A MATRIX THEORY AND NORMS

A.1 Basics

$$\text{Complex Matrix} \quad A \in \mathcal{C}^{\,l\times m}$$
$$\text{Real Matrix} \quad A \in \mathcal{R}^{\,l\times m}$$

elements $a_{ij} = \text{Re } a_{ij} + j \text{ Im } a_{ij}$

$$\begin{aligned} l &= \text{number of rows} \\ &= \text{“outputs” when viewed as an } \textit{operator} \\ m &= \text{number of columns} \\ &= \text{“inputs” when viewed as an } \textit{operator} \end{aligned}$$

- A^T = transpose of A (with elements a_{ji}),
- $\bar{A}$ = conjugate of A (with elements Re $a_{ij} - j$ Im a_{ij}),
- $A^H \triangleq \bar{A}^T$ = conjugate transpose (or Hermitian adjoint) (with elements Re $a_{ji} - j$Im a_{ji}),

Matrix inverse:

$$A^{-1} = \frac{\mathrm{adj}A}{\det A} \tag{A.1}$$

where adj A is the adjugate (or “classical adjoint”) of A which is the transposed matrix of cofactors c_{ij} of A,

$$c_{ij} = [\mathrm{adj}A]_{ji} \triangleq (-1)^{i+j} \det A^{ij} \tag{A.2}$$

Here A^{ij} is a submatrix formed by deleting row i and column j of A.

Example:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}; \quad \det A = a_{11}a_{22} - a_{12}a_{21}$$

$$A^{-1} = \frac{1}{\det A} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} \tag{A.3}$$

Some matrix identities:

$$(AB)^T = B^T A^T, \quad (AB)^H = B^H A^H \tag{A.4}$$

Assuming the inverses exist,

$$(AB)^{-1} = B^{-1} A^{-1} \tag{A.5}$$

A is symmetric if $A^T = A$,

A is Hermitian if $A^H = A$,

A Hermitian matrix is positive definite if $x^H A x > 0$ for any non-zero vector x.

A.1.1 Some determinant identities

The determinant is defined as

$\det A = \sum_{i=1}^{n} a_{ij} c_{ij}$ (expansion along column j) or

$\det A = \sum_{j=1}^{n} a_{ij} c_{ij}$ (expansion along row i),

where c_{ij} is the ij'th cofactor given in (A.2).

1. Let A_1 and A_2 be square matrices of the same dimension. Then

$$\det(A_1 A_2) = \det(A_2 A_1) = \det A_1 \cdot \det A_2 \quad \text{(A.6)}$$

2. Let c be a complex scalar and A an $n \times n$ matrix. Then

$$\det(cA) = c^n \det(A) \quad \text{(A.7)}$$

3. Let A be a non-singular matrix. Then

$$\det A^{-1} = 1/\det A \quad \text{(A.8)}$$

4. Let A_1 and A_2 be matrices of compatible dimensions such that both matrices $A_1 A_2$ and $A_2 A_1$ are square (but A_1 and A_2 need not themselves be square). Then

$$\det(I + A_1 A_2) = \det(I + A_2 A_1) \quad \text{(A.9)}$$

(A.9) is useful in the field of control because it yields $\det(I + GK) = \det(I + KG)$.

5.

$$\det \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} = \det \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix} = \det(A_{11}) \cdot \det(A_{22}) \tag{A.10}$$

6. **Schur's formula** for the determinant of a partitioned matrix:

$$\det \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \det(A_{11}) \cdot \det(A_{22} - A_{21} A_{11}^{-1} A_{12}) = \det(A_{22}) \cdot \det(A_{11} - A_{12} A_{22}^{-1} A_{21}) \tag{A.11}$$

where it is assumed that A_{11} and/or A_{22} are non-singular.

A.2 Eigenvalues and eigenvectors

Definition

Eigenvalues and eigenvectors. Let A be a square $n \times n$ matrix. The eigenvalues λ_i, $i = 1, \ldots, n$, are the n solutions to the n'th order characteristic equation

$$\det(A - \lambda I) = 0 \tag{A.12}$$

The (right) eigenvector t_i corresponding to the eigenvalue λ_i is the nontrivial solution ($t_i \neq 0$) to

$$(A - \lambda_i I)t_i = 0 \quad \Leftrightarrow \quad At_i = \lambda_i t_i \tag{A.13}$$

The corresponding left eigenvectors q_i satisfy

$$q_i^H(A - \lambda_i I) = 0 \quad \Leftrightarrow \quad q_i^H A = \lambda_i q_i^H \tag{A.14}$$

When we just say *eigenvector* we mean the right eigenvector.

Remarks

- The left eigenvectors of A are the (right) eigenvectors of A^H.
- $\rho(A) \stackrel{\Delta}{=} \max_i |\lambda_i(A)|$ is the *spectral radius* of A.
- Eigenvectors corresponding to distinct eigenvalues are always linearly independent.
- Define

$$T = \{t_1, t_2, \ldots, t_n\}; \quad \Lambda = \text{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\} \tag{A.15}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are distinct.

Then we may then write (A.13) in the following form

$$AT = T\Lambda \tag{A.16}$$

From (A.16) we then get that the eigenvector matrix diagonalizes A in the following manner

$$\Lambda = T^{-1}AT \tag{A.17}$$

A.2.1 Eigenvalue properties

1. $\text{tr}A = \sum_i \lambda_i$ where $\text{tr}A$ is the trace of A (sum of the diagonal elements).
2. $\det A = \prod_i \lambda_i$.
3. The eigenvalues of an upper or lower triangular matrix are equal to the diagonal elements of the matrix.
4. For a real matrix the eigenvalues are either real, or occur in complex conjugate pairs.
5. A and A^T have the same eigenvalues (but in general different eigenvectors).
6. The eigenvalues of A^{-1} are $1/\lambda_1, \ldots, 1/\lambda_n$.
7. The matrix $A + cI$ has eigenvalues $\lambda_i + c$.
8. The matrix cA^k where k is an integer has eigenvalues $c\lambda_i^k$.
9. Consider the $l \times m$ matrix A and the $m \times l$ matrix B. Then the $l \times l$ matrix AB and the $m \times m$ matrix BA have the same non-zero eigenvalues.

10. Eigenvalues are invariant under similarity transformations, that is, A and DAD^{-1} have the same eigenvalues.

11. The same eigenvector matrix diagonalizes the matrix A and the matrix $(I + A)^{-1}$.

12. *Gershgorin's theorem.* The eigenvalues of the $n \times n$ matrix A lie in the union of n circles in the complex plane, each with centre a_{ii} and radius $r_i = \sum_{j \neq i} |a_{ij}|$ (sum of off-diagonal elements in row i). They also lie in the union of n circles, each with centre a_{ii} and radius $r'_i = \sum_{j \neq i} |a_{ji}|$ (sum of off-diagonal elements in column i).

13. A symmetric matrix is positive definite if and only if all its eigenvalues are real and positive.

From the above we have, for example, that

$$\lambda_i(S) = \lambda_i((I + L)^{-1}) = \frac{1}{\lambda_i(I + L)} = \frac{1}{1 + \lambda_i(L)} \tag{A.18}$$

A.3 Singular Value Decomposition

Definition: Unitary matrix. A (complex) matrix U is unitary if

$$U^H = U^{-1} \tag{A.19}$$

Note:

$$\|\lambda(U)\| = 1 \;\; \forall i$$

Definition: SVD. Any complex $l \times m$ matrix A may be factorized into a singular value decomposition

$$A = U\Sigma V^H \tag{A.20}$$

where the $l \times l$ matrix U and the $m \times m$ matrix V are unitary, and the $l \times m$ matrix Σ contains a diagonal matrix Σ_1 of real, non-negative singular values, σ_i, arranged in a descending order as in

$$\Sigma = \begin{bmatrix} \Sigma_1 \\ 0 \end{bmatrix}; \quad l \geq m \tag{A.21}$$

or

$$\Sigma = [\, \Sigma_1 \quad 0 \,]; \quad l \leq m \tag{A.22}$$

where

$$\Sigma_1 = \mathrm{diag}\{\sigma_1, \sigma_2, \ldots, \sigma_k\}; \quad k = \min(l, m) \tag{A.23}$$

and

$$\bar{\sigma} \triangleq \sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_k \triangleq \underline{\sigma} \tag{A.24}$$

- The unitary matrices U and V form orthonormal bases for the column (output) space and the row (input) space of A. The column vectors of V, denoted v_i, are called *right* or *input singular vectors* and the column vectors of U, denoted u_i, are called *left* or *output singular vectors*. We define $\bar{u} \equiv u_1$, $\bar{v} \equiv v_1$, $\underline{u} \equiv u_k$ and $\underline{v} \equiv v_k$.
- SVD is not unique since $A = U' \Sigma {V'}^H$, where $U' = US$, $V' = VS$, $S = \mathrm{diag}\{e^{j\theta_i}\}$ and θ_i is any real number, is also an SVD of A. However, the singular values, σ_i, *are* unique.

$$\sigma_i(A) = \sqrt{\lambda_i(A^H A)} = \sqrt{\lambda_i(AA^H)} \qquad \text{(A.25)}$$

The columns of U and V are unit eigenvectors of AA^H and $A^H A$, respectively. To derive (A.25) write

$$\begin{aligned} AA^H &= (U\Sigma V^H)(U\Sigma V^H)^H = (U\Sigma V^H)(V\Sigma^H U^H) \\ &= U\Sigma\Sigma^H U^H \end{aligned} \tag{A.26}$$

or equivalently since U is unitary and satisfies $U^H = U^{-1}$ we get

$$(AA^H)U = U\Sigma\Sigma^H \tag{A.27}$$

$\Rightarrow U$ is the matrix of eigenvectors of AA^H and $\{\sigma_i^2\}$ are its eigenvalues. Similarly, V is the matrix of eigenvectors of $A^H A$.

Definition: The **rank** of a matrix is equal to the number of non-zero singular values of the matrix. Let $\text{rank}(A) = r$, then the matrix A is called rank deficient if $r < k = \min(l, m)$, and we have singular values $\sigma_i = 0$ for $i = r+1, \ldots k$. A rank deficient square matrix is a singular matrix (non-square matrices are always singular).

A.3.3 SVD of a matrix inverse

Provided the $m \times m$ matrix A is non-singular

$$A^{-1} = V\Sigma^{-1}U^H \tag{A.28}$$

Let $j = m - i + 1$. Then it follows from (A.28) that

$$\begin{aligned} \sigma_i(A^{-1}) &= 1/\sigma_j(A), && \text{(A.29)} \\ u_i(A^{-1}) &= v_j(A), && \text{(A.30)} \\ v_i(A^{-1}) &= u_j(A) && \text{(A.31)} \end{aligned}$$

and in particular

$$\bar{\sigma}(A^{-1}) = 1/\underline{\sigma}(A) \tag{A.32}$$

A.3.4 Singular value inequalities

$$\underline{\sigma}(A) \le |\lambda_i(A)| \le \bar{\sigma}(A) \tag{A.33}$$

$$\bar{\sigma}(A^H) = \bar{\sigma}(A) \quad \text{and} \quad \bar{\sigma}(A^T) = \bar{\sigma}(A) \tag{A.34}$$

$$\bar{\sigma}(AB) \le \bar{\sigma}(A)\bar{\sigma}(B) \tag{A.35}$$

$$\underline{\sigma}(A)\bar{\sigma}(B) \le \bar{\sigma}(AB) \quad \text{or} \quad \bar{\sigma}(A)\underline{\sigma}(B) \le \bar{\sigma}(AB) \tag{A.36}$$

$$\underline{\sigma}(A)\underline{\sigma}(B) \le \underline{\sigma}(AB) \tag{A.37}$$

$$\max\{\bar{\sigma}(A), \bar{\sigma}(B)\} \le \bar{\sigma}\begin{bmatrix} A \\ B \end{bmatrix} \le \sqrt{2}\max\{\bar{\sigma}(A), \bar{\sigma}(B)\} \tag{A.38}$$

$$\bar{\sigma}\begin{bmatrix} A \\ B \end{bmatrix} \le \bar{\sigma}(A) + \bar{\sigma}(B) \tag{A.39}$$

$$\bar{\sigma}\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} = \max\{\bar{\sigma}(A), \bar{\sigma}(B)\} \tag{A.40}$$

$$\sigma_i(A) - \bar{\sigma}(B) \le \sigma_i(A+B) \le \sigma_i(A) + \bar{\sigma}(B) \tag{A.41}$$

Two special cases of (A.41) are:

$$|\bar{\sigma}(A) - \bar{\sigma}(B)| \le \bar{\sigma}(A+B) \le \bar{\sigma}(A) + \bar{\sigma}(B) \tag{A.42}$$

$$\underline{\sigma}(A) - \bar{\sigma}(B) \le \underline{\sigma}(A+B) \le \underline{\sigma}(A) + \bar{\sigma}(B) \tag{A.43}$$

(A.43) yields

$$\underline{\sigma}(A) - 1 \le \underline{\sigma}(I+A) \le \underline{\sigma}(A) + 1 \tag{A.44}$$

On combining (A.32) and (A.44) we get

$$\underline{\sigma}(A) - 1 \le \frac{1}{\bar{\sigma}(I+A)^{-1}} \le \underline{\sigma}(A) + 1 \tag{A.45}$$

A.4 Condition number

The **condition number** of a matrix is defined as the ratio

$$\gamma(A) = \sigma_1(A)/\sigma_k(A) = \bar{\sigma}(A)/\underline{\sigma}(A) \tag{A.46}$$

where $k = \min(l, m)$.

A.5 Norms

Definition
A norm of e (which may be a vector, matrix, signal or system) is a real number, denoted $\|e\|$, that satisfies the following properties:

1. Non-negative: $\|e\| \geq 0$.
2. Positive: $\|e\| = 0 \Leftrightarrow e = 0$ (for semi-norms we have $\|e\| = 0 \Leftarrow e = 0$).
3. Homogeneous: $\|\alpha \cdot e\| = |\alpha| \cdot \|e\|$ for all complex scalars α.
4. Triangle inequality:

$$\|e_1 + e_2\| \leq \|e_1\| + \|e_2\| \tag{A.47}$$

We will consider the norms of four different objects (norms on four different vector spaces):

1. e is a constant vector.

2. e is a constant matrix.

3. e is a time dependent signal, $e(t)$, which at each fixed t is a constant scalar or vector.

4. e is a "system", a transfer function $G(s)$ or impulse response $g(t)$, which at each fixed s or t is a constant scalar or matrix.

A.5.1 Vector norms

General:

$$\|a\|_p = (\sum_i |a_i|^p)^{1/p}; \quad p \geq 1 \tag{A.48}$$

Vector 1-norm (or sum-norm)

$$\|a\|_1 \triangleq \sum_i |a_i| \tag{A.49}$$

Vector 2-norm (Euclidean norm).

$$\|a\|_2 \triangleq \sqrt{\sum_i |a_i|^2} \tag{A.50}$$

$$a^H a = \|a\|_2^2 \tag{A.51}$$

Vector ∞ -norm (or max norm)

$$\|a\|_{\max} \equiv \|a\|_\infty \stackrel{\Delta}{=} \max_i |a_i| \tag{A.52}$$

$$\|a\|_{\max} \le \|a\|_2 \le \sqrt{m}\ \|a\|_{\max} \tag{A.53}$$

$$\|a\|_2 \le \|a\|_1 \le \sqrt{m}\ \|a\|_2 \tag{A.54}$$

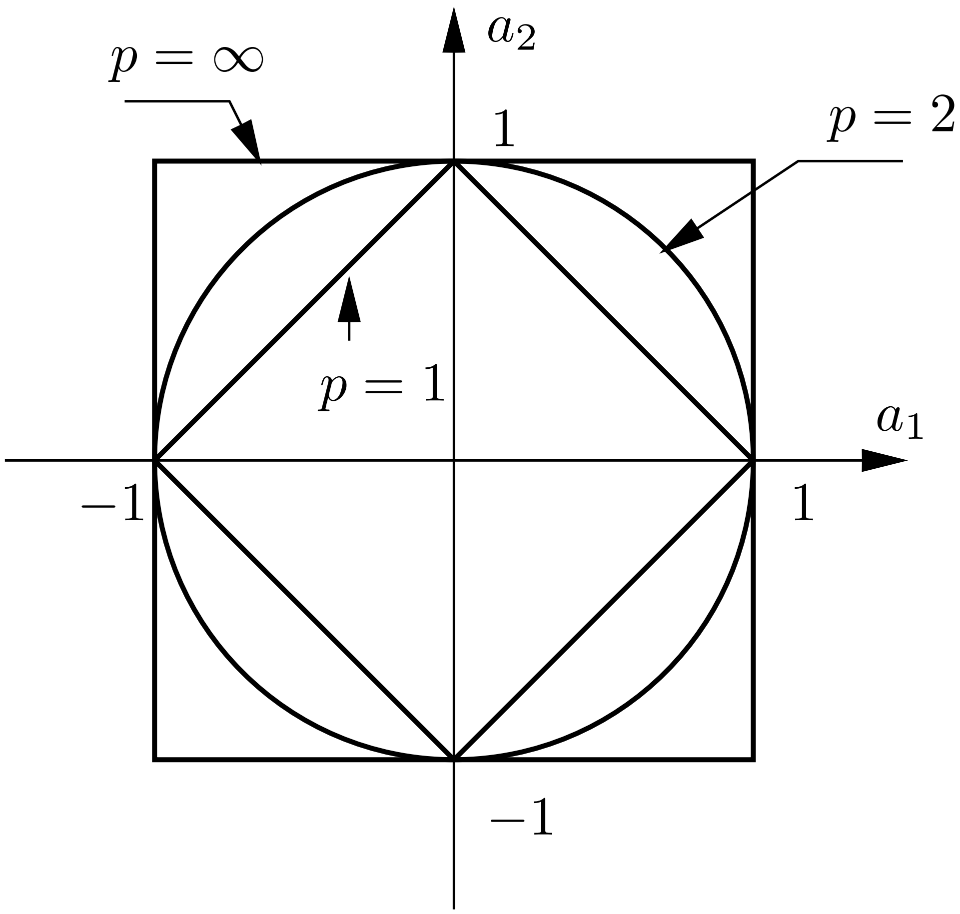

Figure 96: Contours for the vector p-norm, $\|a\|_p = 1$ for $p = 1, 2, \infty$

A.5.2 Matrix norms

Definition

A norm on a matrix $\|A\|$ is a **matrix norm** if, in addition to the four norm properties in Definition A.5, it also satisfies the multiplicative property (also called the consistency condition):

$$\|AB\| \leq \|A\| \cdot \|B\| \tag{A.55}$$

Sum matrix norm.

$$\|A\|_{\text{sum}} = \sum_{i,j} |a_{ij}| \tag{A.56}$$

Frobenius matrix norm (or Euclidean norm).

$$\|A\|_F = \sqrt{\textstyle\sum_{i,j} |a_{ij}|^2} = \sqrt{\text{tr}(A^H A)} \tag{A.57}$$

Max element norm.

$$\|A\|_{\max} = \max_{i,j} |a_{ij}| \tag{A.58}$$

Not a matrix norm as it does not satisfy (A.55). However note that $\sqrt{lm}\ \|A\|_{\max}$ *is* a matrix norm.

Induced matrix norms

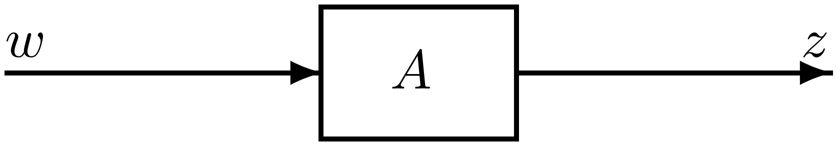

Figure 97: Representation of (A.59)

$$z = Aw \tag{A.59}$$

The *induced norm* is defined as

$$\|A\|_{ip} \triangleq \max_{w \neq 0} \frac{\|Aw\|_p}{\|w\|_p} \tag{A.60}$$

where $\|w\|_p = (\sum_i |w_i|^p)^{1/p}$ denotes the vector p-norm.

- We are looking for a direction of the vector w such that the ratio $\|z\|_p / \|w\|_p$ is maximized.
- The induced norm gives the largest possible "amplifying power" of the matrix. Equivalent definition is:

$$\|A\|_{ip} = \max_{\|w\|_p \leq 1} \|Aw\|_p = \max_{\|w\|_p = 1} \|Aw\|_p \tag{A.61}$$

$$\|A\|_{i1} = \max_j(\textstyle\sum_i |a_{ij}|)$$
"maximum column sum"

$$\|A\|_{i\infty} = \max_i(\textstyle\sum_j |a_{ij}|) \qquad \text{(A.62)}$$
"maximum row sum"

$$\|A\|_{i2} = \bar{\sigma}(A) = \sqrt{\rho(A^H A)}$$
"singular value or spectral norm"

Theorem 14 *All induced norms $\|A\|_{ip}$ are matrix norms and thus satisfy the multiplicative property*

$$\|AB\|_{ip} \leq \|A\|_{ip} \cdot \|B\|_{ip} \qquad \text{(A.63)}$$

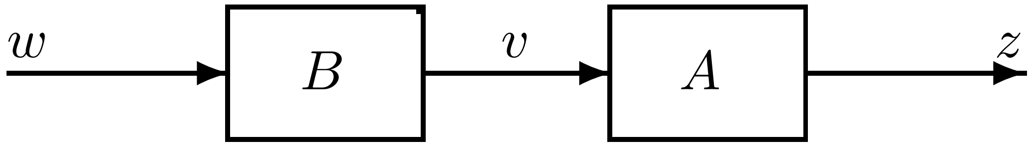

Figure 98:

Implications of the multiplicative property

1. Choose B to be a vector, i.e $B = w$.

$$\|Aw\| \leq \|A\| \cdot \|w\| \tag{A.64}$$

The "matrix norm $\|A\|$ is compatible with its corresponding vector norm $\|w\|$".

2. From (A.64)

$$\|A\| \geq \max_{w \neq 0} \frac{\|Aw\|}{\|w\|} \tag{A.65}$$

For induced norms we have equality in (A.65)
$\|A\|_F \geq \bar{\sigma}(A)$ follows since $\|w\|_F = \|w\|_2$.

3. Choose both $A = z^H$ and $B = w$ as vectors.
Then we derive the Cauchy-Schwarz inequality

$$|z^H w| \leq \|z\|_2 \cdot \|w\|_2 \tag{A.66}$$

A.5.3 The spectral radius $\rho(A)$

$$\rho(A) = \max_i |\lambda_i(A)| \tag{A.67}$$

Not a norm!

Example:

$$A_1 = \begin{bmatrix} 1 & 0 \\ 10 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 10 \\ 0 & 1 \end{bmatrix} \tag{A.68}$$

$$\rho(A_1) = 1, \qquad \rho(A_2) = 1 \tag{A.69}$$

but

$$\rho(A_1 + A_2) = 12, \qquad \rho(A_1 A_2) = 101.99 \tag{A.70}$$

Theorem 15 *For any matrix norm (and in particular for any induced norm)*

$$\rho(A) \leq \|A\| \tag{A.71}$$

A.5.4 Some matrix norm relationships

$$\bar{\sigma}(A) \leq \|A\|_F \leq \sqrt{\min(l,m)}\ \bar{\sigma}(A) \qquad \text{(A.72)}$$

$$\|A\|_{\max} \leq \bar{\sigma}(A) \leq \sqrt{lm}\ \|A\|_{\max} \qquad \text{(A.73)}$$

$$\bar{\sigma}(A) \leq \sqrt{\|A\|_{i1}\|A\|_{i\infty}} \qquad \text{(A.74)}$$

$$\frac{1}{\sqrt{m}}\|A\|_{i\infty} \leq \bar{\sigma}(A) \leq \sqrt{l}\ \|A\|_{i\infty} \qquad \text{(A.75)}$$

$$\frac{1}{\sqrt{l}}\|A\|_{i1} \leq \bar{\sigma}(A) \leq \sqrt{m}\ \|A\|_{i1} \qquad \text{(A.76)}$$

$$\max\{\bar{\sigma}(A), \|A\|_F, \|A\|_{i1}, \|A\|_{i\infty}\} \leq \|A\|_{\text{sum}} \qquad \text{(A.77)}$$

- All these norms, except $\|A\|_{\max}$, are matrix norms and satisfy (A.55).
- The inequalities are tight.
- $\|A\|_{\max}$ can be used as a simple estimate of $\bar{\sigma}(A)$.

The Frobenius norm and the maximum singular value (induced 2-norm) are invariant with respect to unitary transformations.

$$\|U_1 A U_2\|_F = \|A\|_F \tag{A.78}$$

$$\bar{\sigma}(U_1 A U_2) = \bar{\sigma}(A) \tag{A.79}$$

Relationship between Frobenius norm and singular values, $\sigma_i(A)$

$$\|A\|_F = \sqrt{\sum_i \sigma_i^2(A)} \tag{A.80}$$

Perron-Frobenius theorem

$$\min_D \|DAD^{-1}\|_{i1} = \min_D \|DAD^{-1}\|_{i\infty} = \rho(|A|) \tag{A.81}$$

where D is a diagonal "scaling" matrix.

Here:

- $|A|$ denotes the matrix A with all its elements replaced by their magnitudes.
- $\rho(|A|) = \max_i |\lambda_i(|A|)|$ is the Perron root (Perron-Frobenius eigenvalue). Note: $\rho(A) \leq \rho(|A|)$

A.5.5 Matrix and vector norms in MATLAB

$\bar{\sigma}(A) = \|A\|_{i2}$	`norm(A,2)` or `max(svd(A))`
$\|A\|_{i1}$	`norm(A,1)`
$\|A\|_{i\infty}$	`norm(A,'inf')`
$\|A\|_F$	`norm(A,'fro')`
$\|A\|_{\text{sum}}$	`sum (sum(abs(A)))`
$\|A\|_{\max}$	`max(max(abs(A)))` (which is not a matrix norm)

$\rho(A)$	`max(abs(eig(A)))`
$\rho(\|A\|)$	`max(eig(abs(A)))`
$\gamma(A) = \bar{\sigma}(A)/\underline{\sigma}(A)$	`cond(A)`

For vectors:

$\|a\|_1$	`norm(a,1)`
$\|a\|_2$	`norm(a,2)`
$\|a\|_{\max}$	`norm(a,'inf')`

A.5.6 Signal norms

Contrary to spatial norms (vector and matrix norms), choice of temporal norm makes big difference for signals.

Example:

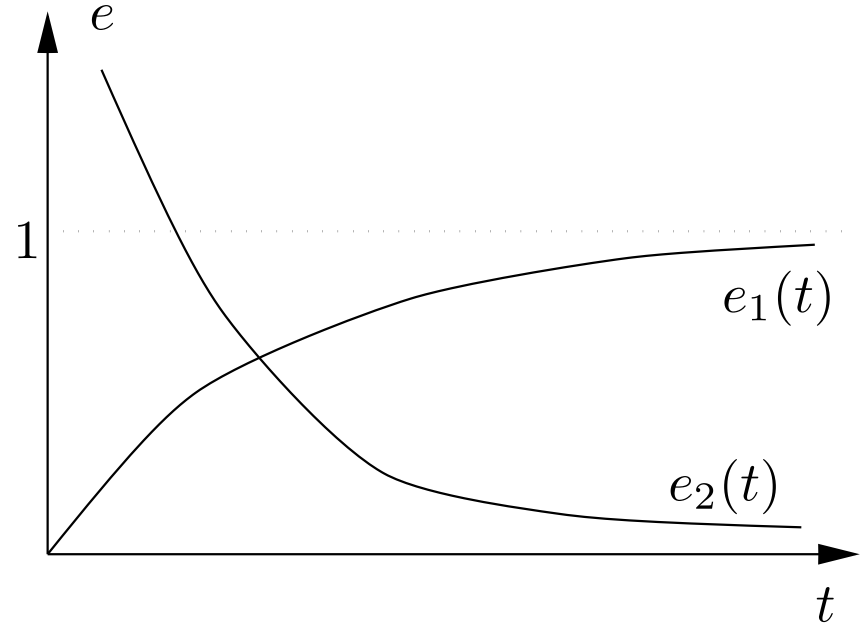

Figure 99: Signals with entirely different 2-norms and ∞-norms.

$$\begin{aligned} \|e_1(t)\|_\infty &= 1, & \|e_1(t)\|_2 &= \infty \\ \|e_2(t)\|_\infty &= \infty, & \|e_2(t)\|_2 &= 1 \end{aligned} \tag{A.82}$$

Compute norm in two steps:

1. "Sum up" the channels at a given time or frequency using a vector norm.
2. "Sum up" in time or frequency using a temporal norm.

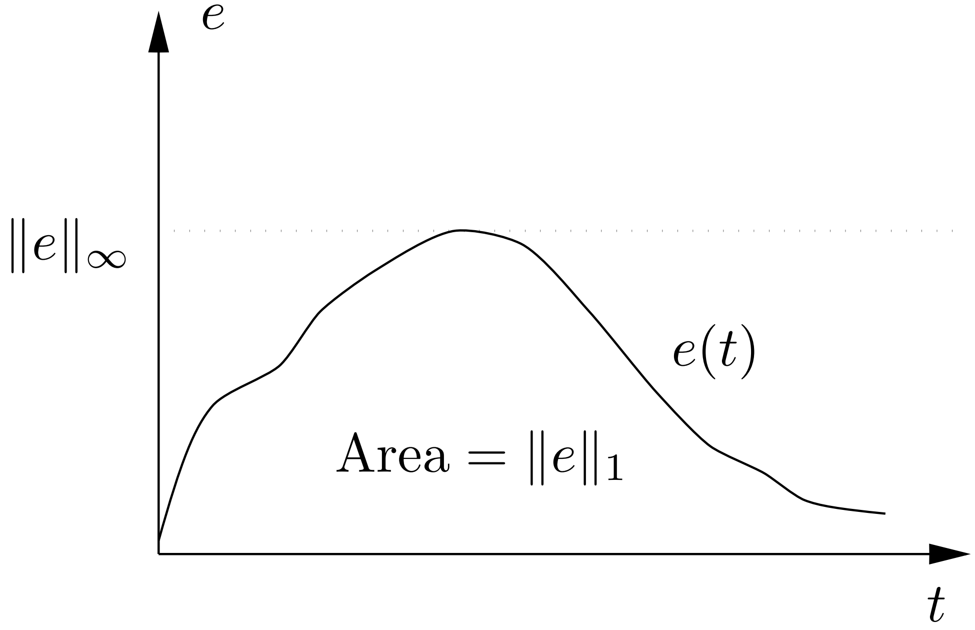

Figure 100: Signal 1-norm and ∞-norm.

General:

$$l_p \text{ norm:} \quad \|e(t)\|_p = \left(\int_{-\infty}^{\infty} \sum_i |e_i(\tau)|^p d\tau \right)^{1/p} \tag{A.83}$$

1-norm in time (integral absolute error (IAE), see Figure 100):

$$\|e(t)\|_1 = \int_{-\infty}^{\infty} \sum_i |e_i(\tau)| d\tau \qquad \text{(A.84)}$$

2-norm in time (quadratic norm, integral square error (ISE), “energy” of signal):

$$\|e(t)\|_2 = \sqrt{\int_{-\infty}^{\infty} \sum_i |e_i(\tau)|^2 d\tau} \qquad \text{(A.85)}$$

∞-norm in time (peak value in time, see Figure 100):

$$\|e(t)\|_\infty = \max_\tau \left(\max_i |e_i(\tau)| \right) \qquad \text{(A.86)}$$

Power-norm or RMS-norm (semi-norm since it does not satisfy property 2)

$$\|e(t)\|_{\text{pow}} = \lim_{T\to\infty} \sqrt{\frac{1}{2T} \int_{-T}^{T} \sum_i |e_i(\tau)|^2 d\tau} \qquad \text{(A.87)}$$