Homework 2 : on website due next Fri 9^{th}

Latex : .tex

SVD singular value decomposition

$$\mathbb{K} = \underline{X}^T \underline{X} = \sum_{i=1}^{n} \vec{x}_i \vec{x}_i^{\,T} \Big\} \text{ outer product rank-1}$$

$(d\times 1)(1\times d) = (d\times d)$

$$\mathbb{E}\Big[\underbrace{\vec{x}\,\vec{x}}^{T}_{}\Big] = \mathbb{C} \quad \text{covariance matrix}$$

random, rank-1, $d \times d$ matrix

expectation

$$\mathbb{C}_{ij} = \mathbb{E}\left[x_i x_j\right]$$

linear regression $\longleftrightarrow$ linear stats.

entries random vector $\vec{x}$

SVD : Any $n \times d$ matrix X decomposes

$$\underset{(n\times d)}{X} = \underset{(n\times n)}{U}\ \underset{(n\times d)}{S}\ \underset{(d\times d)}{V^T}$$

U left singular vectors $\begin{bmatrix} | & | & & | \\ \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_n \\ | & | & & | \end{bmatrix}$ $\vec{u}_i \in \mathbb{R}^n$

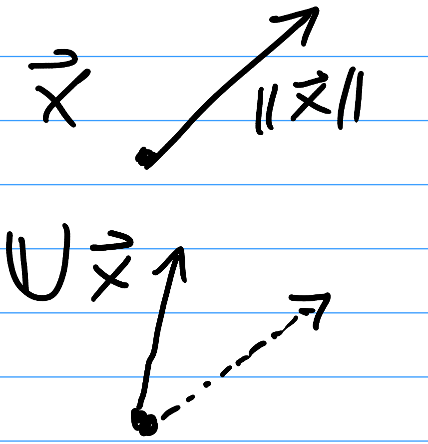
V right singular vectors $\begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_d \\ | & & | \end{bmatrix}$ $\vec{v}_i \in \mathbb{R}^d$

S diagonal $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\min(n,d)} \geq 0$

$n<d$ $[\sigma_1 \cdots \sigma_n | 0]$ $n>d$ $\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_d \\ \hline & 0 & \end{bmatrix}$

$$\mathbb{X} = U S V^T$$

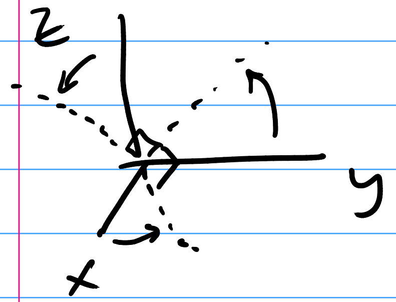
orthogonal matrices, rotating vectors



$$U^T U = I = U U^T$$

$$V^T V = I = V V^T$$

$$\vec{u}_i^{\,T} \vec{u}_j = \begin{cases} 1, & i=j \\ 0, & \text{o.w.} \end{cases}$$



$$\mathbb{X}\vec{v} = U \underset{\text{scale}}{S} \underbrace{V^T \vec{v}}_{\text{put } \vec{v} \text{ into new basis}}$$

$$\begin{bmatrix} 2 & & \\ & 1 & \\ & & 0.1 \end{bmatrix}$$

$\text{rank}(\mathbb{X}) = \#$ nonzero s.v.'s

$$X = U S V^T$$

Beware: numpy returns V^T not V

ex/ $\text{rank}(X) = 2$

$$X = \overbrace{\begin{bmatrix} | & & | \\ \vec{u}_1 & \cdots & \vec{u}_n \\ | & & | \end{bmatrix}} \overbrace{\begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & 0 & \\ & & & 0 \\ & & & & \ddots \end{bmatrix}} \begin{bmatrix} - & \vec{v}_1^T & - \\ - & \vec{v}_2^T & - \\ & \vdots & \\ - & \vec{v}_d^T & - \end{bmatrix}$$

$$= \underbrace{\sigma_1 \vec{u}_1 \vec{v}_1^T}_{} + \underbrace{\sigma_2 \vec{u}_2 \vec{v}_2^T}_{} + \cancel{0 \cdot \vec{u}_3 \vec{v}_3 + 0 \cdots} \rightarrow 0$$

rank-1 matrices

SVD = Unique sum of rank-1 matrices w/ othonormal $\vec{u}_i, \vec{v}_i$

Back to $\mathbb{K} = \mathbb{X}^T \mathbb{X}$

Eigenvalue of $\mathbb{K}$: $\mathbb{K}\vec{w} = \lambda \vec{w}$

only for square matrices!

d

eigenvalue decomposition

- normal matrix → diagonalizable
- symmetric → orthogonal

$$\mathbb{K} = \mathbb{W}\Lambda\mathbb{W}^{-1}$$

diagonalization

diagonal

powers of matrices

$$\mathbb{K}^2 = (W\Lambda W^{-1})(W\Lambda W^{-1})$$
$$= W\Lambda^2 W^{-1}$$

$$\mathbb{K} = \mathbb{X}^T\mathbb{X} = (VS^{\cancel{T}}U^T)(USV^T) = \mathbb{V}S\underbrace{U^TU}_{\text{rotate} \to \text{rotate back}}SV^T$$

I

$$= \mathbb{V}S^2\mathbb{V}^T$$

$$V^TV = I \Leftrightarrow \mathbb{V}^T = V^{-1}$$

Says:

For $\overbrace{X^T X}^{\text{symmetric}}$, right singular vectors V are eigenvectors of

$X^T X$ covariance matrix
$(d \times d)$ coordinate-by-coord.

$X X^T$, left singular vectors
$(n \times n)$ sample-wise covariance

$$X \vec{\beta} \approx \vec{y}$$
$(n \times d)$ $(d \times 1)$ $(n \times 1)$

smallest $\|\vec{\beta}\|$

$n < d$: underdetermined case
$\Rightarrow \exists \vec{\beta}$ that solves if many

BUT w/ noise probably not true $\vec{\beta}$

$n = d$: only one $\vec{\beta}$

$n > d$: overdetermined

solves normal eqns

$$X^T X \vec{\beta} = X^T \vec{y}$$
$$V^T V S^2 V^T \vec{\beta} = V^T V S U^T \vec{y}$$
$$(S^+)^2 S^2 V^T \vec{\beta} = (S^+)^2 S U^T \vec{y}$$
$$V V^T \vec{\beta} = V S^+ U^T \vec{y}$$

$$S^+ = \begin{bmatrix} 1/\sigma_1 & & & \\ & \ddots & & \\ & & 1/\sigma_r & \\ & & & 0 \end{bmatrix}$$

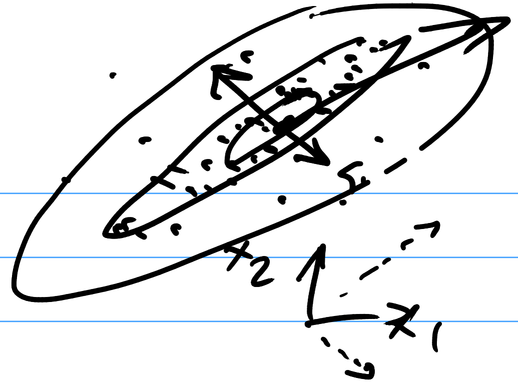
pseudo inverse

$\frac{1}{0}$ isn't defined

Pseudoinverse solution

$$\vec{\beta} = V S^+ U^T \vec{y} = \sum_{i=1}^{\text{rank}(X)} \frac{\vec{v}_i \vec{u}_i^T \vec{y}}{\sigma_i}$$

$$* \quad \vec{\beta} = \sum_{i=1}^{\text{rank}} \vec{v}_i \left(\frac{\overbrace{\vec{u}_i^{\,T} \vec{y}}^{\text{scalar}}}{\sigma_i} \right)$$



$$V S^+ U^T = X^+$$

Small singular values affect solution a <u>lot</u>

Recipe for least squares via pseudoinverse

1) compute SVD
2) use formula *

<u>Next time</u>:
- More interpretation
- PCA