Methods for solving linear systems

Reminders on norms and scalar products of vectors. The application $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ is a norm if

$$1 - \|\underline{v}\| \geq 0, \ \forall \underline{v} \in \mathbb{R}^n \qquad \|\underline{v}\| = 0 \text{ if and only if } \underline{v} = 0;$$
$$2 - \|\alpha \underline{v}\| = |\alpha| \|\underline{v}\| \quad \forall \alpha \in \mathbb{R}, \quad \forall \underline{v} \in \mathbb{R}^n;$$
$$3 - \|\underline{v} + \underline{w}\| \leq \|\underline{v}\| + \|\underline{w}\|, \quad \forall \underline{v}, \underline{w} \in \mathbb{R}^n.$$

Examples of norms of vectors:

$$\|\underline{v}\|_2^2 = \sum_{i=1}^{n} (v_i)^2 \qquad \text{Euclidean norm}$$
$$\|\underline{v}\|_\infty = \max_{1 \leq i \leq n} |v_i| \qquad \text{max norm}$$
$$\|\underline{v}\|_1 = \sum_{i=1}^{n} |v_i| \qquad \text{1-norm}$$

Being in finite dimension, they are all equivalent, with the equivalence constants depending on the dimension n. Ex: $\|v\|_\infty \leq \|v\|_1 \leq n\|v\|_\infty$.

A scalar product is an application $(\cdot,\cdot) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ that verifies:

$1-$ linearity: $(\alpha\underline{v} + \beta\underline{w}, \underline{z}) = \alpha(\underline{v}, \underline{z}) + \beta(\underline{w}, \underline{z})\ \forall \alpha, \beta \in \mathbb{R},\ \forall \underline{v}, \underline{w}, \underline{z} \in \mathbb{R}^n$;

$2-(\underline{v}, \underline{w}) = (\underline{w}, \underline{v})\ \forall \underline{v}, \underline{w} \in \mathbb{R}^n$;

$3-(\underline{v}, \underline{v}) > 0\ \ \forall \underline{v} \neq \underline{0}$ (that is, $(\underline{v}, \underline{v}) \geq 0, (\underline{v}, \underline{v}) = 0$ iff $\underline{v} = \underline{0}$).

To a scalar product we can associate a norm defined as

$$\|\underline{v}\|^2 = (\underline{v}, \underline{v}).$$

Example: $(\underline{v}, \underline{w}) = \sum_{i=1}^{n} v_i w_i, \implies (\underline{v}, \underline{v}) = \sum_{i=1}^{n} v_i v_i = \|\underline{v}\|_2^2.$

(in this case we can write $(\underline{v}, \underline{w}) = \underline{v} \cdot \underline{w}$ or $\underline{v}^T \underline{w}$ for "column" vectors)

Theorem 1 (Cauchy-Schwarz inequality)

Given a scalar product $(\cdot,\cdot)_*$ *and associated norm* $\|\cdot\|_*$*, the following inequality holds:*

$$|(\underline{v},\underline{w})_*| \leq \|\underline{v}\|_*\|\underline{w}\|_* \qquad \forall v, w \in \mathbb{R}^n$$

Proof.

If $\underline{v} = 0$, the claim is immediate. Otherwise, for $t \in \mathbb{R}$, let $t\underline{v} + \underline{w} \in \mathbb{R}^n$. Clearly, $\|t\underline{v} + \underline{w}\|_* \geq 0$. Hence:

$$\|t\underline{v} + \underline{w}\|_*^2 = t^2\|\underline{v}\|_*^2 + 2t(\underline{v},\underline{w})_* + \|\underline{w}\|_*^2 \geq 0$$

The last expression is a non-negative convex parabola in t (No real roots, or 2 coincident). Then the discriminant is non-positive

$$(\underline{v},\underline{w})_*^2 - \|\underline{v}\|_*^2\|\underline{w}\|_*^2 \leq 0$$

and the proof is concluded. □

Reminders on matrices $A \in \mathbb{R}^{n \times n}$

- A is symmetric if $A = A^T$. The eigenvalues of a symmetric matrix are real.
- A symmetric matrix A is positive definite if

$$(A\underline{x}, \underline{x})_2 > 0 \ \forall \underline{x} \in \mathbb{R}^n, \ \underline{x} \neq 0, \quad (A\underline{x}, \underline{x})_2 = 0 \text{ iff } \underline{x} = 0$$

  The eigenvalues of a positive definite matrix are positive.
- if A is non singular, $A^T A$ is symmetric and positive definite

Proof of the last statement:

- $A^T A$ is always symmetric; indeed $(A^T A)^T = A^T (A^T)^T = A^T A$.

To prove that it is also positive definite we have to show that $(A^T A\underline{x}, \underline{x})_2 > 0 \ \forall \underline{x} \in \mathbb{R}^n, \ \underline{x} \neq 0, \quad (A^T A\underline{x}, \underline{x})_2 = 0$ iff $\underline{x} = 0$. We have:

$$(A^T A\underline{x}, \underline{x})_2 = (A\underline{x}, A\underline{x})_2 = \|A\underline{x}\|_2^2 \geq 0, \text{ and } \|A\underline{x}\|_2^2 = 0 \text{ iff } A\underline{x} = \underline{0}$$

If A is non-singular (i.e., $det(A) \neq 0$), the system $A\underline{x} = \underline{0}$ has only the solution $\underline{x} = \underline{0}$, and this ends the proof.

Norms of matrices

Norms of matrices are applications from $\mathbb{R}^{m\times n}$ to $\mathbb{R}$ satisfying the same properties as for vectors. Among the various norms of matrices we will consider the norms associated to norms of vectors, called natural norms, defined as:

$$|||A||| = \sup_{\underline{v}\neq 0} \frac{\|A\underline{v}\|}{\|\underline{v}\|}$$

It can be checked that this is indeed a norm, that moreover verifies:

$$\|A\underline{v}\| \leq |||A|||\,\|\underline{v}\|, \qquad |||AB||| \leq |||A|||||B|||.$$

Examples of natural norms (of square $n \times n$ matrices)

$$\underline{v} \in \mathbb{R}^n; \quad \|\underline{v}\|_\infty \longrightarrow |||A|||_\infty = \max_{i=1,\cdots,n} \sum_{j=1}^{n} |a_{ij}|,$$

$$\underline{v} \in \mathbb{R}^n; \quad \|\underline{v}\|_1 \longrightarrow |||A|||_1 = \max_{j=1,\cdots,n} \sum_{i=1}^{n} |a_{ij}|,$$

$$\underline{v} \in \mathbb{R}^n; \quad \|\underline{v}\|_2 \longrightarrow |||A|||_2 = \sqrt{|\lambda_{\max}(A^T A)|}.$$

If A is symmetric, $|||A|||_\infty = |||A|||_1$, and $|||A|||_2 = |\lambda_{\text{max in abs val}}(A)|$.
Indeed, if $A = A^T$, then
$\max_i \lambda_i(A^T A) = \max_i \lambda_i(A^2) = \max_i \lambda_i(A)^2 = (\max_i |\lambda_i(A)|)^2$.

The *spectral norm* $|||A|||_2$ is the largest singular value of A, i.e., $\sqrt{\lambda_{\max}(A^T A)}$. For symmetric A, it equals $\max_i |\lambda_i(A)|$.

Complex vectors and norms

For $z = a + \mathrm{i}b \in \mathbb{C}$, let

$$\overline{z} = a - \mathrm{i}b, \qquad |z| = \sqrt{z\overline{z}} = \sqrt{a^2 + b^2}.$$

A norm $\|\cdot\| : \mathbb{C}^n \to \mathbb{R}$ satisfies the same positivity and triangle inequalities as in the real case, and

$$\|\alpha \underline{v}\| = |\alpha| \, \|\underline{v}\| \qquad (\alpha \in \mathbb{C},\ \underline{v} \in \mathbb{C}^n).$$

The familiar examples become

$$\|\underline{v}\|_1 = \sum_i |v_i|, \qquad \|\underline{v}\|_2 = \left(\sum_i |v_i|^2 \right)^{1/2}, \qquad \|\underline{v}\|_\infty = \max_i |v_i|.$$

Here the absolute values are complex moduli. For example, $\underline{v} = (1, \mathrm{i})^T$ has $\|\underline{v}\|_2 = \sqrt{2}$, although $\underline{v}^T \underline{v} = 0$: complex conjugation is essential.

Complex scalar products

For $\underline{v} \in \mathbb{C}^n$, write $\underline{v}^* = \overline{\underline{v}}^T$. A complex scalar product is linear in its first argument (our convention), conjugate symmetric and positive definite:

$$\begin{aligned}(\alpha\underline{v} + \beta\underline{w}, \underline{z}) &= \alpha(\underline{v}, \underline{z}) + \beta(\underline{w}, \underline{z}),\\ (\underline{v}, \underline{w}) &= \overline{(\underline{w}, \underline{v})},\\ (\underline{v}, \underline{v}) \in \mathbb{R}, &\quad (\underline{v}, \underline{v}) > 0 \quad \text{if } \underline{v} \neq 0.\end{aligned}$$

Consequently it is conjugate linear in its second argument: $(\underline{v}, \alpha\underline{w}) = \overline{\alpha}(\underline{v}, \underline{w})$. The Euclidean scalar product is

$$(\underline{v}, \underline{w})_2 = \underline{w}^*\underline{v} = \sum_i v_i\overline{w_i}, \qquad (\underline{v}, \underline{v})_2 = \sum_i |v_i|^2.$$

As over $\mathbb{R}$, $\|\underline{v}\| = \sqrt{(\underline{v}, \underline{v})}$ is a norm and

$$|(\underline{v}, \underline{w})| \leq \|\underline{v}\|\,\|\underline{w}\| \qquad \text{(Cauchy–Schwarz)}.$$

Complex matrices: adjoints and positivity

For $A \in \mathbb{C}^{n\times n}$, its adjoint is $A^* = \overline{A}^{\,T}$; in particular $(AB)^* = B^*A^*$.

- A is **Hermitian** if $A = A^*$, the complex analogue of symmetry. Its eigenvalues are real.
- A Hermitian matrix is **positive definite** if

$$\underline{x}^* A\underline{x} = (A\underline{x}, \underline{x})_2 > 0 \qquad (0 \neq \underline{x} \in \mathbb{C}^n).$$

  Equivalently, all its eigenvalues are positive.
- If A is nonsingular, A^*A is Hermitian positive definite:

$$(A^*A)^* = A^*A, \qquad \underline{x}^* A^* A\underline{x} = \|A\underline{x}\|_2^2 > 0 \quad (\underline{x} \neq 0).$$

A real matrix may have complex eigenvalues and eigenvectors. For example, $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ has eigenvalues i and $-\mathrm{i}$, so spectral arguments naturally use $\mathbb{C}^n$.

Natural matrix norms over the complex field

For any vector norm on $\mathbb{C}^n$, its induced (natural) matrix norm is

$$|||A||| = \sup_{0 \neq \underline{v} \in \mathbb{C}^n} \frac{\|A\underline{v}\|}{\|\underline{v}\|}, \qquad A \in \mathbb{C}^{n \times n}.$$

It satisfies $\|A\underline{v}\| \leq |||A||| \, \|\underline{v}\|$ and $|||AB||| \leq |||A||| \, |||B|||$. In particular,

$$|||A|||_1 = \max_j \sum_i |a_{ij}|, \qquad |||A|||_\infty = \max_i \sum_j |a_{ij}|,$$

$$|||A|||_2 = \sqrt{\lambda_{\max}(A^* A)}.$$

For Hermitian A, $|||A|||_2 = \max_i |\lambda_i(A)|$ and $|||A|||_1 = |||A|||_\infty$.

For any A, the **spectral radius** is $\rho(A) = \max_i |\lambda_i(A)|$; it is not in general a matrix norm. The eigenvector proof of $\rho(A) \leq |||A|||$ in slide set 07 uses precisely these complex vector and matrix norms.