Inner products

An **inner product** $(\cdot,\cdot)$ is a mapping from (S,S), where S is a linear space (vector, function, etc.), to $\mathbb{R}$ (more generally to $\mathbb{C}$), with the following properties: For any $x, y, z \in S$,

(i) $(x, x) \geq 0$, and $(x, x) = 0$ iff $x = 0 \in S$.

(ii) $(\alpha x, y) = \alpha (x, y)$, $\forall \alpha \in \mathbb{R}$ ($(\alpha x, y) = \bar{\alpha} (x, y)$, $\forall \alpha \in \mathbb{C}$)

(iii) $(x + y, z) = (x, z) + (y, z)$ ("triangle" equality).

(iv) $(x, y) = (y, x)$ ($(x, y) = \overline{(y, x)}$, for $(x, y) \in \mathbb{C}$)

A commonly used inner product for vectors is the dot product, $(x, y) = x^T \cdot y = \sum_{i=1}^{n} x_i \cdot y_i$; ($(x, y) = x^H \cdot y = \sum_{i=1}^{n} \bar{x}_i \cdot y_i$ for $x, y \in \mathbb{C}^n$).

The following can be proved:

- **Cauchy-Schwarz inequality**: $|(x, y)| \leq (x, x)^{1/2} (y, y)^{1/2}$

Note: An inner product is denoted by $(\cdot,\cdot)$ or by $< \cdot, \cdot >$.

Vector, Matrix and Function Norms

A **norm** $\|\cdot\|$ is a mapping from a space S (vector, function, etc.) to $\mathbb{R}^+ \cup \{0\}$, with the following properties: For any $x, y \in S$,

(i) $\|x\| \geq 0$, and $\|x\| = 0$ iff $x = 0 \in S$.

(ii) $\|\alpha x\| = |\alpha| \|x\|$ for all scalars α.

(iii) $\|x + y\| \leq \|x\| + \|y\|$ (**triangle** or **Minowski's inequality**).

For each inner product $(\cdot,\cdot)$, a respective norm is defined by $\|x\| = (x, x)^{1/2}$.

A norm is a way to quantify (with a single number) a multidimensional quantity. (Consider the Euclidean length of vectors.)

Distance or **error** between vectors x, y: $\|x - y\|$.

Relative error between vectors x, y, with respect to x: $\dfrac{\|x - y\|}{\|x\|}$.

Vector norms

Let $x = (x_1, x_2, \cdots, x_n)^T$. Some common vector norms are:

- p-norm or Hölder norm: $\|x\|_p \equiv (\sum_{i=1}^{n} |x_i|^p)^{1/p}$
- **max (infinity) norm**: $\|x\|_\infty \equiv \max_{i=1}^{n} \{|x_i|\}$
- **Euclidean norm (length, L2, two-norm)**: $\|x\|_2 \equiv \sqrt{(x, x)} \equiv (\sum_{i=1}^{n} x_i^2)^{1/2}$
- **one-norm**: $\|x\|_1 \equiv \sum_{i=1}^{n} |x_i|$

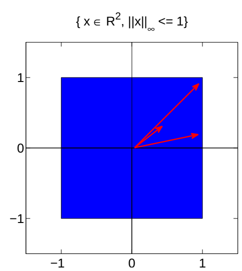
Some relations:

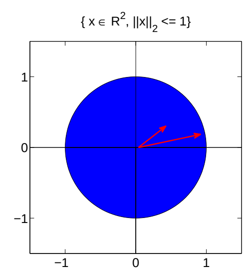
- **Cauchy-Schwarz inequality**: $|(x, y)| \leq (x, x)^{1/2} (y, y)^{1/2}$
- $(x, y) = \|x\|_2 \|y\|_2 \cos\theta$
- $|(x, y)| \leq \|x\|_2 \|y\|_2$
- **Hölder inequality**: If $p, q > 0$ and $\frac{1}{p} + \frac{1}{q} = 1$, then $|(x, y)| \leq \|x\|_p \|y\|_q$.

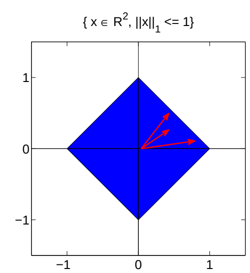
Vector norms -- some relations and some notes

For all vectors $x \in \mathbb{R}^n$, we have $\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1$.

The vectors $x \in \mathbb{R}^2$ for which $\|x\|_\infty \leq 1$ lie in the unit rectangle.
The vectors $x \in \mathbb{R}^2$ for which $\|x\|_2 \leq 1$ lie in the unit circle.
The vectors $x \in \mathbb{R}^2$ for which $\|x\|_1 \leq 1$ lie in the unit rhombus.







For any vector $x \in \mathbb{R}^n$, we have (**equivalence of vector norms**)

- $\|x\|_2 \leq \|x\|_1 \leq \sqrt{n} \|x\|_2$
- $\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n} \|x\|_\infty$
- $\|x\|_\infty \leq \|x\|_1 \leq n \|x\|_\infty$

Matrix norms

Let A be an $m \times n$ matrix. Some common matrix norms are:

- p-**norm** (or **induced** or Hölder norm), $p = 1, 2, \infty$: $\|A\|_p \equiv \max_{x \neq 0} \{ \frac{\|Ax\|_p}{\|x\|_p} \}$

  (or **natural**, or **subordinate**, or **associated** norm)

For the p-norms, properties (iv), (v) and (vi) are also valid:

(iv) $\|Ax\| \leq \|A\|\|x\|$

(v) $\|\mathbf{I}\| = 1$

(vi) $\|AB\| \leq \|A\|\|B\|$

There are other matrix norms, which are not p-norms, and for which the above three properties may not hold.

When referring to matrix norms, a p-norm is assumed, unless otherwise stated.

Matrix norms

The following can be proved:

- $\max_{x \neq 0} \{ \frac{\|Ax\|_p}{\|x\|_p} \} = \max_{\|x\|_p = 1} \{ \|Ax\|_p \}$
- $\|A\|_\infty = \max_{i=1}^{m} \{ \sum_{j=1}^{n} |a_{ij}| \}$ (**row norm**)
- $\|A\|_1 = \max_{j=1}^{n} \{ \sum_{i=1}^{m} |a_{ij}| \}$ (**column norm**)
- Other: There is no easy way to compute the Euclidean norm of a matrix. Requires knowledge on eigenvalues.

For any matrix $A \in \mathbb{R}^{m \times n}$, we have (**equivalence of matrix norms**)

- $\frac{1}{\sqrt{n}} \|A\|_2 \leq \|A\|_1 \leq \sqrt{m} \|A\|_2$
- $\frac{1}{\sqrt{n}} \|A\|_\infty \leq \|A\|_2 \leq \sqrt{m} \|A\|_\infty$
- $\frac{1}{m} \|A\|_1 \leq \|A\|_\infty \leq n \|A\|_1$

- Furthermore, $\|A\|_2^2 \leq \|A\|_1 \|A\|_\infty$

Condition number of a matrix

The **condition number** of a non-singular matrix A is given by $\kappa_a(A) = \|A\|_a \|A^{-1}\|_a$. By convention, the condition number of a singular matrix (with respect to any norm) is infinity.

The condition number of a matrix A is a measure of the relative sensitivity of the solution x of $Ax = b$ to relative changes in A and b. The following can be proved to hold (approximately):

$$\frac{\|x - \hat{x}\|_a}{\|x\|_a} \leq \kappa_a(A) \left(\frac{\|b - \hat{b}\|_a}{\|b\|_a} + \frac{\|A - \hat{A}\|_a}{\|A\|_a} + \frac{\|r\|_a}{\|b\|_a} \right)$$

where $Ax = b$ is the linear system to be solved on a computer, $\hat{A}$ and $\hat{b}$ are the computer representations of A and b, respectively, x and $\hat{x}$ are the (exact) mathematical and the computed solutions, respectively, and $r = \hat{b} - \hat{A}\hat{x}$ is the **residual** corresponding to the computed solution (and reflecting all operations' errors occurring during the process of solution).

- $\kappa_a(A) \geq 1$ (amplification factor), $\kappa_a(\mathbf{I}) = 1$.
- **Well-conditioned** matrix: $\kappa_a(A) \approx 1$.
- **Ill-conditioned** matrix: $\kappa_a(A) \gg 1$. ($\kappa_a(A) \approx \varepsilon_{\text{mach}}^{-1}$.)

Condition number of a matrix -- some notes

- A large condition number is associated with large uncertainty in the solution.

- If $\kappa_a(A) \geq 10^d$, and the computer system has precision of d decimal digits, the solution of a linear system with matrix A in this computer system may give a solution without any correct digits.

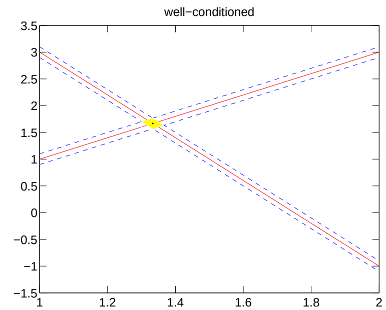
- If $\kappa_a(A) = 10^d$, and the computer system has precision of $e > d$ decimal digits, the solution of a linear system with matrix A in this computer system is expected to give a solution with at least $e - d$ correct digits.

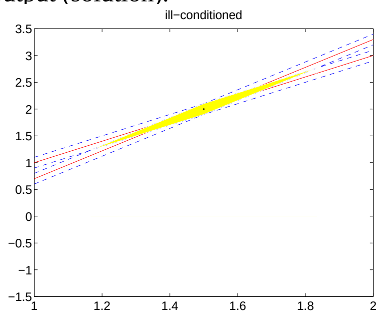
- Small residual ($r = b - A\hat{x}$) does not necessarily imply a small error ($e = x - \hat{x}$). If $\kappa_a(A)$ is large, it may happen that r is small and the error large. (It can be shown that $\kappa^{-1} \frac{\|r\|}{\|b\|} \leq \frac{\|e\|}{\|x\|} \leq \kappa \frac{\|r\|}{\|b\|}$)

- The condition number of a matrix indicates how close to singular (not solvable) a linear system with the given matrix is.

Condition number of a matrix -- geometric interpretation

Geometrically speaking, in the two-dimensional space, a 2×2 linear system is represented by two straight lines, one for each equation. If the two lines are parallel, the system has no solution. If they cross each other, the system has a unique solution. If they overlap, the system has infinitely many solutions. If they cross each other at a small angle (almost parallel lines), the system is ill-conditioned, that is, a small error in the input may give rise to a large error in the output (solution).





The red lines are the exact representation of the equations. The bands within the dashed blue lines are perturbations of the two equations. The exact solution lies on the intersection of the two lines (dot). The inexact solution may lie anywhere in the overlapping area of the bands (yellow).

Ill-conditioned matrix -- example

The matrix

$$A = \begin{bmatrix} 1+\varepsilon & 1 \\ 1 & 1+\varepsilon \end{bmatrix}$$

is ill-conditioned for small ε, i.e. an ε such that $|\varepsilon| \approx 0$.

Consider also the linear systems

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} 1+\varepsilon & 1 \\ 1 & 1+\varepsilon \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

The one to the left does not have a solution, while the one to the right has a unique solution, however, the computations for a specific small ε to compute the solution are unstable and may involve large errors.

Condition number of matrix -- other interpretations

• The condition number of a matrix measures the ratio of the largest relative stretching over the largest relative shrinking that the matrix may do to an arbitrary non-zero vector:

$$\kappa_a(A) = \|A\|_a \|A^{-1}\|_a = \max_{x \neq 0} \frac{\|Ax\|_a}{\|x\|_a} \max_{x \neq 0} \frac{\|A^{-1}x\|_a}{\|x\|_a}$$

$$= \max_{x \neq 0} \frac{\|Ax\|_a}{\|x\|_a} \max_{y \neq 0} \frac{\|y\|_a}{\|Ay\|_a} = \frac{\max_{x \neq 0} \frac{\|Ax\|_a}{\|x\|_a}}{\min_{y \neq 0} \frac{\|Ay\|_a}{\|y\|_a}}$$

• Let $Q \in \mathbb{R}^{m \times n}$ be an orthogonal matrix. Then $\|Q\|_2 = 1$.
• Let $Q \in \mathbb{R}^{n \times n}$ be an orthogonal matrix. Then $\kappa_2(Q) = 1$.

Equivalence of condition numbers of a matrix

- $\frac{1}{n} \kappa_2(A) \le \kappa_1(A) \le n \kappa_2(A)$
- $\frac{1}{n} \kappa_\infty(A) \le \kappa_2(A) \le n \kappa_\infty(A)$
- $\frac{1}{n^2} \kappa_1(A) \le \kappa_\infty(A) \le n^2 \kappa_1(A)$
- In simple words: the value of one condition number of a matrix cannot differ ‘‘very dramatically’’ from the value of another condition number of the same matrix; however, condition numbers of a given matrix may differ more than norms of the same matrix.