

Sensitivity of a linear system

Condition number of a matrix

Solving linear systems in floating point

Given

$$\mathbf{Ax} = \mathbf{b},$$

the formal solution is

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

However,

- ▶ \mathbf{A} is only known approximately
- ▶ \mathbf{b} is only known approximately
- ▶ A numerical algorithm introduces some errors (at least roundoff).

Can only compute an *approximate* solution, \mathbf{x}^* .

Solving linear systems in floating point

Let \mathbf{x}^* is an *approximate* solution of $\mathbf{Ax} = \mathbf{b}$.

We want to have \mathbf{x}^* “close” to the true solution $\hat{\mathbf{x}}$.

- ▶ Deviation $\mathbf{e} = \hat{\mathbf{x}} - \mathbf{x}^*$
- ▶ *Residual* $\mathbf{r} = \mathbf{b} - \mathbf{Ax}^* \equiv \mathbf{Ae}$

Require that either the deviation or the residual is “small”. But what is small.

Need a distance measure for vectors.

Vector norms

A *norm* in the vector space \mathbb{R}^m , $\|\cdot\|$, is a mapping $\mathbb{R}^m \rightarrow \mathbb{R}$, such that

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^m, \forall \alpha \in \mathbb{R}$$

1. $\|\mathbf{x}\| \geq 0$ and $\|\mathbf{x}\| = 0$ iff $\mathbf{x} = \mathbf{0}$
2. $\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|$
3. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ triangle inequality

Vector norms

A *norm* in the vector space \mathbb{R}^m , $\|\cdot\|$, is a mapping $\mathbb{R}^m \rightarrow \mathbb{R}$, such that

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^m, \forall \alpha \in \mathbb{R}$$

1. $\|\mathbf{x}\| \geq 0$ and $\|\mathbf{x}\| = 0$ iff $\mathbf{x} = \mathbf{0}$
2. $\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|$
3. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ triangle inequality

Can define *absolute error*, $\epsilon_{\text{abs}} = \|\hat{\mathbf{x}} - \mathbf{x}^*\|$, *relative error*
 $\epsilon_{\text{rel}} = \|\hat{\mathbf{x}} - \mathbf{x}^*\| / \|\mathbf{x}^*\|$.

Convergence

A sequence $\mathbf{x}^{(n)}$ *converges* to $\hat{\mathbf{x}}$ if

$$\|\mathbf{x}^{(n)} - \hat{\mathbf{x}}\| \rightarrow 0, \quad n \rightarrow \infty$$

Vector norms

p -norms

$$\|\mathbf{x}\|_p = (|x_1|^p + \cdots + |x_m|^p)^{1/p}$$

for $p \geq 1$.

1-norm $\|\mathbf{x}\|_1 = |x_1| + \cdots + |x_m|$

2-norm $\|\mathbf{x}\|_2 = (|x_1|^2 + \cdots + |x_m|^2)^{1/2}$

∞ -norm $\|\mathbf{x}\|_\infty = \max_{1 \leq j \leq m} |x_j|$

Vector norms

p -norms

$$\|\mathbf{x}\|_p = (|x_1|^p + \cdots + |x_m|^p)^{1/p}$$

for $p \geq 1$.

1-norm $\|\mathbf{x}\|_1 = |x_1| + \cdots + |x_m|$

2-norm $\|\mathbf{x}\|_2 = (|x_1|^2 + \cdots + |x_m|^2)^{1/2}$

∞ -norm $\|\mathbf{x}\|_\infty = \max_{1 \leq j \leq m} |x_j|$

All p -norms are equivalent:

$$\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1 \leq m \|\mathbf{x}\|_\infty$$

Matrix norms

Let \mathbf{A} and \mathbf{B} are an m -by- n matrices.

A *norm* in the space $\mathbb{R}^{m,n}$, $\|\cdot\|$, is a mapping $\mathbb{R}^{m,n} \rightarrow \mathbb{R}$, such that $\forall \mathbf{A}, \mathbf{B} \in \mathbb{R}^{m,n}, \forall \alpha \in \mathbb{R}$

1. $\|\mathbf{A}\| \geq 0$ and $\|\mathbf{A}\| = 0$ iff $\mathbf{A} = \mathbf{0}$
2. $\|\alpha \mathbf{A}\| = |\alpha| \|\mathbf{A}\|$
3. $\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$ triangle inequality

Matrix norms

Frobenius norm

$$\|\mathbf{A}\|_F = \sqrt{\sum_{j=1}^m \sum_{k=1}^n |a_{jk}|^2}$$

p -norms

$$\|\mathbf{A}\|_p = \sup_{\mathbf{x} \neq 0} \frac{\|\mathbf{Ax}\|_p}{\|\mathbf{x}\|_p}$$

A matrix p -norm is *subordinate* to (or, is *induced by*) a vector p -norm.

Matrix norms

For p -norms, have

4. $\|\mathbf{Ax}\| \leq \|\mathbf{A}\| \cdot \|\mathbf{x}\|$

5. $\|\mathbf{AB}\| \leq \|\mathbf{A}\| \cdot \|\mathbf{B}\|$

(p -norms are *mutually consistent*)

Matrix norms

For p -norms, have

4. $\|\mathbf{Ax}\| \leq \|\mathbf{A}\| \cdot \|\mathbf{x}\|$

5. $\|\mathbf{AB}\| \leq \|\mathbf{A}\| \cdot \|\mathbf{B}\|$

(p -norms are *mutually consistent*)

NB: Consistency is not guaranteed by properties 1-4. Consider,

$$\|\mathbf{A}\|_{\Delta} = \max_{k,l} |a_{kl}|$$

and

$$\mathbf{A} = \mathbf{B} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Common matrix norms: 1-norm

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix}$$

$p = 1$ (a.k.a. the *column norm*)

$$\|\mathbf{A}\|_1 = \max_{1 \leq k \leq n} \sum_{j=1}^m |a_{jk}|$$

is subordinate to the vector 1-norm.

Common matrix norms: ∞ -norm

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix}$$

$p = \infty$ (a.k.a. the *row norm*)

$$\|\mathbf{A}\|_{\infty} = \max_{1 \leq j \leq m} \sum_{k=1}^n |a_{jk}|$$

is subordinate to the vector ∞ -norm.

Common matrix norms: 2-norm

- ▶ 1- and ∞ -norms are easy to compute, $O(n^2)$ operations
- ▶ 2-norm, subordinate to the vector 2-norm is

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

- ▶ In terms of SVD, $\|\mathbf{A}\|_2 = \sigma_{\max}$
- ▶ Note the naming: *Euclidean norm* typically means the Frobenius norm, not $\|\cdot\|_2$.

Sensitivity of a linear system

Suppose \mathbf{A} is known exactly, and \mathbf{b} is not; we only know an approximation, \mathbf{b}^* . How different are \mathbf{x} and \mathbf{x}^* ?

$$\mathbf{Ax} = \mathbf{b}$$

$$\mathbf{Ax}^* = \mathbf{b}^*$$

$$\|\mathbf{b} - \mathbf{b}^*\|$$

abs error of \mathbf{b} (input)

$$\|\mathbf{x} - \mathbf{x}^*\|$$

abs error of \mathbf{x} (result)

Sensitivity of a linear system

Suppose \mathbf{A} is known exactly, and \mathbf{b} is not; we only know an approximation, \mathbf{b}^* . How different are \mathbf{x} and \mathbf{x}^* ?

$$\mathbf{Ax} = \mathbf{b} \qquad \mathbf{Ax}^* = \mathbf{b}^*$$

$$\|\mathbf{b} - \mathbf{b}^*\| \qquad \text{abs error of } \mathbf{b} \text{ (input)}$$

$$\|\mathbf{x} - \mathbf{x}^*\| \qquad \text{abs error of } \mathbf{x} \text{ (result)}$$

Condition number relates the error of the result to the error of inputs

$$\frac{\|\mathbf{x} - \mathbf{x}^*\|}{\|\mathbf{x}\|} \propto \frac{\|\mathbf{b} - \mathbf{b}^*\|}{\|\mathbf{b}\|}$$

Condition number

- ▶ Assuming $\mathbf{Ax} = \mathbf{b}$, define the *residual*

$$\begin{aligned}\mathbf{r} &= \mathbf{b} - \mathbf{Ax}^* \\ &= \mathbf{b} - \mathbf{b}^* \\ &= \mathbf{A}(\mathbf{x} - \mathbf{x}^*) \quad \Rightarrow \quad \mathbf{x} - \mathbf{x}^* = \mathbf{A}^{-1}\mathbf{r}\end{aligned}$$

Condition number

- ▶ Assuming $\mathbf{Ax} = \mathbf{b}$, define the *residual*

$$\begin{aligned}\mathbf{r} &= \mathbf{b} - \mathbf{Ax}^* \\ &= \mathbf{b} - \mathbf{b}^* \\ &= \mathbf{A}(\mathbf{x} - \mathbf{x}^*) \quad \Rightarrow \quad \mathbf{x} - \mathbf{x}^* = \mathbf{A}^{-1}\mathbf{r}\end{aligned}$$

- ▶
$$\frac{\|\mathbf{b}\|}{\|\mathbf{x}\|} = \frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|} \leq \|\mathbf{A}\|$$

Condition number

- ▶ Assuming $\mathbf{Ax} = \mathbf{b}$, define the *residual*

$$\begin{aligned}\mathbf{r} &= \mathbf{b} - \mathbf{Ax}^* \\ &= \mathbf{b} - \mathbf{b}^* \\ &= \mathbf{A}(\mathbf{x} - \mathbf{x}^*) \quad \Rightarrow \quad \mathbf{x} - \mathbf{x}^* = \mathbf{A}^{-1}\mathbf{r}\end{aligned}$$

- ▶ $\frac{\|\mathbf{b}\|}{\|\mathbf{x}\|} = \frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|} \leq \|\mathbf{A}\|$

- ▶ Define $\delta\mathbf{x} = \frac{\|\mathbf{x} - \mathbf{x}^*\|}{\|\mathbf{x}\|}$ and $\delta\mathbf{b} = \frac{\|\mathbf{b} - \mathbf{b}^*\|}{\|\mathbf{b}\|} \equiv \frac{\|\mathbf{r}\|}{\|\mathbf{b}\|}$

Condition number

$$\begin{aligned}\delta \mathbf{x} &= \frac{\|\mathbf{x} - \mathbf{x}^*\|}{\|\mathbf{x}\|} = \frac{\|\mathbf{A}^{-1} \mathbf{r}\|}{\|\mathbf{x}\|} \\ &\leq \frac{\|\mathbf{A}^{-1}\| \cdot \|\mathbf{r}\|}{\|\mathbf{x}\|} \\ &= \frac{\|\mathbf{A}^{-1}\| \cdot \|\mathbf{r}\|}{\|\mathbf{x}\|} \frac{\|\mathbf{b}\|}{\|\mathbf{b}\|} \\ &\leq \|\mathbf{A}^{-1}\| \|\mathbf{A}\| \delta \mathbf{b}\end{aligned}$$

$$\text{cond } \mathbf{A} = \|\mathbf{A}^{-1}\| \|\mathbf{A}\|$$

Condition number

$$\text{cond } \mathbf{A} = \|\mathbf{A}^{-1}\| \|\mathbf{A}\|$$

- ▶ Condition number is scale invariant

$$\text{cond}(\alpha \mathbf{A}) = \text{cond } \mathbf{A} \quad \forall \alpha \in \mathbb{R}$$

- ▶ Condition number is norm-dependent
- ▶ If $\text{cond } \mathbf{A} \gg 1$, the matrix is *ill-conditioned*

- ▶ For errors in both \mathbf{A} and \mathbf{b} ,

$$\delta \mathbf{x} = \text{cond } \mathbf{A} (\delta \mathbf{A} + \delta \mathbf{b})$$