MA580H Matrix Computations

Lecture 9: Perturbation analysis of linear systems

Rafikul Alam
Department of Mathematics
IIT Guwahati

Outline

- Vector and matrix norms
- Perturbation analysis of linear systems
- Stability analysis of GEPP

Vector norms

Let $\mathcal{V}$ be a vector space over $\mathbb{C}$. Then a function $\|\cdot\| : \mathcal{V} \to \mathbb{R}$ is called a norm on $\mathcal{V}$ if it satisfies the three fundamental properties:

(a) Positive definiteness: $\|v\| \geq 0$ and $\|v\| = 0 \iff v = 0$.

(b) Positively homogeneous: $\|\alpha v\| = |\alpha| \, \|v\|$ for $\alpha \in \mathbb{C}$ and $v \in \mathcal{V}$.

(c) Triangle inequality: $\|u + v\| \leq \|u\| + \|v\|$ for $u, v \in \mathcal{V}$.

Example: Consider $\mathbb{C}^n$ and the vector norms given by

1-norm: $\|x\|_1 := |x_1| + \cdots + |x_n|$.

2-norm: $\|x\|_2 := \sqrt{|x_1|^2 + \cdots |x_n|^2}$.

∞-norm: $\|x\|_\infty := \max_{1 \leq j \leq n} |x_j|$.

Example:

$$\|[1, 1, 3, 5]^\top\|_1 = 10, \|[1, 1, 3, 5]^\top\|_2 = 6 \;\text{ and }\; \|[1, 1, 3, 5]^\top\|_\infty = 5.$$

Matrix norms

Let $A \in \mathbb{C}^{m\times n}$. Then $A : \mathbb{C}^n \longrightarrow \mathbb{C}^m$, $x \longmapsto Ax$, is a linear map. Suppose $\mathbb{C}^n$ and $\mathbb{C}^m$ are equipped with norms. Then

$$\|A\| := \max_{x\neq 0} \frac{\|Ax\|}{\|x\|} = \max_{\|x\|=1} \|Ax\|$$

defines a norm on $\mathbb{C}^{m\times n}$ and is called an induced matrix norm or a subordinate matrix norm.

For the identity matrix $\|Ix\| = \|x\|$ and hence $\|I\| = 1$. Note that

$$\|Ax\| \leq \|A\| \, \|x\|$$

for all $x \in \mathbb{C}^n$.

A matrix norm is said to be sub-multiplicative if $\|AB\| \leq \|A\| \, \|B\|$ holds for all A and B. An induced matrix norm is submultiplicative. Indeed, we have

$$\|ABx\| \leq \|A\| \, \|Bx\| \leq \|A\| \, \|B\| \, \|x\| \Longrightarrow \|AB\| \leq \|A\| \, \|B\|.$$

Matrix norms

The norms $\|A\|_1$, $\|A\|_2$ and $\|A\|_\infty$ induced by 1-norm, 2-norm and ∞-norm are called 1-norm, 2-norm and ∞-norm of A, respectively. Also $\|A\|_2$ is called the spectral norm of A.

Theorem: Let A be an $m \times n$ matrix. Then

$$\begin{aligned}
\|A\|_1 &= \max_{\|x\|_1=1} \|Ax\|_1 = \max_{1\leq j\leq n} \|Ae_j\|_1 = \max_{1\leq j\leq n} \|A(:,j)\|_1 \\
\|A\|_2 &= \max_{\|x\|_2=1} \|Ax\|_2 = \sqrt{\lambda_{\max}(A^*A)} \\
\|A\|_\infty &= \max_{\|x\|_\infty=1} \|Ax\|_\infty = \max_{1\leq i\leq m} \|e_i^\top A\|_1 = \max_{1\leq i\leq m} \|A(i,:)\|_1,
\end{aligned}$$

where $\lambda_{\max}(A^*A)$ denotes that largest eigenvalue of A^*A.

Proof: We have $Ax = x_1 Ae_1 + \cdots + x_n Ae_n \Rightarrow \|Ax\|_1 \leq \max_{1\leq j\leq n} \|Ae_j\|_1 \|x\|_1$. This yields $\|A\|_1 \leq \max_{1\leq j\leq n} \|Ae_j\|_1$. But $\|Ae_j\|_1 \leq \|A\|_1$ for all $j = 1 : n$. Hence we have $\|A\|_1 = \max_{1\leq j\leq n} \|Ae_j\|_1$. ■

Example

Let $A := \begin{bmatrix} 2 & 2 & -4 \\ 1 & 1 & 5 \\ 1 & 3 & 6 \end{bmatrix}$. Then $\|A\|_1 = \max\left(\|Ae_1\|_1, \|Ae_2\|_1, \|Ae_3\|_1\right) = \max(4, 6, 15) = 15.$

We have $\|A\|_\infty = \max(\|e_1^\top A\|_1, \|e_2^\top A\|_1, \|e_3^\top A\|_1) = \max(8, 7, 10) = 10.$

The spectral norm of A is given by $\|A\|_2 = \sqrt{\lambda_{\max}(A^\top A)} = 8.9826.$

Condition number and non-singularity

If A is nonsingular then when is $A + \Delta A$ nonsingular?

Fact: If $\|\Delta A\| \, \|A^{-1}\| < 1$ or equivalently, $\frac{\|\Delta A\|}{\|A\|}\mathrm{cond}(A) < 1$, then $A + \Delta A$ is nonsingular, where $\mathrm{cond}(A) := \|A\| \, \|A^{-1}\|$.

Proof: If possible, suppose that $A + \Delta A$ is singular. Then there is a vector x such that $\|x\| = 1$ and $(A + \Delta A)x = 0$.

Then $x = -A^{-1}\Delta A x \Longrightarrow 1 = \|A^{-1}\Delta A x\| \leq \|A^{-1}\| \, \|\Delta A\|$, which is a contradiction. ■

Remark: There is a ΔA such that $\|\Delta A\| \, \|A^{-1}\| = 1$ and $A + \Delta A$ is singular. In other words, the relative distance to nearest singular matrix $\propto \frac{1}{\mathrm{cond}(A)}$.

Condition number

Definition: Let A be an $n \times n$ nonsingular matrix. Then $\mathrm{cond}(A) := \|A\| \, \|A^{-1}\|$ is called the condition number of A. If $\mathrm{cond}(A)$ is NOT too large then A is said to be well-conditioned. If $\mathrm{cond}(A)$ is large then A is said to be ill-conditioned.

Note that for a subordinate matrix norm, we have $\mathrm{cond}(A) = \|A\| \, \|A^{-1}\| \geq 1$.

Remark: The determinant $\det(A)$ is not a good measure of ill-conditioning of A.

$$A := 10^{-1} I_n \Longrightarrow \det(A) = 10^{-n} \text{ and } \mathrm{cond}(A) = 1.$$

$$B := \begin{bmatrix} 1 & 10^{10} \\ 0 & 1 \end{bmatrix} \Longrightarrow \det(B) = 1 \text{ and } \mathrm{cond}_\infty(B) = (1 + 10^{10})^2 \simeq 10^{20}.$$

Notice that columns A are orthogonal whereas columns of B are nearly linearly dependent. Indeed, $\cos\theta = \langle Be_1, \, Be_2 \rangle / \|Be_1\|_2 \|Be_2\|_2 = 10^{10}/\sqrt{1 + 10^{20}} \simeq 1$.

Sensitivity analysis of linear systems

Consider the linear system

$$\underbrace{\begin{bmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n+1} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{1}{n} & \frac{1}{n+1} & \cdots & \frac{1}{2n-1} \end{bmatrix}}_{\text{Hilbert matrix H}} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

The matrix H is known as a Hilbert matrix, and it is known to be notoriously ill-conditioned.

To see what this means, set $x := \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix}^\top \in \mathbb{R}^n$ and define $b := Hx$. Then x is the solution of $Hx = b$.

Now we use MATLAB to solve the linear system and compare the computed solution with the known solution x.

Sensitivity of solutions of Hilbert systems

```
>> xx = hilb(12)\b; Warning: Matrix is close to singular or
badly scaled. Results may be inaccurate. RCOND = 2.602837e-17.
```

n	$\|\mathtt{x} - \mathtt{xx}\|_\infty$	cond(H)
4	.4130030e-12	2.837500e+04
6	.6964739e-09	2.907028e+07
8	.7311487e-07	3.387279e+10
10	.2047785e-03	3.535233e+13
12	.2476695e-00	3.841961e+16

This would appear to justify the predictions that as n increases, roundoff errors would accumulate and destroy all accuracy in the computed solution of a linear system!

The Hilbert matrix is SPD but the computed solutions differ drastically from true solutions. Is it the fault of the algorithm?

Perturbation of linear system-I

Theorem: Let A be nonsingular and $\mathrm{cond}(A) := \|A\| \, \|A^{-1}\|$. Consider the linear systems $Ax = b$ and $A\hat{x} = b + \Delta b$. Then

$$\frac{\|x - \hat{x}\|}{\|x\|} \leq \mathrm{cond}(A) \frac{\|\Delta b\|}{\|b\|}.$$

Moreover, the upper bound is attained for some Δb.

Proof: We have $\hat{x} - x = A^{-1}\Delta b \Longrightarrow \|x - \hat{x}\| \leq \|A^{-1}\| \, \|\Delta b\|$. Now $Ax = b \Longrightarrow \|b\| \leq \|A\| \, \|x\| \Longrightarrow 1/\|x\| \leq \|A\|/\|b\|$, which yields the bound. ∎

Residual bound: Let $\hat{x} = \mathrm{ALG}(A, b)$. Then the residual $r := b - A\hat{x}$ yields $A\hat{x} = b - r = b + \Delta b$, where $\Delta b := -r$. Hence we have the residual bound $\dfrac{\|x - \hat{x}\|}{\|x\|} \leq \mathrm{cond}(A) \dfrac{\|r\|}{\|b\|}$.

Example

Consider $A := \begin{bmatrix} 1000 & 999 \\ 999 & 998 \end{bmatrix}$. Then $A^{-1} = \begin{bmatrix} -998 & 999 \\ 999 & -1000 \end{bmatrix}$.

Thus $\|A\|_\infty = \|A\|_1 = \|A^{-1}\|_\infty = \|A^{-1}\|_1 = 1999$. Hence $\mathrm{cond}_\infty(A) = \mathrm{cond}_1(A) = (1999)^2 = 3.996 \times 10^6$.

Observe that $A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1999 \\ 1997 \end{bmatrix}$ and $A^{-1} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1997 \\ -1999 \end{bmatrix}$.

Set $b := \begin{bmatrix} 1999 \\ 1997 \end{bmatrix}$ and $\Delta b := 10^{-2} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. Consider system $A\hat{x} = b + \Delta b$. Then $\hat{x} = x + A^{-1}\Delta b = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 19.97 \\ -19.99 \end{bmatrix}$. This shows that

$$\frac{\|x - \hat{x}\|_\infty}{\|x\|_\infty} = 19.99 = (1999)^2 \frac{10^{-2}}{1999} = \mathrm{cond}_\infty(A) \frac{\|\Delta b\|_\infty}{\|b\|_\infty}.$$

Perturbation of linear system-II

Theorem: Consider the systems $Ax = b$ and $(A + \Delta A)\hat{x} = b + \Delta b$. Suppose that A is nonsingular and $\|\Delta A\| \, \|A^{-1}\| < 1$. Then

$$\begin{aligned}\frac{\|x - \hat{x}\|}{\|x\|} &\leq \frac{\mathrm{cond}(A)}{1 - \frac{\|\Delta A\|}{\|A\|}\mathrm{cond}(A)} \left(\frac{\|\Delta A\|}{\|A\|} + \frac{\|\Delta b\|}{\|b\|} \right) \\ &\lesssim \mathrm{cond}(A) \left(\frac{\|\Delta A\|}{\|A\|} + \frac{\|\Delta b\|}{\|b\|} \right).\end{aligned}$$

Proof: We have

$\hat{x} - x = -A^{-1}(\Delta A\hat{x} - \Delta b) \Longrightarrow \|\hat{x} - x\| \leq \|A^{-1}\|(\|\Delta A\| \, \|\hat{x}\| + \|\Delta b\|)$. Now

$\|\hat{x}\| \leq \|\hat{x} - x\| + \|x\| \Longrightarrow (1 - \|A^{-1}\| \, \Delta A\|)\|x - \hat{x}\| \leq \|A^{-1}\|(\|\Delta A\| \, \|x\| + \|\Delta b\|)$.

Now diving both sides by $\|x\|$ and using the fact that $b = Ax \Longrightarrow \|b\| \leq \|A\| \, \|x\| \Longrightarrow \|b\|/\|x\| \leq \|A\|$, we obtain the bound. ■

Example

Consider $A := \begin{bmatrix} 1 & 1+\delta \\ 1-\delta & 1 \end{bmatrix}$, where $\delta > 0$. Then $A^{-1} = \dfrac{1}{\delta^2} \begin{bmatrix} 1 & -1-\delta \\ -1+\delta & 1 \end{bmatrix}$. Hence $\mathrm{cond}_\infty(A) = \dfrac{(2+\delta)^2}{\delta^2}$.

For $\delta := 10^{-2}$, we have $\mathrm{cond}_\infty(A) = (201)^2 = 40401$.

Consider the linear systems $\begin{bmatrix} 1 & 1.01 \\ 0.99 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2.01 \\ 1.99 \end{bmatrix}$ whose solution is $x = \begin{bmatrix} 1 & 1 \end{bmatrix}^\top$ and $\begin{bmatrix} 1 & 1.01 \\ 1 & 1 \end{bmatrix} \hat{x} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$.

Then $\hat{x} = \begin{bmatrix} 2 & 0 \end{bmatrix}^\top$. Note that $\Delta A = 10^{-2} e_2 e_1^\top$ and $\Delta b = 10^{-2} \begin{bmatrix} 1 & 1 \end{bmatrix}^\top$. We have $\|x - \hat{x}\|_\infty / \|x\|_\infty = 1$.