Lecture — L4.1

1.1 Matrix and Vector Norms

Let $x \in \mathbb{R}^n$, i.e., $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, with $x_i \in \mathbb{R}$. A *vector norm on* $\mathbb{R}^n$ *is a function* $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ *satisfying the following properties:*

$$(1)\ \|x\| \geq 0 \ \ \textit{for all } x \in \mathbb{R}^n,$$
$$(2)\ \|x\| = 0 \iff x = 0,$$
$$(3)\ \|\alpha x\| = |\alpha|\,\|x\| \ \ \textit{for all } \alpha \in \mathbb{R} \textit{ and for all } x \in \mathbb{R}^n,$$
$$(4)\ \|x + y\| \leq \|x\| + \|y\| \ \ \textit{for all } x, y \in \mathbb{R}^n.$$

There are many types of norms on $\mathbb{R}^n$*, e.g.,*

$$(1)\ \|x\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2},$$
$$(2)\ \|x\|_\infty = \max_{1 \leq i \leq n} |x_i|.$$

Both the above norms are equivalent.

1.2 Cauchy-Schwarz Inequality

$|\langle x, y\rangle| \equiv |x^T y| = |\sum_{1 \leq i \leq n} x_i y_i| \leq \|x\|_2 \, \|y\|_2.$

What is the use of norms?

(1) *To measure distance between 2 points in space,*

(2) *convergence of sequences, e.g., analysis/of error of iterative methods.*

A sequence $\{x_{(k)}\}_{k=1}^{\infty}$ *of vectors in* $\mathbb{R}^n$ *is said to converge to* x *with respect to the norm* $\|\cdot\|$ *if for any small* $\epsilon > 0$, $\exists\, N(\epsilon)$ *such that* $\left\|x_{(k)} - x\right\| < \epsilon \quad \forall \quad k \geq N(\epsilon)$.

Theorem 1. $\{x_{(k)}\} \to x$ *in* $\mathbb{R}^n$ *w.r.t.* $\|\cdot\| \iff \lim_{k\to\infty} x_{i(k)} = x_i$ *for each* $i = 1, \ldots, n$.

Theorem 2. *For each* $x \in \mathbb{R}^n$, $\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\,\|x\|_\infty$.

1.3 Matrix Norms

Let $A \in \mathbf{M}_{n\times n}(\mathbb{R})$; $\|\cdot\|$ *is a function that maps* A *in* $\mathbf{M}_{n\times n}(\mathbb{R})$ *to a real number.*

Properties of matrix norms

(1) $\|A\| \geq 0$,

(2) $\|A\| = 0 \iff A$ *is a '0' matrix with all entries 0* ,

(3) $\|\alpha A\| = |\alpha|\,\|A\|$ *for all* $\alpha \in \mathbb{R}$ *and for all* $A \in \mathbf{M}_{n\times n}(\mathbb{R})$,

(4) $\|A + B\| \leq \|A\| + \|B\|$ *for all* A, B *in* $\mathbf{M}_{n\times n}(\mathbb{R})$,

(5) $\|AB\| \leq \|A\|\,\|B\|$ *for all* A, B *in* $\mathbf{M}_{n\times n}(\mathbb{R})$.

Theorem 3 (Induced Matrix Norm). *If* $\|\cdot\|$ *is a vector norm on* $\mathbb{R}^n$, *then* $\|A\| = \max_{\|x\|=1} \|Ax\|$ *is a matrix norm.*

Note that, alternatively, $\|A\| = \max_{z\neq 0} \left\|A(\frac{z}{\|z\|})\right\| = \max_{z\neq 0} \frac{\|A(z)\|}{\|z\|}$.

Theorem 4. *For* $A = (a_{ij}) \in \mathbf{M}_{n\times n}(\mathbb{R})$, *we denote* $\|A\|_\infty = \max_{1\leq i\leq n} \sum_{1\leq j\leq n} |a_{ij}|$. *Then* $\|A\|_\infty$ *is a matrix norm.*

Example 5. Let $A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 3 & -1 \\ 5 & -1 & 1 \end{pmatrix}$. Then $\sum_{1\le j\le 3} |a_{1j}| = |1| + |2| + |-1| = 4$, $\sum_{1\le j\le 3} |a_{2j}| = |0| + |3| + |-1| = 4$, $\sum_{1\le j\le 3} |a_{3j}| = |5| + |-1| + |1| = 7$. Hence, $\|A\|_\infty = \max\{4, 4, 7\} = 7$.

1.4 Spectral Radius of a Matrix

Let $A \in \mathbf{M}_{n\times n}(\mathbb{R})$. Define $\rho(A) := \max_{1\le i\le n} |\lambda_i|$; where $\lambda_1, \dots, \lambda_n$ are eigenvalues in $\mathbb{C}$. The following theorem shows how spectral radius is closely related to the norm of the matrix.

Theorem 6. *Let $A \in \mathbf{M}_{n\times n}(\mathbb{R})$. Then*

$$(1)\ \|A\|_2 = \sqrt{\rho(A^T A)},$$

$$(2) \rho(A) \le \|A\|,\ \text{where } \|\cdot\| \text{ is an induced matrix norm.}$$

1.5 A Simple Matrix Decomposition

Consider any matrix, say $A = \begin{pmatrix} 2 & -1 & 5 \\ 0 & 1 & -2 \\ 1 & 5 & -1 \end{pmatrix}$

.

Then we can always write

$$A = \begin{pmatrix} 2 & -1 & 5 \\ 0 & 1 & -2 \\ 1 & 5 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 5 & 0 \end{pmatrix} + \begin{pmatrix} 0 & -1 & 5 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix} = D+L+U = D-(-L)-(-U),$$

where D, L and U respectively denotes the diagonal part, strictly lower triangular part and strictly upper triangular part of the matrix A.

1.6 Iterative Scheme to Solve Systems of Linear Equations

The idea is to solve $Ax = b$ by rewriting it in the form $x = Tx + c$ and then using an iteration scheme of the form $x_{(k)} = Tx_{(k-1)} + c$, where $k = 1, 2, 3, \ldots$.

Jaccobi Iterative Method

Example 7. Let us consider the system of linear equations:

$$\begin{aligned} E_1 :& 10x_1 - x_2 + 2x_3 = 6 \\ E_2 :& - x_1 + 11x_2 - x_3 + 3x_4 = 25 \\ E_3 :& 2x_1 - x_2 + 10x_3 - x_4 = -11 \\ E_4 :& 3x_2 - x_3 + 8x_4 = 15 \end{aligned}$$

This system of equations has a unique solution $x = \begin{pmatrix} 1 \\ 2 \\ -1 \\ 1 \end{pmatrix}$.

Consider the above system of equations in the form $Ax = b$, where $A = \begin{pmatrix} 10 & -1 & 2 & 0 \\ -1 & 11 & -1 & 3 \\ 2 & -1 & 10 & -1 \\ 0 & 3 & -1 & 8 \end{pmatrix}$ and $b = \begin{pmatrix} 6 \\ 25 \\ -11 \\ 15 \end{pmatrix}$.

Let us re-write $Ax = b$ as $x = Tx + c$ in the following way:

$$
\begin{aligned}
x_1 &= \phantom{-\frac{1}{5}x_1 +} \frac{1}{10}x_2 - \frac{1}{5}x_3 \phantom{+ \frac{1}{10}x_4} + \frac{3}{5} \\
x_2 &= \frac{1}{11}x_1 \phantom{+ \frac{1}{10}x_2} + \frac{1}{11}x_3 - \frac{3}{11}x_4 + \frac{25}{11} \\
x_3 &= -\frac{1}{5}x_1 + \frac{1}{10}x_2 \phantom{+ \frac{1}{11}x_3} + \frac{1}{10}x_4 - \frac{11}{10} \\
x_4 &= \phantom{-\frac{1}{5}x_1} -\frac{3}{8}x_2 + \frac{1}{8}x_3 \phantom{+ \frac{1}{10}x_4} + \frac{15}{8}
\end{aligned}
$$

In matrix form the above system of equations can be expressed as

$$
x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{10} & -\frac{1}{5} & 0 \\ \frac{1}{11} & 0 & \frac{1}{11} & -\frac{3}{11} \\ -\frac{1}{5} & \frac{1}{10} & 0 & \frac{1}{10} \\ 0 & -\frac{3}{8} & \frac{1}{8} & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} + \begin{pmatrix} \frac{3}{5} \\ \frac{25}{11} \\ -\frac{11}{10} \\ \frac{15}{8} \end{pmatrix},
$$

which is of the form $x = Tx + c$. We consider the initial guess(say) $x_{(0)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$.

Then $x_{(1)} = \begin{pmatrix} 0.6 \\ 2.2727 \\ -1.1000 \\ 1.8750 \end{pmatrix}$.

Carry forth the computations iteratively we obtain $x_{(10)} = \begin{pmatrix} 1.0001 \\ 1.9998 \\ -0.9998 \\ 0.9998 \end{pmatrix}$,

whence $\frac{\|x_{(10)} - x_{(9)}\|}{\|x_{(10)}\|} < 10^{-3}$, STOP!

Jacobi Iteration in Matrix Form

$$
\begin{aligned}
& Ax = b \\
\implies & (D + L + U)x = b \\
\implies & Dx = -(L + U)x + b \\
\implies & x = -D^{-1}(L + U)x + D^{-1}b
\end{aligned}
$$

Hence the iteration

$$x_{(k)} = -D^{-1}(L + U)x_{(k-1)} + b$$

in the form of iterates we have

$$x_{i_{(k)}} = \frac{\sum_{1 \le j \le n, j \ne i} (-a_{ij} x_{j_{(k-1)}}) + b_i}{a_{ii}};$$

where $i = 1, 2, \ldots, n$ *and* $a_{ii} \neq 0 \implies D$ *is invertible.*

Gauss-Seidel Iterative Scheme

$$x_{i_{(k)}} = \frac{-\sum_{j=1}^{i-1} a_{ij} x_{j_{(k)}} - \sum_{j=i+1}^{n} a_{ij} x_{j_{(k-1)}} + b_i}{a_{ii}};$$

where $i = 1, 2, \ldots, n$.

Why above is a good idea?
Compare w.r.t. the example at the beginning of the lecture.

In matrix form

$$
\begin{aligned}
& Ax = b \\
\implies & (D + L + U)x = b \\
\implies & (D + L)x = -Ux + b \\
\implies & x = -D^{-1}(L + U)x + D^{-1}b
\end{aligned}
$$

iterates as

$$(D + L)x_{(k)} = -Ux_{(k-1)} + b,$$

or

$$x_{(k)} = -(D + L)^{-1}Ux_{(k-1)} + (D + L)^{-1}b; \; k = 1, 2, \ldots.$$

The above iteration is known to be Gauss-Seidel iteration method.

In next lecture, we will talk about convergence of iterative schemes and also SOR scheme.