Sensitivity of Linear Systems

Sensitivity of Linear Systems

Consider the $n \times n$ system of equations

$$Ax = b, \qquad b \neq 0,$$

where A is nonsingular.

Sensitivity of Linear Systems

Consider the $n \times n$ system of equations

$$Ax = b, \qquad b \neq 0,$$

where A is nonsingular.

The sensitivity analysis of the system is a study of the sensitivity of the solution x to changes in A and b.

Sensitivity of Linear Systems

Consider the $n \times n$ system of equations

$$Ax = b, \qquad b \neq 0,$$

where A is nonsingular.

The sensitivity analysis of the system is a study of the sensitivity of the solution x to changes in A and b.

This requires a means of being able to measure matrices and vectors and will be done with the help of matrix and vector norms.

Sensitivity of Linear Systems

Consider the $n \times n$ system of equations

$$Ax = b, \qquad b \neq 0,$$

where A is nonsingular.

The sensitivity analysis of the system is a study of the sensitivity of the solution x to changes in A and b.

This requires a means of being able to measure matrices and vectors and will be done with the help of matrix and vector norms.

The changes/perturbations will be measured in relative sense.

Sensitivity of Linear Systems

Consider the $n \times n$ system of equations

$$Ax = b, \qquad b \neq 0,$$

where A is nonsingular.

The sensitivity analysis of the system is a study of the sensitivity of the solution x to changes in A and b.

This requires a means of being able to measure matrices and vectors and will be done with the help of matrix and vector norms.

The changes/perturbations will be measured in relative sense.

So for a choice of norm $\|\cdot\|$, we measure how

$$\frac{\|\delta x\|}{\|x\|} \to \text{relative change in } x$$

depends on

$$\frac{\|\delta b\|}{\|b\|} \to \text{relative change (perturbation) in } b$$
$$\frac{\|\delta A\|}{\|A\|} \to \text{relative change (perturbation) in } A.$$

Perturbing only b

Suppose

$$Ax = b \qquad (1) \quad \& \quad A\hat{x} = b + \delta b \qquad (2).$$

Perturbing only b

Suppose

$$Ax = b \qquad (1) \quad \& \quad A\hat{x} = b + \delta b \qquad (2).$$

Let $\delta x = \hat{x} - x$. Then, $A(\delta x) = \delta b$ so that

$$\delta x = A^{-1}(\delta b) \qquad (3)$$

For any matrix norm $\|\cdot\|$,

$$(1) \Rightarrow \frac{1}{\|x\|} \leq \frac{\|A\|}{\|b\|} \quad \& \quad (3) \Rightarrow \|\delta x\| \leq \|A^{-1}\| \, \|\delta b\|$$

Perturbing only b

Suppose

$$Ax = b \qquad (1) \quad \& \quad A\hat{x} = b + \delta b \qquad (2).$$

Let $\delta x = \hat{x} - x$. Then, $A(\delta x) = \delta b$ so that

$$\delta x = A^{-1}(\delta b) \qquad (3)$$

For any matrix norm $\|\cdot\|$,

$$(1) \Rightarrow \frac{1}{\|x\|} \leq \frac{\|A\|}{\|b\|} \quad \& \quad (3) \Rightarrow \|\delta x\| \leq \|A^{-1}\| \, \|\delta b\|$$

Combining these inequalities,

$$\frac{\|\delta x\|}{\|x\|} \leq \kappa(A) \frac{\|\delta b\|}{\|b\|} \qquad (4)$$

where $\kappa(A) = \|A\| \, \|A^{-1}\|$ **is called the (normwise) condition number of** A.

Perturbing only b

Suppose

$$Ax = b \qquad (1) \quad \& \quad A\hat{x} = b + \delta b \qquad (2).$$

Let $\delta x = \hat{x} - x$. Then, $A(\delta x) = \delta b$ so that

$$\delta x = A^{-1}(\delta b) \qquad (3)$$

For any matrix norm $\|\cdot\|$,

$$(1) \Rightarrow \frac{1}{\|x\|} \leq \frac{\|A\|}{\|b\|} \quad \& \quad (3) \Rightarrow \|\delta x\| \leq \|A^{-1}\| \, \|\delta b\|$$

Combining these inequalities,

$$\frac{\|\delta x\|}{\|x\|} \leq \kappa(A) \frac{\|\delta b\|}{\|b\|} \qquad (4)$$

where $\kappa(A) = \|A\| \, \|A^{-1}\|$ **is called the (normwise) condition number of** A.

Clearly, $\kappa(A) \geq 1$ for all choices of matrix norms.

Perturbing only b

1. Even small relative perturbation to b can result in large relative change to x if $\kappa(A)$ is large.
2. There are choices of $A, b, \delta b$ for which the inequality (4) is an equality!

Let

$$A = \begin{bmatrix} 1000 & 999 \\ 999 & 998 \end{bmatrix}.$$

Then

$$A^{-1} = \begin{bmatrix} -998 & 999 \\ 999 & -1000 \end{bmatrix},$$

and

$$\|A\|_\infty = \max\{|1000| + |999|, |999| + |998|\} = 1999$$
$$\|A^{-1}\|_\infty = \max\{|-998| + |999|, |999| + |-1000|\} = 1999$$

Therefore $\kappa_\infty(A) = (1999)^2 \approx 4 \times 10^6$.

Perturbing only b

Consider

$$Ax = \underbrace{\begin{bmatrix} 1999 \\ 1997 \end{bmatrix}}_{=:b} \qquad \& \qquad A\hat{x} = \underbrace{\begin{bmatrix} 1998.99 \\ 1997.01 \end{bmatrix}}_{=:b+\delta b}.$$

Perturbing only b

Consider

$$A x = \underbrace{\begin{bmatrix} 1999 \\ 1997 \end{bmatrix}}_{=:b} \qquad \& \qquad A\hat{x} = \underbrace{\begin{bmatrix} 1998.99 \\ 1997.01 \end{bmatrix}}_{=:b+\delta b}.$$

The solutions are

$$x = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } \hat{x} = \begin{bmatrix} 20.97 \\ -18.99 \end{bmatrix},$$

Perturbing only b

Consider

$$A x = \underbrace{\begin{bmatrix} 1999 \\ 1997 \end{bmatrix}}_{=:b} \qquad \& \qquad A\hat{x} = \underbrace{\begin{bmatrix} 1998.99 \\ 1997.01 \end{bmatrix}}_{=:b+\delta b}.$$

The solutions are

$$x = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } \hat{x} = \begin{bmatrix} 20.97 \\ -18.99 \end{bmatrix},$$

and

$$\frac{\|\hat{x} - x\|_\infty}{\|x\|_\infty} = 19.99 \qquad \& \qquad \frac{\|\delta b\|_\infty}{\|b\|_\infty} = \frac{0.01}{1999} \approx 5 \times 10^{-6}.$$

Perturbing only b

Consider

$$Ax = \underbrace{\begin{bmatrix} 1999 \\ 1997 \end{bmatrix}}_{=:b} \qquad \& \qquad A\hat{x} = \underbrace{\begin{bmatrix} 1998.99 \\ 1997.01 \end{bmatrix}}_{=:b+\delta b}.$$

The solutions are

$$x = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } \hat{x} = \begin{bmatrix} 20.97 \\ -18.99 \end{bmatrix},$$

and

$$\frac{\|\hat{x} - x\|_\infty}{\|x\|_\infty} = 19.99 \qquad \& \qquad \frac{\|\delta b\|_\infty}{\|b\|_\infty} = \frac{0.01}{1999} \approx 5 \times 10^{-6}.$$

Clearly

$$\frac{\|\hat{x} - x\|_\infty}{\|x\|_\infty} = (1999)^2 \frac{0.01}{1999} = \kappa_\infty(A) \frac{\|\delta b\|_\infty}{\|b\|_\infty}$$

and the system is very sensitive to small changes in b.

Perturbing A and/or b

Theorem Suppose $A \in \mathbb{F}^{n \times n}$ is nonsingular. Let $\delta A \in \mathbb{F}^{n \times n}$ such that $\frac{\|\delta A\|}{\|A\|} < \frac{1}{\kappa(A)}$. Then $A + \delta A$ is nonsingular.

Perturbing A and/or b

Theorem Suppose $A \in \mathbb{F}^{n\times n}$ is nonsingular. Let $\delta A \in \mathbb{F}^{n\times n}$ such that $\frac{\|\delta A\|}{\|A\|} < \frac{1}{\kappa(A)}$. Then $A + \delta A$ is nonsingular.

Theorem Suppose $A \in \mathbb{F}^{n\times n}$ is nonsingular. Let $\delta A \in \mathbb{F}^{n\times n}$ such that $\frac{\|\delta A\|}{\|A\|} < \frac{1}{\kappa(A)}$. Consider the systems of equations $Ax = b$ and $(A + \delta A)(x + \delta x) = b + \delta b$.

$$\text{Then} \qquad \frac{\|\delta x\|}{\|x\|} \le \kappa(A)\left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|}\right) \Big/ \left(1 - \kappa(A)\frac{\|\delta A\|}{\|A\|}\right).$$

Perturbing A and/or b

Theorem Suppose $A \in \mathbb{F}^{n\times n}$ is nonsingular. Let $\delta A \in \mathbb{F}^{n\times n}$ such that $\frac{\|\delta A\|}{\|A\|} < \frac{1}{\kappa(A)}$. Then $A + \delta A$ is nonsingular.

Theorem Suppose $A \in \mathbb{F}^{n\times n}$ is nonsingular. Let $\delta A \in \mathbb{F}^{n\times n}$ such that $\frac{\|\delta A\|}{\|A\|} < \frac{1}{\kappa(A)}$. Consider the systems of equtions $Ax = b$ and $(A + \delta A)(x + \delta x) = b + \delta b$.

Then $$\frac{\|\delta x\|}{\|x\|} \le \kappa(A)\left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|}\right) \Big/ \underbrace{\left(1 - \kappa(A)\frac{\|\delta A\|}{\|A\|}\right)}_{\approx 1 \text{ if } \frac{\|\delta A\|}{\|A\|} \ll \frac{1}{\kappa(A)}}.$$

Perturbing A and/or b

Theorem Suppose $A \in \mathbb{F}^{n\times n}$ is nonsingular. Let $\delta A \in \mathbb{F}^{n\times n}$ such that $\frac{\|\delta A\|}{\|A\|} < \frac{1}{\kappa(A)}$. Then $A + \delta A$ is nonsingular.

Theorem Suppose $A \in \mathbb{F}^{n\times n}$ is nonsingular. Let $\delta A \in \mathbb{F}^{n\times n}$ such that $\frac{\|\delta A\|}{\|A\|} < \frac{1}{\kappa(A)}$. Consider the systems of equtions $Ax = b$ and $(A + \delta A)(x + \delta x) = b + \delta b$.

Then $$\frac{\|\delta x\|}{\|x\|} \le \kappa(A)\left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|}\right) \Big/ \underbrace{\left(1 - \kappa(A)\frac{\|\delta A\|}{\|A\|}\right)}_{\approx 1 \text{ if } \frac{\|\delta A\|}{\|A\|} \ll \frac{1}{\kappa(A)}}.$$

In such a situation,

$$\frac{\|\delta x\|}{\|x\|} \lesssim \kappa(A)\left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|}\right).$$

ILL conditioned matrices

Another view of $\kappa(A)$:

For an $n \times n$ invertible matrix A, and an induced operator norm $\|\cdot\|$, suppose

$$\operatorname{maxmag} A = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} \quad \& \quad \operatorname{minmag} A = \min_{x \neq 0} \frac{\|Ax\|}{\|x\|}.$$

$\operatorname{maxmag} A \rightarrow$ maximum magnification by A.
$\operatorname{minmag} A \rightarrow$ minimum magnification by A.

ILL conditioned matrices

Another view of $\kappa(A)$:

For an $n \times n$ invertible matrix A, and an induced operator norm $\|\cdot\|$, suppose

$$\operatorname{maxmag} A = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} \quad \& \quad \operatorname{minmag} A = \min_{x \neq 0} \frac{\|Ax\|}{\|x\|}.$$

$\operatorname{maxmag} A \rightarrow$ maximum magnification by A.
$\operatorname{minmag} A \rightarrow$ minimum magnification by A.

There exist $x_0, \hat{x}_0 \neq 0$ such that

$$\operatorname{maxmag} A = \frac{\|Ax_0\|}{\|x_0\|} \quad \& \quad \operatorname{minmag} A = \frac{\|A\hat{x}_0\|}{\|\hat{x}_0\|}.$$

$x_0 \rightarrow$ vector in the direction of maximum maginification by A.
$\hat{x}_0 \rightarrow$ vector in the direction of minimum magnification by A.

ILL conditioned matrices

<u>**Another view of $\kappa(A)$:**</u>

For an $n \times n$ invertible matrix A, and an induced operator norm $\|\cdot\|$, suppose

$$\mathrm{maxmag}\, A = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} \quad \& \quad \mathrm{minmag}\, A = \min_{x \neq 0} \frac{\|Ax\|}{\|x\|}.$$

$$\mathrm{maxmag}\, A \to \text{maximum magnification by } A.$$
$$\mathrm{minmag}\, A \to \text{minimum magnification by } A.$$

There exist $x_0, \hat{x}_0 \neq 0$ such that

$$\mathrm{maxmag}\, A = \frac{\|Ax_0\|}{\|x_0\|} \quad \& \quad \mathrm{minmag}\, A = \frac{\|A\hat{x}_0\|}{\|\hat{x}_0\|}.$$

$x_0 \to$ vector in the direction of maximum maginification by A.
$\hat{x}_0 \to$ vector in the direction of minimum magnification by A.

Theorem: For an $n \times n$ invertible matrix A, $\kappa(A) = \dfrac{\mathrm{maxmag}\, A}{\mathrm{minmag}\, A}$.

ILL conditioned matrices

Another view of $\kappa(A)$:

For an $n \times n$ invertible matrix A, and an induced operator norm $\|\cdot\|$, suppose

$$\mathrm{maxmag}\, A = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} \quad \& \quad \mathrm{minmag}\, A = \min_{x \neq 0} \frac{\|Ax\|}{\|x\|}.$$

$$\mathrm{maxmag}\, A \to \text{maximum magnification by } A.$$
$$\mathrm{minmag}\, A \to \text{minimum magnification by } A.$$

There exist $x_0, \hat{x}_0 \neq 0$ such that

$$\mathrm{maxmag}\, A = \frac{\|Ax_0\|}{\|x_0\|} \quad \& \quad \mathrm{minmag}\, A = \frac{\|A\hat{x}_0\|}{\|\hat{x}_0\|}.$$

$$x_0 \to \text{vector in the direction of maximum maginification by } A.$$
$$\hat{x}_0 \to \text{vector in the direction of minimum magnification by } A.$$

Theorem: For an $n \times n$ invertible matrix A, $\kappa(A) = \dfrac{\mathrm{maxmag}\, A}{\mathrm{minmag}\, A}$.

When A is rectangular, the above is taken as the definition of $\kappa(A)$.

ILL conditioning

A nonsingular matrix A is said to be ill conditioned if $\kappa(A) \gg 1$.

Examples of ill conditioned matrices:

- There is no specific threshold for the value of the condition number beyond which a matrix is considered ill conditioned. This depends on the accuracy of the data, the finite precision arithmetic and the extent of error in solutions that is acceptable.

ILL conditioning

A nonsingular matrix A is said to be ill conditioned if $\kappa(A) \gg 1$.

Examples of ill conditioned matrices:

- There is no specific threshold for the value of the condition number beyond which a matrix is considered ill conditioned. This depends on the accuracy of the data, the finite precision arithmetic and the extent of error in solutions that is acceptable.
- However, matrices with condition numbers of the order of 10^5 or higher are generally considered as ill conditioned. So the matrix $A = \begin{bmatrix} 1000 & 999 \\ 999 & 998 \end{bmatrix}$ with $\kappa_1(A) \approx 4 \times 10^6$ is ill conditioned.

ILL conditioning

A nonsingular matrix A is said to be ill conditioned if $\kappa(A) \gg 1$.

Examples of ill conditioned matrices:

- There is no specific threshold for the value of the condition number beyond which a matrix is considered ill conditioned. This depends on the accuracy of the data, the finite precision arithmetic and the extent of error in solutions that is acceptable.
- However, matrices with condition numbers of the order of 10^5 or higher are generally considered as ill conditioned. So the matrix $A = \begin{bmatrix} 1000 & 999 \\ 999 & 998 \end{bmatrix}$ with $\kappa_1(A) \approx 4 \times 10^6$ is ill conditioned.
- The family of Hilbert matrices $H_n = [h_{ij}]_{n\times n}$ where $h_{ij} = 1/(i+j-1)$ are positive definite matrices which are famous for their ill conditioning as their condition number grows almost exponentially with size. For example $\kappa_2(H_4) \approx 1.6 \times 10^4$ and $\kappa_2(H_8) \approx 1.5 \times 10^{10}$.

ILL conditioning

A nonsingular matrix A is said to be ill conditioned if $\kappa(A) \gg 1$.

Examples of ill conditioned matrices:

- There is no specific threshold for the value of the condition number beyond which a matrix is considered ill conditioned. This depends on the accuracy of the data, the finite precision arithmetic and the extent of error in solutions that is acceptable.
- However, matrices with condition numbers of the order of 10^5 or higher are generally considered as ill conditioned. So the matrix $A = \begin{bmatrix} 1000 & 999 \\ 999 & 998 \end{bmatrix}$ with $\kappa_1(A) \approx 4 \times 10^6$ is ill conditioned.
- The family of Hilbert matrices $H_n = [h_{ij}]_{n\times n}$ where $h_{ij} = 1/(i+j-1)$ are positive definite matrices which are famous for their ill conditioning as their condition number grows almost exponentially with size. For example $\kappa_2(H_4) \approx 1.6 \times 10^4$ and $\kappa_2(H_8) \approx 1.5 \times 10^{10}$.

  (Type `H = hilb(n)` to generate the $n \times n$ Hilbert matrix in Matlab.)

ILL Conditioning

- If A is singular, then $\operatorname{minmag} A = 0$ and $\kappa(A) = \infty$.

ILL Conditioning

- If A is singular, then $\operatorname{minmag} A = 0$ and $\kappa(A) = \infty$.
- So being singular is an extreme case of ill conditioning!

ILL Conditioning

- If A is singular, then $\mathrm{minmag}\, A = 0$ and $\kappa(A) = \infty$.
- So being singular is an extreme case of ill conditioning!
- Ill conditioning is irrespective of the choice of norm.

ILL Conditioning

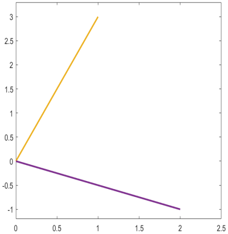
- If A is singular, then $\operatorname{minmag} A = 0$ and $\kappa(A) = \infty$.
- So being singular is an extreme case of ill conditioning!
- Ill conditioning is irrespective of the choice of norm.
- $\kappa(cA) = \kappa(A)$ for any scalar c.

ILL Conditioning

- If A is singular, then $\operatorname{minmag} A = 0$ and $\kappa(A) = \infty$.
- So being singular is an extreme case of ill conditioning!
- Ill conditioning is irrespective of the choice of norm.
- $\kappa(cA) = \kappa(A)$ for any scalar c.
- Sometimes ill conditioning is due to some row/column having much larger norm than the others.

ILL Conditioning

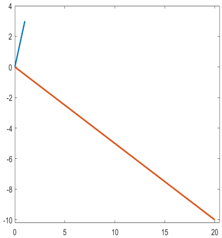
- If A is singular, then $\operatorname{minmag} A = 0$ and $\kappa(A) = \infty$.
- So being singular is an extreme case of ill conditioning!
- Ill conditioning is irrespective of the choice of norm.
- $\kappa(cA) = \kappa(A)$ for any scalar c.
- Sometimes ill conditioning is due to some row/column having much larger norm than the others.



The columns of $A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}$ with `cond(A) = 1.4561`

ILL Conditioning

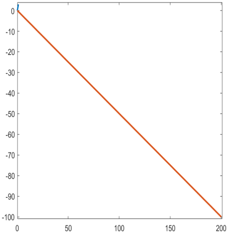
- If A is singular, then $\operatorname{minmag} A = 0$ and $\kappa(A) = \infty$.
- So being singular is an extreme case of ill conditioning!
- Ill conditioning is irrespective of the choice of norm.
- $\kappa(cA) = \kappa(A)$ for any scalar c.
- Sometimes ill conditioning is due to some row/column having much larger norm than the others.



The columns of $A = \begin{bmatrix} 1 & 20 \\ 3 & -10 \end{bmatrix}$ with `cond(A) = 14.2871`

ILL Conditioning

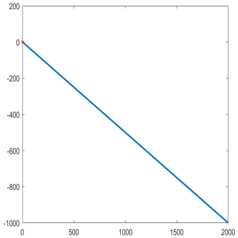
- If A is singular, then $\text{minmag}\, A = 0$ and $\kappa(A) = \infty$.
- So being singular is an extreme case of ill conditioning!
- Ill conditioning is irrespective of the choice of norm.
- $\kappa(cA) = \kappa(A)$ for any scalar c.
- Sometimes ill conditioning is due to some row/column having much larger norm than the others.



The columns of $A = \begin{bmatrix} 1 & 200 \\ 3 & -100 \end{bmatrix}$ with `cond(A) = 142.86`

ILL Conditioning

- If A is singular, then $\operatorname{minmag} A = 0$ and $\kappa(A) = \infty$.
- So being singular is an extreme case of ill conditioning!
- Ill conditioning is irrespective of the choice of norm.
- $\kappa(cA) = \kappa(A)$ for any scalar c.
- Sometimes ill conditioning is due to some row/column having much larger norm than the others.



The columns of $A = \begin{bmatrix} 1 & 2000 \\ 3 & -1000 \end{bmatrix}$ with `cond(A) = 1428.6`

ILL Conditioning

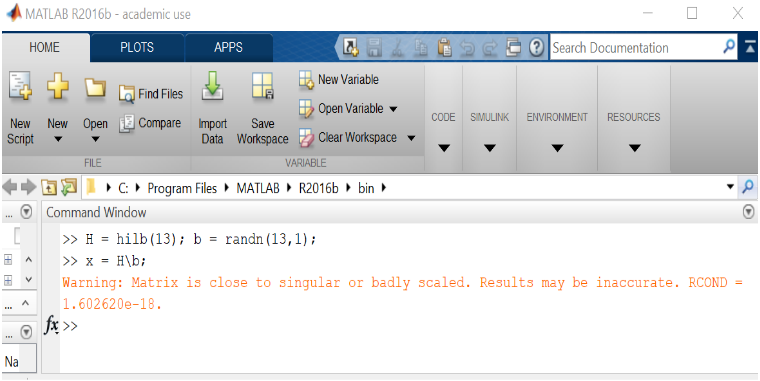
- If A is singular, then $\operatorname{minmag} A = 0$ and $\kappa(A) = \infty$.
- So being singular is an extreme case of ill conditioning!
- Ill conditioning is irrespective of the choice of norm.
- $\kappa(cA) = \kappa(A)$ for any scalar c.
- Sometimes ill conditioning is due to some row/column having much larger norm than the others.
- Ill conditioned matrices have rows/columns that are *'nearly linearly dependent'* and are therefore very close to being singular.

ILL Conditioning

- If A is singular, then $\operatorname{minmag} A = 0$ and $\kappa(A) = \infty$.
- So being singular is an extreme case of ill conditioning!
- Ill conditioning is irrespective of the choice of norm.
- $\kappa(cA) = \kappa(A)$ for any scalar c.
- Sometimes ill conditioning is due to some row/column having much larger norm than the others.
- Ill conditioned matrices have rows/columns that are *'nearly linearly dependent'* and are therefore very close to being singular.
- However, $\det A \approx 0$ is not *always* indicative of ill conditioning!

A geometric view of ill conditioning

Now we know why an attempt to solve a 13×13 system of equations where the coefficient matrix is `hilb(13)` with condition number $\approx 10^{18}$, results in the following:



A geometric view of ill conditioned systems

Consider the system

$$\underbrace{\begin{bmatrix} 1 & 2 \\ 1 & 2.02 \end{bmatrix}}_{:=A} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix},$$

where `cond(A) = 504.018` and the solution is $x_1 = -297$, $x_2 = 150$.

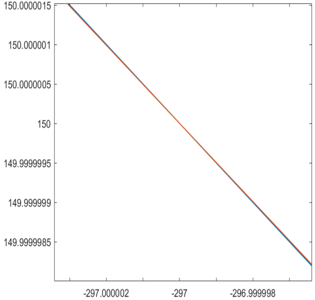
A geometric view of ill conditioned systems

Consider the system

$$\underbrace{\begin{bmatrix} 1 & 2 \\ 1 & 2.02 \end{bmatrix}}_{:=A} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix},$$

where `cond(A) = 504.018` and the solution is $x_1 = -297$, $x_2 = 150$.

Geometrically we have,



A geometric view of ill conditioned systems

For the perturbed system

$$\begin{bmatrix} 1 & 2 \\ 1 & 2.1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix},$$

the solution is $x_1 = -57$, $x_2 = 30$.

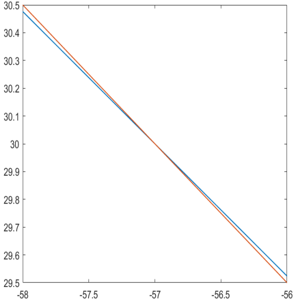
A geometric view of ill conditioned systems

For the perturbed system

$$\begin{bmatrix} 1 & 2 \\ 1 & 2.1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix},$$

the solution is $x_1 = -57, x_2 = 30.$

Geometrically we have,



A geometric view of ill conditioned systems

For the perturbed system

$$\begin{bmatrix} 1 & 2 \\ 1.05 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix},$$

the solution is $x_1 = 60$, $x_2 = -28.5$.

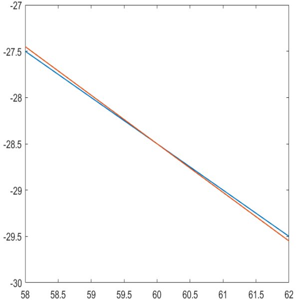
A geometric view of ill conditioned systems

For the perturbed system

$$\begin{bmatrix} 1 & 2 \\ 1.05 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix},$$

the solution is $x_1 = 60,\ x_2 = -28.5.$

Geometrically we have,



Estimating $\kappa_1(A)$

After solving an $n \times n$ system $Ax = b$ often it may be necessary to estimate $\kappa(A)$. For the the norm $\|\cdot\|_1$,

$$\kappa_1(A) = \|A\|_1 \|A^{-1}\|_1 = \max_{1 \le i \le n} \|A(:, i)\|_1 \|A^{-1}\|_1.$$

To estimate $\|A^{-1}\|_1$, solve systems

$$Ax = b_j, \ 1 \le j \le m, \ m \ll n,$$

in $2mn^2$ flops via GEPP using the permutation P, unit lower triangular L and upper triangular U such that $PA = LU$ obtained when solving the system.

If any $b_j, 1 \le j \le m$, is nearly in the direction of $\mathrm{maxmag}\, A^{-1}$ with respect to $\|\cdot\|_1$, then

$$\|A^{-1}\|_1 \approx \frac{\|x_j\|_1}{\|b_j\|_1}.$$

where that $Ax_j = b_j$.

Therefore a cheap estimate of $\kappa_1(A)$ is given by

$$\left(\max_{1 \le j \le m} \frac{\|x_j\|_1}{\|b_j\|_1} \right) \left(\max_{1 \le i \le n} \|A(:, i)\|_1 \right).$$

A posteriori error analysis using the residual

Suppose

$$x_c \rightarrow \text{ computed solution of } Ax = b.$$

How to estimate the relative error $\frac{\|x_c - x\|}{\|x\|}$ in x_c?

Trying to estimate the error in every step of the algorithm is impractical!

Instead the following strategy will be used.

- **Step 1.** Find δA and/or δb such that x_c is the solution of

$$(A + \delta A)x = b + \delta b.$$

- **Step 2.** If $\frac{\|\delta A\|}{\|A\|} < 1/\kappa(A)$, then

$$\frac{\|x_c - x\|}{\|x\|} \leq \kappa(A) \left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|} \right) \Big/ \left(1 - \kappa(A) \frac{\|\delta A\|}{\|A\|} \right).$$

A posteriori error analysis using the residual

Step 1 pushes the error in the solution back into the data. This is called backward error analysis and can be done in three ways:

- **Case 1:** Pushing the error in x_c back into only b.
- **Case 2:** Pushing the error in x_c back into only A.
- **Case 3:** Pushing the error in x_c back into both A and b.

A posteriori error analysis using the residual

Step 1 pushes the error in the solution back into the data. This is called backward error analysis and can be done in three ways:

- **Case 1:** Pushing the error in x_c back into only b.
- **Case 2:** Pushing the error in x_c back into only A.
- **Case 3:** Pushing the error in x_c back into both A and b.

The residual vector $r := b - Ax_c$ plays an important role in each case.

A posteriori error analysis using the residual

Step 1 pushes the error in the solution back into the data. This is called backward error analysis and can be done in three ways:

- **Case 1:** Pushing the error in x_c back into only b.
- **Case 2:** Pushing the error in x_c back into only A.
- **Case 3:** Pushing the error in x_c back into both A and b.

The residual vector $r := b - Ax_c$ plays an important role in each case.

Case 1: $Ax_c = b + \delta b$ where $\delta b = -r$.

By Step 2,

$$\frac{\|x_c - x\|}{\|x\|} \leq \kappa(A) \frac{\|r\|}{\|b\|}.$$

A posteriori error analysis using the residual

Step 1 pushes the error in the solution back into the data. This is called backward error analysis and can be done in three ways:

- **Case 1:** Pushing the error in x_c back into only b.
- **Case 2:** Pushing the error in x_c back into only A.
- **Case 3:** Pushing the error in x_c back into both A and b.

The residual vector $r := b - Ax_c$ plays an important role in each case.

Case 1: $Ax_c = b + \delta b$ where $\delta b = -r$.

By Step 2,

$$\frac{\|x_c - x\|}{\|x\|} \leq \kappa(A) \frac{\|r\|}{\|b\|}.$$

Case 2: $(A + \delta A)x_c = b$ where $\delta A := \frac{r x_c^T}{x_c^T x_c}$. Then

$$\|\delta A\|_2 = \frac{\|r\|_2}{\|x_c\|_2}. \qquad \text{(Exercise!)}$$

If $\frac{\|r\|_2}{\|x_c\|_2 \|A\|_2} < \frac{1}{\kappa_2(A)}$, then

$$\frac{\|x_c - x\|_2}{\|x\|_2} \leq \frac{\kappa_2(A)\, \|r\|_2}{\|x_c\|_2\, \|A\|_2} \Big/ \left(1 - \frac{\kappa_2(A)\, \|r\|_2}{\|x_c\|_2\, \|A\|_2}\right).$$

A posteriori error analysis using the residual

Case 3: Let $r_1 = \alpha b - Ax_c$ and $r_2 = (1-\alpha)b$ where $0 < \alpha < 1$.

Prove that

$$(A + \delta A)x_c = b + \delta b$$

where $\delta A = \frac{r_1 x_c^T}{x_c^T x_c}$ and $\delta b = -r_2$. Also show that

$$\|\delta A\|_2 = \frac{\|r_1\|_2}{\|x_c\|_2}.$$

Further if $\frac{\|r_1\|_2}{\|x_c\|_2 \|A\|_2} < \frac{1}{\kappa_2(A)}$, then show that

$$\frac{\|x_c - x\|_2}{\|x\|_2} \leq \kappa_2(A) \left(\frac{\|r_1\|_2}{\|x_c\|_2 \|A\|_2} + \frac{\|r_2\|_2}{\|b\|_2} \right) \Big/ \left(1 - \frac{\kappa_2(A) \|r_1\|_2}{\|x_c\|_2 \|A\|_2} \right).$$

A posteriori error analysis using the residual

Case 3: Let $r_1 = \alpha b - Ax_c$ and $r_2 = (1-\alpha)b$ where $0 < \alpha < 1$.

Prove that

$$(A + \delta A)x_c = b + \delta b$$

where $\delta A = \frac{r_1 x_c^T}{x_c^T x_c}$ and $\delta b = -r_2$. Also show that

$$\|\delta A\|_2 = \frac{\|r_1\|_2}{\|x_c\|_2}.$$

Further if $\frac{\|r_1\|_2}{\|x_c\|_2 \|A\|_2} < \frac{1}{\kappa_2(A)}$, then show that

$$\frac{\|x_c - x\|_2}{\|x\|_2} \leq \kappa_2(A) \left(\frac{\|r_1\|_2}{\|x_c\|_2 \|A\|_2} + \frac{\|r_2\|_2}{\|b\|_2} \right) \Big/ \left(1 - \frac{\kappa_2(A) \|r_1\|_2}{\|x_c\|_2 \|A\|_2} \right).$$

- If $\max\left\{ \frac{\|r\|}{\|b\|}, \frac{\|r\|}{\|x_c\| \|A\|} \right\} \approx cu$ for a modest constant c, then algorithm to compute x_c has backward stable behaviour with respect to the given problem. (u is unit roundoff $\approx 10^{-16}$ in IEEE double precision.)

A posteriori error analysis using the residual

Case 3: Let $r_1 = \alpha b - Ax_c$ and $r_2 = (1-\alpha)b$ where $0 < \alpha < 1$.

Prove that

$$(A + \delta A)x_c = b + \delta b$$

where $\delta A = \frac{r_1 x_c^T}{x_c^T x_c}$ and $\delta b = -r_2$. Also show that

$$\|\delta A\|_2 = \frac{\|r_1\|_2}{\|x_c\|_2}.$$

Further if $\frac{\|r_1\|_2}{\|x_c\|_2 \|A\|_2} < \frac{1}{\kappa_2(A)}$, then show that

$$\frac{\|x_c - x\|_2}{\|x\|_2} \le \kappa_2(A) \left(\frac{\|r_1\|_2}{\|x_c\|_2 \, \|A\|_2} + \frac{\|r_2\|_2}{\|b\|_2} \right) \Big/ \left(1 - \frac{\kappa_2(A) \, \|r_1\|_2}{\|x_c\|_2 \, \|A\|_2} \right).$$

- If $\max\left\{ \frac{\|r\|}{\|b\|}, \frac{\|r\|}{\|x_c\| \|A\|} \right\} \approx cu$ for a modest constant c, then algorithm to compute x_c has backward stable behaviour with respect to the given problem. (u is unit roundoff $\approx 10^{-16}$ in IEEE double precision.)
- $\kappa(A)$ is not large $\Leftrightarrow$ the problem is well conditioned.

A posteriori error analysis using the residual

Case 3: Let $r_1 = \alpha b - Ax_c$ and $r_2 = (1-\alpha)b$ where $0 < \alpha < 1$.

Prove that

$$(A + \delta A)x_c = b + \delta b$$

where $\delta A = \frac{r_1 x_c^T}{x_c^T x_c}$ and $\delta b = -r_2$. Also show that

$$\|\delta A\|_2 = \frac{\|r_1\|_2}{\|x_c\|_2}.$$

Further if $\frac{\|r_1\|_2}{\|x_c\|_2 \|A\|_2} < \frac{1}{\kappa_2(A)}$, then show that

$$\frac{\|x_c - x\|_2}{\|x\|_2} \le \kappa_2(A) \left(\frac{\|r_1\|_2}{\|x_c\|_2 \, \|A\|_2} + \frac{\|r_2\|_2}{\|b\|_2} \right) \Big/ \left(1 - \frac{\kappa_2(A) \, \|r_1\|_2}{\|x_c\|_2 \, \|A\|_2} \right).$$

- If $\max\left\{ \frac{\|r\|}{\|b\|}, \frac{\|r\|}{\|x_c\| \, \|A\|} \right\} \approx cu$ for a modest constant c, then algorithm to compute x_c has backward stable behaviour with respect to the given problem. (u is unit roundoff $\approx 10^{-16}$ in IEEE double precision.)
- $\kappa(A)$ is not large $\Leftrightarrow$ the problem is well conditioned.

The relative error in x_c is small only if the system is well conditioned and the solution process has been backward stable!