Matrix Inner Product and Norm

CE40282-1: Linear Algebra

Hamid R. Rabiee and Maryam Ramezani
Sharif University of Technology

Inner products

- **Definition 1** (Inner product). *A function* $\langle .,. \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ *is an inner product if*

  1. $\langle x, x \rangle \geq 0, \langle x, x \rangle = 0 \Leftrightarrow x = 0$ *(positivity)*
  2. $\langle x, y \rangle = \langle y, x \rangle$ *(symmetry)*
  3. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$ *(additivity)*
  4. $\langle rx, y \rangle = r\langle x, y \rangle$ *for all* $r \in \mathbb{R}$ *(homogeneity)*

- using properties (2) and (4) and again (2)

$$\langle x, ry \rangle = \langle ry, x \rangle = r\langle y, x \rangle = r\langle x, y \rangle$$

- using properties (2), (3) and again (2).

$$\langle x, y + z \rangle = \langle y + z, x \rangle = \langle y, x \rangle + \langle z, x \rangle = \langle x, y \rangle + \langle x, z \rangle$$

Inner products

- The standard inner product is

$$\langle x, y \rangle = x^T y = \sum x_i y_i, \;\; x, y \in \mathbb{R}^n.$$

- The standard inner product between matrices is

$$\langle X, Y \rangle = \mathrm{Tr}(X^T Y) = \sum_i \sum_j X_{ij} Y_{ij}$$

where $X, Y \in \mathbb{R}^{m \times n}$.

Example

- $p = 3 - x + 2x^2$ and $q = 4x + x^2$
- $U = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}, \quad V = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

Norm

- **Definition** (Norm). *A function* $f : \mathbb{R}^n \to \mathbb{R}$ *is a norm if*

  1. $f(x) \geq 0,\ f(x) = 0 \Leftrightarrow x = 0$ *(positivity)*
  2. $f(\alpha x) = |\alpha| f(x),\ \forall \alpha \in \mathbb{R}$ *(homogeneity)*
  3. $f(x + y) \leq f(x) + f(y)$ *(triangle inequality)*

"Entry-wise" matrix norms

- $$\|A\|_{p,p} = \|\mathrm{vec}(A)\|_p = \left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^p\right)^{1/p}$$
- special case p = 2 is the Frobenius norm, and p = ∞ yields the maximum norm.

$$\|A\|_E = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}(a_{ij})^2}$$

  - Frobenius (Euclidean, Hilbert Schmidt) norm
    - invariant under rotations (unitary operations)

$$\|A\|_{\mathrm{F}} = \|AU\|_{\mathrm{F}} = \|UA\|_{\mathrm{F}}$$

$$\|A+B\|_{\mathrm{F}}^2 = \|A\|_{\mathrm{F}}^2 + \|B\|_{\mathrm{F}}^2 + 2\langle A, B\rangle$$

$$\|A^*A\|_{\mathrm{F}} = \|AA^*\|_{\mathrm{F}} \leq \|A\|_{\mathrm{F}}^2$$

$$\sqrt{\mathrm{trace}(A^*A)} = \sqrt{\sum_{i=1}^{\min\{m,n\}}\sigma_i^2(A)},$$

Singular value of A

  - Max norm

The **max norm** is the elementwise norm with $p = q = \infty$:

$$\|A\|_{\max} = \max_{ij}|a_{ij}|.$$

  - sum-absolute-value norm: $\|A\|_{sav} = \sum_{i,j}|A_{i,j}|$

Frobenius (Euclidean) norm

Let $b_1, \ldots, b_n$ denote the columns of B. Then

$$\|AB\|_{\mathrm{HS}}^2 = \sum_{i=1}^{n} \|Ab_i\|^2 \leq \sum_{i=1}^{n} \|A\|^2 \|b_i\|^2 = \|A\|^2 \|B\|_{\mathrm{HS}}^2.$$

cauchy-Schwarz inequality

Matrix norms induced by vector norms

- $$||A||_p = \max_{\vec{\mathbf{x}} \neq \vec{\mathbf{0}}} \frac{||A\vec{\mathbf{x}}||_p}{||\vec{\mathbf{x}}||_p} = \max_{||\vec{\mathbf{x}}||_p = 1} ||A\vec{\mathbf{x}}||_p$$

- Theorem: $\|Ax\| \leq \|A\| \, \|x\|$ for all vectors $\|x\|$

- Theorem: For all matrices A, B, $\|AB\| \leq \|A\| \, \|B\|$

Matrix norms induced by vector norms

- The norm of a matrix is a real number which is a measure of the magnitude of the matrix.
- $$\|A\|_1 = \max_{1\leq j\leq n}\left(\sum_{i=1}^{n}|a_{ij}|\right)$$
- $$\|A\|_\infty = \max_{1\leq i\leq n}\left(\sum_{j=1}^{n}|a_{ij}|\right)$$
- spectral norm (l2) is the largest singular value (the square root of the largest eigenvalue of the matrix gram A
- Example

$$\|A\|_2 = \sqrt{\max\{\text{eigenvalue}(A^TA)\}} = \max\{\text{sing}(A)\}$$

$$B = \begin{bmatrix} 5 & -4 & 2 \\ -1 & 2 & 3 \\ -2 & 1 & 0 \end{bmatrix}$$

Norms Compare

The 2-norm (spectral norm) of a matrix is the greatest distortion of the unit circle/sphere/hyper-sphere. It corresponds to the largest singular value (or |eigenvalue| if the matrix is symmetric/hermitian).

The Forbenius norm is the "diagonal" between all the singular values.

i.e.

$$||A||_2 = s_1 \quad , \quad ||A||_F = \sqrt{s_1^2 + s_2^2 + \ldots + s_r^2}$$

(r being the rank of A).

Here's a 2D version of it: x is any vector on the unit circle. Ax is the deformation of all those vectors. The length of the red line is the 2-norm (biggest singular value). And the length of the green line is the Forbenius norm (diagonal).