Vector Norms

Matrix Norms
Norm Axioms

Definition

A vector norm \( \| \cdot \| \) is defined by \( \| \cdot \| : \mathbb{R}^n \rightarrow \mathbb{R} \) with

- \( \| x \| \geq 0 \)
- \( \| x \| = 0 \iff x = 0 \)
- \( \| x + y \| \leq \| x \| + \| y \| \) (triangular inequality)
- \( \| \alpha x \| = |\alpha| \| x \| \quad \alpha \in \mathbb{R} \)
Examples

Norms we use in this course:

\[ \|x\|_p := (|x_1|^p + \ldots + |x_n|^p)^{1/p} \]

so-called \textit{p-norm} or \textit{Hölder-norm}.

Example

\[ \|x\|_1 = |x_1| + \ldots + |x_n| \]
\[ \|x\|_2 = \left( |x_1|^2 + \ldots + |x_n|^2 \right)^{1/2} \] (Euklid)
\[ \|x\|_\infty = \max |x_i| \]

Weighted norms:

\[ \|x\|_W := \|Wx\| \]

\(W\) nonsingular, often diagonal. Used in certain proofs and in statistics to weight the different reliability of measurements.
Some inequalities

- \[ \|x\| - \|y\| \leq \|x - y\| \] . This can be deduced from the triangular inequality:

\[ \|x\| = \|(x - y) + y\| \leq \|(x - y)\| + \|y\| . \]

- For the 2-norm we have the Cauchy-Schwarz inequality:

\[ |(x, y)| \leq \|x\| \|y\| \]
Equivalence of Norms in finite dimensions

Theorem

All norms on $\mathbb{R}^n$ are equivalent in the sense:
There are constants $c_1, c_2 > 0$ such that for all $x$

$$c_1 \|x\|_\alpha \leq \|x\|_\beta \leq c_2 \|x\|_\alpha$$

holds.

Proof in three steps:
First we show that if two norms are equivalent to $\|\cdot\|_1$ then they are equivalent to each other (transitivity).
Next we note

$$c_1 \|x\|_1 \leq \|x\|_\beta \leq c_2 \|x\|_1 \quad \forall x \neq 0 \iff c_1 \leq \|u\|_\beta \leq c_2 \quad \forall u \text{ with } \|u\|_1 = 1$$

Finally we show that

$$f : x \mapsto \|x\|_\beta$$

is continuous with respect to the $\|\cdot\|_1$ norm. The rest follows from the fact, that a continuous function takes its maximum and minimum on a compact set, here the unit sphere.
Vector Norms

Matrix Norms
Matrix norms as operator norms

We consider a matrix as a linear operator:

\[ A : x \in \mathbb{R}^n \rightarrow y \in \mathbb{R}^m \]

We define a matrix norm by relating the size of the image to the size of its preimage:

**Definition**

A matrix norm can be defined by

\[
\| A \|_{(p,q)} := \operatorname{sub}_{x \in \mathbb{R}^n / \{0\}} \frac{\| Ax \|_p}{\| x \|_q} = \max_{x \in S_q} \| Ax \|_p
\]

with \( S_q := \{ x \in \mathbb{R}^n : \| x \|_q = 1 \} \) (unit sphere).

We note \( Q \) orthogonal \( \rightarrow \| Q \|_{(2,2)} = 1. \)
Special matrix norms

Theorem

- \( \|A\|_1 = \|A\|_{(1,1)} = \max_j \sum_i \|A_{ij}\| \) 1-norm
- \( \|A\|_\infty = \|A\|_{(\infty,\infty)} = \max_i \sum_j \|A_{ij}\| \) \(\infty\)-norm
- \( \|A\|_2 = \|A\|_{(2,2)} = \sqrt{\rho(A^TA)} \) 2-norm

with \( \rho \) being the spectral radius:

\[ \rho(B) = \max_i |\lambda_i(B)| \] largest eigenvalue

(Proof: first two statements in the exercises, third exercise in lecture)
Examples

- Let $A$ be a $1 \times n$ matrix. Let $a$ be a vector with the same entries. Compare the 2-norm of $A$ (matrix norm) with the 2-norm of $a$ (vector norm). (Hint: Apply Cauchy-Schwarz)
- Let $u$, $v$ be two vectors: $A := uv^T$ is called its outer product. Compute its 2-norm.
Frobenius Norm

Definition

The Frobenius norm of a matrix \( A \) is

\[
\| A \|_F = \left( \sum_i \sum_j |A_{ij}|^2 \right)^{\frac{1}{2}}
\]

it is not an operator norm.

Equivalent expressions

\[\begin{align*}
\| A \|_F &= \left( \sum_i |A_i| \| \frac{2}{2} \right)^{\frac{1}{2}} \\
\| A \|_F &= \left( \sum_j |A_{.j}| \| \frac{2}{2} \right)^{\frac{1}{2}} \\
\| A \|_F &= \left( \text{tr}(A^T A) \right)^{\frac{1}{2}} = \left( \text{tr}(AA^T) \right)^{\frac{1}{2}}
\end{align*}\]
Bounding products of matrices

**Theorem**

\[ \|AB\|_F \leq \|A\|_F \|B\|_F \]

**Proof:**

Consider \( C := AB \) and note \( C_{ij} = (A_i, B_j) \).

Apply Cauchy-Schwartz inequality:

\[ |C_{ij}|^2 \leq |A_i|^2 |B_j|^2 \]

and sum up

\[ \|C\|_F^2 = \sum_i \sum_j |C_{ij}|^2 \leq \sum_i |A_i|^2 \left( \sum_j |B_j|^2 \right) = \|A\|_F^2 \|B\|_F^2 \]
Invariance under Orthogonal Multiplication

A matrix $Q$ is orthogonal if $Q^TQ = QQ^T = I$

**Theorem**

*The 2-norm and the Frobenius norm are invariant under multiplication with orthogonal matrices, i.e.*

$$\|QA\|_2 = \|A\|_2 \quad \text{and} \quad \|QA\|_F = \|A\|_F$$

The proof uses in the Frobenius case the representation of the Frobenius norm by the trace operator.