Matrix norms

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

One can define a norm for matrices, too.

Definition

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

Idea: look for a value of $\|A\|$ that ensures $\|Ax\| \leq \|A\| \|x\|$.

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

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: $\|A\|_2$. It satisfies $\|A\|_2 = \|AU\|_2 = \|UA\|_2$ for each orthogonal U.

(People often omit the subscript 2.)

Frobenius norm

Other matrix norm of a different kind: Frobenius norm

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

Satisfies all the properties in the previous slide.

Norm and SVD

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

$$\|A\|_2 = \|U\Sigma V^T\|_2 = \|\Sigma\|_2 = \sigma_1.$$

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

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.

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 $\|A - A_k\|_F$?