Introduction to Numerical Analysis I
Handout 14

1 Numerical Linear Algebra (cont)

1.2 Norm of Matrix

Reminder: The vector norms we use are

$$\|v\|_1 = \sum_{i=1}^{n} |v_i|, \quad \|v\|_2 = \sqrt{\sum_{i=1}^{n} |v_i|^2}, \quad \|v\|_\infty = \max_i |v_i|$$

Definition 1.1 (Spectral Radius). Let A be $n \times n$ matrix and λ_j be j'th eigenvalue of A, then the spectral radius ρ reads for

$$\rho\left(A_{n,n}\right) = \max_{1 \leqslant j \leqslant n} |\lambda_j|$$

Definition 1.2 (An induced norm of matrix). A matrix norm induced from a vector norm $\|\cdot\}$ is given by

$$\|A\| = \max_{0 \neq \vec{v} \in \Re^n} \frac{\|A\vec{v}\|}{\|\vec{v}\|}$$

Theorem 1.3. $\|A\| = \max_{\substack{\vec{v} \in \Re^n \\ \|\vec{v}\|=1}} \|A\vec{v}\|$

Proof:

$$\|A\| = \max_{0 \neq \vec{v} \in \Re^n} \frac{\|A\vec{v}\|}{\|\vec{v}\|} =$$

$$\max_{0 \neq \vec{v} \in \Re^n} \left\| \frac{A\vec{v}}{\|\vec{v}\|} \right\| = \max_{0 \neq \vec{v} \in \Re^n} \left\| A \frac{\vec{v}}{\|\vec{v}\|} \right\| = \max_{\substack{\vec{v} \in \Re^n \\ \|\vec{v}\|=1}} \|A\vec{v}\|$$

Properties:

- $\|A\vec{v}\| \leqslant \|A\| \|\vec{v}\|, \forall \vec{v} \in \Re^n$
  Proof:
  $$\forall \vec{v} \in \Re^n \backslash \{0\} : \frac{\|A\vec{v}\|}{\|\vec{v}\|} \leqslant \|A\| \Rightarrow \|A\vec{v}\| \leqslant \|A\| \|\vec{v}\|$$

- $\|AB\| \leqslant \|A\| \|B\|$
  Proof:
  $$\|AB\| = \max_{\|\vec{v}\|=1} \|AB\vec{v}\| \leqslant \max_{\|\vec{v}\|=1} \|A\| \|B\vec{v}\| \leqslant$$
  $$\leqslant \max_{\|\vec{v}\|=1} \|A\| \|B\| \|\vec{v}\| = \|A\| \|B\|$$

The norms we will use:

1. $\|A\|_1 = \max_{\text{columns } j} \|A_{\downarrow j}\|_1 = \max_j \sum_i |a_{ij}|$

   Proof:

   $$\|Av\|_1 = \left\| \sum_{j=1}^{n} v_j A_{\downarrow j} \right\|_1 \leq \sum \|A_{\downarrow j}\|_1 |v_j| \leq \max_{\text{columns } j} \|A_{\downarrow j}\|_1 \|v\|_1$$

   Which implies

   $$\|A\|_1 = \max_{\|v\|=1} \|Av\|_1 \leq \max_{\text{columns } j} \|A_{\downarrow j}\|_1 .$$

   From the other side, by definition $\|A\|_1 \geq \|Av\|_1$ for any $\|v\| = 1$. Choose $v = e_m = (0, ...1..., 0)^T$, where m satisfies

   $$A_{\downarrow m} = \max_{\text{columns } j} \left\{ \|A_{\downarrow j}\|_1 \right\},$$

   to get $Ae_m = A_{\downarrow m}$, that is $\|A\|_1 \geq \|A_{\downarrow m}\|_1$.

2. $\|A\|_\infty = \max_{\text{rows } i} \{\|A_{i \rightarrow}\|_1\} = \max_i \left\{ \sum_j |a_{ij}| \right\}$, the proof is similar.

3. $\|A\|_2 = \sqrt{\rho\left(AA^T\right)}$
   Proof: Since $A^T A$ is symmetric matrix one writes $A^T A = Q^T \Lambda Q$, where Q is orthogonal matrix and $\Lambda = \text{diag}(\lambda_1^2, \ldots, \lambda_n^2)$ is the diagonal matrix of eigenvalues of $A^T A$ which are $\text{eig}(A^T A) = \text{eig}(A)^2$. Thus, $\|Av\|^2 = (Av, Av) = Av \cdot Av = v^T \left(A^T A\right) v = v^T \left(Q^T \Lambda Q\right) v \underset{y=Qv}{=} y^T \Lambda y$ and therefore $\|A\|_2 = \max_{\|v\|=1} \|Av\|_2 = \max_{\|v\|=1} \sqrt{v^T \Lambda v} = \sqrt{\rho\left(AA^T\right)}$.

Example 1.4. Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, then

1. $\|A\|_1 = \max\left\{ \|(1,3)^T\|_1, \|(2,4)^T\|_1 \right\} = 6$
2. $\|A\|_\infty = \max\left\{ \|(1,2)\|_1, \|(3,4)\|_1 \right\} = 7$
3. $\|A\|_2 = \sqrt{\rho\left(AA^T\right)} =$

$$= \sqrt{\rho\left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \right)} = \sqrt{\rho\left(\begin{bmatrix} 5 & 11 \\ 11 & 25 \end{bmatrix} \right)}$$

$$\left|\lambda - AA^T\right| = \left| \begin{bmatrix} \lambda - 5 & 11 \\ 11 & \lambda - 25 \end{bmatrix} \right| =$$

$$(\lambda - 5)(\lambda - 25) - 121 = \lambda^2 - 30\lambda + 4$$

$$\lambda_{1,2} = \frac{30 \pm \sqrt{900 - 16}}{2} = 15 \pm \sqrt{221}$$

Thus,

$$\|A\|_2 = \sqrt{\rho\left(AA^T\right)} = \sqrt{15 + \sqrt{221}} \approx 5.465$$

1.2.1 Matrix Condition Number

Consider linear system $Ax = b$, where b were obtained with error. That is, instead of original system we solve

$$A\tilde{x} = A(x + \delta x) = \tilde{b} = (b + \delta b)$$

We want to analyze the sensitivity of the solution x to the small changes in the data b. Thus, the relative error in x is given by

$$\frac{\|\delta x\|}{\|x\|} = \frac{\|\tilde{x} - x\|}{\|x\|} = \frac{\left\|A^{-1}\tilde{b} - A^{-1}b\right\|}{\|x\|} = \frac{\left\|A^{-1}\delta b\right\|}{\|x\|} \leqslant$$
$$\leqslant \frac{\left\|A^{-1}\right\| \|\delta b\|}{\|x\|}\frac{\|A\|}{\|A\|} \leqslant \frac{\left\|A^{-1}\right\| \|\delta b\|}{\|Ax\|}\|A\| =$$
$$= \left\|A^{-1}\right\| \|A\| \frac{\|\delta b\|}{\|b\|}$$

Definition 1.5 (Matrix Condition Number). Let $A_{n\times n}$ be non singular matrix, then the condition number of matrix is given by

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

Example 1.6. Let
$$A = \begin{bmatrix} 100 & 99 \\ 99 & 98 \end{bmatrix}, \; b = \begin{bmatrix} 1.005 \\ 0.995 \end{bmatrix}, \; \delta b = \begin{bmatrix} 0.1 \\ -0.1 \end{bmatrix}$$

1. Find the solution to

$$A\tilde{x} = b + \delta b = \tilde{b} = \begin{bmatrix} 1.1 \\ 0.9 \end{bmatrix}$$

$$\tilde{x} = \frac{\begin{bmatrix} 100 \cdot 0.995 - 99 \cdot 1.005 \\ 1.005 \cdot 98 - 99 \cdot 0.995 \end{bmatrix}}{100 \cdot 98 - 99^2} = \begin{bmatrix} -18.7 \\ 18.9 \end{bmatrix}$$

2. Compute the relative error in input and output.

$$x = \begin{bmatrix} 0.015 \\ -0.005 \end{bmatrix} \qquad \delta x = \begin{bmatrix} -19 \\ 19 \end{bmatrix}$$

$$\frac{\|\delta b\|_\infty}{\|b\|_\infty} \approx 0.1 \qquad \frac{\|\delta x\|_\infty}{\|x\|_\infty} = \frac{19}{0.015} \approx 1266.6$$

Note that the problem is hard, that is the problem is ill-conditioned, a little error in the input caused to very big error in output. This is usually the case in inverse problems.

3. Bound the error in output as a function of the error in input.

$$A^{-1} = \begin{bmatrix} 100 & 99 \\ 99 & 98 \end{bmatrix}^{-1} = \begin{bmatrix} -98 & 99 \\ 99 & -100 \end{bmatrix}$$

Thus $\text{cond}(A, \infty) = \left\|A^{-1}\right\|_\infty \|A\|_\infty = 199^2 \approx 40000$, therefore $1267 = \frac{\|\delta x\|}{\|x\|} \leqslant cond(M) \frac{\|\delta y\|}{\|y\|} = 40000 \cdot 0.1 = 4000$

Another way to see the problem is via eigenvalues. For the matrix A in this example we have a big gap between the two eigenvalues

$$\lambda_1 \approx -\frac{1}{200} = -0.005, \; v_1 \approx \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

and

$$\lambda_2 \approx 200, \; v_2 \approx \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

This cause the following problem

$$Ax = v_2 \Rightarrow x = \tfrac{1}{\lambda_2} v_2$$

but if we add a small error, to the RHS, that is

$$Ax \approx 0.01 v_1 + v_2$$

we will get

$$x = 0.01 A^{-1} v_1 + A^{-1} v_2 = 0.01 \tfrac{1}{\lambda_1 v_1} + \tfrac{1}{\lambda_2} v_2 \approx$$
$$\approx -200 \cdot 0.01 v_1 + \tfrac{1}{200} v_2 = -2v_1 + \tfrac{1}{200} v_2 \approx -2v_1$$

That is the error were magnified more then the exact part.