Review of Basics

Lecture Notes

Ahmet Ademoglu, *PhD*
Bogazici University
Institute of Biomedical Engineering

Unit sample function

$$\delta(n) = \begin{cases} 1 & ; \quad n = 0 \\ 0 & ; \quad otherwise \end{cases}$$

Sifting Property

$$x(n) = \sum_{k=-\infty}^{\infty} x(k)\delta(n-k)$$

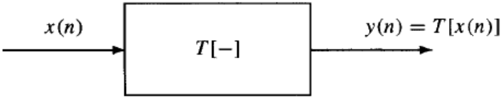
Unit step function

$$u(n) = \begin{cases} 1 & ; \quad n \geq 0 \\ 0 & ; \quad otherwise \end{cases}$$

The relation between $\delta(n)$ and $u(n)$

$$u(n) = \sum_{k=-\infty}^{n} \delta(k) \longleftrightarrow \delta(n) = u(n) - u(n-1)$$

Linearity



$$T[ax_1(n) + bx_2(n)] = aT[x_1(n)] + bT[x_2(n)]$$

If the input is $x(n) = \sum_{k=-\infty}^{\infty} x(k)\delta(n-k)$

then the output is

$$y(n) = \sum_{k=-\infty}^{\infty} x(k)T[\delta(n-k)] = \sum_{k=-\infty}^{\infty} x(k)h_k(n)$$

where $h_k(n)$ is the response of the system to the unit pulse located at k, *i.e.* $\delta(n-k)$.

Shift Invariance

If $T[\delta(n)] = h(n)$ and $T[\delta(n-k)] = h[n-k]$ then
$$h_k(n) = h(n-k)$$

Convolution Sum

$$y(n) = \sum_{k=-\infty}^{\infty} x(k)h(n-k) = \sum_{k=-\infty}^{\infty} x(n-k)h(k) = x(n) * h(n)$$

Causality

The response of the system at time n_0 depends only on the input values at $n \leq n_0$ *i.e.*

$$y[n] = T[x(n), x(n-1), \ldots]$$

For a linear and shift invariant system $h(n) = 0$ for $n < 0$.

Stability

If a bounded input yields a bounded output, the system is BIBO stable

$$|x(n)| < A < \infty \xrightarrow{T} |y(n)| < B < \infty$$

For a linear and shift invariant system $\sum_{k=-\infty}^{\infty} |h(n)| < \infty$.

Time Domain Descriptions of LSI Systems

$$y(n) + \sum_{k=1}^{p} a(k)y(n-k) = \sum_{k=0}^{q} b(q)x(n-k)$$

IIR System

$$y(n) = \sum_{k=0}^{q} b(q)x(n-k) - \sum_{k=1}^{p} a(p)y(n-k)$$

FIR System

$$y(n) = \sum_{k=0}^{q} b(q)x(n-k)$$

Discrete Time Fourier Transform Pair

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x(n)e^{-j\omega n} \longleftrightarrow x(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} X(e^{j\omega})e^{j\omega n}d\omega$$

Convolution Theorem for an LSI system

$$y(n) = x(n) * h(k) = h(n) * x(n) \longleftrightarrow Y(e^{j\omega}) = X(e^{j\omega})H(e^{j\omega})$$

Parseval's Theorem

$$\sum_{n=-\infty}^{\infty} |x(n)|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |X(e^{j\omega})|^2 d\omega$$

$z-$**Transform**

$$X(z) = \sum_{n=-\infty}^{\infty} x(n) z^{-n}$$

where $z = re^{j\omega}$ and when $r = 1$, the $z-$transform becomes the Fourier transform.
The obtain a finite expression for $X(z)$, the $z-$transform requires a region of convergence where $z \subset \mathcal{Z}$.

Properties of $z-$Transform

Property	Sequence	Transform
	$x(n)$	$X(e^{j\omega})$
Delay	$x(n-n_0)$	$e^{-j\omega n_0}X(e^{j\omega})$
Modulation	$e^{j\omega n}x(n)$	$X(e^{j(\omega-\omega_0)})$
Conjugation	$x^*(n)$	$X^*(e^{-j\omega})$
Time reversal	$x(-n)$	$X(e^{-j\omega})$
Convolution	$x(n)*y(n)$	$X(e^{j\omega})Y(e^{j\omega})$
Multiplication	$nx(n)$	$j\frac{d}{d\omega}X(e^{j\omega})$

Properties of $z-$Transform

Property	Sequence	Transform
	$x(n)$	$X(z)$
Delay	$x(n-n_0)$	$z^{-n_0}X(z)$
Modulation	$\alpha^n x(n)$	$X(z/\alpha)$
Conjugation	$x^*(n)$	$X^*(z^*)$
Time reversal	$x(-n)$	$X(z^{-1})$
Convolution	$x(n)*y(n)$	$X(z)Y(z)$
Multiplication	$nx(n)$	$-z\frac{d}{dz}X(z)$

Commonly encountered series and their closed forms

$$\sum_{n=0}^{N-1} a^n = \frac{1-a^N}{1-a}$$

$$\sum_{n=0}^{\infty} a^n = \frac{1}{1-a} \ : \ |a| < 1$$

$$\sum_{n=0}^{N-1} na^n = \frac{(N-1)a^{N+1} - Na^N + a}{(1-a)^2}$$

$$\sum_{N=0}^{\infty} na^n = \frac{a}{(1-a)^2} \ : \ |a| < 1$$

$$\sum_{n=0}^{N-1} n = \frac{1}{2}N(N-1)$$

$$\sum_{n=0}^{N-1} n^2 = \frac{1}{6}N(N-1)(2N-1)$$

Transfer function of an LSI system

$$\mathcal{Z}\left[y(n)+\sum_{k=1}^{p}a(k)y(n-k)=\sum_{k=0}^{k}b(q)x(n-k)\right]$$

$$\frac{Y(z)}{X(z)}=H(z)=\frac{\sum_{k=0}^{q}b(k)z^{-k}}{1+\sum_{k=1}^{p}a(k)z^{-k}}=b(0)\frac{\prod_{k=1}^{q}(1-z_k z^{-1})}{\prod_{k=1}^{p}(1-p_k z^{-1})}$$

z_k : zeros

p_k : poles

Linear Phase Filters and Conjugate Symmetry

$$H(e^{j\omega})=A(e^{j\omega})e^{j(\beta-\alpha\omega)}$$
$$h^*(n)=\pm h(N-1-n)$$

Minimum phase filter

A stable and causal filter having a rational system function with all of its poles and zeros inside the unit circle is called a *minimum phase filter*.

Some Useful $z-$Transform Pairs

Sequence	Transform	ROC
$\delta(\mathrm{n})$	1	All z
$\alpha^n u(n)$	$\frac{1}{1-\alpha z^{-1}}$	$\lvert z\rvert > \alpha$
$-\alpha^n u(-n-1)$	$\frac{1}{1-\alpha z^{-1}}$	$\lvert z\rvert < \alpha$
$\alpha^{\lvert n\rvert}$	$\frac{1-\alpha^2}{(1-\alpha z^{-1})(1-\alpha z)}$	$\alpha < \lvert z\rvert < 1/\alpha$

Discrete Time Fourier Transform and DFT
For an N point finite sequence $x(n)$, the discrete Fourier transform (DFT) pair is

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j2\pi kn/N} \longleftrightarrow x(n) = \frac{1}{N} \sum_{k=1}^{N-1} X(k) e^{j2\pi kn/N}$$

Both $x(n)$ and are $X(k)$ are assumed as discrete and periodic.

Circular Convolution

$$y(n) = \sum_{k=0}^{N-1} x(k)_N h(n-k)_N$$

$$Y(k) = X(k)H(k)$$

Vectors

$$\mathbf{x} = [x_1, x_2, \ldots, x_N]^T$$

Hermitian Transpose

$$\mathbf{x}^H = (\mathbf{x}^T)^* = [x_1^*, x_2^*, \ldots, x_N^*]$$

Signal Vector

$$\mathbf{x}(n) = [x(n), x(n-1), \ldots, x(n-N+1)]^T$$

Euclidian Norm of a vector

$$|\mathbf{a}| = \sqrt{a_1^2 + a_2^2 + \ldots + a_N^2}$$

Inner Product

$$< \mathbf{a}, \mathbf{b} >= \mathbf{a}^H \mathbf{b} = \sum_{i=1}^{N} a_i^* b_i = |\mathbf{a}||\mathbf{b}| \cos(\theta)$$

Orthogonality,

$$< \mathbf{a}, \mathbf{b} >= 0$$

If **a** and **b** have unit length then $< \mathbf{a}, \mathbf{b} >= 0$ implies **orthonormality**.

Cauchy-Schwarz Inequality

$$| < \mathbf{a}, \mathbf{b} > | \leq |\mathbf{a}||\mathbf{b}|$$

$$2| < \mathbf{a}, \mathbf{b} > | \leq |\mathbf{a}|^2 + |\mathbf{b}|^2$$

$$|\mathbf{a} \pm \mathbf{b}| = |\mathbf{a}|^2 \pm 2 < \mathbf{a}, \mathbf{b} > +|\mathbf{b}|^2$$

Vector Spaces, Basis Vectors, Linear Independence
For a set of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N$, if any vector $\mathbf{v}$ can be expressed as

$$\mathbf{v} = \sum_{i=1}^{N} \alpha_i \mathbf{v}_i$$

then $\mathbf{v}_i$ are called the **basis vectors** that span a vector space.

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_N \mathbf{v}_N = \mathbf{0}$$

Matrices

$$\mathbf{A}_{n\times m} = \{a_{ij}\} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

Convolution Matrix

$$y(n) = \sum_{k=0}^{N-1} h(n-k)x(k)$$

can be expressed as vector-matrix operation as

$$\underbrace{\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{N-1} \end{bmatrix}}_{\mathbf{y}} = \underbrace{\begin{bmatrix} h_0 & 0 & \cdots & 0 \\ h_1 & h_0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ h_{N-1} & h_{N-2} & \cdots & h_{N-M} \end{bmatrix}}_{\mathbf{H}} \underbrace{\begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{M-1} \end{bmatrix}}_{\mathbf{x}}$$

Determinant

$$det(\mathbf{A}) = |\mathbf{A}| = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} |\mathbf{A}_{ij}|$$

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$|\mathbf{A}| = a_{11}a_{22} - a_{12}a_{21}$$

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$|\mathbf{A}| = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

$$= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{31}a_{23}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})$$

Hermitian Matrix
A matrix is Hermitian if

$$\mathbf{A} = \mathbf{A}^H = (\mathbf{A}^*)^T$$

Rank of a matrix

$$rank(\mathbf{A})_{m\times n} = \rho(\mathbf{A}) \leq \min(m, n)$$

is the number of independent columns in $\mathbf{A}$.
For a **full rank** matrix $\rho(\mathbf{A}) = \min(m, n)$.
If $\mathbf{A}$ is full rank and $m = n$ then it has a unique inverse so that

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$$

Properties of Determinant

1. $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$
2. $|\mathbf{A}^T| = |\mathbf{A}|$
3. $|\alpha\mathbf{A}| = \alpha^n|\mathbf{A}|$
4. $|\mathbf{A}^{-1}| = \dfrac{1}{|\mathbf{A}|}$

Inverse and Determinant

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} adj(\mathbf{A})$$

$$adj(\mathbf{A})^T = cofactor(\mathbf{A})$$

$$c_{ij} : cofactor(\mathbf{A})_{ij}$$

$$c_{ij} = (-1)^{i+j} m_{ij}$$

$$c_{ij} : Minor(\mathbf{A})_{ij}$$

$c_{ij} = |\mathbf{A}|$ with ith row and jth column deleted

$$|\mathbf{A}| = a_{i1}c_{i1} + a_{i2}c_{i2} + \cdots + a_{in}c_{in} = \sum_{j=1}^{n} a_{ij}c_{ij}$$

Matrix Inversion Lemma

$$(\mathbf{A}+\mathbf{BCD})^{-1}=\mathbf{A}^{-1}-\mathbf{A}^{-1}(\mathbf{I}+\mathbf{BCDA}^{-1})^{-1}\mathbf{BCDA}^{-1}$$

$$(\mathbf{A}+\mathbf{BCD})^{-1}=\mathbf{A}^{-1}-\mathbf{A}^{-1}\mathbf{B}(\mathbf{I}+\mathbf{CDA}^{-1}\mathbf{B})^{-1}\mathbf{CDA}^{-1}$$

Woodbury Identity

In a special case where $\mathbf{B}\rightarrow\mathbf{u}$, $\mathbf{C}=1$, $\mathbf{D}=\mathbf{v}^H$
the Matrix Inversion Lemma leads to

$$(\mathbf{A}+\mathbf{uv}^H)^{-1}=\mathbf{A}^{-1}-\frac{\mathbf{A}^{-1}\mathbf{uv}^H\mathbf{A}^{-1}}{1+\mathbf{v}^H\mathbf{A}^{-1}\mathbf{u}}$$

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m &= b_2 \\
\vdots \qquad\qquad & \vdots \quad \vdots \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m &= b_n
\end{aligned}
$$

$$
\mathbf{A}_{n\times m}\mathbf{x}_{m\times 1} = \mathbf{b}_{n\times 1} = \mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \ldots + \mathbf{a}_m x_m
$$

$$
\{\mathbf{AB}\}_{ij} = \sum_k a_{ik} b_{kj}
$$

$$\mathbf{a}^T\mathbf{B}\mathbf{c} = \sum_i \sum_j a_i b_{ij} c_j$$

$$\mathbf{ABC} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots \\ \downarrow & \downarrow & \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \dots \\ b_{21} & b_{22} & \dots \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \mathbf{c}_1 \longrightarrow \\ \mathbf{c}_2 \longrightarrow \\ \vdots \end{bmatrix}$$

$$= b_{11}\underset{\big\downarrow}{\mathbf{a}_1}\mathbf{c}_1 \longrightarrow \quad + \quad b_{12}\underset{\big\downarrow}{\mathbf{a}_1}\mathbf{c}_2 \longrightarrow \quad + \quad b_{21}\underset{\big\downarrow}{\mathbf{a}_2}\mathbf{c}_1 \longrightarrow \quad + \quad \dots$$

Eigenvalue Problem : $\mathbf{Av} = \lambda\mathbf{v}$

$$\mathbf{A}\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_N \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_N \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_N \end{bmatrix}$$

$$\mathbf{AV} = \mathbf{V\Lambda} \longrightarrow \mathbf{A} = \mathbf{V\Lambda V}^{-1}$$

$$f(\mathbf{A}) = \mathbf{V}f(\mathbf{\Lambda})\mathbf{V}^{-1} \qquad e^{\mathbf{A}} = \mathbf{I} + \frac{1}{1!}\mathbf{A} + \frac{1}{2!}\mathbf{A}^2 + \cdots$$

Characteristic polynomial

$$|\lambda\mathbf{I} - \mathbf{A}| = \lambda^N + a_1\lambda^{N-1} + \cdots + a_{N-1}\lambda + a_N = 0$$

Every matrix satisfies its characteristic equation

$$\mathbf{A}^N + a_1\mathbf{A}^{N-1} + \cdots + a_{N-1}\mathbf{A} + a_N = \mathbf{0}$$

TRACE of **A**

$$Trace(\mathbf{A}) = \sum_i a_{ii}$$

Useful Identities

- $Trace(\mathbf{AB}) = Trace(\mathbf{BA})$
- $Trace(\mathbf{ABC}) = Trace(\mathbf{CAB}) = Trace(\mathbf{BCA})$
- $\mathbf{x}^T\mathbf{A}\mathbf{y} = Trace(\mathbf{y}\mathbf{x}^T\mathbf{A})$
- $Trace(\mathbf{A}) = \sum_i \lambda_i$

Singular Value Decomposition

Every matrix $\mathbf{A}$ can be decomposed into

$\mathbf{A}_{m\times n} = \mathbf{U}_{m\times m}\mathbf{\Sigma}_{m\times n}\mathbf{V}^T_{n\times n}$

$eig(\mathbf{A}\mathbf{A}^T) = \mathbf{U}\mathbf{D}\mathbf{U}^T \longrightarrow \mathbf{U}$: left eigenvectors

$\mathbf{u}_i^T\mathbf{u}_j = \delta_{ij} \longrightarrow \mathbf{U}^{-1} = \mathbf{U}^T$

$eig(\mathbf{A}^T\mathbf{A}) = \mathbf{V}\mathbf{D}\mathbf{V}^T \longrightarrow \mathbf{V}$: right eigenvectors

$\mathbf{v}_i^T\mathbf{v}_j = \delta_{ij} \longrightarrow \mathbf{V}^{-1} = \mathbf{V}^T$

$\mathbf{\Sigma}$: singular values $\longrightarrow \mathbf{\Sigma}^2 = \mathbf{D}$

$$\mathbf{A} = \begin{bmatrix} \underbrace{\begin{matrix}\mathbf{u}_1 & \dots & \mathbf{u}_r \\ \downarrow & \vdots & \downarrow\end{matrix}}_{Col(A)} & \underbrace{\begin{matrix}\mathbf{u}_{r+1} & \dots & \mathbf{u}_m \\ \downarrow & \vdots & \downarrow\end{matrix}}_{Null\ Space} \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \mathbf{0} \\ & \ddots & & \\ & & \sigma_r & \\ \mathbf{0} & & & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^T & \longrightarrow \\ \vdots & \\ \mathbf{v}_r^T & \longrightarrow \\ \mathbf{v}_{r+1}^T & \longrightarrow \\ \vdots & \\ \mathbf{v_n}^T & \longrightarrow \end{bmatrix}$$

Block Diagonal Matrices

$$\begin{bmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{B} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{F_A}^{-1} & -\mathbf{A}^{-1}\mathbf{C}\mathbf{F_D}^{-1} \\ -\mathbf{F_D}^{-1}\mathbf{B}\mathbf{A}^{-1} & \mathbf{F_D}^{-1} \end{bmatrix}$$

$$\mathbf{F_A} = [\mathbf{A} - \mathbf{C}\mathbf{D}^{-1}\mathbf{B}] \qquad \mathbf{F_D} = [\mathbf{D} - \mathbf{B}\mathbf{A}^{-1}\mathbf{C}]$$

where $\mathbf{F_D}$ and $\mathbf{F_D}$ are the Schur's complement of $\mathbf{A}$ and $\mathbf{D}$, respectively.
Determinant of a block diagonal matrix is

$$\begin{vmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{B} & \mathbf{D} \end{vmatrix} = \left| \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{B} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{A}^{-1}\mathbf{C} \\ \mathbf{0} & \mathbf{D} - \mathbf{B}\mathbf{A}^{-1}\mathbf{C} \end{bmatrix} \right| = |\mathbf{A}| \cdot |\mathbf{D} - \mathbf{B}\mathbf{A}^{-1}\mathbf{C}| = |\mathbf{A}| \cdot |\mathbf{F_D}|$$

$$\begin{vmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{B} & \mathbf{D} \end{vmatrix} = \left| \begin{bmatrix} \mathbf{I} & \mathbf{C} \\ \mathbf{0} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{A} - \mathbf{C}\mathbf{D}^{-1}\mathbf{B} & \mathbf{0} \\ \mathbf{D}^{-1}\mathbf{B} & \mathbf{I} \end{bmatrix} \right| = |\mathbf{D}| \cdot |\mathbf{A} - \mathbf{C}\mathbf{D}^{-1}\mathbf{B}| = |\mathbf{D}| \cdot |\mathbf{F_A}|$$

It follows by the assignment $\begin{bmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{B} & \mathbf{D} \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{X} \\ \mathbf{X}^T & -\mathbf{D}^{-1} \end{bmatrix}$ that
$|\mathbf{A} + \mathbf{X}\mathbf{D}\mathbf{X}^T| = |\mathbf{A}| \cdot |\mathbf{D}| \cdot |\mathbf{D}^{-1} + \mathbf{X}^T\mathbf{A}^{-1}\mathbf{X}|$

Norm of a matrix $\mathbf{A}$ or a vector $\mathbf{a}$

- L_0 Norm : $||\mathbf{A}||_0 = \sum_{ij} |a_{ij}|^0$
- L_1 Norm : $||\mathbf{A}||_1 = \sum_{ij} |a_{ij}|$
- L_2 Norm of a vector $\mathbf{a} \rightarrow ||\mathbf{a}||_2 = \sqrt{\sum_i a_i^2}$
- Frobenius Norm of a matrix $\mathbf{A} \rightarrow ||\mathbf{A}||_F = \sqrt{\sum_{ij} a_{ij}^2}$
- L_∞ Norm : $||\mathbf{A}||_\infty = \max_{ij} |a_{ij}|$
- $L_{2,1}$ Norm of a matrix $\mathbf{A} = \sum_i ||\mathbf{A}(i,:)||_2$
- $L_{1,2}$ Norm of a matrix $\mathbf{A} = \sqrt{\sum_i ||\mathbf{A}(i,:)||_1^2}$
- Nuclear norm of a matrix $\mathbf{A}$ $||\mathbf{A}||_* = \sum_i \lambda_i$ $\lambda_i = \{singular\ values\ of\ \mathbf{A}\}$

Solution of Linear System of Equations $\mathbf{Ax} = \mathbf{b}$

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m &= b_2 \\ \vdots \qquad\qquad & \vdots \quad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m &= b_n \end{aligned}$$

$$\mathbf{A}_{m\times n}\mathbf{x}_n = \mathbf{b}_m$$

- Least Squares Solution
- Minimum Norm Solution
- Regularized Least Squares Solution

Solution of Linear System of Equations $\mathbf{Ax} = \mathbf{b}$

- $m < n$ Overdetermined Case $\longrightarrow$ Least Squares Solution
  $\frac{\partial}{\partial \mathbf{x}} |\mathbf{Ax} - \mathbf{b}|^2 = 0,$
  $\frac{\partial}{\partial \mathbf{x}} \left(\mathbf{x}^T \mathbf{A}^T \mathbf{Ax} - \mathbf{x}^T \mathbf{A}^T \mathbf{b} - \mathbf{b}^T \mathbf{Ax} + \mathbf{b}^T \mathbf{b} \right) = 2\mathbf{A}^T \mathbf{Ax} - \mathbf{A}^T \mathbf{b} - \mathbf{A}^T \mathbf{b} = 0$
  $$\mathbf{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b}$$
  If $\frac{\partial}{\partial \mathbf{x}} (\mathbf{Ax} - \mathbf{b})^T \mathbf{W} (\mathbf{Ax} - \mathbf{b}) = 0$ then it is Weighted Least Squares Solution
  $$\mathbf{x} = (\mathbf{A}^T \mathbf{WA})^{-1} \mathbf{A}^T \mathbf{Wb}$$
- $m > n$ Underdetermined Case $\longrightarrow$ Minimum Norm Solution
  $\frac{\partial}{\partial \mathbf{x}} \left(\mathbf{x}^T \mathbf{x} + \lambda^T (\mathbf{Ax} - \mathbf{b}) \right) = 2\mathbf{x} + \mathbf{A}^T \lambda = 0$ and $\lambda = -2(\mathbf{AA}^T)^{-1} \mathbf{b}$
  $$\mathbf{x} = \mathbf{A}^T (\mathbf{AA}^T)^{-1} \mathbf{b}$$
- $m > n \longrightarrow$ Regularized Least Squares Solution
  $\frac{\partial}{\partial \mathbf{x}} \left(|\mathbf{Ax} - \mathbf{b}|^2 + \mu \mathbf{x}^T \mathbf{x} \right) = 0$
  $$\mathbf{x}_\mu = (\mathbf{A}^T \mathbf{A} + \mu \mathbf{I})^{-1} \mathbf{A}^T \mathbf{b}$$
- $m > n \longrightarrow$ Smoothness Regularization
  $\frac{\partial}{\partial \mathbf{x}} \left(|\mathbf{Lx}|^2 + \lambda^T (\mathbf{Ax} - \mathbf{b}) \right) = 0$, $\mathbf{L}$: The Laplacian Operator
  $$\mathbf{x} = (\mathbf{L}^T \mathbf{L})^{-1} \mathbf{A}^T (\mathbf{A} (\mathbf{L}^T \mathbf{L})^{-1} \mathbf{A}^T)^{-1} \mathbf{b}$$
- $m > n \rightarrow$ General Minimum Norm Solution with Constraint $\mathbf{Cx} - \mathbf{d}$
  $\frac{\partial}{\partial \mathbf{x}} \left(|\mathbf{Ax} - \mathbf{b}|^2 + \lambda^T (\mathbf{Cx} - \mathbf{d}) \right) = 0 \longrightarrow \mathbf{x} = ?$

Some Properties

- A matrix $\mathbf{A}$ is *positive definite*, if for any vector $\mathbf{x}$ the quadratic form always maintains the property that $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$.
- The nonzero eigenvectors of $\mathbf{A}$, $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly independent if λ_i are distinct. This implies that if $\mathbf{A}\mathbf{v} = \mathbf{0}$ then $\mathbf{A}$ has $n - rank(\mathbf{A})$ eigenvalues λ_i which are equal to 0.
- Eigenvalues of a Hermitian matrix are real.
- A Hermitian matrix is positive definite if and only if $\lambda_i > 0$.
- The eigenvectors of a Hermitian matrix corresponding to distinct eigenvalues are orthogonal.
- A Hermitian matrix $\mathbf{A}$ may be decomposed using *spectral theorem* as

  $$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^H = \sum_{i=1}^{n} \lambda_i \mathbf{v}_i \mathbf{v_i}^H$$ where $\mathbf{v}_i$ are orthonormal.

If $\mathbf{A} = \mathbf{B} + \alpha\mathbf{I}$ where $\mathbf{B} = \lambda\mathbf{u}_1\mathbf{u}_1^H$ and $|\mathbf{u}| = 1$.
$\mathbf{B}\mathbf{u}_1 = \lambda(\mathbf{u}_1\mathbf{u}_1^H)\mathbf{u}_1 = \lambda\mathbf{u}_1$
$\mathbf{u}_1$ is an eigenvector of $\mathbf{B}$ with eigenvalue λ.
Since $\mathbf{A}$ is Hermitian, eigenvectors of $\mathbf{A} = \{\mathbf{u}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and $\lambda_i = \{\lambda + \alpha, \alpha, \ldots, \alpha\}$.

$$\mathbf{A}^{-1} = \frac{1}{\lambda + \alpha}\mathbf{v}_1\mathbf{v}_1^T + \sum_{i=2}^{n} \frac{1}{\alpha}\mathbf{v}_i\mathbf{v}_i^H$$

Optimization of Complex Functions

The minimum for convex functions : $\frac{d}{dx}f(x) = 0$ and $\frac{d^2}{dx^2}f(x) > 0$.

$$\nabla_x f(\mathbf{x}) = \frac{d}{d\mathbf{x}}f(\mathbf{x}) = \begin{bmatrix} \frac{\partial}{\partial x_1}f(\mathbf{x}) \\ \frac{\partial}{\partial x_2}f(\mathbf{x}) \\ \vdots \\ \frac{\partial}{\partial x_n}f(\mathbf{x}) \end{bmatrix}$$

For an extremum point of $f(\mathbf{x})$,

$\nabla_x f(\mathbf{x}) = \mathbf{0}$ and $\mathbf{H}_x$ must be positive definite where

$\{\mathbf{H}_\mathbf{x}\}_{ij} = \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j}$.

$\frac{d}{dz}z = 1 \qquad \frac{d}{dz}|z|^2 = z^* \qquad \frac{d}{dz}\mathbf{a}^H\mathbf{z} = \mathbf{a}^* \qquad \frac{d}{dz}\mathbf{z}^H\mathbf{a} = \mathbf{0} \qquad \frac{d}{dz}\mathbf{z}^H\mathbf{A}\mathbf{z} = (\mathbf{A}\mathbf{z})^*$

$\frac{d}{dz^*}z = 0 \qquad \frac{d}{dz^*}|z|^2 = z^* \qquad \frac{d}{dz^*}\mathbf{a}^H\mathbf{z} = \mathbf{0} \qquad \frac{d}{dz^*}\mathbf{z}^H\mathbf{a} = \mathbf{a} \qquad \frac{d}{dz^*}\mathbf{z}^H\mathbf{A}\mathbf{z} = \mathbf{A}\mathbf{z}$

Minimize $Q = \frac{1}{2}\mathbf{z}^H\mathbf{R}\mathbf{z}$ with the constraint $\mathbf{z}^H\mathbf{a} = 1$.

$$Q_R(\mathbf{z}, \lambda) = \frac{1}{2}\mathbf{z}^H\mathbf{R}\mathbf{z} + \lambda(1 - \mathbf{z}^H\mathbf{a})$$

$$\nabla_{z^*} Q_R(\mathbf{z}, \lambda) = \mathbf{R}\mathbf{z} - \lambda\mathbf{a} = \mathbf{0} \qquad \frac{\partial Q_R(\mathbf{z},\lambda)}{\partial \lambda} = 1 - \mathbf{z}^H\mathbf{a}$$

$$\mathbf{z} = \frac{\mathbf{R}^{-1}\mathbf{a}}{\mathbf{a}^H\mathbf{R}^{-1}\mathbf{a}}$$

The minimum value of Q is

$$\min\{Q\} = \frac{\mathbf{z}^H\mathbf{a}}{\mathbf{a}^H\mathbf{R}^{-1}\mathbf{a}} = \frac{1}{\mathbf{a}^H\mathbf{R}^{-1}\mathbf{a}}$$