7.1 Norms of Vectors and Matrices

Column vector: $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, or $\boldsymbol{x} = [x_1, x_2, \dots, x_n]^t$.

Motivation: Consider to solve the linear system

$$3.3330x_1 + 15920x_2 - 10.333x_3 = 15913$$
$$2.2220x_1 + 16.710x_2 + 9.6120x_3 = 28.544$$
$$1.5611x_1 + 5.1791x_2 + 1.6852x_n = 8.4254$$

by Gaussian elimination with 5-digit rounding arithmetic and partial pivoting. The system has exact solution $\boldsymbol{x} = [x_1, x_2, x_3]^t = [1,1,1]^t$**.** The approximate solution is $\tilde{\boldsymbol{x}} = [1.2001, 0.99991, 0.92538]^t$. How to quantify the approximation error?

Definition. A **vector norm** on R^n is a function, $||\cdot||$, from R^n to R with the properties:

(i) $||\boldsymbol{x}|| \geq 0$ for all $\boldsymbol{x} \in R^n$

(ii) $||\boldsymbol{x}|| = 0$ if and only if $\boldsymbol{x} = \boldsymbol{0}$

(iii) $||\alpha\boldsymbol{x}|| = |\alpha|\,||\boldsymbol{x}||$ for all $\alpha \in R$ and $\boldsymbol{x} \in R^n$

(iv) $||\boldsymbol{x} + \boldsymbol{y}|| \leq ||\boldsymbol{x}|| + ||\boldsymbol{y}||$ for all $\boldsymbol{x}, \boldsymbol{y} \in R^n$

Definition. The **Euclidean norm** l_2 and the **infinity norm** l_∞ for the vector $\boldsymbol{x} = [x_1, x_2, \dots, x_n]^t$ are defined by

$$||\boldsymbol{x}||_2 = \{\sum_{i=1}^{n} x_i^2\}^{1/2}$$

and

$$||\boldsymbol{x}||_\infty = \max_{1 \leq i \leq n} |x_i|$$

Example. Compute l_2 norm and l_∞ norm of the vector $\boldsymbol{x} = [-1, 1, -2]^t$.

Theorem 7.3 Cauchy-Schwarz Inequality for Sums. For each $\boldsymbol{x} = [x_1, x_2, \dots, x_n]^t$ and $\boldsymbol{y} = [y_1, y_2, \dots, y_n]^t$ in R^n,

$$\boldsymbol{x}^t \boldsymbol{y} = \sum_{i=1}^{n} x_i y_i \leq \left\{\sum_{i=1}^{n} {x_i}^2\right\}^{\frac{1}{2}} \left\{\sum_{i=1}^{n} {y_i}^2\right\}^{\frac{1}{2}} = \left|\left|\boldsymbol{x}\right|\right|_2 \cdot \left|\left|\boldsymbol{y}\right|\right|_2$$

Remark: $||\boldsymbol{x} + \boldsymbol{y}||_2 \leq \left|\left|\boldsymbol{x}\right|\right|_2 + \left|\left|\boldsymbol{y}\right|\right|_2$

Definition. The **distance** between two vectors $\boldsymbol{x} = [x_1, x_2, \dots, x_n]^t$ and $\boldsymbol{y} = [y_1, y_2, \dots, y_n]^t$ is the norm of the difference of the vectors. The l_2 and l_∞ distances are:

$$||\boldsymbol{x} - \boldsymbol{y}||_2 = \{\sum_{i=1}^{n} (x_i - y_i)^2\}^{1/2}$$
$$||\boldsymbol{x} - \boldsymbol{y}||_\infty = \max_{1 \leq i \leq n} |x_i - y_i|$$

Example.

$$3.3330x_1 + 15920x_2 - 10.333x_3 = 15913$$
$$2.2220x_1 + 16.710x_2 + 9.6120x_3 = 28.544$$
$$1.5611x_1 + 5.1791x_2 + 1.6852x_n = 8.4254$$

has exact solution $\boldsymbol{x} = [x_1, x_2, x_3]^t = [1,1,1]^t$**.** The Gaussian elimination with 5-digit rounding arithmetic and partial pivoting produces approximate solution $\widetilde{\boldsymbol{x}} = [1.2001, 0.99991, 0.92538]^t$. Determine l_2 and l_∞ distances between exact and approximate solutions.

Solution:

$$||\boldsymbol{x} - \widetilde{\boldsymbol{x}}||_\infty = \max\{|1 - 1.2001|, |1 - 0.99991|, |1 - 0.92538|\} = \max\{0.2001, 0.00009, 0.07462\} = 0.2001$$

$$\left|\left|\boldsymbol{x} - \widetilde{\boldsymbol{x}}\right|\right|_2 = \sqrt{(1 - 1.2001)^2 + (1 - 0.99991)^2 + (1 - 0.92538)^2} = 0.21356$$

Definition. A sequence $\{\boldsymbol{x}^{(k)}\}_{k=1}^{\infty}$ of vectors in R^n is said to **converge** to $\boldsymbol{x}$ with respect to the norm $||\cdot||$ if, given any $\varepsilon > 0$, there exists an integer $N(\varepsilon)$ such that

$$\left|\left|\boldsymbol{x}^{(k)} - \boldsymbol{x}\right|\right| < \varepsilon, \qquad \text{for all } k \geq N(\varepsilon).$$

Theorem 7.6. The sequence of vectors $\{\boldsymbol{x}^{(k)}\}_{k=1}^{\infty}$ converges to $\boldsymbol{x}$ in R^n with respect to the norm $||\cdot||_\infty$ if and only if

$$\lim_{k\to\infty} x_i^{(k)} = x_i.$$

Example. Show that $\boldsymbol{x}^{(k)} = (x_1^{(k)}, x_2^{(k)}, x_3^{(k)}, x_4^{(k)})^t = (1, 2 + \frac{1}{k}, \frac{3}{k^2}, e^{-k}\sin(k))^t$ converges to $\boldsymbol{x} = (1,2,0,0)^t$.

Solution:

$$\lim_{k\to\infty} x_1^{(k)} = \lim_{k\to\infty} 1 = 1$$
$$\lim_{k\to\infty} x_2^{(k)} = \lim_{k\to\infty} 2 + \frac{1}{k} = 2$$
$$\lim_{k\to\infty} x_3^{(k)} = \lim_{k\to\infty} \frac{3}{k^2} = 0$$
$$\lim_{k\to\infty} x_4^{(k)} = \lim_{k\to\infty} e^{-k}\sin(k) = 0$$

By **Theorem 7.6**, the sequence $\{\boldsymbol{x}^{(k)}\}$ converges to $(1,2,0,0)^t$.

Theorem 7.7. For each $\boldsymbol{x} \in R^n$, $||\boldsymbol{x}||_\infty \leq ||\boldsymbol{x}||_2 \leq \sqrt{n}||\boldsymbol{x}||_\infty$

Remark: All norms in R^n are equivalent with respect to convergence, that is, if $||\cdot||$ and $||\cdot||'$ are any two norms on R^n and $\{\boldsymbol{x}^{(k)}\}_{k=1}^{\infty}$ converges to $\boldsymbol{x}$ in $||\cdot||$, then $\{\boldsymbol{x}^{(k)}\}_{k=1}^{\infty}$ converges to $\boldsymbol{x}$ in $||\cdot||'$.

Definition. A **matrix norm** $||\cdot||$ on $n \times n$ matrices is a real-valued function satisfying

(i) $||A|| \geq 0$

(ii) $||A|| = 0$ if and only if $A = 0$

(iii) $||\alpha A|| = |\alpha|\,||A||$

(iv) $||A + B|| \leq ||A|| + ||B||$

(v) $||AB|| \leq ||A||\,||B||$

The **distance** between $n \times n$ matrices A and B with respect to a matrix norm is $||A - B||$.

Theorem 7.9. If $||\cdot||$ is a vector norm, the **induced** (or **natural**) **matrix norm** is given by

$$||A|| = \max_{||\boldsymbol{x}||=1} ||A\boldsymbol{x}||$$

Example. $||A||_\infty = \max_{||\boldsymbol{x}||_\infty=1} ||A\boldsymbol{x}||_\infty$, the l_∞ induced norm.
$||A||_2 = \max_{||\boldsymbol{x}||_2=1} ||A\boldsymbol{x}||_2$, the l_2 induced norm.

Alternative definition: For any vector $\boldsymbol{z} \neq \boldsymbol{0}$, the vector $\boldsymbol{x} = \boldsymbol{z}/||\boldsymbol{z}||$ has $||\boldsymbol{x}|| = 1$.
Since $\max_{||\boldsymbol{x}||=1} ||A\boldsymbol{x}|| = \max_{\boldsymbol{z}\neq\boldsymbol{0}} ||A(\frac{\boldsymbol{z}}{||\boldsymbol{z}||})|| = \max_{\boldsymbol{z}\neq\boldsymbol{0}} \frac{||A\boldsymbol{z}||}{||\boldsymbol{z}||}$,
we can alternatively define $||A|| = \max_{\boldsymbol{z}\neq\boldsymbol{0}} \frac{||A\boldsymbol{z}||}{||\boldsymbol{z}||}$.

Corollary 7.10. For any vector $\boldsymbol{z} \neq \boldsymbol{0}$, matrix A and induced matrix norm $||\cdot||$,

$$||A\boldsymbol{z}|| \leq ||A|| \cdot ||\boldsymbol{z}||$$

Theorem 7.11. If $A = [a_{ij}]$ is an $n \times n$ matrix, then

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

Example. Determine $||A||_\infty$ for the matrix $A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & -1 \\ 5 & -1 & 1 \end{bmatrix}$

7.2 Eigenvalues and Eigenvectors

Definition. If A is an $n \times n$ matrix, the **characteristic polynomial** of A is

$$p(\lambda) = \det(A - \lambda I).$$

Definition. If $p(\lambda)$ is the characteristic polynomial of the matrix A, the zeros of $p(\lambda)$ are **eigenvalues** of the matrix A. If λ is an eigenvalue of A and $\boldsymbol{x} \neq \boldsymbol{0}$ satisfies $(A - \lambda I)\boldsymbol{x} = \boldsymbol{0}$, then $\boldsymbol{x}$ is an **eigenvector** corresponding to λ.

Geometric interpretation of **eigenvector $\boldsymbol{x}$** corresponding to λ.

Example. Find eigenvalues and eigenvectors of the matrix $A = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 1 & 2 \\ 1 & -1 & 4 \end{bmatrix}$.

Definition. The **spectral radius** $\rho(A)$ of a matrix A is defined by

$$\rho(A) = \max|\lambda|, \qquad \text{where } \lambda \text{ is an eigenvalue of } A.$$

Remark: For complex $\lambda = a + bj$, we define $|\lambda| = \sqrt{a^2 + b^2}$.

Theorem 7.15. If A is an $n \times n$ matrix, then

(i) $||A||_2 = [\rho(A^t A)]^{1/2}$

(ii) $\rho(A) \leq ||A||$, for any induced matrix norm $||\cdot||$.

Example. Determine l_2 induced norm of $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix}$

Solution

$$A^t A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 2 & -1 \\ 2 & 6 & 4 \\ -1 & 4 & 5 \end{bmatrix}$$

$$\text{Solve } \det(A^t A - \lambda I) = 0$$

$$0 = -\lambda(\lambda^2 - 14\lambda + 42)$$

$$\text{Then } \lambda = 0, \lambda = 7 \pm \sqrt{7}$$

$$||A||_2 = [\rho(A^t A)]^{1/2} = \sqrt{\max(0, 7 + \sqrt{7}, 7 - \sqrt{7})} = \sqrt{7 + \sqrt{7}}$$

Convergent Matrices

Definition. An $n \times n$ matrix A is **convergent** if $\lim_{k\to\infty}(A^k)_{ij} = 0$ for each $i = 1,2,,\ldots n$ and $j = 1,2,,\ldots n$.

Example. Show that $A = \begin{bmatrix} 1/2 & 0 \\ 1/4 & 1/2 \end{bmatrix}$ is a convergent matrix.

Theorem 7.17 The following statements are equivalent.

(i) A is **convergent** matrix.
(ii) $\lim_{n\to\infty} ||A^n|| = 0$ for some natural norm.
(iii) $\lim_{n\to\infty} ||A^n|| = 0$ for all natural norm.
(iv) $\rho(A) < 1$
(v) $\lim_{n\to\infty} A^n \boldsymbol{x} = \boldsymbol{0}$ for every $\boldsymbol{x}$.