LA

Linear Algebra

LA.4 NORMS OF VECTORS AND MATRICES

The definition of norm for vectors and matrices provides a proper language to define notions like “small perturbations”, “near rank deficiency” and “distance” in vector spaces.

Definition LA.4.1 (Vector norms) – A *vector norm* on $\mathcal{R}^n$, denoted as $\|\cdot\|$, is a function $\mathcal{R}^n \to \mathcal{R}$, that satisfies the following properties:

i) $\|x\| \geq 0 \; \forall x \in \mathcal{R}^n$, $\|x\| = 0$ if and only if $x = 0$;

ii) $\|\gamma x\| = |\gamma| \|x\| \;\; \forall \, \gamma \in \mathcal{R}, \; x \in \mathcal{R}^n$;

iii) $\|x + y\| \leq \|x\| + \|y\| \;\; \forall \, x, y \in R^n$.

A useful class of norms is given by Holder or p–norms defined as

$$\|x\|_p = \left(|x_1^p| + \ldots + |x_n^p|\right)^{1/p}, \qquad p \geq 1 \tag{LA.4.1}$$

assuming that a finite limit exists for $p \to \infty$. The most common norms can be obtained taking $p = 1, \; 2$ and ∞ in (LA.4.1):

- $\|x\|_1 = |x_1| + \ldots + |x_n|$;
- $\|x\|_2 = \sqrt{|x_1^2| + \ldots + |x_n^2|} = \sqrt{x^T x}$ (l_2 or Euclidean norm defining the length of x);
- $\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$ (l_∞ or infinity–norm).

p-norms satisfy the following Holder inequality

$$|\, x^T y \,| \leq \|x\|_p \, \|y\|_q \quad \text{with} \quad \frac{1}{p} + \frac{1}{q} = 1\,; \tag{LA.4.2}$$

taking $p = q = 2$ in (LA.4.2) we obtain the Cauchy-Schwartz inequality

$$|x^T y| \leq \|x\|_2 \, \|y\|_2 \,. \tag{LA.4.3}$$

For any nonzero vector x and norm, the normalized vector u having the same direction as x and unitary norm is given by

$$u = \frac{1}{\|x\|}\, x. \tag{LA.4.4}$$

Vector norms define distances in vector spaces; thus if $\hat{x}$ denotes an approximation of x then, for a given vector norm, the absolute error is given by

$$\epsilon = \|\hat{x} - x\| \tag{LA.4.5}$$

while the relative error is defined as

$$\epsilon = \frac{\|\hat{x} - x\|}{\|x\|}\,. \tag{LA.4.6}$$

Definition LA.4.2 (Matrix norms) – A *matrix norm*, denoted as in the vector case by $\|\cdot\|$, is a scalar function that satisfies the following properties:

i) $\|A\| \geq 0\ \forall A \in \mathcal{R}^{m\times n}$, $\|A\| = 0$ if and only if $A = 0$;

ii) $\|\gamma A\| = |\gamma|\|A\|\ \forall\, \gamma \in \mathcal{R},\ A \in \mathcal{R}^{m\times n}$;

iii) $\|A + B\| \leq \|A\| + \|B\|\ \forall\, A, B \in R^{m\times n}$.

A further property of matrix norms (consistency) is given by the following relation

iv) $\|A\ B\| \leq \|A\|\|B\|$.

For any vector norm $\|\cdot\|_p$, the induced matrix norm $\|A\|_p$ is defined as

$$\|A\|_p = \sup_{x\neq 0} \frac{\|Ax\|_p}{\|x\|_p}. \tag{LA.4.7}$$

The matrix norms corresponding to one, two and infinity vector norms are:

- $\|A\|_1 = \sup_{1\leq j\leq n} \sum_{i=1}^{m} |a_{ij}|$;
- $\|A\|_2 = \sigma_{max}$, where σ_{max} denotes the largest singular value of A;
- $\|A\|_\infty = \sup_{1\leq i\leq m} \sum_{j=1}^{n} |a_{ij}|$.

p-norms satisfy the following important property

$$\|Ax\|_p \leq \|A\|_p\ \|x\|_p\ \forall\, x \in \mathcal{R}^n,\ A \in \mathcal{R}^{m\times n} \tag{LA.4.8}$$

where the equality must hold for at least one nonzero vector x. A vector norm $\|\cdot\|$ and a matrix norm $\|\cdot\|'$ are defined as compatible when

$$\|Ax\| \leq \|A\|'\|x\|\ \forall x \in \mathcal{R}^n,\ A \in \mathcal{R}^{m\times n} \tag{LA.4.9}$$

so that any vector norm and its induced matrix norm are always compatible. The Frobenius norm or F–norm of a matrix is defined as

$$\|A\|_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2} \ . \tag{LA.4.10}$$

Frobenius and Euclidean norms are compatible, i.e.

$$\|Ax\|_2 \le \|A\|_F \|x\|_2 \ \forall\, x \in \mathcal{R}^n,\ A \in \mathcal{R}^{m\times n} \tag{LA.4.11}$$

and satisfy the following relations:

$$\|A\|_F = \mathrm{trace}\big(A^T A\big) \tag{LA.4.12}$$

$$\|A\|_2 \le \|A\|_F. \tag{LA.4.13}$$

Frobenius and Euclidean norms are invariant with respect to orthogonal transformations; thus for all orthogonal matrices Q and Z of appropriate dimensions

$$\|QAZ\|_2 = \|A\|_2 \tag{LA.4.14}$$

$$\|QAZ\|_F = \|A\|_F. \tag{LA.4.15}$$

If $\sigma_1 \ge \sigma_2 \ge \ldots \sigma_p \ge 0$ are the singular values of A, then

$$\|A\|_F = \sqrt{\sigma_1^2 + \ldots + \sigma_p^2} \tag{LA.4.16}$$

$$\|A\|_2 = \sigma_1. \tag{LA.4.17}$$

If $\rho(A) = r > 0$ and therefore $\sigma_1 \ge \sigma_2 \ge \sigma_r > \sigma_{r+1} = \ldots = \sigma_p = 0$, then

$$\|A^+\|_2 = \frac{1}{\sigma_r} \ . \tag{LA.4.18}$$

Definition LA.4.3 (Condition numbers) – The condition number of a matrix A is given by

$$C(A) = \|A\|_2 \, \|A^+\|_2 = \frac{\sigma_1}{\sigma_r}. \tag{LA.4.19}$$

Definition LA.4.4 (Quadratic Forms) – A quadratic form (over $\mathcal{R}^n$) is a function q : $\mathcal{R}^n \to \mathcal{R}$ of the type

$$q(x) = x^T A\, x = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \tag{LA.4.20}$$

where A is a $(n \times n)$ symmetric matrix. A quadratic form is defined as

- *positive definite* if $q(x) > 0 \;\; \forall x \neq 0$
- *positive semi–definite* if $q(x) \geq 0 \;\; \forall x \neq 0$
- *negative definite* if $q(x) < 0 \;\; \forall x \neq 0$
- *negative semi–definite* if $q(x) \leq 0 \;\; \forall x \neq 0$

The following theorem characterizes quadratic forms by means of the eigenvalues $\lambda_i, \; i = 1, \ldots, n$ of A.

Theorem LA.4.1 – A quadratic form $q(x) = x^T A \, x$ is

- positive definite if and only if $\lambda_i > 0, \; i = 1, \ldots, n$,
- positive semi–definite if and only if $\lambda_i \geq 0, \; i = 1, \ldots, n$,
- negative definite if and only if $\lambda_i < 0, \; i = 1, \ldots, n$,
- negative semi–definite if and only if $\lambda_i \leq 0, \; i = 1, \ldots, n$.

These definitions are extended to matrices as follows.

Definition LA.4.5 – A symmetric matrix is defined as positive (negative) definite if all its eigenvalues are positive (negative), positive (negative) semidefinite if all its eigenvalues are nonnegative (nonpositive).

SECTIONS	MODULES	QUESTIONS	HOME PAGE
PREV. MODULE	FAQ	TUTOR	NEXT MODULE