Vector Norms

Definition: A vector norm on R^n is a function

$$\tau \; : \; R^n \; \rightarrow \; R^+ = \{x \geq 0 | \; x \in R\}$$

that satisfies

(1) $\tau(\mathbf{x}) > 0 \;\; \forall \; \mathbf{x} \neq \mathbf{0}, \; \tau(\mathbf{0}) = 0$

(2) $\tau(c\mathbf{x}) = |c|\tau(\mathbf{x}) \; \forall \; c \in R, \; \mathbf{x} \in R^n$

(3) $\tau(\mathbf{x} + \mathbf{y}) \leq \tau(\mathbf{x}) + \tau(\mathbf{y}) \; \forall \; \mathbf{x}, \mathbf{y} \in R^n$

Hölder norm (p-norm) $\|\mathbf{x}\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p} \;\; for \; p \geq 1.$

(p=1) $\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|$ (Mahattan or City-block distance)

(p=2) $\|\mathbf{x}\|_2 = (\sum_{i=1}^n |x_i|^2)^{1/2}$ (Euclidean distance)

(p=∞) $\|\mathbf{x}\|_\infty = max_{1 \leq i \leq n}\{|x_i|\}$ (∞-norm)

Matrix Norms

Definition: A matrix norm on $R^{m\times n}$ is a function

$$\tau \ : \ R^{m\times n} \ \rightarrow \ R^{+} = \{x \geq 0 |\ x \in R\}$$

that satisfies

(1) $\tau(A) > 0 \ \ \forall\ A \neq O,\ \tau(O) = 0$

(2) $\tau(cA) = |c|\tau(A)\ \forall\ c \in R,\ \ A \in R^{m\times n}$

(3) $\tau(A+B) \leq \tau(A) + \tau(B) \ \ \forall\ A, B \in R^{m\times n}$

Consistency Property: $\tau(AB) \leq \tau(A)\tau(B) \ \ \forall\ A, B$

(a) $\tau(A) = max\{|a_{ij}| \mid 1 \leq i \leq m,\ 1 \leq j \leq n\}$

(b) $\|A\|_F = \left[\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2\right]^{1/2}$ (Fröbenius norm)

Subordinate Matrix Norm: $\|A\| = max_{\|\mathbf{x}\|\neq\mathbf{0}}\{\|A\mathbf{x}\|/\|\mathbf{x}\|\}$

(1) If $A \in R^{m\times n}$, then $\|A\|_1 = max_{1\leq j\leq n}\left(\sum_{i=1}^{m}|a_{ij}|\right)$

(2) If $A \in R^{m\times n}$, then $\|A\|_\infty = max_{1\leq i\leq m}\left(\sum_{j=1}^{n}|a_{ij}|\right)$

(3) Let $A \in R^{n\times n}$ be real symmetric, then $\|A\|_2 = max_{1\leq i\leq n}|\lambda_i|$, where $\lambda_i \in \lambda(A)$

Theorem: Let $A = [a_{ij}] \in R^{m\times n}$, and define $\|A\|_1 = Max_{\|\mathbf{u}\|_1=1}\{\|A\mathbf{u}\|_1\}$. Show that

$$\|A\|_1 = Max_{1\le j\le n}\left\{\sum_{i=1}^{m}|a_{ij}|\right\}$$

(**Proof**) Let $\sum_{i=1}^{m}|a_{iK}| = Max_{1\le j\le n}\{\sum_{i=1}^{m}|a_{ij}|\}$, for any $\mathbf{x}\in R^n$ with $\|\mathbf{x}\|_1 = 1$, we have

$$\begin{aligned}
\|A\mathbf{x}\|_1 &= \textstyle\sum_{i=1}^{m}|\sum_{j=1}^{n}a_{ij}x_j| \\
&\le \textstyle\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}x_j| \\
&= \textstyle\sum_{j=1}^{n}\sum_{i=1}^{m}|a_{ij}|\cdot|x_j| \\
&= \textstyle\sum_{j=1}^{n}|x_j|\{\sum_{i=1}^{m}|a_{ij}|\} \\
&\le \textstyle\sum_{j=1}^{n}|x_j|\{\sum_{i=1}^{m}|a_{iK}|\} \\
&= \textstyle\{\sum_{j=1}^{n}|x_j|\}\{\sum_{i=1}^{m}|a_{iK}|\} \\
&= \textstyle\|\mathbf{x}\|_1\{\sum_{i=1}^{m}|a_{iK}|\} \\
&= \textstyle\sum_{i=1}^{m}|a_{iK}|
\end{aligned}$$

Thus,

$$Max_{\|\mathbf{u}\|_1=1}\{\|A\mathbf{u}\|_1\} \le Max_{1\le j\le n}\left\{\sum_{i=1}^{m}|a_{ij}|\right\} = \sum_{i=1}^{m}|a_{iK}| \quad for\ a\ K\ , \quad 1\le K\le m.$$

In particular, when $\mathbf{x}\in R^n$ is selected as $\mathbf{x} = \mathbf{e}_K$, that is, $x_K = 1$, and $x_i = 0\ \forall\ 1\le i\le n,\ i\ne K$, then the above equality holds, which completes the proof.

Theorem: Let $A = [a_{ij}] \in R^{m\times n}$, and define $\|A\|_\infty = Max_{\|\mathbf{u}\|_\infty=1}\{\|A\mathbf{u}\|_\infty\}$. Show that

$$\|A\|_\infty = Max_{1\le i\le m}\left\{\sum_{j=1}^{n}|a_{ij}|\right\}$$

(**Proof**) Let $\sum_{j=1}^{n}|a_{Kj}| = Max_{1\le i\le m}\left\{\sum_{j=1}^{n}|a_{ij}|\right\}$, for any $\mathbf{x} \in R^n$ with $\|\mathbf{x}\|_\infty = 1$, we have

$$\begin{aligned}
\|A\mathbf{x}\|_\infty &= Max_{1\le i\le m}\left\{|\textstyle\sum_{j=1}^{n} a_{ij}x_j|\right\} \\
&\le Max_{1\le i\le m}\left\{\textstyle\sum_{j=1}^{n} |a_{ij}|\cdot|x_j|\right\} \le Max_{1\le i\le m}\left\{\textstyle\sum_{j=1}^{n} |a_{ij}|\|\mathbf{x}\|_\infty\right\} \\
&\le Max_{1\le i\le m}\left\{\textstyle\sum_{j=1}^{n} |a_{ij}|\right\} = \textstyle\sum_{j=1}^{n} |a_{Kj}|
\end{aligned}$$

Thus, $\|A\|_\infty = Max_{\|\mathbf{u}\|_\infty=1}\{\|A\mathbf{u}\|_\infty\} \le Max_{1\le i\le m}\left\{\sum_{j=1}^{n}|a_{ij}|\right\}$.

In particular, if we pick up $\mathbf{y} \in R^n$ such that $y_j = sign(a_{Kj})$, $\forall\ 1 \le j \le n$, then $\|\mathbf{y}\|_\infty = 1$, and $\|A\mathbf{y}\|_\infty = \sum_{j=1}^{n} |a_{Kj}|$, which completes the proof.

Theorem: Let $A = [a_{ij}] \in R^{n\times n}$, and define $\|A\|_2 = Max_{\|\mathbf{x}\|_2=1}\{\|A\mathbf{x}\|_2\}$. Show that

$$\|A\|_2 = \sqrt{\rho(A^tA)} = \sqrt{maximum\ eigenvalue\ of\ A^tA}\ \ (spectral\ radius)$$

(**Proof**) Let $\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_n$ be eigenvalues and their corresponding unit eigenvectors $\mathbf{u}_1,\ \mathbf{u}_2,\ \cdots,\ \mathbf{u}_n$ of matrix A^tA, that is,

$$(A^tA)\mathbf{u}_i = \lambda_i\mathbf{u}_i \ \ and \ \ \|\mathbf{u}_i\|_2 = 1 \ \ \forall\ 1 \le i \le n.$$

Since $\mathbf{u}_1,\ \mathbf{u}_2,\ \cdots,\ \mathbf{u}_n$ must be an orthonormal basis based on *spectrum decomposition theorem*, for any $\mathbf{x} \in R^n$, we have $\mathbf{x} = \sum_{i=1}^{n} c_i\mathbf{u}_i$. Then

$$\begin{aligned}
\|A\|_2 &= Max_{\|\mathbf{x}\|_2=1}\{\|A\mathbf{x}\|_2\} \\
&= \sqrt{Max_{\|\mathbf{x}\|_2=1}\{\|A\mathbf{x}\|_2^2\}} \\
&= Max_{\|\mathbf{x}\|_2=1}\{\sqrt{\mathbf{x}^tA^tA\mathbf{x}}\} \\
&= Max_{\|\mathbf{x}\|_2=1}\{\sqrt{|\sum_{i=1}^{n}\lambda_i c_i^2|}\} \\
&\le Max_{1\le j\le n}\{\sqrt{|\lambda_j|}\}
\end{aligned}$$

The equality holds if $|\lambda_1| = Max_{1\le j\le n}|\lambda_j|$ and $\mathbf{u}_1 = \mathbf{e}_1$ is selected and $\mathbf{u}_j = \mathbf{0}$ for $2 \le j \le n$.

Theorem: Let $A \in R^{n\times n}$ and $A^t = A$, show that the eigenvectors corresponding to dintinct eigenvalues are orthogonal.

(**Proof**) Let λ and μ be two distinct eigenvalues of A with corresponding eigenvectors $\mathbf{x}$ and $\mathbf{y}$, then we have

$$A\mathbf{x} = \lambda\mathbf{x} \quad \rightarrow \quad \mathbf{y}^t A\mathbf{x} = \lambda\mathbf{y}^t\mathbf{x} = \lambda\langle\mathbf{y},\mathbf{x}\rangle = \lambda\langle\mathbf{x},\mathbf{y}\rangle$$

and

$$A\mathbf{y} = \mu\mathbf{y} \quad \rightarrow \quad \mathbf{x}^t A\mathbf{y} = \mu\mathbf{x}^t\mathbf{y} = \mu\langle\mathbf{x},\mathbf{y}\rangle = \mu\langle\mathbf{y},\mathbf{x}\rangle$$

Since $A^t = A$, then $(\mathbf{y}^t A\mathbf{x})^t = \mathbf{x}^t A^t\mathbf{y} = \mathbf{x}^t A\mathbf{y}$, thus $\lambda\langle\mathbf{x},\mathbf{y}\rangle = \mu\langle\mathbf{x},\mathbf{y}\rangle$, which implies that $(\lambda - \mu)\langle\mathbf{x},\mathbf{y}\rangle = 0$ because $\lambda \neq \mu$, and hence $\langle\mathbf{x},\mathbf{y}\rangle = 0$ or say, $\mathbf{x}$ and $\mathbf{y}$ are orthogonal.

7. Let $b = \begin{bmatrix} 1 \\ -5 \\ 4 \end{bmatrix}$, $C = \begin{bmatrix} -2 & 1 & 0 \\ 1 & -2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$. *Writing* **a single Matlab command** *to solve each of the following questions* for **a** $\sim$ **h** and answer the questions for **i** $\sim$ **h**.

(a) Randomly generate a 3 by 3 matrix A whose elements are integers in [0, 10). (A = **fix(10*random('unif',0,1,3,3))**)

(b) Input vector b.

`(b = [1; -5; 4])`

(c) Solve the linear system $Ax = b$ for x.

`(x = A\b)`

(d) Input matrix C given above.

`(C = [-2,1,0; 1,-2,0; 0,0,2])`

(e) Compute the characteristic polynomial for C. **p=poly(C)**

(f) Compute the eigenvalues and eigenvectors of C. **[U, D]=eig(C)**

(g) Compute the trace of matrix C. **trace(C)**

(h) Compute the rank of matrix C. **rank(C)**

(i) Compute the $LU - decomposition$ of the matrix C. **[L,U,P]=lu(C)**

(j) Compute the $QR - factorization$ of the matrix C. **[Q, R]=qr(C)**

(k) What is the result of **(e)**? **p=[1,2,-5,-6]**

(l) What is the result of **(f)**? **-3, -1, 2**, also see problem 4 for the corresponding eigenvectors.

(m) What is the result of **(g)**? **-2**

(n) What is the result of **(h)**? **3**

(o) What is the result of **(i)**?

(p) What is the result of **(j)**?