HW 6 due Tues Dec 13.

Office Hours, by appointment.

Singular Value Decomposition.

Recall, for normal matrices $A^*A = AA^*$ we have the spectral Theorem:

$$A = U \Lambda U^*$$

where $U^*U = I$ (unitary) & Λ is diagonal.

If $A^*A \neq AA^*$ then A might not even be diagonalizable:

$$A \neq X \Lambda X^{-1}$$

The Jordan form is one way out.

The other important Theorem is the SVD which applies even to non-square matrices:

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

Diagonal pos. real entries

$U^* U = I$
$V^* V = I$.

<u>Proof</u>: Let A $m \times n$, rank r.

Consider $A^* A$ $n \times n$, rank r.

Since $A^* A$ pos. semi-def, the eigenvalues are

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > \underbrace{0, 0, \ldots, 0}_{n-r \text{ zeros}}.$$

write $\sigma_i^2 = \lambda_i$
$\sigma_i > 0$.

Since $A^* A$ is normal (in fact, Hermitian), Spectral Theorem

$\Longrightarrow$ $\exists$ orthonormal eigenvectors.

$\vec{v}_1, \ldots, \vec{v}_r, \vec{v}_{r+1}, \ldots, \vec{v}_n \in \mathbb{R}^n$

$$A^* A \vec{v}_i = \sigma_i^2 \vec{v}_i \qquad i \le r$$

$$A^* A \vec{v}_i = 0 \qquad i > r.$$

TRICK: for $1 \le \boxed{i \le r}$, define vectors $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r \in \mathbb{R}^m$ by

$$\vec{u}_i = \frac{1}{\sigma_i} A \vec{v}_i . \qquad A\vec{v}_i = \sigma_i \vec{u}_i$$

$\neq 0$

Then automatically have

$$A\left(\vec{v}_1 | \cdots | \vec{v}_r\right)$$

$$= \left(A\vec{v}_1 | \cdots | A\vec{v}_r\right)$$

$$= \left(\sigma_1 \vec{u}_1 | \cdots | \sigma_r \vec{u}_r\right)$$

$$= \left(\vec{u}_1 \middle| \cdots \middle| \vec{u}_r \right) \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix}$$

$$A V = U \Sigma$$

$$\underset{m \times n}{A} \; \underset{n \times r}{\left(\vec{v}_1 \middle| \cdots \middle| \vec{v}_r \right)} = \underset{m \times r}{\left(\vec{u}_1 \middle| \cdots \middle| \vec{u}_r \right)} \; \underset{r \times r.}{\Sigma}$$

Know $V^* V = I_r$ since $\vec{v}_i$ are orthonormal: $\vec{v}_i^{\,*} \vec{v}_j = \delta_{ij}$.

Hence

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

Still need to check that $U^* U = I_r$ i.e. that $\vec{u}_i^{\,*} \vec{u}_j = \delta_{ij}$.

By definition,

$$\vec{u}_i^{\,*} \vec{u}_j = \left(\frac{1}{\sigma_i} A \vec{v}_i \right)^* \left(\frac{1}{\sigma_j} A \vec{v}_j \right)$$

$$= \frac{1}{\sigma_i \sigma_j} \vec{v}_i^{\,*} A^* A \vec{v}_j$$

$$= \frac{1}{\sigma_i \sigma_j} \vec{v}_i^{\,*} \left(\sigma_j^2 \vec{v}_j \right)$$

$$= \frac{\sigma_j}{\sigma_i} \vec{v}_i^{\,*} \vec{v}_j = \begin{cases} 0 & i \neq j \\ \sigma_i / \sigma_i = 1 & i = j. \end{cases}$$

Hence $\vec{u}_i^{\,*} \vec{u}_j = \delta_{ij}$ (orthonormal).

✓

But more is true. We started with $\vec{v}_i$ to get $\vec{u}_i$. But it's really a symmetric theorem.

In fact we have

$$AA^* = (U \Sigma V^*)(U \Sigma V^*)^*$$

$$= U \Sigma \underbrace{\cancel{V^* V}}_{I} \Sigma^* U^*$$

$$= U \Sigma^2 U^* \qquad \Sigma^* = \Sigma.$$

$$\Longrightarrow \quad AA^* \vec{u}_i = \sigma_i^2 \vec{u}_i$$

For all $i \leq r$.

That is, A^*A & AA^* have the same nonzero eigenvalues $\sigma_1^2 \geq \cdots \geq \sigma_r^2$ and $\exists$ orthonormal sets of e. vectors satisfying

$$\boxed{\begin{aligned} A^*A\vec{v}_i &= \sigma_i^2 \vec{v}_i \\ AA^*\vec{u}_i &= \sigma_i^2 \vec{u}_i \\ A\vec{v}_i &= \sigma_i \vec{u}_i \end{aligned}}$$

SVD.

SVD gives the Moore-Penrose pseudo-inverse of a rectangular matrix. A $m \times n$.

$$\underset{m\times n}{A} = \underset{m\times r}{U} \, \underset{r\times r}{\Sigma} \, \underset{r\times n.}{V^*}$$

Σ is diagonal and invertible:

$$\begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix}^{-1} = \begin{pmatrix} 1/\sigma_1 & & \\ & \ddots & \\ & & 1/\sigma_r \end{pmatrix}.$$

Define the pseudo-inverse

$$\underset{m\times n}{A^+} = \underset{m\times r}{V}\ \underset{r\times r}{\Sigma^{-1}}\ \underset{r\times n}{U^*}.$$

Satisfies 4 defining properties:

$$\boxed{\begin{aligned} AA^+A &= A \\ A^+AA^+ &= A^+ \\ (AA^+)^* &= AA^+ \\ (A^+A)^* &= A^+A. \end{aligned}}$$

$\exists$ unique matrix with these properties & SVD allows you to find it.

If A has independent columns then

$$A^+ = (A^*A)^{-1}A^*,$$

which (as we know) is a left inverse of A.

$$A^+A = \cancel{(A^*A)^{-1}A^*A} = I.$$

If A has independent rows then

$$A^+ = A^*(AA^*)^{-1}$$

which is a right inverse:

$$AA^+ = \cancel{AA^*(AA^*)^{-1}} = I.$$

If A is square with ind rows & columns, then

$$AA^+ = A^+A = I.$$

$$\Longrightarrow \quad A^+ = A^{-1}.$$

What is SVD used for?

Eckart - Young - Mirsky etc...

Rediscovered many times...

Given $m \times n$ matrix A, rank r,
consider SVD:

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

$$= \left(\vec{u}_1 \middle| \cdots \middle| \vec{u}_r \right) \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix} \begin{pmatrix} \vec{v}_1^{\,*} \\ \hline \vdots \\ \hline \vec{v}_r^{\,*} \end{pmatrix}.$$

$$= \sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^{\,*}$$

rank one matrix.

<u>Sort</u> $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$.

For any $k \leq r$, define

$$A_k = \sum_{i=1}^{k} \sigma_i \vec{u}_i \vec{v}_i^{\,*}$$

Then (theorem) for any $m \times n$ matrix B of rank $\leq k$, have

$$\|A - A_k\| \leq \|A - B\|$$

best rank k approximation to A.

for any "unitarily invariant" norm.
Includes Frobenius & Operator norms.

Special Case on HW6.
Rayleigh Quotient.
Given $m \times n$ matrix A, consider

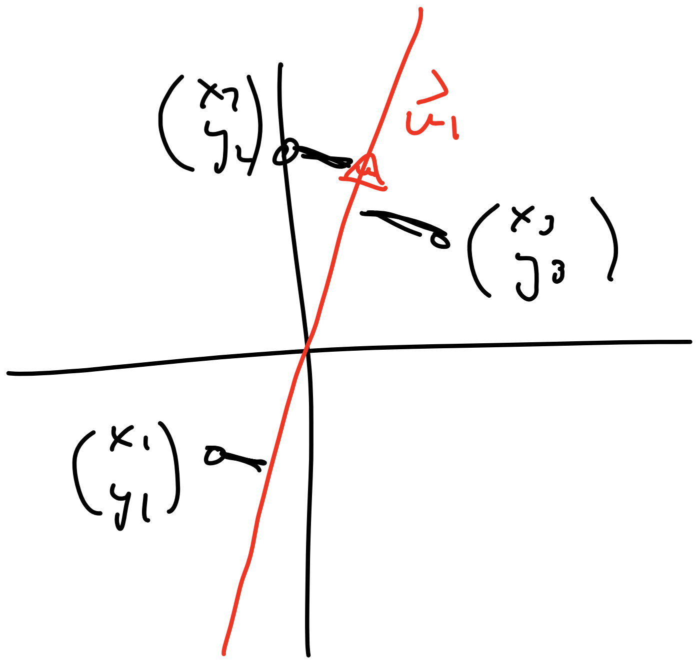
$$r : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$$

$$r(\vec{x}) = \|A\vec{x}\|^2 / \|\vec{x}\|^2$$

Problem: Find $\vec{x}$ to maximize $r(\vec{x})$.

Answer: vector $\vec{u}_1$ from the SVD of A.

Apply to T.L.S.



$$X = \left(\begin{array}{c|c|c} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{array} \right) \quad 2 \times 3.$$

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

$$= \left(\vec{u}_1 \,\middle|\, \vec{u}_2 \right) \begin{pmatrix} \sigma_1 & \\ & \sigma_2 \end{pmatrix} \left(\begin{array}{c} \vec{v}_1^{\,*} \\ \hline \vec{v}_2^{\,*} \end{array} \right).$$

Apply SVD to image compression.

$m \times n$ pixel image $=$ $m \times n$ matrix.
greyscale.

A $m \times n$ image (rank r). certainly $r = \min\{m,n\}$.

$$A = \sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^*$$

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0.$$

Truncation $(k \leq r)$

$$A_k = \sum_{i=1}^{k} \sigma_i \vec{u}_i \vec{v}_i^*$$

is a compressed version of the image.

Works well if the singular values drop off: