

Conditioning of Square Linear Systems

Contents

① Vector and Matrix Norms

② Conditioning

Vector and Matrix Norms

Vector Norms

← generalization of the absolute value function.

The “length” of a vector \mathbf{v} can be measured by its **norm**.

Definition 1 (p -Norm of a Vector)

Let $p \in [1, \infty)$. The p -norm of $\mathbf{v} \in \mathbb{R}^m$ is denoted by $\|\mathbf{v}\|_p$ and is defined by

$$\|\mathbf{v}\|_p = \left(\sum_{i=1}^m |v_i|^p \right)^{1/p}.$$

When $p = \infty$,

$$\|\mathbf{v}\|_\infty = \max_{1 \leq i \leq m} |v_i|.$$

The most commonly used p values are 1, 2, and ∞ :

$$\|\mathbf{v}\|_1 = \sum_{i=1}^m |v_i|, \quad \|\mathbf{v}\|_2 = \sqrt{\sum_{i=1}^m |v_i|^2}.$$

Pythagorean then.

Vector Norms

vectors
nonnegative #.

In general, any function $\|\cdot\| : \mathbb{R}^m \rightarrow \mathbb{R}^+ \cup \{0\}$ is called a **vector norm** if it satisfies the following three properties:

- 1 $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = \mathbf{0}$.
- 2 $\|\alpha\mathbf{x}\| = |\alpha| \|\mathbf{x}\|$ for any constant α and any $\mathbf{x} \in \mathbb{R}^m$.
- 3 $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$. This is called the *triangle inequality*.

Unit Vectors

- A vector \mathbf{u} is called a **unit vector** if $\|\mathbf{u}\| = 1$.
- Depending on the norm used, unit vectors will be different.
- For instance:

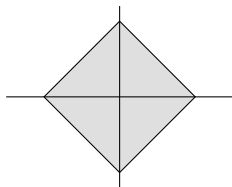


Figure 1: 1-norm

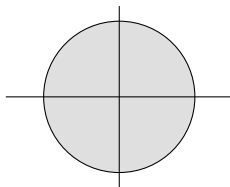


Figure 2: 2-norm

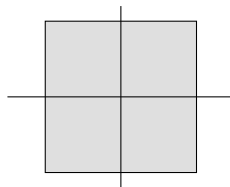


Figure 3: ∞ -norm

Matrix Norms

The “size” of a matrix $A \in \mathbb{R}^{m \times n}$ can be measured by its **norm** as well. As above, we say that a function $\|\cdot\| : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^+ \cup \{0\}$ is a **matrix norm** if it satisfies the following three properties:

- 1 $\|A\| = 0$ if and only if $A = 0$.
- 2 $\|\alpha A\| = |\alpha| \|A\|$ for any constant α and any $A \in \mathbb{R}^{m \times n}$.
- 3 $\|A + B\| \leq \|A\| + \|B\|$ for any $A, B \in \mathbb{R}^{m \times n}$. This is called the *triangle inequality*.

Matrix Norms (cont')

- If, in addition to satisfying the three conditions, it satisfies

$$\|AB\| \leq \|A\| \|B\| \quad \text{for all } A \in \mathbb{R}^{m \times n} \text{ and all } B \in \mathbb{R}^{n \times p},$$

it is said to be **consistent**.

- If, in addition to satisfying the three conditions, it satisfies

$$\|A\mathbf{x}\| \leq \|A\| \|\mathbf{x}\| \quad \text{for all } A \in \mathbb{R}^{m \times n} \text{ and all } \mathbf{x} \in \mathbb{R}^n,$$

then we say that it is **compatible** with a vector norm.

Induced Matrix Norms

Definition 2 (p -Norm of a Matrix)

Let $p \in [1, \infty]$. The p -norm of $A \in \mathbb{R}^{m \times n}$ is given by

$$\|A\|_p = \max_{\mathbf{x} \neq 0} \frac{\|A\mathbf{x}\|_p}{\|\mathbf{x}\|_p} = \max_{\|\mathbf{x}\|_p=1} \|A\mathbf{x}\|_p .$$

- The definition of this particular matrix norm is **induced** from the vector p -norm.
- By construction, matrix p -norm is a compatible norm.
- Induced norms describe how the matrix stretches unit vectors with respect to the vector norm.

Induced Matrix Norms

The commonly used p -norms (for $p = 1, 2, \infty$) can also be calculated by

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^m |a_{ij}|,$$

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^T A)} = \sigma_{\max}(A),$$

$$\|A\|_{\infty} = \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}|.$$

In words,

- The 1-norm of A is the maximum of the 1-norms of all column vectors.
- The 2-norm of A is the square root of the largest eigenvalue of $A^T A$.
- The ∞ -norm of A is the maximum of the 1-norms of all row vectors.

Non-Induced Matrix Norm – Frobenius Norm

Definition 3 (Frobenius Norm of a Matrix)

The Frobenius norm of $A \in \mathbb{R}^{m \times n}$ is given by

$$\|A\|_F = \left(\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \right)^{1/2}.$$

- This is not induced from a vector p -norm.
- However, both p -norm and the Frobenius norm are consistent and compatible.
- For compatibility of the Frobenius norm, the vector norm must be the 2-norm, that is, $\|A\mathbf{x}\|_2 \leq \|A\|_F \|\mathbf{x}\|_2$.

Norms in MATLAB

- Vector p -norms can be easily computed:

```
norm(v, 1)      % = sum(abs(v))  
norm(v, 2)      % = sqrt(v'*v)   if v is a column  
norm(v, 'inf')  % = max(abs(v))
```

- The same function `norm` is used to calculate matrix p -norms:

```
norm(A, 1)      % = max(sum(abs(A), 1))  
norm(A, 2)      % = max(sqrt(eig(A'*A)))  
norm(A, Inf)    % = max(sum(abs(A), 2))
```

- To calculate the Frobenius norm:

```
norm(A, 'fro')  % = sqrt(A(:)'*A(:))  
               % = norm(A(:), 2)
```

Conditioning

Conditioning of Solving Linear Systems: Overview

- Analyze how robust (or sensitive) the solutions of $A\mathbf{x} = \mathbf{b}$ are to perturbations of A and \mathbf{b} .
- For simplicity, consider separately the cases where

- 1 \mathbf{b} changes to $\mathbf{b} + \delta\mathbf{b}$, while A remains unchanged, that is

$$A\mathbf{x} = \mathbf{b} \longrightarrow A(\mathbf{x} + \delta\mathbf{x}) = \mathbf{b} + \delta\mathbf{b}.$$

- 2 A changes to $A + \delta A$, while \mathbf{b} remains unchanged, that is

$$A\mathbf{x} = \mathbf{b} \longrightarrow (A + \delta A)(\mathbf{x} + \delta\mathbf{x}) = \mathbf{b}.$$

Sensitivity to Perturbation of RHS

Case 1. $A\mathbf{x} = \mathbf{b} \rightarrow A(\mathbf{x} + \delta\mathbf{x}) = \mathbf{b} + \delta\mathbf{b}$

- Bound $\|\delta\mathbf{x}\|$ in terms of $\|\delta\mathbf{b}\|$:

$$A\mathbf{x} + A\delta\mathbf{x} = \mathbf{b} + \delta\mathbf{b}$$

$$A\delta\mathbf{x} = \delta\mathbf{b} \quad \implies \quad \|\delta\mathbf{x}\| \leq \|A^{-1}\| \|\delta\mathbf{b}\|.$$

$$\delta\mathbf{x} = A^{-1}\delta\mathbf{b}$$

- Sensitivity in terms of relative errors:

$$\frac{\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|}}{\frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|}} = \frac{\|\delta\mathbf{x}\| \|\mathbf{b}\|}{\|\delta\mathbf{b}\| \|\mathbf{x}\|} \leq \frac{\|A^{-1}\| \|\delta\mathbf{b}\| \cdot \|A\| \|\mathbf{x}\|}{\|\delta\mathbf{b}\| \|\mathbf{x}\|} = \|A^{-1}\| \|A\|.$$

Sensitivity to Perturbation of Matrix

Case 2. $A\mathbf{x} = \mathbf{b} \rightarrow (A + \delta A)(\mathbf{x} + \delta \mathbf{x}) = \mathbf{b}$

- Bound $\|\delta \mathbf{x}\|$ now in terms of $\|\delta A\|$:

$$\begin{aligned} A\mathbf{x} + A\delta \mathbf{x} + (\delta A)\mathbf{x} + (\delta A)\delta \mathbf{x} &= \mathbf{b} \\ A\delta \mathbf{x} &= -(\delta A)\mathbf{x} - (\delta A)\delta \mathbf{x} \\ \delta \mathbf{x} &= -A^{-1}(\delta A)\mathbf{x} - A^{-1}(\delta A)\delta \mathbf{x} \end{aligned} \quad \Rightarrow \quad \begin{aligned} \|\delta \mathbf{x}\| &\lesssim \|A^{-1}\| \|\delta A\| \|\mathbf{x}\|. \\ &\text{(first-order truncation)} \end{aligned}$$

- Sensitivity in terms of relative errors:

$$\frac{\frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|}}{\frac{\|\delta A\|}{\|A\|}} = \frac{\|\delta \mathbf{x}\| \|A\|}{\|\delta A\| \|\mathbf{x}\|} \lesssim \frac{\|A^{-1}\| \|\delta A\| \|\mathbf{x}\| \cdot \|A\|}{\|\delta A\| \|\mathbf{x}\|} = \|A^{-1}\| \|A\|.$$

Matrix Condition Number

- Motivated by the previous estimations, we define the **matrix condition number** by

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

where the norms can be any p -norm or the Frobenius norm.

- A subscript on κ such as 1, 2, ∞ , or F(robenius) is used if clarification is needed.

Matrix Condition Number (cont')

- We can write

$$\frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq \kappa(A) \frac{\|\delta \mathbf{b}\|}{\|\mathbf{b}\|}, \quad \frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq \kappa(A) \frac{\|\delta A\|}{\|A\|},$$

where the second inequality is true only in the limit of infinitesimal perturbations δA .

- The matrix condition number $\kappa(A)$ is equal to the condition number of solving a linear system of equation $A\mathbf{x} = \mathbf{b}$.
- The exponent of $\kappa(A)$ in scientific notation determines the approximate number of digits of accuracy that will be lost in calculation of \mathbf{x} .
- Since $1 = \|I\| = \|A^{-1}A\| \leq \|A^{-1}\|\|A\| = \kappa(A)$, a condition number of 1 is the best we can hope for.
- If $\kappa(A) > \boxed{\text{eps}}^{-1}$, then for computational purposes the matrix is singular.

Condition Numbers in MATLAB

- Use `cond` to calculate various condition numbers:

```
cond(A)           % the 2-norm; or  cond(A, 2)
cond(A, 1)        % the 1-norm
cond(A, Inf)      % the infinity-norm
cond(A, 'fro')    % the Frobenius norm
```

- A condition number estimator (in 1-norm)

```
condest(A)        % faster than cond
```

- The fastest method to estimate the condition number is to use `linsolve` function as below:

```
[x, inv_condest] = linsolve(A, b);
fast_condest = 1/inv_condest;
```