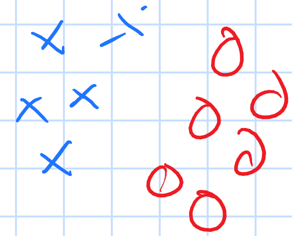
# Recap From 02/10/2021

## Principal Component Analysis

Goal: Project data to a lower dimensional space

$$\boxed{T : \mathbb{R}^D \mapsto \mathbb{R}^d}$$

• Linear



$D \geq d$ [usually $D > d$]

- 2 or 3D for plotting
- Assume that information is in the variance

## Set up

Write a vector as $w \in \mathbb{R}^D$

$$\hat{x}_k = m + \alpha_k w \quad \rightarrow \quad \hat{x}_k - m = \alpha_k w$$

Need to find $\alpha_k$

$$m = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \|w\|_2^2 = 1$$

$$w^T(x_k - m) = \alpha_k w^T w \quad (\|w\|_2^2)$$

$$\Rightarrow \alpha_k = w^T(x_k - m)$$

## Optimization Task [Distortion minimization]

$$J(w, \alpha_1, \ldots, \alpha_n) = \sum_{i=1}^{n} \|\underbrace{(m + \alpha_i w)}_{\hat{x}_i} - x_i\|_2^2$$

$(D\times 1)(1\times D)$ ... (after some reductions)

$$\propto -\sum_{i=1}^{n} w^T(x_i - m)(x_i - m)^T w$$

Def $S = \sum_{i=1}^{n} (x_i - m)(x_i - m)^T$ [scatter matrix]

$$\boxed{= -w^T S w}$$

$\boxed{[\text{Also } \|w\|_2^2 = 1]}$

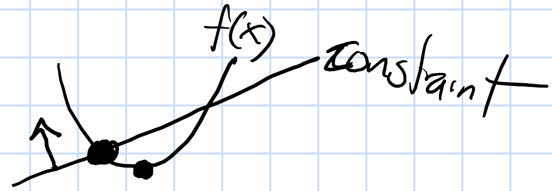
- Positive semi-Definite ($z^T S z \geq 0$)
- Symmetric

Goal

$$w^{*} = \arg\min_{w \in \mathbb{R}^D} \{ -w^T S w \} \qquad S = \sum_{i=1}^{n} (x_i - m)(x_i - m)^T$$

$$\text{s.t.} \quad \|w\|_2^2 - 1 = 0 \quad (w^T w = 1)$$

$$\boxed{\begin{aligned} &\arg\min J(w) \\ &\text{s.t.} \quad h_1(w) = 0 \end{aligned}}$$




$h_1(w) = \|w\|_2^2 - 1$

$$L(w, \lambda_1) = J(w) + \lambda_1 h_1(w)$$

Lagrangian

$\Rightarrow \frac{\partial L}{\partial w} = 0$ solve for $w$

$\frac{\partial L}{\partial \lambda_1} = 0$

$$L(w, \lambda) = -w^T S w + \lambda (w^T w - 1)$$

$$\frac{\partial L}{\partial w} = \boxed{-2Sw + 2\lambda w = 0}$$

$$h(w) = w^T w = w_1^2 + w_2^2 + w_3^2$$

$$[w_1 \ w_2 \cdots w_D] \begin{bmatrix} S_{11} & S_{12} & \cdots \\ \vdots & S_{22} & \\ \vdots & & \ddots \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_D \end{bmatrix}$$

$$\frac{\partial h}{\partial w_1} = 2w_1$$

$$\frac{\partial h}{\partial w_2} = 2w_2 \qquad \frac{\partial h}{\partial w} = 2w$$

$$\frac{\partial h}{\partial w_3} = 2w_3$$

$$\boxed{Sw = \lambda w}$$

What is $\lambda$? $w$?

$\lambda$ & $w$ are the eigenvalues and eigenvectors of $S$

Goal was to reduce the dimensionality

$w \to$ eigenvector of $S$

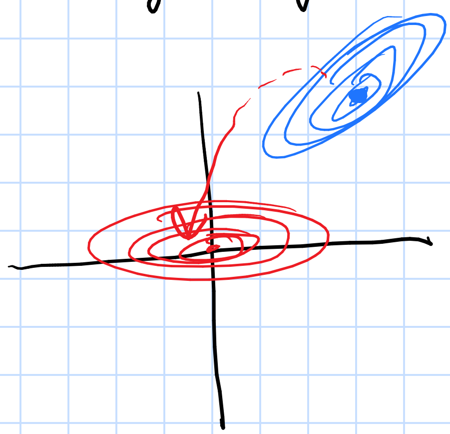
$\lambda \to$ eigenvalue of $S$

What is a property of the eigenvalue?

Choose the eigenvectors that have the largest eigenvalues

$$\bar{W} = [w_1 \;\; w_2] \in \mathbb{R}^{d \times 2}$$

$$z_k = \bar{w}^T (x_k - m)$$



$$\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_D \geq 0$$

Method: How much variation is retained after the projection

$$F(k) = \frac{\sum_{i=1}^{k} \lambda_i}{\sum_{j=1}^{D} \lambda_j}$$

$F(1) = 0.7$

$F(2) = 0.85$

$F(3) = 0.91$

90%