E2 212: Matrix Theory

Topic: Matrix Norms



Prof. Sudhan Majhi

TA: Misanolla Vamsi

SPWiCOM LAB
Department of Electrical Communication Engineering
Indian Institute of Science, Bangalore, India.

Matrix Norm:

A function $\|\cdot\| : M_n \to \mathbb{R}$ is a *matrix norm* if $\forall\, A, B \in M_n$ it satisfies:

(1)	$\|A\| \geq 0$	Non-negativity
(1a)	$\|A\| = 0$ iff $A = 0$	Positivity
(2)	$\|cA\| = \|c\|\,\|A\| \quad \forall\, c \in \mathbb{F}$	Homogeneity
(3)	$\|A + B\| \leq \|A\| + \|B\|$	Sub-additivity
(4)	$\|AB\| \leq \|A\| \cdot \|B\|$	Sub-multiplicativity

Omitting (1a) $\to$ *Semi-matrix norm.*
Omitting (4) $\to$ *Generalized matrix norms.*

★ $\|A^2\| \leq \|A\|^2$.

For $A^2 = A$ (idempotent), we have

$$\|A^2\| = \|A\| \;\leq\; \|A\|^2 \qquad \Longrightarrow \qquad \|A\| \geq 1.$$

★ If A is non-singular, $I = AA^{-1}$.

$$\|I\| = \|AA^{-1}\| \;\leq\; \|A\| \cdot \|A^{-1}\| \qquad \Longrightarrow \qquad \|A^{-1}\| \;\geq\; \frac{\|I\|}{\|A\|}$$

$$\|I\| \;\geq\; 1.$$

Example: The ℓ_1-norm defined for $A \in M_n$ by

$$\|A\|_1 = \sum_{i,j=1}^{n} |a_{ij}| \qquad \text{(in the vector sense).}$$

Is it a matrix norm?

$$\begin{aligned}\|AB\|_1 = \sum_{i,j=1}^{n} \Big| \sum_{k=1}^{n} a_{ik} b_{kj} \Big| \;&\leq\; \sum_{i,j,k=1}^{n} |a_{ik} b_{kj}| \\ &\leq \sum_{i,j,k,m=1}^{n} |a_{ik}|\,|b_{mj}| = \left(\sum_{i,k=1}^{n} |a_{ik}| \right) \left(\sum_{j,m=1}^{n} |b_{mj}| \right) \\ &= \|A\|_1 \, \|B\|_1.\end{aligned}$$

Example: The ℓ_2-norm (Frobenius norm, Schur norm, or Hilbert–Schmidt norm) defined for $A \in M_n$ by

$$\|A\|_2 = \left| \operatorname{tr} AA^* \right|^{1/2} = \left(\sum_{i,j=1}^{n} |a_{ij}|^2 \right)^{1/2}$$

is a matrix norm because

$$\begin{aligned}\|AB\|_2 = \left(\sum_{i,j=1}^{n} \Big| \sum_{k=1}^{n} a_{ik} b_{kj} \Big|^2 \right)^{1/2} &\leq \left(\sum_{i,j=1}^{n} \Big(\sum_{k=1}^{n} |a_{ik}|^2 \Big) \Big(\sum_{m=1}^{n} |b_{mj}|^2 \Big) \right)^{1/2} \\ = \left(\sum_{i,k=1}^{n} |a_{ik}|^2 \right)^{1/2} \left(\sum_{m,j=1}^{n} |b_{mj}|^2 \right)^{1/2} &= \|A\|_2 \, \|B\|_2.\end{aligned}$$

Example: The ℓ_∞-norm defined for $A \in M_n$ by

$$\|A\|_\infty = \max_{1 \leq i,j \leq n} |a_{ij}|$$

is a norm on the vector space M_n but is **not** a matrix norm.

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

Example: Consider the matrix

$$J = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \qquad J^2 = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} = 2J,$$

$$\|J\|_\infty = 1, \qquad \|J^2\|_\infty = \|2J\|_\infty = 2\,\|J\|_\infty = 2.$$

Since $\|J^2\|_\infty \geq \|J\|_\infty^2$ fails, $\|\cdot\|_\infty$ is **not sub-multiplicative**. However, if we define

$$N(A) = n\,\|A\|_\infty, \qquad A \in M_n,$$

then we have

$$\begin{aligned} N(AB) = n \max_{1\leq i,j\leq n} \Big|\sum_{k=1}^{n} a_{ik}b_{kj}\Big| \;&\leq\; n \max_{1\leq i,j\leq n} \sum_{k=1}^{n} |a_{ik}b_{kj}| \\ &\leq n \max_{1\leq i,j\leq n} \sum_{k=1}^{n} \|A\|_\infty \|B\|_\infty = n\,\|A\|_\infty \cdot n\,\|B\|_\infty = N(A)\,N(B). \end{aligned}$$

Hence *Scaled ∞-norm is a matrix norm but ∞-norm is not a matrix norm.*

Induced matrix norm:

Def^n: Let $\|\cdot\|$ be a norm on $\mathbb{C}^n$. Define $|\!|\!| \cdot |\!|\!|$ on M_n by

$$|\!|\!|A|\!|\!| = \max_{\|x\|=1} \|Ax\|,$$

alternatively: $$|\!|\!|A|\!|\!| = \max_{\|x\|\leq 1} \|Ax\| = \max_{x\neq 0} \frac{\|Ax\|}{\|x\|} = \max_{\|x\|_\alpha = 1} \frac{\|Ax\|_\alpha}{\|x\|_\alpha}, \text{ in which } \|\cdot\|_\alpha \in \mathbb{R}^n.$$

The function defined above satisfies the following:

(a) $\quad \|\!|I|\!\| = 1$

(b) $\quad \|Ay\| \leq \|\!|A|\!\| \, \|y\|$ for any $A \in M_n$ and $y \in \mathbb{C}^n$

(c) $\quad \|\!|\cdot|\!\|$ is a matrix norm on M_n

(d) $\quad \|\!|A|\!\| = \max_{\|x\|=\|y\|^D=1} \left|y^H A x\right|$

Proof:

(a)

$$\|\!|I|\!\| = \max_{\|x\|=1} \|Ix\| = \max_{\|x\|=1} \|x\| = 1.$$

(b) $y = 0$, nothing to prove. If $y \neq 0$, $y/\|y\|$ is a unit vector,

$$\|\!|A|\!\| = \max_{\|x\|=1} \|Ax\| \;\geq\; \left\| A \frac{y}{\|y\|} \right\| = \|Ay\|/\|y\| \qquad \Longrightarrow \qquad \|\!|A|\!\| \, \|y\| \;\geq\; \|Ay\|.$$

(c) Positivity, non-negativity, homogeneity, sub-additivity are true (these are in book).

$$\begin{aligned} \|ABx\| &\leq \|\!|A|\!\| \cdot \|Bx\| \qquad \text{(using (b))} \\ &\leq \|\!|A|\!\| \cdot \|\!|B|\!\| \cdot \|x\|. \end{aligned}$$

$$\|\!|AB|\!\| = \max_{\|x\|=1} \|ABx\| \;\leq\; \max_{\|x\|=1} \|\!|A|\!\| \cdot \|\!|B|\!\| \cdot \|x\| = \|\!|A|\!\| \cdot \|\!|B|\!\| \qquad \text{(Sub-multiplicativity) proved.}$$

(d) **Two-variable optimization problem:**

$$\begin{aligned} \max_{\|x\|=\|y\|^D=1} \left|y^H A x\right| &= \max_{\|x\|=1} \; \max_{\|y\|^D=1} \left|y^H A x\right| \\ &= \max_{\|x\|} \|Ax\|^{DD} \\ &= \max_{\|x\|=1} \|Ax\| = \|\!|A|\!\|. \end{aligned}$$

8th class

The maximum column sum matrix norm $|||\cdot|||_1$

$$|||A|||_1 = \max_{1\le j\le n} \sum_{i=1}^{n} |a_{ij}| = \max_{1\le j\le n} \|a_j\|_1, \qquad A = [a_1, \ldots, a_n].$$

$$\begin{aligned}
\|Ax\|_1 &= \|a_1x_1 + \cdots + a_nx_n\|_1 \\
&\le \sum_{i=1}^{n} \|a_ix_i\|_1 = \sum_{i=1}^{n} |x_i|\,\|a_i\|_1 \\
&\le \sum_{i=1}^{n} |x_i| \Big(\max_{1\le k\le n} \|a_k\|_1 \Big) = \|x\|_1\,|||A|||_1 \\
\implies &\max_{\|x\|_1=1} \|Ax\|_1 \;\le\; |||A|||_1 \qquad (*)
\end{aligned}$$

If we now choose $x = e_k$, then for any $k = 1, \ldots, n$ we have

$$\max_{\|x\|_1=1} \|Ax\|_1 \;\ge\; \|1 \cdot a_k\|_1 = \|a_k\|_1$$

and hence

$$\max_{\|x\|_1=1} \|Ax\|_1 \;\ge\; \max_{1\le k\le n} \|a_k\|_1 = |||A|||_1.$$

Combined with $(*)$, $|||\cdot|||_1$ is induced by $\|\cdot\|_1$.

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

Maximum row sum matrix norm: $\|\cdot\|_\infty$ is defined on M_n by

$$|||A|||_\infty = \max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}| = \max_{\|x\|_\infty=1}\|Ax\|_\infty.$$

$$\|Ax\|_\infty \;\le\; |||A|||\,\|x\|_\infty$$

We claim that $|||\cdot|||_\infty$ is induced by the ℓ_∞-norm on $\mathbb{C}^n$ and hence is a matrix norm.

$$\begin{aligned}\|Ax\|_\infty = \max_{1\le i\le n}\Big|\sum_{j=1}^{n}a_{ij}x_j\Big| \;&\le\; \max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}x_j| \\ &\le \max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}|\,\|x\|_\infty = |||A|||_\infty\,\|x\|_\infty \qquad \big[\|x\|_\infty = \max_j|x_j|,\ |x_j|\le\|x\|_\infty\big]\end{aligned}$$

and hence $\max\limits_{\|x\|_\infty=1}\|Ax\|_\infty \le |||A|||_\infty$.

If $A=0$, nothing to prove. For $A\neq 0$, suppose that the kth row of A is non-zero and define the vector $z=[z_j]\in\mathbb{C}^n$ by

$$z_j = \frac{\overline{a_{kj}}}{|a_{kj}|} \text{ if } a_{kj}\neq 0, \qquad z_j = 1 \text{ if } a_{kj}=0. \qquad (z_i = e^{-\theta_i})$$

Then $\|z\|_\infty = 1$, $a_{kj}z_j = |a_{kj}|$ for all $j=1,2,\ldots,n$ and

$$\begin{aligned}\max_{\|x\|_\infty=1}\|Ax\|_\infty \ge \|Az\|_\infty = \max_{1\le i\le n}\Big|\sum_{j=1}^{n}a_{ij}z_j\Big| \;&\ge\; \Big|\sum_{j=1}^{n}\bar{a}_{kj}z_j\Big| \\ &= \sum_{j=1}^{n}|a_{kj}|.\end{aligned}$$

Thus

$$\max_{\|x\|_\infty=1}\|Ax\|_\infty \;\ge\; \max_{1\le k\le n}\sum_{j=1}^{n}|a_{kj}| = |||A|||_\infty.$$

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

Spectrum norm: The spectral norm $\|\!|\cdot|\!\|_2$ is defined on M_n by

$$\|\!|A|\!\|_2 = \sigma_1(A), \quad \text{the largest singular value of } A.$$

Let $A = V\Sigma W^H$ where $\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_n)$ and $\sigma_1 \geq \cdots \geq \sigma_n \geq 0$.

$$\begin{aligned}\max_{\|x\|_2=1} \|Ax\|_2 &= \max_{\|x\|_2=1} \|V\Sigma W^H x\|_2 = \max_{\|x\|_2=1} \|\Sigma W^H x\|_2 \qquad \text{(Euclidean norm, } \|Vv\|_2 = \|v\|_2) \\ &= \max_{\|Wy\|_2=1} \|\Sigma y\|_2 = \max_{\|y\|_2=1} \|\Sigma y\|_2 \;\leq\; \max_{\|y\|_2=1} \|\sigma_1 y\| = \sigma_1\end{aligned}$$

$$\implies \max_{\|x\|_2=1} \|Ax\|_2 = \sigma_1(A) \qquad \begin{bmatrix} y = e_1 = (1, 0, \ldots, 0)^T \\ \|y\|_2 = 1,\ \Sigma y = \sigma_1 e_1 \\ \|\Sigma y\|_2 = \sigma_1 \end{bmatrix} \qquad \begin{matrix} y = W^H x \\ \Rightarrow x = Wy \end{matrix}$$

Theorem: (9th class)

Suppose that $\|\!|\cdot|\!\|$ is a matrix norm on M_n and $S \in M_n$ is non-singular. Then the function

$$\|\!|A|\!\|_S = \|\!|SAS^{-1}|\!\| \qquad \text{for all } A \in M_n$$

is a matrix norm. Moreover, if $\|\cdot\|$ is induced by the norm $\|\cdot\|$ on $\mathbb{C}^n$, then the matrix norm $\|\!|\cdot|\!\|_S$ is induced by the norm $\|\cdot\|_S$ on $\mathbb{C}^n$.

Proof: The axioms (1), (1a), (2), and (3) are straightforward.

Submultiplicativity:

$$\begin{aligned}\|\!|AB|\!\|_S &= \|\!|SABS^{-1}|\!\| \\ &= \|\!|(SAS^{-1})(SBS^{-1})|\!\| \\ &\leq \|\!|SAS^{-1}|\!\| \cdot \|\!|SBS^{-1}|\!\| \\ &= \|\!|A|\!\|_S \, \|\!|B|\!\|_S.\end{aligned}$$

2nd:

$$\|x\|_S = \|Sx\| \;\Rightarrow\; \max_{\|x\|_S=1} \|Ax\|_S = \max_{\|Sx\|=1} \|SAx\| = \max_{\|y\|=1} \|SAS^{-1}y\| = \|\!|SAS^{-1}|\!\|.$$

Theorem. Let $\|\!|\cdot|\!\|$ be a matrix norm on M_n. Then

(a) $\quad |\lambda| \le \rho(A) \le \|\!|A|\!\|$

(b) $\quad$ For non-singular A : $\quad \rho(A) \;\ge\; |\lambda| \;\ge\; \dfrac{1}{\|\!|A^{-1}|\!\|}$

(c) $\quad \rho(A)$ is the greatest lower bound for all the values of matrix norm.

Proof:

$$Ax = \lambda x, \qquad X = x\,\mathbf{1}^T = [x, x, \ldots, x], \qquad AX = \lambda X.$$

$$\begin{aligned} \|\!|\lambda X|\!\| = |\lambda| \cdot \|\!|X|\!\| = \|\!|AX|\!\| \;&\le\; \|\!|A|\!\| \cdot \|\!|X|\!\| \\ \Longrightarrow \quad |\lambda| \;&\le\; \|\!|A|\!\| \quad \Longrightarrow \quad \|\!|A|\!\| \ge |\lambda| \\ \Longrightarrow \quad \|\!|A|\!\| \;&\ge\; \rho(A) \end{aligned}$$

We have $\rho(A) \ge |\lambda|$.

$$\therefore \quad |\lambda| \;\le\; \rho(A) \;\le\; \|\!|A|\!\|.$$

(b) For non-singular A:

$$\|\!|A^{-1}|\!\| \;\ge\; |\lambda^{-1}| \qquad \|\!|A^{-1}|\!\| \;\ge\; \frac{1}{|\lambda|} \quad \Longrightarrow \quad |\lambda| \ge \frac{1}{\|\!|A^{-1}|\!\|} \text{ and } \rho(A) \ge |\lambda|$$

$$\therefore \quad \rho(A) \;\ge\; |\lambda| \;\ge\; \frac{1}{\|\!|A^{-1}|\!\|}.$$

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

(c) From (a) we can say that $\rho(A)$ is the greatest lower bound for all the values of the matrix norm.

Theorem. Let $A \in M_n$ and $\varepsilon > 0$ be given. There exists a $\|\!|\cdot|\!\|$ on M_n so that

$$\rho(A) \;\leq\; |\!|\!|A|\!|\!| \;\leq\; \rho(A) + \varepsilon.$$

Proof: For unitary $U \in M_n$ and an upper triangular $\Delta \in M_n$ such that $A = U\Delta U^H$, set $D_t = \mathrm{diag}(t, t^2, t^3, \ldots, t^n)$.

$$D_t \Delta D_t^{-1} = D_t \begin{bmatrix} \lambda_1 & d_{12} & \cdots & d_{1n} \\ & \lambda_2 & \cdots & d_{2n} \\ & & \ddots & \vdots \\ & & & \lambda_n \end{bmatrix} \begin{bmatrix} t^{-1} & & & \\ & t^{-2} & & \\ & & \ddots & \\ & & & t^{-n} \end{bmatrix} = \begin{bmatrix} a_1 & t^{-1}d_{12} & \cdots & t^{-n+1}d_{1n} \\ & a_2 & t^{-1}d_{23} & \cdots \\ & 0 & & \ddots \end{bmatrix}$$

$$\|a_i'\|_1 = |a_i| + \sum_{\substack{j=1 \\ j<i}}^{n} |t^{-d} d_{ji}|$$

Suppose $t > 0$ is chosen large enough. Then

$$|\!|\!|D_t \Delta D_t^{-1}|\!|\!|_1 \;\leq\; \rho(A) + \varepsilon \qquad (\because \Delta = U^H A U)$$

$$|\!|\!|D_t U^H A U D_t^{-1}|\!|\!|_1 = |\!|\!|A|\!|\!|_{D_t U^H}$$

$$\implies \quad |\!|\!|A|\!|\!| \;\leq\; \rho(A) + \varepsilon$$

$|\!|\!|A|\!|\!| \;\geq\; \rho(A)$ for any arbitrary convergent matrix.

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

$$\|\tilde{A}^k\| \;\leq\; 1 \qquad \text{for all } k \geq N$$

where

$$\tilde{A} = \frac{A}{\rho(A)+\varepsilon} \qquad \Longrightarrow \qquad \left\|\left(\frac{A}{\rho(A)+\varepsilon}\right)^k\right\| \;\leq\; 1$$

$$\Longrightarrow \quad \frac{\|A^k\|}{(\rho(A)+\varepsilon)^k} \;\leq\; 1$$

$$\Longrightarrow \quad \|A^k\| \;\leq\; (\rho(A)+\varepsilon)^k$$

$$\Longrightarrow \quad \|A^k\|^{1/k} \;\leq\; \rho(A)+\varepsilon.$$

$A \in M_n$, $\lim_{k\to\infty} A^k = 0$; then A is **convergent**.

Theorem. Let $A \in M_n$. Then

$$\lim_{k\to\infty} A^k = 0 \qquad \text{iff} \qquad \rho(A) < 1.$$

Proof:

$$Ax = \lambda x, \qquad A^k x = \lambda^k x \to 0 \text{ only if } |\lambda| < 1 \;\; \forall\, \lambda \in \sigma(A).$$

$$A^k x = \lambda^k x \to 0 \text{ only if } \rho(A) < 1.$$

Conversely: if $\rho(A) < 1$,

$$\left.\begin{array}{l} \rho(A) < 1 \;\Rightarrow\; 1-\rho(A) > 0 \\ \Rightarrow\; 1-\rho(A) > \varepsilon \\ \Rightarrow\; \varepsilon + \rho(A) < 1 \;\Rightarrow\; \|A\| < 1 \end{array}\right\} \qquad \rho(A) \;\leq\; \|A\| \;\leq\; \rho(A)+\varepsilon$$

$$\Longrightarrow \quad \|A\| \;\geq\; \rho(A) \qquad \Longrightarrow \quad \|A\| < 1 \qquad (a)\; \|A\| \leq \rho(A)+\varepsilon$$

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

If $\|A\| < 1$, then $\|A^k\| \leq \|A\|^k \to 0$. 10th class

Corollary. Let $\|\cdot\|$ be a matrix norm on M_n and let $A \in M_n$. Then

$$\rho(A) = \lim_{k\to\infty} \|A^k\|^{1/k}.$$

Since $\rho(A)^k = \rho(A^k) \leq \|A^k\|$ we have

$$\rho(A) \;\leq\; \|A^k\|^{1/k} \qquad \text{for } \forall\, k.$$

$\tilde{A} = [\rho(A) + \varepsilon]^{-1} A$ has a spectral radius less than 1 for $\varepsilon > 0$; hence it is convergent.
Thus $\|\tilde{A}^k\| \to 0$ as $k \to \infty$, and there is some $N = N(\varepsilon, A)$ such that $\|\tilde{A}^k\| \leq 1$ for all $k \geq N \Rightarrow \|A^k\| \leq (\rho(A) + \varepsilon)^k$

$$\Longrightarrow \quad \|A^k\|^{1/k} \;\leq\; \rho(A) + \varepsilon, \quad \text{and} \quad \rho(A) \leq \|A^k\|^{1/k} \quad \Longrightarrow \quad \lim_{k\to\infty} \|A^k\|^{1/k} \text{ exists and equals } \rho(A).$$

Radius of Convergence: (RoC)

$$\sum_{k=0}^{\infty} a_k z^k \text{ has RoC } |R| \qquad \text{if} \quad \begin{cases} |z| < R \to \text{absolutely convergent} \\ |z| > R \to \text{absolutely divergent} \end{cases}$$

Let R be the RoC of a scalar power series $\sum_{k=0}^{\infty} a_k z^k$ and $A \in M_n$ be given. Then the matrix power series $\sum_{k=0}^{\infty} a_k A^k$ converges if $\rho(A) < R$.

This condition is satisfied iff there is a $\|\!|\cdot|\!\|$ on M_n s.t. $\|\!|A|\!\| < R$.

$$\left\|\sum_k a_k A^k\right\| \le \sum_k \|\!|a_k A^k|\!\| = \sum_k |a_k|\,\|\!|A^k|\!\|$$
$$\le \sum_k |a_k|\,\|\!|A|\!\|^k \qquad \Longrightarrow \qquad \sum_k a_k A^k \text{ is convergent if } \|\!|A|\!\| < R$$
$$\Longrightarrow \quad \rho(A) < R.$$

Theorem. A matrix $A \in M_n$ is non-singular if there is a matrix norm $\|\!|\cdot|\!\|$ such that $\|\!|I - A|\!\| < 1$. If this condition is satisfied,

$$A^{-1} = \sum_{k=0}^{\infty} (I - A)^k.$$

Proof: If $\|\!|I - A|\!\| < 1$, then $\sum_{k=0}^{\infty}(I-A)^k \to C$ as RoC of $\sum z^k$ is 1.

$$A\sum_{k=0}^{N}(I-A)^k = \big(I - (I-A)\big)\sum_{k=0}^{N}(I-A)^k = I - (I-A)^{N+1} \;\to\; I$$

as $N \to \infty$, we conclude that $C = A^{-1}$.

Theorem (Lévy–Desplanques theorem)**.** Let $A = [a_{ij}] \in M_n$. If $|a_{ii}| > \sum_{j\neq i} |a_{ij}|$ for all $i = 1, 2, \ldots, n$, then A is non-singular.

Proof: A is non-singular $\Leftarrow$ let $D = \text{diag}(a_{11}, \ldots, a_{nn})$. $B = I - D^{-1}A = [b_{ij}]$ has all 0s on the main diagonal and $b_{ij} = -a_{ij}/a_{ii}$ if $i \neq j$.

$$\Longrightarrow \quad \|\!|B|\!\|_\infty < 1 \quad \Longrightarrow \quad \rho(B) < 1 \quad \Longrightarrow \quad 1 \notin \sigma(B)$$
$$\Longrightarrow \quad I - B = D^{-1}A \text{ non-singular} \quad \Longrightarrow \quad A \text{ is non-singular.}$$

Equivalent matrix norm: Any two norms on $\mathbb{R}^n$ or $\mathbb{C}^n$ are equivalent:

$$|||A|||_\alpha \leq c_{\alpha\beta} |||A|||_\beta, \qquad c_{\alpha\beta} = \max_{A\neq 0} \frac{|||A|||_\alpha}{|||A|||_\beta}.$$

Adjoint: If $\|\cdot\|$ is a norm on M_n then $\|\cdot\|'$ is the adjoint norm,

$$\|A\|' = \|A^H\| \qquad \text{and} \qquad \big(\|A\|'\big)' = \|A\|.$$

- $\|A\|_2' = \|A^H\|_2 = \|A\|_2$
- $\|A\|_1' = \|A^H\|_1 = \|A\|_1$
- $|||\cdot|||_1' = |||\cdot|||_\infty$
- $\|A\| = \|A\|'$ for all $A \in M_n$ is said to be **self-adjoint**.

Theorem. Let $|||\cdot|||$ be a the matrix norm on M_n that is induced by a norm $\|\cdot\|$ on $\mathbb{C}^n$. Then

$$(a) \qquad |||\cdot|||' \text{ is induced by the norm } \|\cdot\|^D$$

$$|||A|||' = |||A^H||| = \max_{\|x\|=\|y\|^D=1} \big|y^H A^H x\big| = \max_{\|y\|^D=1} \max_{\|x\|=1} \big|x^H A y\big| = \max_{\|y\|^D=1} \|Ay\|^D$$

Induced norm: $|||A^H||| = \max_{\|x\|=1} \|A^H x\|$

Dual norm: $\|z\| = \max_{\|y\|^D=1} |y^H z| \;\Rightarrow\; \|A^H x\| = \max_{\|y\|^D=1} |y^H A^H x| \;\Rightarrow\; |||A^H|||$

Definition of Dual matrix norm: Let $\|\cdot\|$ be a norm on M_n. Its dual norm

$$\|A\|^D = \max_{\|B\|=1} \big|\mathrm{tr}(B^H A)\big|.$$

Condition number of A: Let $\|\cdot\|$ be a given matrix norm and suppose that $A \in M_n$ is non-singular. We wish to compute the inverse of A, but instead, we need to compute

$$B = A + \Delta A,$$

in which we assume that $\|A^{-1}\Delta A\| < 1$ to ensure that B is non-singular.
Since $B = A(I + A^{-1}\Delta A)$ and $\rho(A^{-1}\Delta A) \leq \|A^{-1}\Delta A\| < 1$,

$$\|A^{-1}\Delta A\| < 1 \;\Rightarrow\; -1 \notin \sigma(A^{-1}\Delta A) \;\Rightarrow\; B \text{ is non-singular.}$$

We have:

$$A^{-1}(\Delta A)B^{-1} = A^{-1}(B - A)B^{-1} = A^{-1} - B^{-1}$$
$$\|A^{-1} - B^{-1}\| = \|A^{-1}\Delta A B^{-1}\| \;\leq\; \|A^{-1}\Delta A\| \cdot \|B^{-1}\|$$

Since $B^{-1} = A^{-1} - A^{-1}(\Delta A)B^{-1}$

$$\|B^{-1}\| \;\leq\; \|A^{-1}\| + \|A^{-1}\Delta A B^{-1}\| \;\leq\; \|A^{-1}\| + \|A^{-1}\Delta A\|\,\|B^{-1}\|$$

$$\|B^{-1}\| = \|(A + \Delta A)^{-1}\| \;\leq\; \frac{\|A^{-1}\|}{1 - \|A^{-1}\Delta A\|}$$

$$\Longrightarrow \quad \|A^{-1} - B^{-1}\| \;\leq\; \frac{\|A^{-1}\| \cdot \|A^{-1}\Delta A\|}{1 - \|A^{-1}\Delta A\|} \;\leq\; \frac{\|A^{-1}\| \cdot \|A^{-1}\| \cdot \|\Delta A\|}{1 - \|A^{-1}\Delta A\|}$$

Relative error:

$$\Longrightarrow \quad \frac{\|A^{-1} - (A + \Delta A)^{-1}\|}{\|A^{-1}\|} \;\leq\; \frac{\|A^{-1}\|\,\|A\|\,\|\Delta A\|}{1 - \|A^{-1}\Delta A\|\,\|A\|}$$

The quantity

$$\kappa(A) = \begin{cases} \|A^{-1}\|\,\|A\| & \text{if } A \text{ is non-singular} \\ \infty & \text{if } A \text{ is singular} \end{cases}$$

is the **condition number** for matrix inversion with respect to the matrix norm $\|\cdot\|$.

$$\kappa(A) = \|A^{-1}\|\,\|A\| \;\geq\; \|A^{-1}A\| = \|I\| \;\geq\; 1 \qquad \text{for any matrix norm.}$$

Bound:

$$\frac{\|A^{-1}-(A+\Delta A)^{-1}\|}{\|A^{-1}\|} \leq \frac{\kappa(A)\,\|\Delta A\|}{1-\|A^{-1}\Delta A\|\,\|A\|} \leq \frac{\kappa(A)\,\|\Delta A\|}{\|A\|}\Big/\left(1-\frac{\kappa(A)\,\|\Delta A\|}{\|A\|}\right), \qquad \text{where } \|A^{-1}\Delta A\|<1.$$

Note:

1. A is ill-conditioned or poorly conditioned if $\kappa(A)$ is large, if $\kappa(A)$ is small (near 1), A is well conditioned. If $\kappa(A)=1$, A is perfectly conditioned.
2. Suppose we use spectral norm:
$$\kappa(A)=\frac{\sigma_1(A)}{\sigma_n(A)} \rightarrow \frac{\text{largest singular value}}{\text{smallest singular value}}.$$
3. $\kappa(AB) \leq \kappa(A)\cdot\kappa(B)$.
4. For unitary U and V
$$\begin{aligned}\kappa(UAV) &= \|VA^{-1}\bar{V}^{-1}\|\,\|UAV\| \\ &= \|A^{-1}\|\,\|A\| \qquad * \text{ due to isometry of unitary matrices, } \|UAV\|=\|A\| \\ &= \kappa(A)\end{aligned}$$

A priori bound: Suppose we want to solve a linear system

$$Ax=b, \qquad A\in M_n,\ |A|\neq 0 \text{ and } b\neq 0\in\mathbb{C}^n.$$

Because of computational error, we solve a perturbed system

$$(A+\Delta A)\tilde{x}=b+\Delta b, \qquad A,\Delta A\in M_n,\ b,\Delta b\in\mathbb{C}^n,\ \tilde{x}=x+\Delta x.$$

$$\begin{aligned}(A+\Delta A)(x+\Delta x) &= Ax+(\Delta A)x+(A+\Delta A)\Delta x \\ &= b+(\Delta A)x+(A+\Delta A)\Delta x = b+\Delta b \\ \implies \quad (\Delta A)x+(A+\Delta A)\Delta x &= \Delta b.\end{aligned}$$

$$\Delta x = (A + \Delta A)^{-1}\big(\Delta b - (\Delta A)x\big)$$

$$\begin{aligned}
\|\Delta x\| &= \|(A + \Delta A)^{-1}(\Delta b - \Delta A x)\| \\
&\leq \|(A + \Delta A)^{-1}\|\big(\|\Delta b\| + \|(\Delta A)x\|\big) \\
&= \frac{\|A^{-1}\|}{1 - \|A^{-1}\Delta A\|}\big(\|\Delta b\| + \|\Delta A\|\,\|x\|\big) \\
\implies \quad \frac{\|\Delta x\|}{\|x\|} &\leq \frac{\|A^{-1}\|\,\|A\|}{1 - \|A^{-1}\Delta A\|}\left(\frac{\|\Delta b\|}{\|A\|\,\|x\|} + \frac{\|\Delta A\|}{\|A\|}\right) \\
&= \frac{\kappa(A)}{1 - \|A^{-1}\Delta A\|}\left(\frac{\|\Delta b\|}{\|b\|} + \frac{\|\Delta A\|}{\|A\|}\right) \\
&\leq \frac{\kappa(A)}{1 - \kappa(A)\dfrac{\|\Delta A\|}{\|A\|}}\left(\frac{\|\Delta b\|}{\|b\|} + \frac{\|\Delta A\|}{\|A\|}\right).
\end{aligned}$$

Extra notes: