Matrix norms

Recall: $\|x\|_2 = \boxed{\sqrt{x^T x},}$ and $\|Ux\| = \|x\|$ for orthogonal U.

One can define a norm for matrices, too.

Definition

$$\|A\|_2 := \max_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \max_{\|z\|=1} \|Az\|.$$

Idea: look for ~~a~~ smallest value of $\|A\|$ that ensures $\boxed{\|Ax\| \leq \|A\| \|x\|.}$

The construction works for every vector norm ($\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty \dots$)

$$\|Ax\| \leq \|A\| \cdot \|x\|$$

$$\|Ax\| = 2 \qquad \|x\| = 1$$

$$\|Ay\| = \underline{3.5} \qquad \|y\| = 1$$

$$\|Az\| = 4.7 \qquad \|z\| = 1 \leftarrow$$

$$\|A\| \geq 4.7\ldots$$

$$\frac{\|Ax\|}{\|x\|} = \|Az\| \qquad z = \frac{1}{\|x\|} \cdot x$$

<u>Remark</u>: if $Av = \lambda v$

(eigvl-eigvec pair),

$$\|Av\| = \|\lambda v\| = |\lambda| \cdot \|v\|$$

$$\rightsquigarrow \|A\| \geq \max_{(\lambda_i \text{ eigvls of } A)} |\lambda_i|$$

Norm properties

Properties

For each choice of matrices A, B and vector x for which the operations make sense,

- $\|A\| \geq 0$, with equality iff A is all-zeros;
- $\|\alpha A\| = |\alpha| \|A\|$ for each $\alpha \in \mathbb{R}$;
- $\|A + B\| \leq \|A\| + \|B\|$;
- $\|AB\| \leq \|A\| \|B\|$;
- $\|Av\| \leq \|A\| \|x\|$ (same norm for matrices and vectors).

Our favorite norm: $\|A\|_{\boxed{2}}$. It satisfies $\underline{\|A\|_2} = \underline{\|AU\|_2} = \underline{\|UA\|_2}$ for each orthogonal U.

(People often omit the subscript 2.)

Norm computations:

$$\left\| U_1 \sigma_1 V_1^T + U_2 \sigma_2 V_2^T \right\| \leq \|U_1\| \cdot |\sigma_1| \cdot \|V_1^T\| + \|U_2\| \cdot |\sigma_2| \cdot \|V_2^T\|$$

$$\|UA\| = \max_{x \neq 0} \frac{\|UAx\|}{\|x\|} = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \|A\|$$

$$\|AU\| = \max_{x \neq 0} \frac{\|AUx\|}{\|x\|} = \max_{y \neq 0} \frac{\|Ay\|}{\|y\|} = \|A\|$$

$y = Ux \quad \|y\| = \|x\|$

Frobenius norm

Other matrix norm of a different kind: Frobenius norm

$$\|A\|_F = \left\| \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \right\|_F = \sqrt{a_{11}^2 + a_{12}^2 + \dots + a_{mn}^2}.$$

Satisfies all the properties in the previous slide.

Norm and SVD

Since orthogonal matrices do not change $\|\cdot\|_2$,

$$\underline{\|A\|_2} = \|U\Sigma V^T\|_2 = \underline{\|\Sigma\|_2} = \boxed{\sigma_1}.$$

(Why is $\|\Sigma\|_2 = \sigma_1$ for the diagonal matrix Σ in SVD? Similar argument to the one we used for $\lambda_{\text{min}} x^T x \le x^T A x \le \lambda_{\text{max}} x^T x$.)

$$\Sigma = \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_n \end{pmatrix} \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad \Sigma_1 x = \begin{bmatrix} \sigma_1 x_1 \\ \sigma_2 x_2 \\ \vdots \\ \sigma_n x_n \end{bmatrix}$$

$$\|\Sigma_1\| = \max \frac{\|\Sigma_1 x\|}{\|x\|} = \frac{\sqrt{\sigma_1^2 x_1^2 + \dots + \sigma_n^2 x_n^2}}{\sqrt{x_1^2 + x_2^2 + \dots + x_n^2}} \leq$$

$$\leq \frac{\sqrt{\sigma_1^2 x_1^2 + \sigma_1^2 x_2^2 + \dots + \sigma_1^2 x_n^2}}{\sqrt{x_1^2 + x_2^2 + \dots + x_n^2}} = |\sigma_1|$$

Eckart-Young theorem

Theorem

For a matrix A with SVD $A = U\Sigma V^T$, the solution of

$$\min_{\operatorname{rank} X \leq k} \|A - X\|$$

is given by truncated SVD:

$$X = \begin{bmatrix} u_1 & u_2 & \cdots & u_k \\ & & & \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_k \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_k \end{bmatrix}^T$$

$$= u_1\sigma_1 v_1^T + u_2\sigma_2 v_2^T + \cdots + u_k\sigma_k v_k^T.$$

Works both for $\|\cdot\|_2$ and $\|\cdot\|_F$.

Geometric/application meaning: we will see in the lab.

A has rank $r \iff Ax$ is always linear combination of r lin. ind. vectors

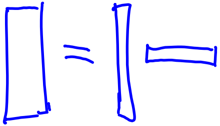
Ex: if A has rank 1,

Ax, Ay, Az are all multiple of the same vector

Theo: if $A \in \mathbb{R}^{m \times n}$ has rank $\leq r$ we can write $A = \underline{B}\,\underline{C}$ with

$B \in \mathbb{R}^{m \times r}$

$C \in \mathbb{R}^{r \times n}$



mn $\qquad$ $mr + rn$

<u>Ex</u>: A has rank 1 $\Rightarrow$ we can write it as $A = xy^T$ x, y vectors

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix} = \begin{bmatrix} x_1y_1 & x_1y_2 & \cdots & x_1y_n \\ x_2y_1 & x_2y_2 & \cdots & x_2y_n \\ \vdots & \vdots & & \vdots \\ x_my_1 & x_my_2 & \cdots & x_my_n \end{bmatrix}$$

<u>Problem</u>: given A <u>not</u> of low rank, can I find a matrix with small rank s.t. $\|A - X\|_2$ is small?

Exercises

1. Show that $\|A\| \geq \|c\|$, where c is one of the columns of A.
2. Show that $\|UA\|_2 = \|A\|_2$ for each orthogonal U.
3. Show that $\|AU\|_2 = \|A\|_2$ for each orthogonal U.
4. Let A_k be the best rank-k approximation of A (computed through SVD/Eckart-Young theorem). What is the value of $\|A - A_k\|_2$? Of $\|A - A_k\|_F$?