

Inner product of 2 vectors (Read sec. 2.2)

- Inner product of 2 vectors x and y in \mathbb{R}^n :

$$x_1y_1 + x_2y_2 + \dots + x_ny_n \text{ in } \mathbb{R}^n$$

Notation: (x, y) or $y^T x$

- For complex vectors

$$(x, y) = x_1\bar{y}_1 + x_2\bar{y}_2 + \dots + x_n\bar{y}_n \text{ in } \mathbb{C}^n$$

Note: $(x, y) = y^H x$

An important property: Given $A \in \mathbb{C}^{m \times n}$ then

$$(Ax, y) = (x, A^H y) \quad \forall x \in \mathbb{C}^n, \forall y \in \mathbb{C}^m$$

✎ Show that when Q is orthogonal then $\|Qx\|_2 = \|x\|_2$

2-1

Csci 5304 – September 8, 2013

Norms are needed to measure lengths of vectors and closeness of two vectors. Examples of use: Estimate convergence rate of an iterative method; Estimate the error of an approximation to a given solution; ...

- A vector norm on a vector space \mathbb{X} is a real-valued function on \mathbb{X} , which satisfies the following three conditions:

1. $\|x\| \geq 0$, $\forall x \in \mathbb{X}$, and $\|x\| = 0$ iff $x = 0$.
2. $\|\alpha x\| = |\alpha| \|x\|$, $\forall x \in \mathbb{X}$, $\forall \alpha \in \mathbb{C}$.
3. $\|x + y\| \leq \|x\| + \|y\|$, $\forall x, y \in \mathbb{X}$.

- 3. is called the triangle inequality.

2-2

Csci 5304 – September 8, 2013

Important example: Euclidean norm on $\mathbb{X} = \mathbb{C}^n$,

$$\|x\|_2 = (x, x)^{1/2} = \sqrt{|x_1|^2 + |x_2|^2 + \dots + |x_n|^2}$$

- Most common vector norms in numerical linear algebra: special cases of the Hölder norms

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}.$$

✎ Find out (bbl search) how to show that these are indeed norms for any $p \geq 1$ (Not easy for 3rd requirement!)

A few properties:

- The limit of $\|x\|_p$ when p tends to infinity exists:

$$\lim_{p \rightarrow \infty} \|x\|_p = \max_{i=1}^n |x_i|$$

2-3

Csci 5304 – September 8, 2013

- Defines a norm denoted by $\|\cdot\|_\infty$.

- The cases $p = 1$, $p = 2$, and $p = \infty$ lead to the most important norms in practice. These are:

$$\begin{aligned} \|x\|_1 &= |x_1| + |x_2| + \dots + |x_n|, \\ \|x\|_2 &= [|x_1|^2 + |x_2|^2 + \dots + |x_n|^2]^{1/2}, \\ \|x\|_\infty &= \max_{i=1, \dots, n} |x_i|. \end{aligned}$$

- The Cauchy-Schwartz inequality (important) is:

$$|(x, y)| \leq \|x\|_2 \|y\|_2.$$

- The Hölder inequality (less important for $p \neq 2$) is:

$$|(x, y)| \leq \|x\|_p \|y\|_q, \text{ with } \frac{1}{p} + \frac{1}{q} = 1$$

2-4

Csci 5304 – September 8, 2013

Equivalence of norms:

In finite dimensional spaces ($\mathbb{R}^n, \mathbb{C}^n, \dots$) all norms are 'equivalent': if ϕ_1 and ϕ_2 are two norms then there is a constant α such that,

$$\phi_1(x) \leq \alpha \phi_2(x)$$

✎ How can you prove this result?

➤ We can bound one norm in terms of the other:

$$\beta \phi_2(x) \leq \phi_1(x) \leq \alpha \phi_2(x)$$

✎ Show that for any x : $\frac{1}{\sqrt{n}} \|x\|_1 \leq \|x\|_2 \leq \|x\|_1$

✎ What are the "unit balls" $B_p = \{x \mid \|x\|_p \leq 1\}$ associated with the norms $\|\cdot\|_p$ for $p = 1, 2, \infty$, in \mathbb{R}^2 ?

Convergence of vector sequences

A sequence of vectors $x^{(k)}, k = 1, \dots, \infty$ converges to a vector x with respect to the norm $\|\cdot\|$ if, by definition,

$$\lim_{k \rightarrow \infty} \|x^{(k)} - x\| = 0$$

➤ **Important point:** because all norms in \mathbb{R}^n are equivalent, the convergence of $x^{(k)}$ w.r.t. a given norm implies convergence w.r.t. any other norm.

➤ **Notation:**

$$\lim_{k \rightarrow \infty} x^{(k)} = x$$

Example: The sequence

$$x^{(k)} = \begin{pmatrix} 1 + 1/k \\ \frac{k}{k + \log_2 k} \\ \frac{1}{k} \end{pmatrix}$$

converges to

$$x = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

➤ **Note:** Convergence of $x^{(k)}$ to x is the same as the convergence of each individual component $x_i^{(k)}$ of $x^{(k)}$ to the corresponding component x_i of x .

Matrix norms

➤ See Sec. 2.3 of text

➤ Can define matrix norms by considering $m \times n$ matrices as vectors in \mathbb{R}^{mn} . These norms satisfy the usual properties of vector norms, i.e.,

1. $\|A\| \geq 0$, $\forall A \in \mathbb{C}^{m \times n}$, and $\|A\| = 0$ iff $A = 0$
2. $\|\alpha A\| = |\alpha| \|A\|$, $\forall A \in \mathbb{C}^{m \times n}$, $\forall \alpha \in \mathbb{C}$
3. $\|A + B\| \leq \|A\| + \|B\|$, $\forall A, B \in \mathbb{C}^{m \times n}$.

➤ However, these will lack (in general) the right properties for composition of operators (product of matrices).

➤ The case of $\|\cdot\|_2$ yields the Frobenius norm of matrices.

- Given a matrix A in $\mathbb{C}^{m \times n}$, define the set of **matrix norms**

$$\|A\|_p = \max_{x \in \mathbb{C}^n, x \neq 0} \frac{\|Ax\|_p}{\|x\|_p}.$$

- These norms satisfy the usual properties of vector norms (see previous page).
- The matrix norm $\|\cdot\|_p$ is **induced** by the vector norm $\|\cdot\|_p$.
- Again, important cases are for $p = 1, 2, \infty$.

Consistency / sub-multiplicativity of matrix norms

- A fundamental property of matrix norms is consistency

$$\|AB\|_p \leq \|A\|_p \|B\|_p.$$

[Also termed “sub-multiplicativity”]

- Consequence: $\|A^k\|_p \leq \|A\|_p^k$
- A^k converges to zero if **any** of its p -norms is < 1

[Note: sufficient but not necessary condition]

Frobenius norms of matrices

- The **Frobenius norm** of a matrix is defined by

$$\|A\|_F = \left(\sum_{j=1}^n \sum_{i=1}^m |a_{ij}|^2 \right)^{1/2}.$$

- Same as the 2-norm of the column vector in \mathbb{C}^{mn} consisting of all the columns (respectively rows) of A .
- This norm is also consistent [but not induced from a vector norm]

- ✎ Compute the Frobenius norms of the matrices

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 3 & 2 \end{pmatrix} \quad \begin{pmatrix} 1 & 2 & -1 \\ -1 & \sqrt{5} & 0 \\ -1 & 1 & \sqrt{2} \end{pmatrix}$$

- ✎ Prove that the Frobenius norm is consistent [Hint: Use Cauchy-Schwartz]
- ✎ Define the ‘vector 1-norm’ of a matrix A as the 1-norm of the vector of stacked columns of A . Is this norm a consistent matrix norm? [Hint: Result is true – Use Cauchy-Schwarz to prove it.]

Expressions of standard matrix norms

► Recall the notation: (for square $n \times n$ matrices)

$\rho(A) = \max |\lambda_i(A)|$; $Tr(A) = \sum_{i=1}^n a_{ii} = \sum_{i=1}^n \lambda_i(A)$
where $\lambda_i(A)$, $i = 1, 2, \dots, n$ are all eigenvalues of A

$$\|A\|_1 = \max_{j=1, \dots, n} \sum_{i=1}^m |a_{ij}|,$$

$$\|A\|_\infty = \max_{i=1, \dots, m} \sum_{j=1}^n |a_{ij}|,$$

$$\|A\|_2 = [\rho(A^H A)]^{1/2} = [\rho(AA^H)]^{1/2},$$

$$\|A\|_F = [Tr(A^H A)]^{1/2} = [Tr(AA^H)]^{1/2}.$$

► Eigenvalues of $A^H A$ are real ≥ 0 . Their square roots are **singular values** of A . To be covered later.

► $\|A\|_2$ == the largest singular value of A and $\|A\|_F$ = the 2-norm of the vector of all singular values of A .

✎ Compute the p -norm for $p = 1, 2, \infty, F$ for the matrix

$$A = \begin{pmatrix} 0 & 2 \\ 0 & 1 \end{pmatrix}$$

✎ Show that $\rho(A) \leq \|A\|$ for any matrix norm.

✎ Is $\rho(A)$ a norm?

1. $\rho(A) = \|A\|_2$ when A is Hermitian ($A^H = A$). ► True for this particular case...

2. ... However, not true in general. For

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

we have $\rho(A) = 0$ while $A \neq 0$. Also, triangle inequality not satisfied for the pair A , and $B = A^T$. Indeed, $\rho(A + B) = 1$ while $\rho(A) + \rho(B) = 0$.