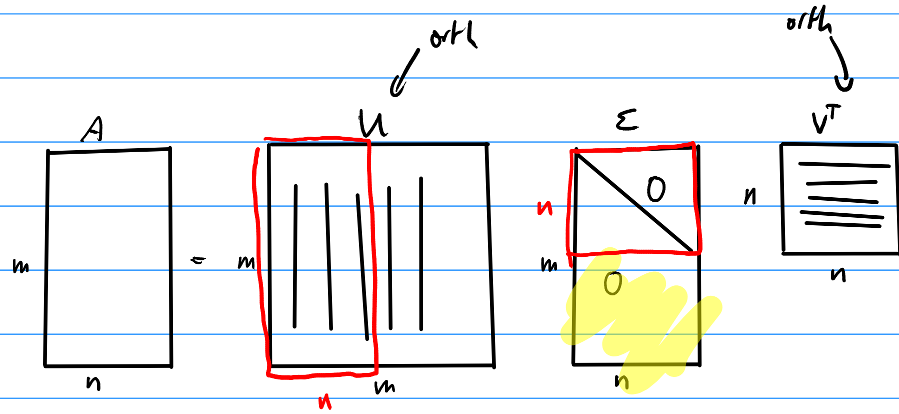
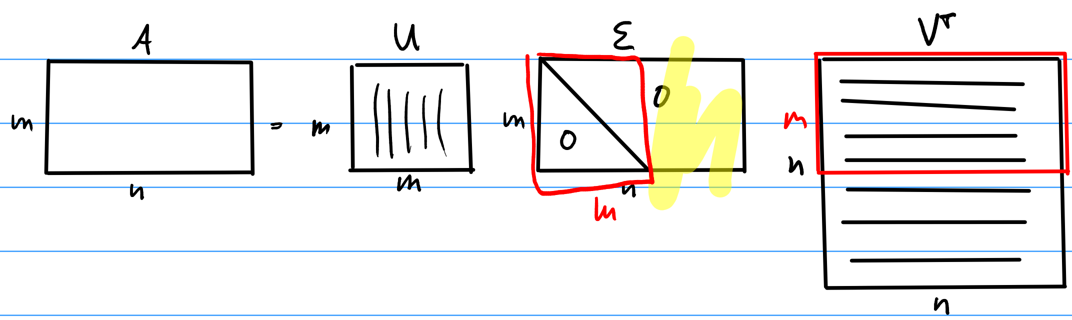
Midterm 2: April 9

Outline: SVD

- low-rank approx
- computing norm/cond
- PCA
- SVD → least squares

What was the SVD again? And: is there a reduced form?





e

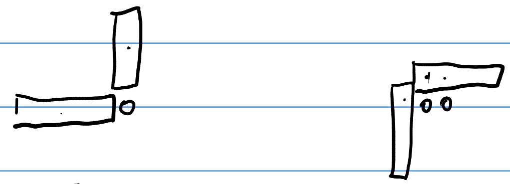
What's another way of writing the SVD?

Starting from $A = U\Sigma V^T = \begin{pmatrix} | & & | \\ u_1 & \cdots & u_m \\ | & & | \end{pmatrix} \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \sigma_n \\ 0 & & \end{pmatrix} \begin{pmatrix} - & v_1 & - \\ & \vdots & \\ - & v_n & - \end{pmatrix}$

$$A = \sigma_1 u_1 v_1^T + \cdots + \sigma_n u_n v_n^T$$

outer product

inner :



$x^T y$ $\qquad$ xy^T

If $\sigma_1 \geq \sigma_2 \geq \cdots \sigma_n \geq 0$, we can use the above form of A, but stop early:

$$A \approx \sigma_1 u_1 v_1^T + \cdots + \sigma_k u_k v_k^T + \cancel{\sigma_n u_n v_n^T}$$

What do the singular values mean? (in particular the first/largest one)

$$\|A\|_2 = \max_{\|x\|_2=1} \|Ax\|_2 = \max_{\|y\|_2=1} \|\Sigma y\|_2 = \sigma_1$$

Say I want just a rank-1 approximation of A:

$$A_1 = \sigma_1 u_1 v_1^T \quad \rightsquigarrow \quad \|A_1\|_2 = \sigma_1$$

$$= U \begin{pmatrix} \sigma_1 & 0 \\ 0 & 0 \end{pmatrix} V^T$$

How expensive is it to compute the SVD?

$O(n^3)$

So why bother with the SVD if it is so expensive? I.e. what makes the SVD special?

$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_n u_n v_n^T$ $\quad \sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \cdots \geq \sigma_n \geq 0$

$A_k = \sigma_1 u_1 v_1^T + \cdots + \sigma_k u_k v_k^T$ $\quad (k < n)$

Then $\|A - B\|_2$ where rank(B)=k is minimized by A_k.

("Eckart-Young theorem")

A_k is called the rank-k best-approximation

What is the Frobenius norm of a matrix?

$$\|A\|_F = \sqrt{a_{11}^2 + a_{12}^2 + \cdots a_{1n}^2 + a_{21}^2 + \cdots + a_{mn}^2}$$

⚠ not really a matrix norm in our sense--
as in, not produced by any vector norm

How about rank-k best-approximation in the Frobenius norm?

Then $\|A-B\|_{\cancel{2}\,F}$ where rank(B)=k is minimized by A_k.

("~~Eckart Young theorem~~")

A_k is called the rank-k best-approximation

9 Applications of the SVD

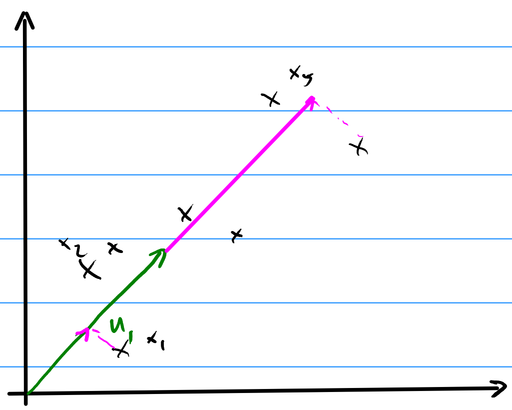
(1) Rank-k approximation

Rank-1 approximation

$$X = \underbrace{\begin{pmatrix} | & & | \\ x_1 & \cdots & x_n \\ | & & | \end{pmatrix}}_{n} \Big\} 2$$

$$X = \underset{\substack{2\times 2 \\ \text{(reduced)}}}{U} \underset{2\times 2}{\Sigma} \underset{2\times n}{V^T}$$

$$U = \begin{pmatrix} | & | \\ u_1 & u_2 \\ | & | \end{pmatrix} \qquad V = \begin{pmatrix} | & | \\ v_1 & v_2 \\ | & | \end{pmatrix}$$



$$X_1 = \sigma_1 u_1 v_1^T \approx X$$

break down by columns

$$\underbrace{\|X - X_1\|_F^2}_{\text{as small as it gets}} = \sum_i \| \vec{x}_i - \sigma_1 \vec{u}_1 v_{1i} \|_2^2$$

$$X \approx \underbrace{\sigma_1 u_1 v_1^T}_{\text{bigger}} + \cancel{\sigma_2 u_2 v_2^T}$$

So now how about rank-k approximation?

Give an example of where rank-k approximation does something useful.

((Image compression))

(2) Computing the 2-norm

$$A = U \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{pmatrix} V^*$$

$$\|A\|_2 = \sigma_1$$

(3) Computing the 2-norm condition number

$$A = U \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{pmatrix} V^T$$

Assume A is invertible.

2-norm condition number= $\|A\|_2 \|A^{-1}\|_2$

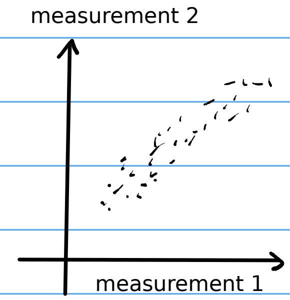
$$= \sigma_1 \|A^{-1}\|_2$$

$$= \sigma_1 \cdot \frac{1}{\sigma_n} = \frac{\sigma_1}{\sigma_n}$$

$$A^{-1} = A^{+} = V \begin{pmatrix} 1/\sigma_1 & & \\ & \ddots & \\ & & 1/\sigma_n \end{pmatrix} U^T$$

If A has a nullspace, i.e. if $\sigma_n = 0$ then the condition number is infinity.

(4) Principal Component Analysis ("PCA")



Have: a pile of "data"

More precisely: m 'measurements' from n 'trials'

each resulting in a real number

x_{ij} $\quad i = 1 \ldots m \quad j = 1 \ldots n$

Data matrix: $X = (x_{ij}) = \begin{pmatrix} | & | & | & | & | & | \\ | & | & | & | & | & | \end{pmatrix}$ $\downarrow$ i: measurements

$\rightarrow$ j: trials

How do I compute a PCA?

How do I compute a PCA? (cont'd)

(5) Least squares for underdetermined and singular systems

(6) "Total" least squares

For a given matrix A, find the vector x so that

- $\|Ax\|_2$ is minimal
- $\|x\|_2 = 1$