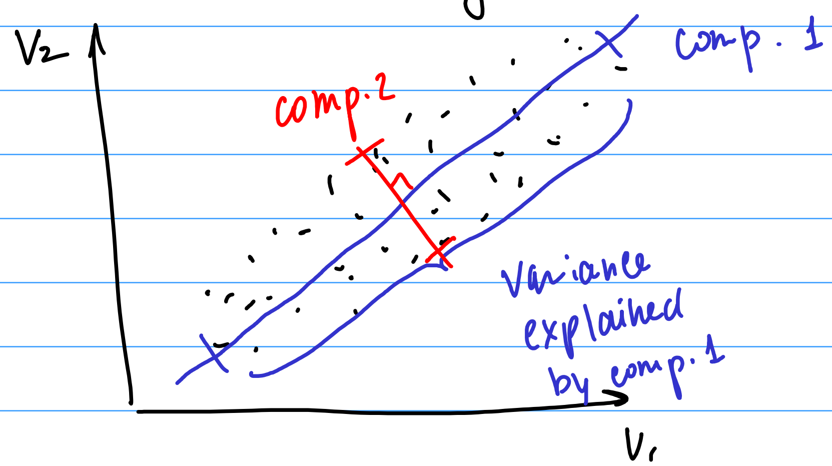
PCA

Variable (Dimension) Reduction

Goal - maximising variance of new components



$$z = Xu$$

$$\max_u \ \mathrm{Var}(z)$$

$$\text{s.t. } \|u\| = 1$$

PCA Algorithms

(I) Sequential algorithm

$\{x_1, \dots, x_m\}$ - centered data

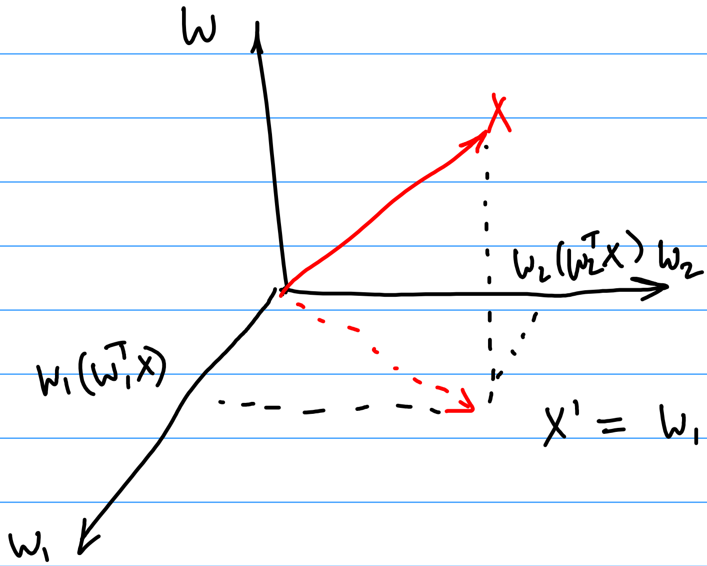
1) 1^{st} PCA component :

$$w_1 = \underset{\|w\|=1}{\operatorname{argmax}} \frac{1}{m} \sum_{i=1}^{m} (w^T x_i)^2$$

2) k^{th} PCA component:

$$w_k = \arg\max \frac{1}{m} \sum_{i=1}^{m} \Big(w^T \big(x_i - \underbrace{\sum_{j=1}^{k-1} w_j w_j^T x_i}_{} \big) \Big)^2$$

x' – PCA reconstruction



$$x' = w_1 (w_1^T x) + w_2 (w_2^T x)$$

(II) Sample Covariance matrix

$\{x_1, \ldots, x_m\}$

$$\Sigma = \frac{1}{m} \sum_{i=1}^{m} (x_i - \bar{x})(x_i - \bar{x})^T \qquad \bar{x} = \frac{1}{m} \sum x_i$$

Eigenvectors of Σ – PCA basis vectors

Eigenvalues of Σ – more important eigen vectors

$$u^T X X^T u \to \max$$

s.t. $u^T u = 1$

$$\mathcal{L} = u^T X X^T u - \lambda u^T u$$

$$\frac{\partial \mathcal{L}}{\partial u} = X X^T u - \lambda u = 0$$

$$(X X^T - \lambda \cdot \mathbb{1}) u = 0$$

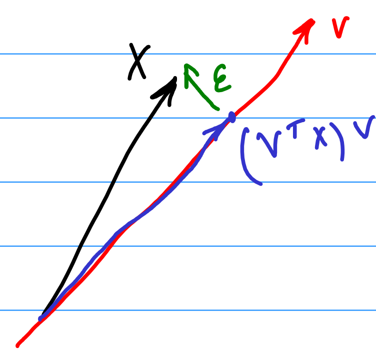
$u \neq 0 \Rightarrow u$ – eigenvector of XX^T

λ – eigenvalues

PCA ↗ max variance $\frac{1}{n} \sum_{i=1}^{n} (v^T x)^2 = v^T X X^T v$

↘ min reconstruction error

$$\frac{1}{n} \sum_{i=1}^{n} \| x_i - (v^T x_i) v \|^2$$

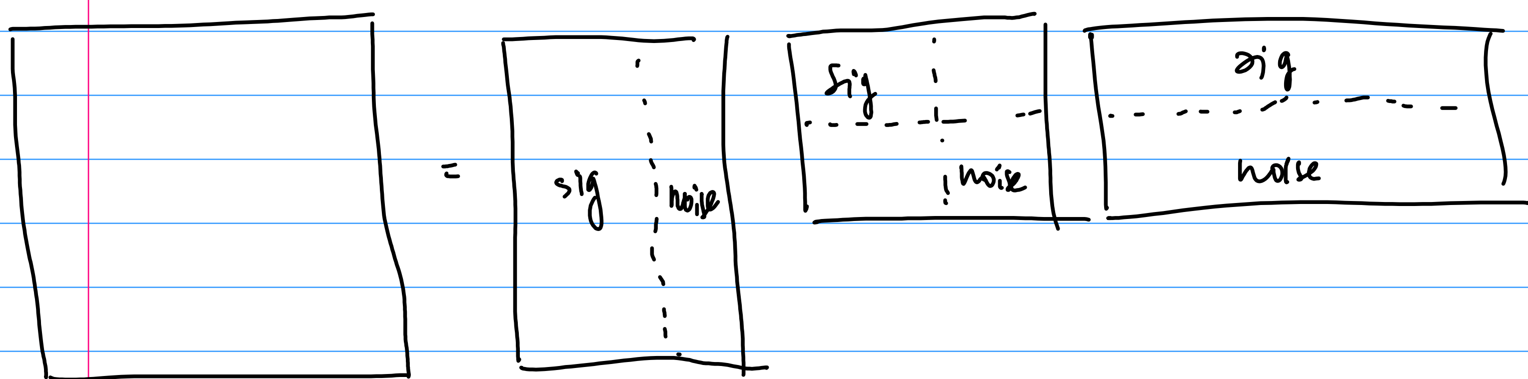


PCA Algorithm #3

III SVD

$$X = [x_1, \dots, x_m] \in \mathbb{R}^{N \times m}$$

$$X = U \cdot S \cdot V^T$$



Sample

Columns: principal vectors $\{u^{(1)}, \dots, u^{(k)}\}$

Diagonal: shows the importance of each vector

Columns: coefs to reconstruct sample x

VIF:

$$y_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_k \cdot x_{ki} + \varepsilon_i$$

1) $X_j \mid X_{-j} \Rightarrow R_j^2$

2) $$\mathrm{VIF}(\hat{\beta}_j) = \frac{1}{1 - R_j^2}$$

$\mathrm{VIF} > 5 \Rightarrow$ multicollinearity