The SVD

Let A be an $m \times n$ matrix. Then A can always be written

$$A = U \Sigma V^*$$

where

$U \in \mathbb{C}^{m \times m}$ $\quad UU^* = I \quad$ (unitary)

$V \in \mathbb{C}^{n \times n}$ $\quad VV^* = I \quad$ (")

Notation: let $U = [u^{(1)} | \dots | u^{(m)}]$

$V = [v^{(1)} | \dots | v^{(n)}]$

$$AV = U\Sigma$$

$$Av^{(i)} = \sigma_i u^{(i)}$$

$$A^T u^{(i)} = \sigma_i v^{(i)}$$

$\{u^{(i)}, v^{(i)}, \sigma_i\}$ triplets.

If A is tall

$$\underset{m\times n}{\begin{bmatrix} A \end{bmatrix}} = \underset{m\times m}{\begin{bmatrix} | & | & & | \\ u^{(1)} & u^{(2)} & \cdots & u^{(m)} \\ | & | & & | \end{bmatrix}} \underset{m\times n}{\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix}} \underset{n\times n}{\begin{bmatrix} - & v^{(1)T} & - \\ - & v^{(n)T} & - \end{bmatrix}}$$

$$= \begin{bmatrix} \sigma_1 u^{(1)} + 0u^{(2)} + \cdots + 0u^{(m)} & \cdots & 0u^{(1)} + \cdots + \sigma_n u^{(n)} + 0 \cdots 0 \\ \text{first col of product } U\Sigma & & n\text{'th col of product} \end{bmatrix} \begin{bmatrix} - & v^{(1)T} & - \\ & \vdots & \\ - & v^{(n)T} & - \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_1 u_1 | \cdots | \sigma_n u_n \end{bmatrix} \begin{bmatrix} - & v^{(1)T} & - \\ - & v^{(n)T} & - \end{bmatrix}$$

$$A = \sigma_1 u^{(1)} v^{(1)T} + \cdots + \sigma_n u^{(n)} v^{(n)T}$$

In particular, $A_i = \sigma_1 \underline{u}^{(1)} v_i^{(1)T} + \cdots + \sigma_n \underline{u}^{(n)} v_i^{(n)T}$

$$A_i = \sum_{j=1}^{n} \sigma_j \underline{u}^{(j)} v_i^{(j)T}$$

Thin SVD

$$A = \hat{U} \hat{\Sigma} \hat{V}^T.$$

Establishing SVD?

Connect to the eigenvector problems AA^*, A^*A.

Constructive proof:

$$A = U \Sigma V^* \implies A A^* = U \Sigma V^* V \Sigma^* U^*$$
$$= U \underbrace{\Sigma \Sigma^*}_{m \times m} U^*$$

But $(A A^*)^* = A^{**} A^*$
$$= A A^* \quad \text{s.a.}$$

So there exist real eigenvalues $\lambda_1, \dots, \lambda_n$

$$\underset{m \times m}{\Lambda} = \begin{bmatrix} \lambda_1 & & & & \\ & \ddots & & & \\ & & \lambda_n & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix}$$

eigenvectors: $[u^{(1)} | \dots | u^{(n)}]$

Same argument for $A^* A = V \underset{n \times n}{\Lambda} V^*$

$$V = [v^{(1)} | \dots | v^{(n)}]$$

The Snapshot Method.

$$C = \frac{1}{P} \sum_{\mu}^{1} x^{(\mu)} x^{(\mu)T} \qquad N \times N$$

$$C u = \lambda u. \qquad N \times N \text{ problem}$$

$$C = \frac{1}{P} A A^T$$

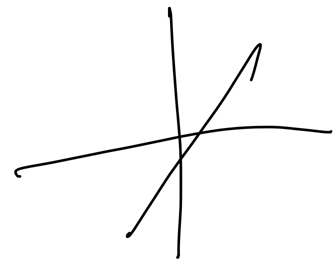
$$A = U \Sigma V^T$$

$$A^T A V = V \Lambda \qquad P \times P \text{ problem.}$$

$$U = A V \Sigma^{\dagger}$$ ← pseudo-inverse

find $u^{(i)}$ assoc. w/ non-zero λ_i
this way

Add discussion of rank revealing $\lambda_i = 0$



$$[x]_\beta = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_D \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad \lambda_{D+1} = 0.$$

200 pics in $\mathbb{R}^{200}$

noise $\lambda_i + \alpha^2$?

Implications of zero eigenvalues

200 images in $\mathbb{R}^{10^6}$ gives a data matrix of rank $(\underline{X}) \leq 200$.

$$[x^{(\mu)}]_\beta = \begin{bmatrix} \alpha_1^{(\mu)} \\ \alpha_2^{(\mu)} \\ \vdots \\ \alpha_D^{(\mu)} \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \qquad \lambda_i = \frac{1}{P} \sum_{\mu=1}^{P} |\alpha_i^{(\mu)}|^2$$

$$\lambda_{D+1} = 0 \Rightarrow \alpha_{D+1}^{(\mu)} = 0 \text{ all } \mu$$

Data must live in 200 D linear space.

Rank of a Matrix.

$\text{rank}(\underset{\sim}{X}) = $ # non-zero singular values

$\text{rank}(\underset{\sim}{X}\underset{\sim}{X}^*) = $ # non-zero eigenvalues $\underset{\sim}{X}\underset{\sim}{X}^*$

$\text{rank}(\underset{\sim}{X}^*\underset{\sim}{X}) = $ # " " $\underset{\sim}{X}^*\underset{\sim}{X}$.

by the arguments above these are all the same.

How do we "estimate" r?

i) scree plot

ii) local svd

iii) cross validatory SVD.

Cross Validatory SVD

goal: determine data dimension.

$$\underset{m\times n}{\underline{X}} = U \Sigma V^T.$$

Assume $m > n \geq \operatorname{rank}(\underline{X})$

delete column j

$$\underline{X}^j = \begin{pmatrix} & \cancel{} & \\ & \cancel{} & \\ & \cancel{} & \end{pmatrix} \qquad m \times (n-1)$$

j

$$= \underset{m\times(n-1)}{\tilde{U}} \; \underset{(n-1)\times(n-1)}{\tilde{\Sigma}} \; \underset{(n-1)\times(n-1)}{\tilde{V}^T}$$

$\tilde{U}$ still serves as a basis for $R(A)$.

$${}^{i}\underline{X} = \begin{pmatrix} \\ \text{- - - - - - - -} \\ \\ \end{pmatrix} \; i \text{ delete } i\text{th row}$$

$$\underset{(m-1)\times n}{{}^{i}\bar{X}} = \underset{m-1\times n}{{}^{i}\bar{U}} \; \underset{n\times n}{{}^{i}\bar{\Sigma}} \; \underset{n\times n}{{}^{i}\bar{V}^{T}}$$

Thus ${}^{i}\bar{V}$ still serves as a basis for $R(A^T)$.

SVD with point $\bar{X}_{ij}$ removed:

$${}^{i}\bar{X}^{j} = \tilde{U}^{j}\, S^{j}\, {}^{i}\bar{V}$$

where
$$\left({}^{i}S^{ij}\right)_{kk} = \begin{cases} \sqrt{{}^{j}\tilde{\sigma}_{kk}\; {}^{i}\bar{\sigma}_{kk}} & kk = 1 \dots n-1 \\ \bar{\sigma}_{nn} & kk = nn. \end{cases}$$

$$X^M_{ij} = \sum_{k=1}^{M} \tilde{v}_{ik} s_{kk} \bar{v}_{kj}$$

parity check: $\text{sign}(\tilde{u}_{ik} s_{kk} \bar{v}_{kj})$ assign

$= \text{sign}(u_{ik} s_{kk} v_{kj}) \leftarrow$ given

$$\text{PRESS}(M) = \frac{1}{mn} \sum_{ij}^{mn} (x^M_{ij} - x_{ij})^2$$

$$R(M) = \text{PRESS}(M) \Big/ \frac{1}{mn} \sum_{ij} (x_{ij} - \text{mean}(x_{ij}))^2$$