KTH ROYAL INSTITUTE OF TECHNOLOGY

Lecture 5

Ch. 5, Norms for vectors and matrices

Emil Björnson/Magnus Jansson/Mats Bengtsson

April 27, 2016

Norms for vectors and matrices — Why?

Problem: Measure size of vector or matrix.
What is "small" and what is "large"?

Problem: Measure distance between vectors or matrices.
When are they "close together" or "far apart"?

Answers are given by norms.

Also: Tool to analyze convergence and stability of algorithms.

Vector norm — axiomatic definition

Definition: Let V be a vector space over a field $\mathbf{F}$ ($\mathbf{R}$ or $\mathbf{C}$).
A function $||\cdot|| : V \to \mathbf{R}$ is a vector norm if for all $x, y \in V$

(1)	$\|\|x\|\| \geq 0$	nonnegative
(1a)	$\|\|x\|\| = 0$ iff $x = 0$	positive
(2)	$\|\|cx\|\| = \|c\|\ \|\|x\|\|$ for all $c \in \mathbf{F}$	homogeneous
(3)	$\|\|x + y\|\| \leq \|\|x\|\| + \|\|y\|\|$	triangle inequality

A function not satisfying (1a) is called a vector **seminorm**.

Interpretation: Size/length of vector.

Inner product — axiomatic definition

Definition: Let V be a vector space over a field $\mathbf{F}$ ($\mathbf{R}$ or $\mathbf{C}$).
A function $\langle\cdot,\cdot\rangle : V \times V \to \mathbf{F}$ is an inner product if for all $x, y, z \in V$,

(1)	$\langle x, x\rangle \geq 0$	nonnegative
(1a)	$\langle x, x\rangle = 0$ iff $x = 0$	positive
(2)	$\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$	additive
(3)	$\langle cx, y\rangle = c\langle x, y\rangle$ for all $c \in \mathbf{F}$	homogeneous
(4)	$\langle x, y\rangle = \overline{\langle y, x\rangle}$	Hermitian property

Interpretation: "Angle" (distance) between vectors.

Connections between norm and inner products

Corollary: If $\langle \cdot, \cdot \rangle$ is an inner product, then $||x|| = (\langle x, x \rangle)^{1/2}$ is a vector norm.
Called: **Vector norm derived from an inner product**.
Satisfies parallelogram identity (Necessary and sufficient condition):

$$\frac{1}{2}(||x+y||^2 + ||x-y||^2) = ||x||^2 + ||y||^2$$

Theorem (Cauchy-Schwarz inequality):

$$|\langle x, y \rangle|^2 \leq \langle x, x \rangle \langle y, y \rangle$$

We have equality iff $x = cy$ for some $c \in \mathbf{F}$ (i.e., linearly dependent)

Examples

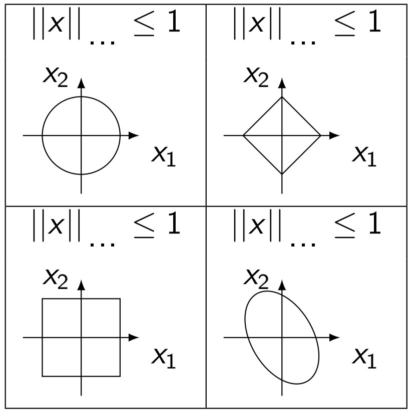
- The **Euclidean norm** (l_2) on $\mathbf{C}^n$:
$$||x||_2 = (|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2)^{1/2}.$$
- The **sum norm** (l_1), also called one-norm or Manhattan norm:
$$||x||_1 = |x_1| + |x_2| + \cdots + |x_n|.$$
- The **max norm** (l_∞):
$$||x||_\infty = \max\{|x_1|, |x_2|, \ldots, |x_n|\}$$

The sum and max norms cannot be derived from an inner product!

Unit balls for different norms

The shape of the unit ball characterizes the norm.

Fill in which norm corresponds to which unit ball!



Properties: Convex and compact (for finite dimensions), includes the origin.

Examples cont'd

- The l_p-norm on $\mathbf{C}^n$ is ($p \geq 1$):
$$||x||_p = (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{1/p}$$
- Norms may also be constructed from others, e.g.,:
$$||x|| = \max\{||x||_{p_1}, ||x||_{p_2}\}$$
or let nonsingular $T \in M_n$ and $|| \cdot ||$ be a given, then
$$||x||_T = ||Tx||.$$
(same notation sometimes used for $||x||_W = x^* W x$)
- Norms on infinite-dimensional vector spaces (e.g., all continuous functions on an interval $[a, b]$): "similarly" defined (sums become integrals)

Convergence

Assume: Vector space V over $\mathbf{R}$ or $\mathbf{C}$.

Definition: The sequence $\{x^{(k)}\}$ of vectors in V converges to $x \in V$ with respect to $||\cdot||$ iff

$$||x^{(k)} - x|| \to 0 \quad \text{as} \quad k \to \infty.$$

Infinite dimension:

- Sequence can converge in one norm, but not another.
- Important to state choice of norm.

Convergence: Finite dimension

Corollary: For any vector norms $||\cdot||_\alpha$ and $||\cdot||_\beta$ on a finite-dimensional V, there exists $0 \le C_m < C_M < \infty$ such that

$$C_m||x||_\alpha \le ||x||_\beta \le C_M||x||_\alpha \qquad \forall x \in V$$

Conclusion: Convergence in one norm $\Rightarrow$ convergence in all norms.

Note: Result also holds for **pre-norms**, without the triangle inequality.

Definition: Two norms are **equivalent** if convergence in one of the norms always implies convergence in the other.

Conclusion: All norms are equivalent in the finite dimensional case.

Convergence: Cauchy sequence

Definition: A sequence $\{x^{(k)}\}$ in V is a Cauchy sequence with respect to $||\cdot||$ if for every $\epsilon > 0$ there is a $N_\epsilon > 0$ such that

$$||x^{(k_1)} - x^{(k_2)}|| \le \epsilon$$

for all $k_1, k_2 \ge N_\epsilon$.

Theorem: A sequence $\{x^{(k)}\}$ in a finite dimensional V converges to a vector in V iff it is a Cauchy sequence.

Dual norms

Definition: The dual norm of $\|\cdot\|$ is

$$\|y\|^D = \max_{x:\|x\|=1} \operatorname{Re} y^*x = \max_{x:\|x\|=1} |y^*x| = \max_{x\neq 0} \frac{|y^*x|}{\|x\|}$$

Examples:

Norm	Dual norm
$\|\cdot\|_2$	$\|\cdot\|_2$
$\|\cdot\|_1$	$\|\cdot\|_\infty$
$\|\cdot\|_\infty$	$\|\cdot\|_1$

- Dual of dual norm is the original norm.
- Euclidean norm is its own dual.
- Generalized Cauchy-Schwarz: $|y^*x| \le \|x\|\|y\|^D$

Vector norms applied to matrices

M_n is a vector space (of dimension n^2)

Conclusion: We can apply vector norms to matrices.

Examples: The l_1 norm: $||A||_1 = \sum_{i,j} |a_{ij}|$.
The l_2 norm (Euclidean/Frobenius norm):
$||A||_2 = \left(\sum_{i,j} |a_{ij}|^2\right)^{1/2}$.
The l_∞ norm: $||A||_\infty = \max_{i,j} |a_{ij}|$.

Observation: Matrices have certain properties (e.g., multiplication).
May be useful to define particular **matrix norms**.

Matrix norm — axiomatic definition

Definition: $|||\cdot||| : M_n \to \mathbf{R}$ is a matrix norm if for all $A, B \in M_n$,

(1)	$|||A||| \geq 0$	nonnegative
(1a)	$|||A||| = 0$ iff $A = 0$	positive
(2)	$|||cA||| = |c|\,|||A|||$ for all $c \in \mathbf{C}$	homogeneous
(3)	$|||A+B||| \leq |||A||| + |||B|||$	triangle inequality
(4)	$|||AB||| \leq |||A|||\,|||B|||$	submultiplicative

Observations:
- All vector norms satisfy (1)-(3), some may satisfy (4).
- **Generalized matrix norm** if not satisfying (4).

Which vector norms are matrix norms?

$||A||_1$ and $||A||_2$ are matrix norms.

$||A||_\infty$ is not a matrix norm (but a generalized matrix norm).

However, $|||A||| = n||A||_\infty$ is a matrix norm.

Induced matrix norms

Definition: Let $||\cdot||$ be a vector norm on $\mathbf{C}^n$. The matrix norm

$$|||A||| = \max_{||x||=1} ||Ax||$$

is **induced** by $||\cdot||$.

Properties of induced norms $|||\cdot|||$:
- $|||I||| = 1$.
- The only matrix norm $N(A)$ with $N(A) \leq |||A|||$ for all $A \in M_n$ is $N(\cdot) = |||\cdot|||$.

Last property called **minimal matrix norm**.

Examples

The maximum column sum (induced by l_1):

$$|||A|||_1 = \max_j \sum_i |a_{ij}|$$

The spectral norm (induced by l_2):

$$|||A|||_2 = \max\{\sqrt{\lambda} : \lambda \in \sigma(A^*A)\}$$

The maximum row sum (induced by l_∞):

$$|||A|||_\infty = \max_i \sum_j |a_{ij}|$$

Application: Computing Spectral radius

Recall: Spectral radius: $\rho(A) = \max\{|\lambda| : \lambda \in \sigma(A)\}$.
Not a matrix norm, but very related.

Theorem: For any matrix norm $|||\cdot|||$ and $A \in M_n$,

$$\rho(A) \leq |||A|||.$$

Lemma: For any $A \in M_n$ and $\epsilon > 0$, there is $|||\cdot|||$ such that

$$\rho(A) \leq |||A||| \leq \rho(A) + \epsilon$$

Corollary: For any matrix norm $|||\cdot|||$ and $A \in M_n$,

$$\rho(A) = \lim_{k\to\infty} |||A^k|||^{1/k}$$

Application: Convergence of A^k

Lemma: If there is a matrix norm with $|||A||| < 1$ then $\lim_{k\to\infty} A^k = 0$.

Theorem: $\lim_{k\to\infty} A^k = 0$ iff $\rho(A) < 1$.

Matrix extension of $\lim_{k\to\infty} x^k = 0$ iff $|x| < 1$.

Application: Power series

Theorem: $\sum_{k=0}^{\infty} a_k A^k$ converges if there is a matrix norm such that $\sum_{k=0}^{\infty} |a_k|\, |||A|||^k$ converges.

Corollary: If $|||A||| < 1$ for some matrix norm, then $I - A$ is invertible and

$$(I - A)^{-1} = \sum_{k=0}^{\infty} A^k$$

Matrix extension of $(1 - x)^{-1} = \sum_{k=0}^{\infty} x^k$ for $|x| < 1$.

Useful to compute "error" between A^{-1} and $(A + E)^{-1}$.

Unitarily invariant and condition number

Definition: A matrix norm is unitarily invariant if $|||UAV||| = |||A|||$ for all $A \in M_n$ and all unitary matrices $U, V \in M_n$.

Examples: Frobenius norm $||\cdot||_2$ and spectral norm $|||\cdot|||_2$.

Definition: **Condition number for matrix inversion** with respect to the matrix norm $|||\cdot|||$ of nonsingular $A \in M_n$ is

$$\kappa(A) = |||A^{-1}|||\ |||A|||$$

Frequently used in perturbation analysis in numerical linear algebra.

Observation: $\kappa(A) \geq 1$ (from submultiplicative property).
Observation: For unitarily invariant norms: $\kappa(UAV) = \kappa(A)$.