin a point

$\vec{x}_i \in \bar{D}$

$$\vec{x}_i' = \frac{\vec{u}\cdot x_i}{\vec{u}\cdot\vec{u}}\,\vec{u} = \underbrace{(\vec{u}\cdot x_i)}_{a_j}\vec{u} = a_i\vec{u}$$

$\vec{u}\cdot\vec{u}$ — 1

a_i — as projected point

$$a_i = \vec{u}\cdot\vec{x}_i = \vec{u}^T\vec{x}_i$$

$\bar{\mu} = 0 \Rightarrow$ mean of projected points is 0 $\quad \bar{\mu}_a = 0$

$$\bar{\mu}_a = \frac{1}{n}\sum_{i=1}^{n} a_i = \frac{1}{n}\sum_{i=1}^{n}\vec{u}^T\vec{x}_i = \vec{u}^T\underbrace{\left(\frac{1}{n}\sum_{i=1}^{n}\vec{x}_i\right)}_{\bar{\mu}} = u^T\cdot\vec{0} = \vec{0} \checkmark$$

Goal is to comput $\vec{u}$ s.t variance of the projected pts is maximized

Get a rep for variance

$$\sigma_u^2 = \frac{1}{n}\sum_{i=1}^{n}(a_i - \underset{0}{\mu_a})^2 = \frac{1}{n}\sum_{i=1}^{n} a_i^2 = \frac{1}{n}\sum_{i=1}^{n}(\vec{u}\cdot\vec{x}_i)^2$$

$\alpha^T = \alpha$ for a scalar

$$= \frac{1}{n}\sum_{i=1}^{n}(\vec{u}^T\vec{x}_i)(\vec{u}^T\vec{x}_i)^T = \frac{1}{n}\sum_{i=1}^{n}\vec{u}^T\vec{x}_i x_i^T\vec{u} = \vec{u}^T\underbrace{\left(\frac{1}{n}\sum_{i=1}^{n}\vec{x}_i\vec{x}_i^T\right)}_{\Sigma}\vec{u}$$

Σ cov matrix of $\bar{D}$

$$\Rightarrow \boxed{\sigma_u^2 = \vec{u}^T\Sigma\vec{u}}$$

$\Leftarrow$ we will come back to this