Norms on $\mathbb{R}^{n\times n}$ (matrix norms)

Induced Matrix Norms

Given a norm $\|\cdot\|$ on $\mathbb{R}^n$,
the corresponding induced (or subordinate) matrix norm

$$\|A\| := \max_{v\in\mathbb{R}^n,\, v\neq 0} \frac{\|Av\|}{\|v\|} = \max_{w\in\mathbb{R}^n,\, \|w\|=1} \|Aw\|$$

Induced Matrix Norms

Given a norm $\|\cdot\|$ on $\mathbb{R}^n$,
the corresponding induced (or subordinate) matrix norm

$$\|A\| \quad := \quad \max_{v \in \mathbb{R}^n,\, v \neq 0} \frac{\|Av\|}{\|v\|} \quad = \quad \max_{w \in \mathbb{R}^n,\, \|w\|=1} \|Aw\|$$

Matrix p-norm

$$\|A\|_p \quad := \max_{w \in \mathbb{R}^n,\, \|w\|_p=1} \|Aw\|_p,$$

where $\|w\|_p = \sqrt[p]{|w_1|^p + |w_2|^p + \cdots + |w_n|^p}$

Characterizations of the 1-norm, ∞-norm

1-norm is the maximal column sum

$$\|A\|_1 := \max_{w \in \mathbb{R}^n,\, \|w\|_1 = 1} \|Aw\|_1 = \max_{k=1,\dots,n} \sum_{j=1}^{n} |a_{jk}| = \max_{k=1,\dots,n} \|a_k\|_1$$

∞-norm is the maximal row sum

$$\|A\|_\infty := \max_{w \in \mathbb{R}^n,\, \|w\|_\infty = 1} \|Aw\|_\infty = \max_{j=1,\dots,n} \sum_{k=1}^{n} |a_{jk}|$$

Characterization of the 2-norm (the spectral norm)

Theorem

For every $A \in \mathbb{R}^{n \times n}$, we have

$$\|A\|_2 \;=\; \sqrt{\lambda_n}$$

where λ_n denotes the largest eigenvalue of $A^T A$.

Characterization of the 2-norm (the spectral norm)

Theorem

For every $A \in \mathbb{R}^{n\times n}$, we have

$$\|A\|_2 \;=\; \sqrt{\lambda_n}$$

where λ_n denotes the largest eigenvalue of $A^T A$.

Example.

$$A \;=\; \begin{bmatrix} 2 & 5 \\ -6 & 1 \end{bmatrix}, \quad A^T A = \begin{bmatrix} 2 & -6 \\ 5 & 1 \end{bmatrix} \begin{bmatrix} 2 & 5 \\ -6 & 1 \end{bmatrix} = \begin{bmatrix} 40 & 4 \\ 4 & 26 \end{bmatrix}$$

Characterization of the 2-norm (the spectral norm)

Theorem

For every $A \in \mathbb{R}^{n\times n}$, we have

$$\|A\|_2 \;=\; \sqrt{\lambda_n}$$

where λ_n denotes the largest eigenvalue of A^TA.

Example.

$$A \;=\; \begin{bmatrix} 2 & 5 \\ -6 & 1 \end{bmatrix}, \quad A^TA = \begin{bmatrix} 2 & -6 \\ 5 & 1 \end{bmatrix}\begin{bmatrix} 2 & 5 \\ -6 & 1 \end{bmatrix} = \begin{bmatrix} 40 & 4 \\ 4 & 26 \end{bmatrix}$$

Characteristic polynomial of A^TA

$$p(\lambda) = \det(A^TA - \lambda I) = (40-\lambda)(26-\lambda) - 16 = \lambda^2 - 66\lambda + 1024$$

Characterization of the 2-norm (the spectral norm)

Theorem

For every $A \in \mathbb{R}^{n\times n}$, we have

$$\|A\|_2 = \sqrt{\lambda_n}$$

where λ_n denotes the largest eigenvalue of A^TA.

Example.

$$A = \begin{bmatrix} 2 & 5 \\ -6 & 1 \end{bmatrix}, \quad A^TA = \begin{bmatrix} 2 & -6 \\ 5 & 1 \end{bmatrix}\begin{bmatrix} 2 & 5 \\ -6 & 1 \end{bmatrix} = \begin{bmatrix} 40 & 4 \\ 4 & 26 \end{bmatrix}$$

Characteristic polynomial of A^TA

$$p(\lambda) = \det(A^TA - \lambda I) = (40-\lambda)(26-\lambda) - 16 = \lambda^2 - 66\lambda + 1024$$

Eigenvalues of A^TA: $33 \pm \sqrt{65}$, $\quad \|A\|_2 = \sqrt{33 + \sqrt{65}}$

Characterization of the 2-norm (the spectral norm)

First observe that $A^T A$ has real nonnegative eigenvalues

Characterization of the 2-norm (the spectral norm)

First observe that $A^T A$ has real nonnegative eigenvalues , as

$$A^T A q = \lambda q \quad \Longrightarrow \quad \lambda = \frac{q^T A^T A q}{q^T q} = \frac{\|Aq\|_2^2}{\|q\|_2^2} \geq 0$$

Characterization of the 2-norm (the spectral norm)

First observe that $A^T A$ has real nonnegative eigenvalues , as

$$A^T A q = \lambda q \quad \Longrightarrow \quad \lambda = \frac{q^T A^T A q}{q^T q} = \frac{\|Aq\|_2^2}{\|q\|_2^2} \geq 0$$

- Sort them from smallest to largest, i.e.,

$$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n.$$

Characterization of the 2-norm (the spectral norm)

First observe that $A^T A$ has real nonnegative eigenvalues , as

$$A^T Aq = \lambda q \quad \Longrightarrow \quad \lambda = \frac{q^T A^T Aq}{q^T q} = \frac{\|Aq\|_2^2}{\|q\|_2^2} \geq 0$$

- Sort them from smallest to largest, i.e.,

$$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n.$$

- Denote with q_j the eigenvector corresponding to λ_j s.t. $\{q_1, \ldots, q_n\}$ is orthonormal.

Characterization of the 2-norm (the spectral norm)

Expand any $w \in \mathbb{R}^n$, $\|w\|_2 = 1$ in terms of $\{q_1, \ldots, q_n\}$, that is

$$w = \alpha_1 q_1 + \alpha_2 q_2 + \cdots + \alpha_n q_n$$

for some $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$

Characterization of the 2-norm (the spectral norm)

Expand any $w \in \mathbb{R}^n$, $\|w\|_2 = 1$ in terms of $\{q_1, \ldots, q_n\}$, that is

$$w = \alpha_1 q_1 + \alpha_2 q_2 + \cdots + \alpha_n q_n$$

for some $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$ s.t.

$$\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_n^2 = 1 \tag{1}$$

- (1) follows from

$$\begin{aligned} 1 = w^T w &= (\alpha_1 q_1 + \cdots + \alpha_n q_n)^T (\alpha_1 q_1 + \cdots + \alpha_n q_n) \\ &= \alpha_1^2 + \cdots + \alpha_n^2. \end{aligned}$$

Characterization of the 2-norm (the spectral norm)

Consequently,

$$\|Aw\|_2 = \sqrt{w^T A^T A w}$$

Characterization of the 2-norm (the spectral norm)

Consequently,

$$\begin{aligned}\|Aw\|_2 &= \sqrt{w^T A^T A w} \\ &= \sqrt{(\alpha_1 q_1 + \cdots + \alpha_n q_n)^T (\alpha_1 \lambda_1 q_1 + \cdots + \alpha_n \lambda_n q_n)}\end{aligned}$$

Characterization of the 2-norm (the spectral norm)

Consequently,

$$
\begin{aligned}
\|Aw\|_2 &= \sqrt{w^T A^T A w} \\
&= \sqrt{(\alpha_1 q_1 + \cdots + \alpha_n q_n)^T (\alpha_1 \lambda_1 q_1 + \cdots + \alpha_n \lambda_n q_n)} \\
&= \sqrt{\alpha_1^2 \lambda_1 + \cdots + \alpha_n^2 \lambda_n} \le \sqrt{(\alpha_1^2 + \cdots + \alpha_n^2)\lambda_n} = \sqrt{\lambda_n}.
\end{aligned}
$$

Characterization of the 2-norm (the spectral norm)

Consequently,

$$
\begin{aligned}
\|Aw\|_2 &= \sqrt{w^T A^T A w} \\
&= \sqrt{(\alpha_1 q_1 + \cdots + \alpha_n q_n)^T (\alpha_1 \lambda_1 q_1 + \cdots + \alpha_n \lambda_n q_n)} \\
&= \sqrt{\alpha_1^2 \lambda_1 + \cdots + \alpha_n^2 \lambda_n} \;\leq\; \sqrt{(\alpha_1^2 + \cdots + \alpha_n^2)\lambda_n} \;=\; \sqrt{\lambda_n}.
\end{aligned}
$$

Specifically, for $w = q_n$ we have

$$
\|Aq_n\|_2 = \sqrt{q_n^T A^T A q_n} = \sqrt{q_n^T (\lambda_n q_n)} = \sqrt{\lambda_n}
$$

hence the result.