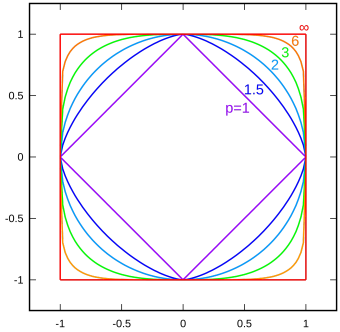
Algorithms and Data Structures

Numerics

Matrix Norm



Learning goals

- Definition of matrix norm
- Inequalities of matrix norm

REMINDER: MATRIX NORM

Motivation:
In statistics we are often confronted with matrices (e.g. design matrix $\mathbf{X}$). To perform an error analysis for related LES, meaning we want to know if the given problem is well-conditioned, the matrix norm is used.

Definition:
$\|\cdot\| : \mathbb{R}^n \to \mathbb{R}_0^+$ is called norm, if:

- $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$ (positive definite),
- $\|a\mathbf{x}\| = |a|\|\mathbf{x}\|$ (homogeneity),
- $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ (triangle inequality).

General p norm of a vector $\mathbf{x} \in \mathbb{R}^n$:

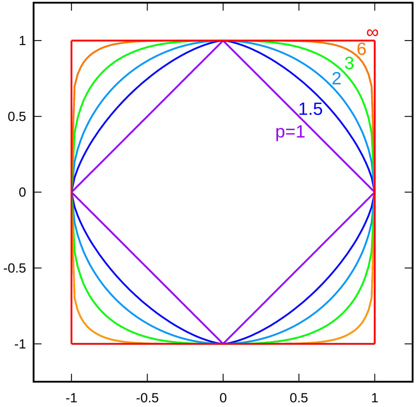
$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p},$$

where $\|\mathbf{x}\|_\infty = \max_i(|x_i|)$.

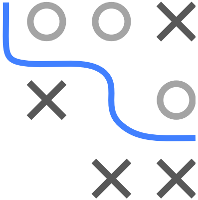
REMINDER: MATRIX NORM / 2

Examples:

$$\|\mathbf{x}\|_1 = \sum_i |x_i| \qquad \|\mathbf{x}\|_2 = \sqrt{\mathbf{x}^T\mathbf{x}} \qquad \|\mathbf{x}\|_\infty = \max_i |x_i|$$



REMINDER: MATRIX NORM / 3



Corresponding **matrix norm** (for $\mathbf{A} \in \mathbb{R}^{n\times n}$) is defined as

$$\|\mathbf{A}\|_p := \sup_{\mathbf{x}\neq\mathbf{0}} \left(\frac{\|\mathbf{A}\mathbf{x}\|_p}{\|\mathbf{x}\|_p} \right) = \sup_{\|\mathbf{x}\|_p=1} \left(\|\mathbf{A}\mathbf{x}\|_p \right).$$

Examples for matrix norms induced by vector norms:

- $\|\mathbf{A}\|_1 = \max_j \left(\sum_i |A_{ij}| \right)$ (maximum absolute column sum norm)

$$\mathbf{A} = \begin{pmatrix} 1 & -2 & -3 \\ 2 & 3 & -1 \end{pmatrix} \Rightarrow \|\mathbf{A}\|_1 \begin{aligned}[t] &= \max(\|A_1\|_1, \|A_2\|_1, \|A_3\|_1) \\ &= \max(3, 5, 4) = 5 \end{aligned}$$

- $\|\mathbf{A}\|_2 = \left(\text{largest eigenvalue of } \mathbf{A}^\top\mathbf{A}\right)^{1/2}$ (spectral norm)
- $\|\mathbf{A}\|_\infty = \max_i \left(\sum_j |A_{ij}| \right)$ (maximum absolute row sum norm)

REMINDER: MATRIX NORM / 4

Another common matrix norm is the **Frobenius norm** which can be interpreted as an extension of the Euclidean norm for vectors to matrices. It is defined as follows:

$$\|\mathbf{A}\|_F = \sqrt{\mathrm{trace}(\mathbf{A}^\top \mathbf{A})} = \sqrt{\sum_i \sum_j A_{ij}^2}$$

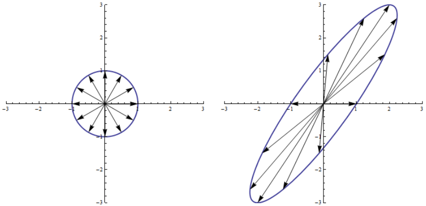
It is: $\quad \|\mathbf{A}\|_2 \leq \|\mathbf{A}\|_F$

Most important to us is $\|.\| := \|.\|_2$

REMINDER: MATRIX NORM / 5

Intuition matrix norm:

- Longest possible "stretch" of a vector of length 1 when multiplied by $\mathbf{A}$.
- For spectral norm: longest possible "stretch" in direction of the eigenvector of $\mathbf{A}^T\mathbf{A}$ (major axis of the ellipse) belonging to the largest absolute eigenvalue.



Left: Vectors of length 1. Right: Vectors after multiplication by A.

REMINDER: MATRIX NORM / 6

1. $||\mathbf{Ax}||_p \leq ||\mathbf{A}||_p ||\mathbf{x}||_p$,
   i.e., $||\mathbf{A}||_p$ is the smallest number to which this applies, because
   $||\mathbf{A}||_p \geq \frac{||\mathbf{Ax}||_p}{||\mathbf{x}||_p}$ for every $\mathbf{x} \neq 0$.
2. $||\boldsymbol{AB}||_p \leq ||\mathbf{A}||_p ||\boldsymbol{B}||_p$

   Proof: Let $\mathbf{x}$ be arbitrary with $||\mathbf{x}||_p = 1$ Then

   $$||\boldsymbol{ABx}||_p \leq ||\mathbf{A}||_p ||\boldsymbol{B}\mathbf{x}||_p \leq ||\mathbf{A}||_p ||\boldsymbol{B}||_p ||\mathbf{x}||_p = ||\mathbf{A}||_p ||\boldsymbol{B}||_p$$

   and thus

   $$||\boldsymbol{AB}||_p = \sup_{||\mathbf{x}||_p = 1} ||\boldsymbol{ABx}||_p \leq ||\mathbf{A}||_p ||\boldsymbol{B}||_p$$