Matrix norms

Recall: $\|\mathbf{v}\|_2 = \sqrt{\mathbf{v}^\top \mathbf{v}}$, and $\|U\mathbf{v}\|_2 = \|\mathbf{v}\|_2$ for orthogonal U.

One can define a norm for matrices, too.

Definition (induced matrix norm)

Given a norm on vectors (e.g., $\|\cdot\|_2, \|\cdot\|_\infty, \dots$), we can define a corresponding norm on matrices:

$$\|A\| := \max_{\mathbf{v} \neq \mathbf{0}} \frac{\|A\mathbf{v}\|}{\|\mathbf{v}\|} = \max_{\|\mathbf{u}\|=1} \|A\mathbf{u}\|.$$

Idea: it's the smallest value of $\|A\|$ that ensures $\|A\mathbf{v}\| \leq \|A\|\|\mathbf{v}\|$ for all $\mathbf{v}$.

This general 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 $\mathbf{v}$ for which the operations make sense,

- $\|A\| \geq 0$, with equality iff A is all-zeros;
- $\|\alpha A\| = |\alpha|\|A\|$ for each $\alpha \in \mathbb{R}$;
- $\|A + B\| \leq \|A\| + \|B\|$;
- $\|AB\| \leq \|A\|\|B\|$;
- $\|A\mathbf{v}\| \leq \|A\|\|\mathbf{v}\|$ (if 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

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

It satisfies all the properties in the previous slide (reducing to $\|\mathbf{v}\|_F = \|\mathbf{v}\|_2$ on vectors); in particular, $\|AU\|_F = \|UA\|_F = \|A\|_F$. However, it does not come from the 'induced' construction.

Norm and SVD

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

$$\|A\|_2 = \|USV^T\|_2 = \|S\|_2 = \sigma_1.$$

(Why is $\|S\|_2 = \sigma_1$ for the diagonal matrix S in SVD? By a similar argument to the one we used for $\lambda_{\min}\mathbf{x}^T\mathbf{x} \le \mathbf{x}^T A\mathbf{x} \le \lambda_{\max}\mathbf{x}^T\mathbf{x}$.)

Similarly, $\|A\|_F^2 = \sum_{i=1}^{\min(m,n)} \sigma_i^2$.

Eckart-Young theorem

Theorem

For a matrix A with SVD $A = USV^T$, the solution of

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

for both $\|\cdot\|_2$ and $\|\cdot\|_F$ is given by truncated SVD:

$$X = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_k \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_k \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_k \end{bmatrix}^T$$

$$= \mathbf{u}_1\sigma_1\mathbf{v}_1^T + \mathbf{u}_2\sigma_2\mathbf{v}_2^T + \cdots + \mathbf{u}_k\sigma_k\mathbf{v}_k^T.$$

Geometric/application meaning: we will see experimentally in the next lectures!

Exercises

1. Show that $\|A\| \geq \|\mathbf{c}\|$, where $\mathbf{c}$ is one of the columns of A.
2. SHow that for each eigenvalue λ of A we have $|\lambda| \leq \|A\|$.
3. Show that $\|UA\|_2 = \|A\|_2$ for each orthogonal U.
4. Show that $\|AU\|_2 = \|A\|_2$ for each orthogonal U.
5. Show that (for a square matrix) $\|A^{-1}\|_2 = \frac{1}{\sigma_m}$, where σ_m is the smallest singular value of A. (Hint: in a previous exercise, we asked you to compute the SVD of A^{-1} from that of A.)
6. 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 $\|A - A_k\|_F$?

References: Trefethen-Bau book, Lectures 3 and 5.