MA60053 - Computational Linear Algebra
Sensitivity analysis

M. Rajesh Kannan

Department of Mathematics,
Indian Institute of Technology Kharagpur,
email: rajeshkannan1.m@gmail.com, rajeshkannan@maths.iitkgp.ac.in

January 23, 2020

A closer look at linear systems

Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^{m \times 1}$.

Observation

The linear system $Ax = b$ has a solution if and only if b is a linear combination of columns, $a_1, \ldots, a_n$, of A,

$$b = a_1 x_1 + \cdots + a_n x_n,$$

where

$$A = (a_1 \ldots a_n), \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Sensitivity

Example

Define the function $f : \mathbb{R} \to \mathbb{R}$ as $f(x) = 9^x$. Consider the effect of a small perturbation to the input of $f(50) = 9^{50}$, such as

$$f(50.5) = \sqrt{9} \times 9^{50} = 3f(50).$$

Here a 1 percent change in the input causes a 300 percent change of the output.

Sensitivity of linear systems

Example

The linear system $Ax = b$ with

$$A = \begin{pmatrix} 1/3 & 1/3 \\ 1/3 & 0.3 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

has the solution

$$x = \begin{pmatrix} -27 \\ 30 \end{pmatrix}.$$

Example cont.

Example

However, a small change of the $(2,2)^{th}$ element of the matrix A from 0.3 to $1/3$ results in the total loss of the solution, because the system $\tilde{A}x = b$ with

$$\tilde{A} = \begin{pmatrix} 1/3 & 1/3 \\ 1/3 & 1/3 \end{pmatrix}$$

has no solution. Since,

$$b = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

does belong to range space of A.

Example

The linear system $Ax = b$ with

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1+\epsilon \end{pmatrix}, b = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, 0 < \epsilon \ll 1,$$

has the solution

$$x = \tfrac{1}{\epsilon} \begin{pmatrix} -2-\epsilon \\ 2 \end{pmatrix}.$$

Example

The linear system $Ax = b$ with

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1+\epsilon \end{pmatrix}, b = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, 0 < \epsilon \ll 1,$$

has the solution

$$x = \tfrac{1}{\epsilon} \begin{pmatrix} -2-\epsilon \\ 2 \end{pmatrix}.$$

But changing the $(2,2)^{th}$ element of A from $1+\epsilon$ to 1 results in the loss of the solution, because the linear system $\tilde{A}x = b$ with

$$\tilde{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

has no solution. This happens regardless of how small ϵ is.

Absolute and relative error

Definition

If the scalar $\tilde{x}$ is an approximation to the scalar x, then we call $|x - \tilde{x}|$ an absolute error. If $x \neq 0$, then we call $\frac{|x-\tilde{x}|}{|x|}$ a relative error. If $\tilde{x} \neq 0$, then $\frac{|x-\tilde{x}|}{|\tilde{x}|}$ is also a relative error.

Absolute and relative error

Definition

If the scalar $\tilde{x}$ *is an approximation to the scalar* x*, then we call* $|x-\tilde{x}|$ *an absolute error. If* $x \neq 0$*, then we call* $\frac{|x-\tilde{x}|}{|x|}$ *a relative error. If* $\tilde{x} \neq 0$*, then* $\frac{|x-\tilde{x}|}{|\tilde{x}|}$ *is also a relative error.*

How about matrices?

Absolute and relative errors(using norm)

Definition

If $\tilde{x}$ is an approximation to a vector $x \in \mathbb{R}^n$, then $\|x - \tilde{x}\|$ is a normwise absolute error. If $x \neq 0$ or $\tilde{x} \neq 0$, then $\frac{\|x-\tilde{x}\|}{\|x\|}$ and $\frac{\|x-\tilde{x}\|}{\|\tilde{x}\|}$ are normwise relative errors.

Sensitivity of linear systems

Example

Consider the linear system $Ax = b$, where $A = \begin{pmatrix} 1000 & 999 \\ 999 & 998 \end{pmatrix}$ and $b = \begin{pmatrix} 1999 \\ 1997 \end{pmatrix}$.

Sensitivity of linear systems

Example

Consider the linear system $Ax = b$, where $A = \begin{pmatrix} 1000 & 999 \\ 999 & 998 \end{pmatrix}$ and $b = \begin{pmatrix} 1999 \\ 1997 \end{pmatrix}$. Then, $x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is the unique solution to the above system.

Sensitivity of linear systems

Example

Consider the linear system $Ax = b$, where $A = \begin{pmatrix} 1000 & 999 \\ 999 & 998 \end{pmatrix}$ and $b = \begin{pmatrix} 1999 \\ 1997 \end{pmatrix}$. Then, $x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is the unique solution to the above system.
Now, let us consider a slightly perturbed linear system $Ax = b$, where $A = \begin{pmatrix} 1000 & 999 \\ 999 & 998 \end{pmatrix}$ and $b = \begin{pmatrix} 1998.99 \\ 1997.01 \end{pmatrix}$. Then $x = \begin{pmatrix} 20.97 \\ -18.99 \end{pmatrix}$ is the unique solution to the above system.

Condition number

Definition

For an invertible matrix A, the condition number of A with respect to a norm $\|.\|$, denoted by $\kappa(A)$, is defined to be

$$k(A) = \|A\|\|A^{-1}\|.$$

Condition number

Definition

For an invertible matrix A, the condition number of A with respect to a norm $\|.\|$, denoted by $\kappa(A)$, is defined to be

$$k(A) = \|A\|\|A^{-1}\|.$$

Notation: For $1 \leq p \leq \infty$**,** $\kappa_p(A) = \|A\|_p\|A^{-1}\|_p$.

Condition number

Definition

For an invertible matrix A, the condition number of A with respect to a norm $\|.\|$, denoted by $\kappa(A)$, is defined to be

$$k(A) = \|A\|\|A^{-1}\|.$$

Notation: For $1 \leq p \leq \infty$**,** $\kappa_p(A) = \|A\|_p\|A^{-1}\|_p$.

Example

If $A = \begin{pmatrix} 1000 & 999 \\ 999 & 998 \end{pmatrix}$, then $A^{-1} = \begin{pmatrix} -998 & 999 \\ 999 & -1000 \end{pmatrix}$
Then, $||A||_1 = ||A||_\infty = 1999$, and $||A^{-1}||_1 = ||A^{-1}||_\infty = 1999$. Thus $\kappa_1(A) = \kappa_\infty(A) = 1999 \times 1999$.

Theorem

Let $A \in \mathbb{R}^{n \times n}$ *be an invertible matrix. Then*

- $\kappa(A) = \kappa(A^{-1})$.
- $\kappa(A) = \kappa(cA)$ *for any non zero real number* c.
- $\kappa(A) \geq 1$.

Theorem

Let $A \in \mathbb{R}^{n \times n}$ be an invertible matrix. Then

- $\kappa(A) = \kappa(A^{-1})$.
- $\kappa(A) = \kappa(cA)$ *for any non zero real number* c.
- $\kappa(A) \geq 1$.

Remark

1. *Condition number of a singular matrix is defined to be infinity.*
2. *In general, there is no relationship between the condition number and the determinant. E.g. For the matrix* $A = \begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix}$, *where* $\alpha \neq 0$, $\det(A_\alpha) = \alpha^2$ *and* $\kappa(A_\alpha) = 1$.

Condition number - measure of sensitivity of linear systems

Theorem

Let A be non-singular, and let x and $\tilde{x} = x + \Delta x$ be the solutions of $Ax = b$ and $A\tilde{x} = b + \delta b$. Then

$$\frac{||\Delta x||}{||x||} \leq \kappa(A) \frac{||\Delta b||}{||b||}.$$

Condition number - measure of sensitivity of linear systems

Theorem

Let A be non-singular, and let x and $\tilde{x} = x + \Delta x$ be the solutions of $Ax = b$ and $A\tilde{x} = b + \delta b$. Then

$$\frac{||\Delta x||}{||x||} \leq \kappa(A) \frac{||\Delta b||}{||b||}.$$

Remark

If we perturb the coefficient matrix A, then, also, we can bound the error in the solution. Note that, perturbed matrix need not be invertible.

Condition number - measure of sensitivity of linear systems

Theorem

Let A be an invertible matrix. If $\dfrac{\|\Delta A\|}{\|A\|} < \dfrac{1}{\kappa(A)}$*, then* $A + \Delta A$ *is invertible.*

Theorem

Let A be an invertible matrix. If x and $\tilde{x} = x + \Delta x$ are the solutions to the systems $Ax = b$ and $(A + \Delta A)\tilde{x} = b$, and $\dfrac{\|\Delta A\|}{\|A\|} < \dfrac{1}{\kappa(A)}$*, then*

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{k(A)\frac{\|\Delta A\|}{\|A\|}}{1 - \frac{\|\Delta A\|}{\|A\|}\kappa(A)}.$$

Condition number - measure of sensitivity of linear systems

Theorem

Let A be an invertible matrix. If $Ax = b$ and

$$(A + \Delta A)(x + \Delta x) = (b + \Delta b);\ b + \Delta b \neq 0,$$

then

$$\frac{\|\Delta x\|}{\|\tilde{x}\|} \leq \kappa(A)\left(\frac{\|\Delta A\|}{\|A\|} + \frac{\|\Delta b\|}{\|b + \Delta b\|} + \frac{\|\Delta A\|\|\Delta b\|}{\|A\|\|b + \Delta b\|}\right).$$

Theorem

Let A be an invertible matrix, and $\frac{\|\Delta A\|}{\|A\|} < \frac{1}{\kappa(A)}$. If $Ax = b$ and

$$(A + \Delta A)(x + \Delta x) = (b + \Delta b);\ b \neq 0,$$

then

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{\kappa(A)\left(\frac{\|\Delta A\|}{\|A\|} + \frac{\|\Delta b\|}{\|b\|}\right)}{1 - \frac{\|\Delta A\|}{\|A\|}\kappa(A)}.$$

Example

Example

If $A = \begin{pmatrix} 1000 & 999 \\ 999 & 998 \end{pmatrix}$, then $A^{-1} = \begin{pmatrix} -998 & 999 \\ 999 & -1000 \end{pmatrix}$
Then, $||A||_1 = ||A||_\infty = 1999$, and $||A^{-1}||_1 = ||A^{-1}||_\infty = 1999$. Thus $\kappa_1(A) = \kappa_\infty(A) = 1999 \times 1999$.

The condition number of the matrix A is high, so the solutions of the perturbed system in the previous example changed drastically.

Geometric meaning of condition number

Definition

The maximum and minimum magnification by the matrix A are defined, respectively, by

- $\text{maxmag}(A) = \max_{\|x\|=1} \|Ax\|$,
- $\text{minmag}(A) = \min_{\|x\|=1} \|Ax\|$.

Geometric meaning of condition number

Definition

The maximum and minimum magnification by the matrix A are defined, respectively, by

- $\text{maxmag}(A) = \max\limits_{\|x\| = 1} \|Ax\|$,
- $\text{minmag}(A) = \min\limits_{\|x\| = 1} \|Ax\|$.

Theorem

If A is nonsingular matrix, then

1. $\text{maxmag}(A) = \dfrac{1}{\text{minmag}(A^{-1})}$, *and*
2. $\text{minmag}(A) = \dfrac{1}{\text{maxmag}(A^{-1})}$.

Geometric meaning of condition number

Theorem

If A is a nonsingular matrix, then

$$\kappa(A) = \frac{\operatorname{maxmag}(A)}{\operatorname{minmag}(A)}.$$

Observations

- Consider $A = \begin{pmatrix} 1000 & 999 \\ 999 & 998 \end{pmatrix}$.

Observations

- Consider $A = \begin{pmatrix} 1000 & 999 \\ 999 & 998 \end{pmatrix}$.
- As, $\mathrm{maxmag}(A) = \|A\|_\infty$, so $\mathrm{maxmag}(A) = 1999$. For the vector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, it is easy to see that, $\begin{pmatrix} 1000 & 999 \\ 999 & 998 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1999 \\ 1997 \end{pmatrix}$.

Observations

- Consider $A = \begin{pmatrix} 1000 & 999 \\ 999 & 998 \end{pmatrix}$.
- As, $\mathtt{maxmag}(A) = \|A\|_\infty$, so $\mathtt{maxmag}(A) = 1999$. For the vector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, it is easy to see that, $\begin{pmatrix} 1000 & 999 \\ 999 & 998 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1999 \\ 1997 \end{pmatrix}$.
- So, with respect $\|.\|_\infty$, the vector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is magnified maximally by A, and hence it gives a direction of maximum magnification. Also, the vector $\begin{pmatrix} 1999 \\ 1997 \end{pmatrix}$ is in the direction of minimum magnification A^{-1}.

Observations

- Similarly for the matrix $A^{-1} = \begin{pmatrix} -998 & 999 \\ 999 & -1000 \end{pmatrix}$ the vector $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ is in a direction of maximum magnification of A^{-1}, and the vector $\begin{pmatrix} 1997 \\ -1999 \end{pmatrix}$ is in the direction of minimum magnification of A.

Using these observations, let us construct an interesting example.

Spectacular example(Watkins)

Example

Consider the linear system $Ax = b$, where $A = \begin{pmatrix} 1000 & 999 \\ 999 & 998 \end{pmatrix}$ and $b = \begin{pmatrix} 1999 \\ 1997 \end{pmatrix}$.

Spectacular example(Watkins)

Example

Consider the linear system $Ax = b$, where $A = \begin{pmatrix} 1000 & 999 \\ 999 & 998 \end{pmatrix}$ and $b = \begin{pmatrix} 1999 \\ 1997 \end{pmatrix}$. Then, $x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is the unique solution to the above system.

Spectacular example(Watkins)

Example

Consider the linear system $Ax = b$, where $A = \begin{pmatrix} 1000 & 999 \\ 999 & 998 \end{pmatrix}$ and $b = \begin{pmatrix} 1999 \\ 1997 \end{pmatrix}$. Then, $x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is the unique solution to the above system.

Now, let us consider a slightly perturbed linear system $A(x + \Delta x) = b + \Delta b$, where $\Delta b = \begin{pmatrix} -0.01 \\ 0.01 \end{pmatrix}$, a vector in the direction of maximum magnification by A^{-1}. Then

$$x + \Delta x = A^{-1}\begin{pmatrix} 1999 \\ 1997 \end{pmatrix} + A^{-1}\Delta b = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 19.97 \\ -19.99 \end{pmatrix} = \begin{pmatrix} 20.97 \\ -18.99 \end{pmatrix}.$$

Scaling

Example

Consider the linear system $Ax = b$, where $A = \begin{pmatrix} 1 & 0 \\ 0 & \epsilon \end{pmatrix}$, where $0 < \epsilon \ll 1$ and, $b = \begin{pmatrix} 1 \\ \epsilon \end{pmatrix}$.

Scaling

Example

Consider the linear system $Ax = b$, where $A = \begin{pmatrix} 1 & 0 \\ 0 & \epsilon \end{pmatrix}$, where $0 < \epsilon \ll 1$ and, $b = \begin{pmatrix} 1 \\ \epsilon \end{pmatrix}$. Then, $x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is the unique solution to the above system.

Scaling

Example

Consider the linear system $Ax = b$, where $A = \begin{pmatrix} 1 & 0 \\ 0 & \epsilon \end{pmatrix}$, where $0 < \epsilon \ll 1$ and, $b = \begin{pmatrix} 1 \\ \epsilon \end{pmatrix}$. Then, $x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is the unique solution to the above system. $\kappa_1(A) = \kappa_\infty(A) = \frac{1}{\epsilon}$, so the matrix is ill-conditioned.

Take $\Delta b = \begin{pmatrix} 0 \\ \epsilon \end{pmatrix}$, then $x + \Delta x = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $\frac{\|\Delta b\|_\infty}{\|b\|_\infty} = \epsilon$, and $\frac{\|\Delta x\|_\infty}{\|x\|_\infty} = 1$.

Multiply the second row of the system by $\frac{1}{\epsilon}$, then we get a well conditioned system, with $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Theorem

Let A be any nonsingular matrix, and let $a_1, a_2, \ldots, a_n$ be the columns of A. Then for any i and j,

$$\kappa_p(A) \geq \frac{\|a_i\|_p}{\|a_j\|_p},$$

for $1 \leq p \leq \infty$.

Theorem

Let A be any nonsingular matrix, and let $a_1, a_2, \ldots, a_n$ be the columns of A. Then for any i and j,

$$\kappa_p(A) \geq \frac{\|a_i\|_p}{\|a_j\|_p},$$

for $1 \leq p \leq \infty$.

Remark

1. *If the columns of the matrix A have different orders of magnitude, then A is ill-conditioned.*

Theorem

Let A be any nonsingular matrix, and let $a_1, a_2, \ldots, a_n$ be the columns of A. Then for any i and j,

$$\kappa_p(A) \geq \frac{\|a_i\|_p}{\|a_j\|_p},$$

for $1 \leq p \leq \infty$.

Remark

1. *If the columns of the matrix A have different orders of magnitude, then A is ill-conditioned. Similarly for the rows.*

Theorem

Let A be any nonsingular matrix, and let $a_1, a_2, \ldots, a_n$ be the columns of A. Then for any i and j,

$$\kappa_p(A) \geq \frac{\|a_i\|_p}{\|a_j\|_p},$$

for $1 \leq p \leq \infty$.

Remark

1. *If the columns of the matrix A have different orders of magnitude, then A is ill-conditioned. Similarly for the rows. s*
2. *Necessary condition for a matrix to be well-conditioned is that all its rows and columns are of roughly the same magnitude.*

Theorem

Let A be any nonsingular matrix, and let $a_1, a_2, \ldots, a_n$ be the columns of A. Then for any i and j,

$$\kappa_p(A) \geq \frac{\|a_i\|_p}{\|a_j\|_p},$$

for $1 \leq p \leq \infty$.

Remark

1. *If the columns of the matrix A have different orders of magnitude, then A is ill-conditioned. Similarly for the rows. s*
2. *Necessary condition for a matrix to be well-conditioned is that all its rows and columns are of roughly the same magnitude. It is not sufficient!*

Theorem

Let A be any nonsingular matrix, and let $a_1, a_2, \ldots, a_n$ be the columns of A. Then for any i and j,

$$\kappa_p(A) \geq \frac{\|a_i\|_p}{\|a_j\|_p},$$

for $1 \leq p \leq \infty$.

Remark

1. *If the columns of the matrix A have different orders of magnitude, then A is ill-conditioned. Similarly for the rows. s*
2. *Necessary condition for a matrix to be well-conditioned is that all its rows and columns are of roughly the same magnitude. It is not sufficient! Example?*