Notes on Linear Algebra and Matrix Analysis

Maxim Neumann
May 2006, Version 0.1.1

1 Matrix Basics

Literature to this topic: [1–4].
$\mathbf{x}^\dagger\mathbf{y} \iff <\mathbf{y},\mathbf{x}>$: standard inner product.
$\mathbf{x}^\dagger\mathbf{x} = 1 : x$ is normalized
$\mathbf{x}^\dagger\mathbf{y} = 0 : x, y$ are orthogonal
$\mathbf{x}^\dagger\mathbf{y} = 0,\ \mathbf{x}^\dagger\mathbf{x} = 1,\ \mathbf{y}^\dagger\mathbf{y} = 1 : x, y$ are orthonormal
$\mathbf{Ax} = \mathbf{y}$ is uniquely solvable if $\mathbf{A}$ is linear independent (nonsingular).
Majorization: Arrange $\mathbf{b}$ and $\mathbf{a}$ in increasing order $(\mathbf{b}_m, \mathbf{a}_m)$ then:

$$\mathbf{b} \text{ majorizes } \mathbf{a} \iff \sum_{i=1}^{k} b_{m_i} \geq \sum_{i=1}^{k} a_{m_i} \quad \forall\, k \in [1,\dots,n] \tag{1}$$

The collection of all vectors $\mathbf{b} \in \mathbb{R}^n$ that majorize a given vector $\mathbf{a} \in \mathbb{R}^n$ may be obtained by forming the convex hull of $n!$ vectors, which are computed by permuting the n components of $\mathbf{a}$.
Direct sum of matrices $\mathbf{A} \in \mathbf{M}_{n1}, \mathbf{B} \in \mathbf{M}_{n2}$:

$$\mathbf{A} \oplus \mathbf{B} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{bmatrix} \in \mathbf{M}_{n1+n2} \tag{2}$$

$[\mathbf{A}, \mathbf{B}] = \operatorname{trace} \mathbf{AB}^\dagger$: matrix inner product.

1.1 Trace

$$\begin{aligned} \operatorname{trace} \mathbf{A} &= \sum_i^n \lambda_i && (3) \\ \operatorname{trace}(\mathbf{A}+\mathbf{B}) &= \operatorname{trace} \mathbf{A} + \operatorname{trace} \mathbf{B} && (4) \\ \operatorname{trace} \mathbf{AB} &= \operatorname{trace} \mathbf{BA} && (5) \end{aligned}$$

1.2 Determinants

The determinant $\det(\mathbf{A})$ expresses the volume of a matrix A.

$$\det(\mathbf{A}) = 0 \iff \begin{cases} \mathbf{A} \text{ is } \textit{singular.} \\ \text{Linear equation is not solvable.} \\ \mathbf{A}^{-1} \text{ does not exists} \\ \text{vectors in } \mathbf{A} \text{ are linear dependent} \end{cases} \tag{6}$$

$$\det(\mathbf{A}) \neq 0 \iff A \text{ is } \textit{regular/nonsingular.}$$

$$A_{ij} \in \mathbb{R} \to \det(\mathbf{A}) \in \mathbb{R} \qquad A_{ij} \in \mathbb{C} \to \det(\mathbf{A}) \in \mathbb{C}$$

If $\mathbf{A}$ is a square matrix$(\mathbf{A}_{n\times n})$ and has the eigenvalues λ_i, then $\det(\mathbf{A}) = \prod \lambda_i$

$$\begin{aligned} \det \mathbf{A}^T &= \det \mathbf{A} && (7) \\ \det \mathbf{A}^\dagger &= \overline{\det \mathbf{A}} && (8) \\ \det \mathbf{AB} &= \det \mathbf{A}\ \det \mathbf{B} && (9) \end{aligned}$$

Elementary operations on matrix and determinant:

Interchange of two rows :	$\det \mathbf{A}$	$*=$	-1
Multiplication of a row by a nonzero scalar c :	$\det \mathbf{A}$	$*=$	c
Addition of a scalar multiple of one row to another row :	$\det \mathbf{A}$	$=$	$\det \mathbf{A}$

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \tag{10}$$

2 Eigenvalues, Eigenvectors, and Similarity

$\sigma(\mathbf{A}_{n\times n}) = \{\lambda_1, \ldots, \lambda_n\}$ is the set of eigenvalues of $\mathbf{A}$, also called the spectrum of $\mathbf{A}$.
$\rho(\mathbf{A}) = \max\{|\lambda_i|\}$ is the **spectral radius** of matrix $\mathbf{A}$.
$\Lambda = \mathbf{D}_{n\times n} = \text{diag}\{\lambda_1, \ldots, \lambda_n\}$ is the diagonal matrix of eigenvalues.
$\mathbf{V} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is the matrix of eigenvectors.

2.1 Characteristic Polynomial

$$p_{\mathbf{A}}(t) = \det(t\mathbf{I} - \mathbf{A}) \tag{11}$$

$$p_{\mathbf{A}}(t) = \prod_{i=1}^{m} (t - \lambda_i)^{s_i}, \quad 1 \le s_i \le n, \quad \sum_i^m s_i = n, \quad \lambda_i \in \sigma(\mathbf{A}) \tag{12}$$

2.2 Eigendecomposition

$$\mathbf{M}\mathbf{v} = \lambda \mathbf{v} \tag{13}$$

If $\mathbf{M} = \frac{1}{N}[\text{Re}(\mathbf{x}), \text{Im}(\mathbf{x})]^T [\text{Re}(\mathbf{x}), \text{Im}(\mathbf{x})]$, then λ is the variance of $\mathbf{M}$ along $\mathbf{v}$. $sdev(\mathbf{M})$ along $\mathbf{v}$ is then $= \sqrt{\lambda}$.
Characteristic function: $|\lambda \mathbf{I} - \mathbf{M}| = 0$

$$\text{trace } \mathbf{M} = \sum \lambda_i \tag{14}$$

$$\det \mathbf{M} = |\mathbf{M}| = \prod \lambda_i \tag{15}$$

$$\sigma(\mathbf{AB}) = \sigma(\mathbf{BA}) \qquad \text{even if } \mathbf{AB} \neq \mathbf{BA} \tag{16}$$

$$\mathbf{M} = \mathbf{M}^{T*} \quad \rightarrow \quad \lambda \in \mathbf{R} \tag{17}$$

$$\mathbf{M}^k \quad \rightarrow \quad \lambda^k \tag{18}$$

$\mathbf{M}$ is squared matrix. With $\mathbf{D} = \text{diag } \lambda_i$ and $\mathbf{V} = [\mathbf{v_0}, \mathbf{v_1}, \ldots]$:

$$\mathbf{MV} = \mathbf{DV} \quad \Longleftrightarrow \quad \mathbf{M} = \mathbf{VDV^{-1}} \tag{19}$$

$$\mathbf{M}^2 = (\mathbf{VDV^{-1}})(\mathbf{VDV^{-1}}) = \mathbf{VD^2V^{-1}} \tag{20}$$

$$\mathbf{M}^n = \mathbf{VD}^n\mathbf{V^{-1}} \qquad with\ n \in \mathbf{N}^{>0} \tag{21}$$

$$\mathbf{M}^{-1} = \mathbf{VD^{-1}V^{-1}} \qquad with\ \mathbf{D}^{-1} = \text{diag } \lambda_i^{-1} \tag{22}$$

$$e^{\mathbf{M}} = \mathbf{V}e^{\mathbf{D}}\mathbf{V^{-1}} \qquad with\ e^{\mathbf{D}} = \text{diag } e^{\lambda_i} \tag{23}$$

$$\mathbf{M}^{1/2} = \mathbf{VD^{1/2}V^{-1}} \qquad with\ \mathbf{D}^{1/2} = \text{diag } \lambda_i^{1/2} \tag{24}$$

(24) is also known as square root of the matrix $\mathbf{M}$ with $\mathbf{M} = \mathbf{M^{1/2}M^{1/2\dagger}}$.
$\mathbf{A}$ triangular $\rightarrow \sigma(\mathbf{A}) = \{a_{ii} | 1 \le i \le n\}$

2.3 Similarity

$\mathbf{A}$ is *similar* to $\mathbf{B}$ if $\exists$ $\mathbf{S}$ such that:

$$\mathbf{B} = \mathbf{S}^{-1}\mathbf{AS} \qquad \rightarrow \mathbf{A} \sim \mathbf{B} \iff \mathbf{B} \sim \mathbf{A} \tag{25}$$

$$\mathbf{A} \sim \mathbf{B} \rightarrow \begin{cases} \sigma(\mathbf{A}) = \sigma(\mathbf{B}) \\ \det \mathbf{A} = \det \mathbf{B} \\ \text{trace } \mathbf{A} = \text{trace } \mathbf{B} \\ \text{rank } \mathbf{A} = \text{rank } \mathbf{B} \\ p_{\mathbf{A}}(t) = p_{\mathbf{B}}(t) \quad p_{\mathbf{A}}(t)\text{: characteristic polynomial of } \mathbf{A} \end{cases} \tag{26}$$

$\forall$ $\mathbf{A} \in \mathbf{M}_n$ $\mathbf{A} \sim \mathbf{A^T}$!! $\rightarrow \text{rank } \mathbf{A} = \text{rank } \mathbf{A}^T$ (row rank = column rank).
$\forall$ $\mathbf{A} \in \mathbf{M}_n$ $\mathbf{A} \sim \mathbf{S}$ $\mathbf{S} = \mathbf{S}^T := \mathbf{S}$ is a symmetric matrix
$\mathbf{A}$ is diagonalizable $\iff \exists \mathbf{D} : \mathbf{A} \sim \mathbf{D}$
$\mathbf{A} \sim \mathbf{D} \iff \mathbf{A}$ has n linearly independent eigenvectors $\mathbf{V} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$:

$$\mathbf{A} \sim \mathbf{D} \iff \mathbf{D} = \mathbf{S}^{-1}\mathbf{AS} \quad \text{with } \mathbf{S} = \mathbf{V}_a,\ \mathbf{D} = \Lambda_a \tag{27}$$

$\mathbf{A}, \mathbf{B}$ ($\mathbf{A} \not\sim \mathbf{B}$) are simultaneously diagonalizable, if:

$$\exists \quad \mathbf{S} : \quad \mathbf{D}_a = \mathbf{S}^{-1}\mathbf{AS},\ \mathbf{D}_b = \mathbf{S}^{-1}\mathbf{BS} \tag{28}$$

$\mathbf{A}, \mathbf{B}$ commute $\iff$ $\mathbf{A}, \mathbf{B}$ are *simultaneously diagonalizable*

Commuting family $\mathcal{W} \subseteq \mathbb{M}_n$: each pair of matrices in $\mathcal{W}$ is commutative under multiplication $\iff$ $\mathcal{W}$ is a *simultaneously diagonalizable* family.
If $\mathcal{W}$ is a commuting family, then there is a vector $\mathbf{x}$ that is an eigenvector of every $\mathbf{A} \in \mathcal{W}$.

2.4 Eigenvectors, Eigenspace

Eigenspace is the the set of eigenvectors corresponding to the eigenvalue (one or more) of **A**.
Geometric multiplicity of λ is the dimension of the eigenspace of **A** corresponding to the eigenvalue λ. It is the maximum number of linearly independent eigenvectors associated with an eigenvalue.
Algebraic multiplicity of λ is the multiplicity of the eigenvalue λ as a zero in the characteristic polynom. It is the amount of same values of λ in $\sigma(\mathbf{A})$. In general, the term *multiplicity* usually references the algebraic multiplicity.
Geometric multiplicity of $\lambda \quad \leq \quad$ algebraic multiplicity of λ.

$$\exists\, \lambda \in \sigma(\mathbf{A})\text{: geometric mult.} < \text{algebraic mult.} \quad \rightarrow \quad \mathbf{A} \text{ is } defective. \tag{29}$$

$$\forall\, \lambda \in \sigma(\mathbf{A})\text{: geometric mult.} = \text{algebraic mult.} \quad \rightarrow \quad \mathbf{A} \text{ is } nondefective. \tag{30}$$

$$\forall\, \lambda \in \sigma(\mathbf{A})\text{: geometric mult. of } \lambda = 1 \quad \rightarrow \quad \mathbf{A} \text{ is } nonderogatory. \tag{31}$$

If $\operatorname{rank}(\mathbf{A}) = k$ then $\exists \mathbf{x}_i, \mathbf{y}_i : \quad \mathbf{A} = \mathbf{x}_1 \mathbf{y}_1^\dagger + \ldots + \mathbf{x}_k \mathbf{y}_k^\dagger$

3 Unitary Equivalence, Normal and Unitary Matrices (real orthogonal matrices in $\mathbb{R}^n$)

$$\mathbf{U} \text{ is unitary} \iff \begin{cases} \iff \mathbf{U}\mathbf{U}^\dagger = \mathbf{I} \\ \iff \mathbf{U}^\dagger \text{ is unitary} \\ \iff \mathbf{U} \text{ is nonsingular and } \mathbf{U}^{-1} = \mathbf{U}^\dagger \\ \iff \mathbf{U} \text{ columns/rows form an orthonormal set} \\ \iff \mathbf{U}\mathbf{x} = \mathbf{y} \rightarrow \mathbf{y}^\dagger \mathbf{y} = \mathbf{x}^\dagger \mathbf{x} \iff |\mathbf{y}| = |\mathbf{x}| \forall \mathbf{x} \in \mathbb{C}^{\ltimes} \end{cases} \tag{32}$$

$\mathbf{U}, \mathbf{V}$ are unitary $\rightarrow \mathbf{UV}$ is also unitary.
A is *unitary equivalent* to **B** if $\exists$ **U**(unitary matrix) such that:

$$\mathbf{B} = \mathbf{U}^\dagger \mathbf{A} \mathbf{U} \tag{33}$$

Unitary equivalence implies similarity, but not conversely. It corresponds, like similarity, to a change of basis, but of a special type – a change from one *orthonormal* basis to another. An orthonormal change of basis leaves unchanged the sum of squares of the absolute values of the entries ($\operatorname{trace} \mathbf{A}^\dagger \mathbf{A} = \operatorname{trace} \mathbf{B}^\dagger \mathbf{B}$ if **A** is unitary equivalent to **B**).
Schur's theorem: Every matrix **A** is unitary equivalent to a upper(lower) triangular matrix **T**:

$$\forall \mathbf{A}\ \exists \mathbf{U}, \mathbf{T} \ : \quad \mathbf{U}^\dagger \mathbf{A} \mathbf{U} = \mathbf{T} \tag{34}$$

Where **U** is unitary, **T** is triangular. **U**, **T** are not unique.

3.1 Normal matrices

Normal matrices generalize the unitary, real summetric, Hermitian, and *skew*– Hermitian matrices(and other). The class of normal matrices is closed under unitary equivalence. Condition:

$$\mathbf{A}^\dagger \mathbf{A} = \mathbf{A}\mathbf{A}^\dagger \tag{35}$$

$$\{\text{Unitary, Hermitian, Skew–Hermitian}\} \subseteq \text{Normal} \tag{36}$$

$$\mathbf{A} \text{ normal} \iff \begin{cases} \qquad \mathbf{A} \text{ is normal} \\ \iff \mathbf{A} \text{ is unitarily diagonalizable} \\ \iff \sum_{i,j}^{n} |a_{ij}|^2 = \sum_{i}^{n} |\lambda_i|^2 \\ \iff \exists \text{ orthonormal set of } n \text{ eigenvectors of } \mathbf{A} \\ \iff \exists\, \mathbf{U} : \ \mathbf{A}^\dagger = \mathbf{A}\mathbf{U} \qquad \mathbf{U} \text{ is unitary} \end{cases} \tag{37}$$

$$\mathbf{A} \text{ normal} \rightarrow \mathbf{A} \text{ is nondefective (geom.mult.=algb.mult.)} \tag{38}$$

4 Canonical Forms

Jordan block $\mathbf{J}_k(\lambda) \in \mathbf{M}_k$ and Jordan matrix $\mathbf{J} \in \mathbf{M}_n$:

$$\mathbf{J}_k(\lambda) = \begin{bmatrix} \lambda & 1 & & \mathbf{0} \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ \mathbf{0} & & & \lambda \end{bmatrix} \qquad \mathbf{J} = \begin{bmatrix} \mathbf{J}_{n_1}(\lambda_1) & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{J}_{n_k}(\lambda_k) \end{bmatrix}, \quad \sum_i^k n_i = n \tag{39}$$

Jordan canonical form theorem: $\forall \mathbf{A} \in \mathbf{M}_n \ \exists \, \mathbf{S}, \mathbf{J} \ : \quad \mathbf{A} = \mathbf{S}\mathbf{J}\mathbf{S}^{-1}$
Where $\mathbf{J}$ is unique up to permutations of the diagonal Jordan blocks. By convention, the Jordan blocks corresponding to each distinct eigenvalue are presented in decreasing order with the largest block first.
Minimal Polynomial $q_{\mathbf{A}}(t)$ of $\mathbf{A}$ is the unique monic polynomial $q_{\mathbf{A}}(t)$ of minimum degree that annihilates $\mathbf{A}$. (monic: a polynomial whose highest–order term has coefficient +1; annihilates: a polynomial whose value is the $\mathbf{0}$ matrix).
Similar matrices have the same minimal polynomial.
Minimal polynomial $q_{\mathbf{A}}(t)$ divides the characteristic polynomial $p_{\mathbf{A}}(t) = q_{\mathbf{A}}(t)h_{\mathbf{A}}(t)$. Also, $q_{\mathbf{A}}(\lambda) = \mathbf{0} \quad \forall \, \lambda \in \sigma(\mathbf{A})$.

$$q_{\mathbf{A}}(t) = \prod_{i=1}^{m}(t-\lambda_i)^{r_i}, \quad \lambda_i \in \sigma(\mathbf{A}) \tag{40}$$

Where r_i is the order of the largest Jordan block of $\mathbf{A}$ corresponding to λ_i.
Polar decomposition: $(\operatorname{rank}\mathbf{P} = \operatorname{rank}\mathbf{A})$

$$\begin{array}{llll} \forall \mathbf{A}: & \mathbf{A} = & \mathbf{P}\mathbf{U} & \mathbf{P} \text{ positive semidifinite, } \mathbf{U} \text{ unitary} \\ \forall \mathbf{A} \text{ nonsingular}: & \mathbf{A} = & \mathbf{G}\mathbf{Q} & \mathbf{G} = \mathbf{G}^{\mathbf{T}},\ \mathbf{Q}\mathbf{Q}^T = \mathbf{I} \end{array} \tag{41}$$

Singular value decomposition:

$$\forall \mathbf{A}: \mathbf{A} = \mathbf{V}\Sigma\mathbf{W}^\dagger \quad \mathbf{V}, \mathbf{W} \text{ unitary, } \Sigma \text{ nonnegative diagonal, } \operatorname{rank}\Sigma = \operatorname{rank}\mathbf{A} \tag{42}$$

Triangular factorization:

$$\forall \mathbf{A}: \ \mathbf{A} = \mathbf{U}\mathbf{R}\mathbf{U}^\dagger \quad \mathbf{U} \text{ unitary, } \mathbf{R} \text{ upper triangular} \tag{43}$$

Others:

$$\forall \mathbf{A} \text{ Hermitian}: \ \mathbf{A} = \mathbf{S}\mathbf{I}_{\mathbf{A}}\mathbf{S}^\dagger \quad \mathbf{S} \text{ nonsingular, } \mathbf{I}_{\mathbf{A}} = \operatorname{diag} \in \{-1, 0, +1\} \tag{44}$$

Where $\mathbf{I}_{\mathbf{A}}$ is a diagonal matrix with entries $\in \{-1, 0, +1\}$. The number of $+1(-1)$ entries in $\mathbf{I}_{\mathbf{A}}$ is the same as the number of positive (negative) eigenvalues of $\mathbf{A}$, the number of 0 entries is equal to $(n - \operatorname{rank}\mathbf{A})$.

$$\begin{array}{llll} \forall \mathbf{A} \text{ normal}: & \mathbf{A} = & \mathbf{U}\Lambda\mathbf{U}^\dagger & U \text{ unitary, } \Lambda = \operatorname{diag}\lambda_i \in \sigma(\mathbf{A}) \\ \forall \mathbf{A} \text{ symmetric}: & \mathbf{A} = & \mathbf{S}\mathbf{K}_{\mathbf{A}}\mathbf{S}^{\mathbf{T}} & \mathbf{S} \text{ nonsingular, } \mathbf{K}_{\mathbf{A}} = \operatorname{diag} \in \{0, +1\}, \\ & & & \operatorname{rank}\mathbf{K}_{\mathbf{A}} = \operatorname{rank}\mathbf{A} \\ \forall \mathbf{A} \text{ symmetric}: & \mathbf{A} = & \mathbf{U}\Sigma\mathbf{U}^{\mathbf{T}} & \mathbf{U} \text{ unitary, } \Sigma = \operatorname{diag} \geq 0,\ \operatorname{rank}\Sigma = \operatorname{rank}\mathbf{A} \\ \forall \mathbf{U} \text{ unitary}: & \mathbf{U} = & \mathbf{Q}e^{i\mathbf{E}} & \mathbf{Q} \text{ real orthogonal, } \mathbf{E} \text{ real symmetric} \\ \forall \mathbf{P} \text{ orthogonal}: & \mathbf{P} = & \mathbf{Q}e^{i\mathbf{F}} & \mathbf{Q} \text{ real orthogonal, } \mathbf{E} \text{ real skew–symm.} \\ \forall \mathbf{A}: & \mathbf{A} = & \mathbf{S}\mathbf{U}\Sigma\mathbf{U}^T\mathbf{S}^{-1} & \mathbf{S} \text{ nonsingular, } \mathbf{U} \text{ unitary, } \Sigma = \operatorname{diag} \geq 0 \end{array} \tag{45}$$

LU factorization: is not unique. Only if all upper left principal submatrices of $\mathbf{A}$ are nonsingular (and $\mathbf{A}$ is nonsingular):

$$\mathbf{A} = \mathbf{L}\mathbf{U} \quad \mathbf{L} \text{ lower triangular, } \mathbf{U} \text{ upper triangular, } \det\mathbf{A}(\{1,..,k\}) \neq 0,\ k = 1,..,n \tag{46}$$

$$\begin{array}{llll} \forall \mathbf{A} \text{ nonsingular}: & \mathbf{A} = & \mathbf{P}\mathbf{L}\mathbf{U} & \mathbf{P} \text{ permutation, } \mathbf{L}, \mathbf{U} \text{ lower/upper triangular} \\ \forall \mathbf{A}: & \mathbf{A} = & \mathbf{P}\mathbf{L}\mathbf{U}\mathbf{Q} & \mathbf{P}, \mathbf{Q} \text{ permutation, } \mathbf{L}, \mathbf{U} \text{ lower/upper triangular} \end{array} \tag{47}$$

5 Hermitian and Symmetric Matrices

$$\forall \, \mathbf{A} \begin{cases} \mathbf{A} + \mathbf{A}^\dagger, \mathbf{A}\mathbf{A}^\dagger, \mathbf{A}^\dagger\mathbf{A} & \text{Hermitian} \\ \mathbf{A} - \mathbf{A}^\dagger & \text{skew–Hermitian} \\ \mathbf{A} = \mathbf{H}_{\mathbf{A}} + \mathbf{S}_{\mathbf{A}} & \mathbf{H}_{\mathbf{A}} = \frac{\mathbf{A}+\mathbf{A}^\dagger}{2} \text{ Hermitian, } \mathbf{S}_{\mathbf{A}} = \frac{\mathbf{A}-\mathbf{A}^\dagger}{2} \text{ skew–Hermitian} \\ \mathbf{A} = \mathbf{E}_{\mathbf{A}} + i\mathbf{F}_{\mathbf{A}} & \mathbf{E}_{\mathbf{A}}, \mathbf{F}_{\mathbf{A}} \text{ Hermitian, } \quad \mathbf{E}_{\mathbf{A}} = \mathbf{H}_{\mathbf{A}},\ \mathbf{F}_{\mathbf{A}} = -i\mathbf{S}_{\mathbf{A}} \end{cases} \tag{48}$$

$\mathbf{H}_{\mathbf{A}}, \mathbf{S}_{\mathbf{A}}, \mathbf{E}_{\mathbf{A}}, \mathbf{F}_{\mathbf{A}}$ are unique $\forall \mathbf{A}$. $\mathbf{H}_{\mathbf{A}}, \mathbf{E}_{\mathbf{A}}, \mathbf{F}_{\mathbf{A}} \in \mathbf{M}(\mathbb{R})$.

$$\mathbf{H}_{\mathbf{A}}\mathbf{S}_{\mathbf{A}} = \mathbf{S}_{\mathbf{A}}\mathbf{H}_{\mathbf{A}} \iff \mathbf{A} \text{ is normal} \tag{49}$$

$$\forall\, \mathbf{A},\mathbf{B} \text{ Hermitian} \rightarrow \begin{cases} \mathbf{A}^k & \text{Hermitian } \forall\, k \in \mathbb{N}^{\geq 0} \\ \mathbf{A}^{-1} & \text{Hermitian} \quad (\mathbf{A} \text{ nonsingular}) \\ a\mathbf{A}+b\mathbf{B} & \text{Hermitian } \forall\, a,b \in \mathbb{R} \\ i\mathbf{A} & \text{skew–Hermitian} \\ \mathbf{x}^\dagger\mathbf{A}\mathbf{x} \in \mathbb{R} & \forall\, \mathbf{x} \in \mathbb{C}^n \\ \sigma(\mathbf{A}) \in \mathbb{R} & \\ \mathbf{S}^\dagger\mathbf{A}\mathbf{S} & \text{Hermitian } \forall\, \mathbf{S} \\ \mathbf{A} & \text{normal (see 3.1)} \\ \mathbf{A}\mathbf{A}^\dagger = \mathbf{A}^2 = \mathbf{A}^\dagger\mathbf{A} & \\ \mathbf{A} = \mathbf{U}\Lambda\mathbf{U}^\dagger & \mathbf{U} \text{ unitary, } \Lambda = diag \in \sigma(\mathbf{A}) \in \mathbb{R} \\ \mathbf{D} = \mathbf{U}^\dagger\mathbf{A}\mathbf{U} & \mathbf{A} \text{ is unitarily diagonalizable} \\ \mathbf{v}_i^\dagger\mathbf{v}_j = 0 \quad i \neq j & \text{eigenvectors are othonormal} \\ \text{trace}(\mathbf{A}\mathbf{B})^2 \leq \text{trace}\,\mathbf{A}^2\mathbf{B}^2 & \\ \text{rank}(\mathbf{A}) = & \text{number of nonzero eigenvalues} \\ \text{rank}(\mathbf{A}) \geq \frac{(\text{trace}\,\mathbf{A})^2}{\text{trace}\,\mathbf{A}^2} & \end{cases} \tag{50}$$

$$\forall\, \mathbf{A},\mathbf{B} \text{ skew–Hermitian} \rightarrow \begin{cases} a\mathbf{A}+b\mathbf{B} & \text{skew–Hermitian } \forall\, a,b \in \mathbb{R} \\ i\mathbf{A} & \text{Hermitian} \\ \sigma(\mathbf{A}) \in Imaginary & \end{cases} \tag{51}$$

Commutivity of Hermitian matrices! Let $\mathcal{W}$ be a given family of Hermitian matrices:

$$\exists\, \mathbf{U} \text{ unitary} : \mathbf{D}_\mathbf{A} = \mathbf{U}\mathbf{A}\mathbf{U}^\dagger,\ \mathbf{D}_\mathbf{A} \text{ is diagonal } \forall \mathbf{A} \in \mathcal{W} \iff \mathbf{A}\mathbf{B} = \mathbf{B}\mathbf{A}\ \forall \mathbf{A},\mathbf{B} \in \mathcal{W} \tag{52}$$

$$\mathbf{A},\mathbf{B} \text{ Hermitian} \rightarrow \mathbf{A}\mathbf{B} \text{ Hermitian} \iff \mathbf{A}\mathbf{B} = \mathbf{B}\mathbf{A} \tag{53}$$

But in general: $\mathbf{AB} \neq \mathbf{BA}$!!! and $\mathbf{AB}$ is in general not Hermitian!!!
A class of matrices where $\mathbf{A}$ is similar to its Hermitian ($\mathbf{A} \sim \mathbf{A}^\dagger$):

$$\mathbf{A} \sim \mathbf{A}^\dagger \iff \begin{cases} \iff \mathbf{A} \sim \mathbf{B} & \mathbf{B} \in \mathbf{M}_n(\mathbb{R}) \\ \iff \mathbf{A} \sim \mathbf{A}^\dagger & \\ \iff \mathbf{A} \sim \mathbf{A}^\dagger & \text{via a Hermitian similarity transformation} \\ \iff \mathbf{A} = \mathbf{H}\mathbf{K} & \mathbf{H},\mathbf{K} \text{ Hermitian} \\ \iff \mathbf{A} = \mathbf{H}\mathbf{K} & \mathbf{H},\mathbf{K} \text{ Hermitian with at least one nonsingular} \end{cases} \tag{54}$$

A matrix $\mathbf{A} \in \mathbf{M}_n$ is uniquely determined by the Hermitian (sesquilinear) form $\mathbf{x}^\dagger\mathbf{A}\mathbf{x}$:

$$\mathbf{x}^\dagger\mathbf{A}\mathbf{x} = \mathbf{x}^\dagger\mathbf{B}\mathbf{x} \quad \forall\, \mathbf{x} \in \mathbb{C}^n \iff \mathbf{A} = \mathbf{B} \tag{55}$$

The class of hermitian matrices is closed under unitary equivalence.

$$\mathbf{A} = \mathbf{A}^\dagger, \quad \forall\, \mathbf{x} : \quad \mathbf{x}^\dagger\mathbf{A}\mathbf{x} = \begin{cases} \geq 0 & \rightarrow \sigma(\mathbf{A}) \geq 0 \quad \mathbf{A} \text{ is positive semidifinite} \\ > 0 & \qquad\qquad\quad \mathbf{A} \text{ is positive difinite} \end{cases} \tag{56}$$

5.1 Variational characterization of eigenvalues of Hermitian matrices

Matrix $\mathbf{A}$ is a Hermitian matrix with eigenvalues $\lambda_i \in \mathbb{R}$.
Convention for eigenvalues of Hermitian matrices: $\lambda_{min} = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n = \lambda_{max}$
Rayleigh–Ritz ratio: $\frac{\mathbf{x}^\dagger\mathbf{A}\mathbf{x}}{\mathbf{x}^\dagger\mathbf{x}}$

$$\lambda_1\mathbf{x}^\dagger\mathbf{x} \leq \mathbf{x}^\dagger\mathbf{A}\mathbf{x} \leq \lambda_n\mathbf{x}^\dagger\mathbf{x} \quad \forall\, x \in \mathbb{C}^n \tag{57}$$

$$\lambda_{max} = \lambda_n = \max_{\mathbf{x}\neq 0} \frac{\mathbf{x}^\dagger\mathbf{A}\mathbf{x}}{\mathbf{x}^\dagger\mathbf{x}} = \max_{\mathbf{x}^\dagger\mathbf{x}=1} \mathbf{x}^\dagger\mathbf{A}\mathbf{x} \tag{58}$$

$$\lambda_{min} = \lambda_1 = \min_{\mathbf{x}\neq 0} \frac{\mathbf{x}^\dagger\mathbf{A}\mathbf{x}}{\mathbf{x}^\dagger\mathbf{x}} = \min_{\mathbf{x}^\dagger\mathbf{x}=1} \mathbf{x}^\dagger\mathbf{A}\mathbf{x} \tag{59}$$

Geometrical interpretation: λ_{max} is the largest value of the function $\mathbf{x}^\dagger\mathbf{A}\mathbf{x}$ as $\mathbf{x}$ ranges over the unit sphere in $\mathbb{C}^n$, a compact set. Analog is λ_{min} the smallest (negative) value.

$$\lambda_{n-1} = \inf_{\mathbf{w}\in\mathbb{C}^n} \sup_{\substack{\mathbf{x}^\dagger\mathbf{x}=1 \\ \mathbf{x}\perp\mathbf{w}}} \mathbf{x}^\dagger\mathbf{A}\mathbf{x} \qquad\qquad \lambda_2 = \sup_{\mathbf{w}\in\mathbb{C}^n} \inf_{\substack{\mathbf{x}^\dagger\mathbf{x}=1 \\ \mathbf{x}\perp\mathbf{w}}} \mathbf{x}^\dagger\mathbf{A}\mathbf{x} \tag{60}$$

Courant–Fischer:

$$\min_{\mathbf{w}_1,\mathbf{w}_2,\ldots,\mathbf{w}_{n-k}\in\mathbb{C}^n} \max_{\substack{\mathbf{x}^\dagger\mathbf{x}=1,\mathbf{x}\in\mathbb{C}^n \\ \mathbf{x}\perp\mathbf{w}_1,\mathbf{w}_2,\ldots,\mathbf{w}_{n-k}}} \mathbf{x}^\dagger\mathbf{A}\mathbf{x} = \lambda_k \tag{61}$$

$$\max_{\mathbf{w}_1,\mathbf{w}_2,\ldots,\mathbf{w}_{k-1}\in\mathbb{C}^n} \min_{\substack{\mathbf{x}^\dagger\mathbf{x}=1,\mathbf{x}\in\mathbb{C}^n \\ \mathbf{x}\perp\mathbf{w}_1,\mathbf{w}_2,\ldots,\mathbf{w}_{k-1}}} \mathbf{x}^\dagger\mathbf{A}\mathbf{x} = \lambda_k \tag{62}$$

$$\forall\, \mathbf{B}\ (\text{even } \mathbf{B}\neq\mathbf{B}^\dagger) \quad\rightarrow\quad \min_{\mathbf{x}^\dagger\mathbf{x}=1} |\mathbf{x}^\dagger\mathbf{B}\mathbf{x}| \leq |\lambda_i| \leq \max_{\mathbf{x}^\dagger\mathbf{x}-1} |\mathbf{x}^\dagger\mathbf{B}\mathbf{x}|,\ 1\leq i\leq n \tag{63}$$

The diagonal entries of **A** majorizes the vector of eigenvalues of **A** (comp. (1)):

$$\forall\, \mathbf{A},\mathbf{B} \text{ Hermitian} \rightarrow \begin{cases} \operatorname{diag}(\mathbf{A}) \text{ majorizes } \operatorname{diag}(\Lambda) \\ \prod\operatorname{diag}(\mathbf{A}) \geq \prod\operatorname{diag}(\Lambda) = \prod\lambda_i = \det(\mathbf{A}) \\ \lambda_i(\mathbf{A}+\mathbf{B}) \leq \min\{\lambda_i(\mathbf{A})+\lambda_j(\mathbf{B}) : i+j=k+n\} \end{cases} \tag{64}$$

5.2 Complex symmetric matrices

A complex symmetric matrix need not to be normal.

$$\mathbf{A} \text{ symmetric} \rightarrow \mathbf{A} = \mathbf{U}\Sigma\mathbf{U}^T \tag{65}$$

U is unitary and contains an orthonormal set of eigenvectors of $\mathbf{A}\overline{\mathbf{A}}$; Σ is a diagonal matrix with the nonnegative square roots of eigenvalues of $\mathbf{A}\overline{\mathbf{A}}$.

$$\mathbf{A} = \mathbf{A}^T \iff \exists\,\mathbf{B} : \mathbf{A} = \mathbf{B}\mathbf{B}^T \tag{66}$$

$$\forall\, \mathbf{A}\in\mathbf{M}_n \rightarrow \begin{cases} \mathbf{A}\sim\mathbf{S} & \mathbf{S} \text{ symmetric!!} \\ \mathbf{A}=\mathbf{B}\mathbf{C} & \mathbf{B},\mathbf{C} \text{ symmetric!!} \end{cases} \tag{67}$$

5.3 Congruence and simultaneous diagonalization of Hermitian and symmetric matrices

General quadratic form: $Q_{\mathbf{A}}(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x} = \sum_{i,j=1}^n a_{ij}x_ix_j, \quad \mathbf{x}\in\mathbb{C}^n$

(General) Hermitian form: $H_{\mathbf{B}}(\mathbf{x}) = \mathbf{x}^\dagger\mathbf{B}\mathbf{x} = \sum_{i,j=1}^n b_{ij}\overline{x_i}x_j, \quad \mathbf{x}\in\mathbb{C}^n$

$Q_{\mathbf{A}}(\mathbf{S}\mathbf{x}) = (\mathbf{S}\mathbf{x})^T\mathbf{A}(\mathbf{S}\mathbf{x}) = \mathbf{x}^T(\mathbf{S}^T\mathbf{A}\mathbf{S})\mathbf{x} = Q_{\mathbf{S}^T\mathbf{A}\mathbf{S}}(\mathbf{x})$

$H_{\mathbf{B}}(\mathbf{S}\mathbf{x}) = (\mathbf{S}\mathbf{x})^\dagger\mathbf{B}(\mathbf{S}\mathbf{x}) = \mathbf{x}^\dagger(\mathbf{S}^\dagger\mathbf{B}\mathbf{S})\mathbf{x} = H_{\mathbf{S}^\dagger\mathbf{B}\mathbf{S}}(\mathbf{x})$

$$if\,\exists\,\mathbf{S} : \mathbf{B} = \mathbf{S}\mathbf{A}\mathbf{S}^\dagger \qquad \mathbf{B} \text{ is } {}^*congruent (\text{“star–congruent”}) \text{ to } \mathbf{A} \tag{68}$$

$$if\,\exists\,\mathbf{S} : \mathbf{B} = \mathbf{S}\mathbf{A}\mathbf{S}^T \qquad \mathbf{B} \text{ is } {}^T congruent (\text{“tee–congruent”}) \text{ to } \mathbf{A} \tag{69}$$

Congruence is an equivalence relation.

$$\mathbf{A} = \mathbf{A}^\dagger \rightarrow \mathbf{S}\mathbf{A}\mathbf{S}^\dagger = (\mathbf{S}\mathbf{A}\mathbf{S}^\dagger)^\dagger \quad \text{congruence preserves type of matrix} \tag{70}$$

$$\mathbf{A} \text{ congruent to } \mathbf{B} \rightarrow \begin{cases} \operatorname{rank}\mathbf{A} = \operatorname{rank}\mathbf{B} \\ \mathbf{B} \text{ congruent to } \mathbf{A} \end{cases} \tag{71}$$

Inertia of Hermitian $\mathbf{A} : i(\mathbf{A}) = (i_+(\mathbf{A}), i_-(\mathbf{A}), i_0(\mathbf{A}))$, where i_+ is the number of positive eigenvalues, i_- of negative eigenvalues, and i_0 is the number of zero eigenvalues.
Signature of Hermitian $\mathbf{A} : i_+(\mathbf{A}) - i_-(\mathbf{A})$
Inertia matrix: $\mathbf{I}(\mathbf{A}) = diag(1_1, 1_2, \ldots, 1_{i_+}, -1_{i_++1}, \ldots, -1_{i_++i_-}, 0_{i_++i_-+1}, \ldots, 0_n)$
Sylvester's law of inertia: $\mathbf{A},\mathbf{B}$ Hermitian $\rightarrow \exists\,\mathbf{S} : \mathbf{A} = \mathbf{S}\mathbf{B}\mathbf{S}^\dagger \iff i(\mathbf{A}) = i(\mathbf{B}) \rightarrow \mathbf{A}$ is congruent to **B**.
$\mathbf{A},\mathbf{B}$ are *simultaneously diagonalizable by congruence* if $\exists\,\mathbf{U}$ unitary : $\mathbf{U}\mathbf{A}\mathbf{U}^\dagger$ and $\mathbf{U}\mathbf{B}\mathbf{U}^\dagger$ are both diagonal, or $\exists\,\mathbf{S}$: $\mathbf{S}\mathbf{A}\mathbf{S}^\dagger$ and $\mathbf{S}\mathbf{B}\mathbf{S}^\dagger$ are both diagonal. The same works for symmetric matrices with the transpose operator.

5.4 Consimilarity and condiagonalization

$\mathbf{A},\mathbf{B}$ are *consimilar* if $\exists\,\mathbf{S}$: $\mathbf{A} = \mathbf{S}\mathbf{B}\overline{\mathbf{S}}^{-1}$. If **S** can be taken to be unitary, **A** and **B** are *unitarily consimilar*. Special cases of consimilarity include ^{T}congruence, *congruence, and ordinary similarity. Consimilarity is an equivalence relation.
A is *contriangularizable* if $\exists\,\mathbf{S}$: $\mathbf{S}^{-1}\mathbf{A}\overline{\mathbf{S}}$ is upper triangular. **A** is *condiagonalizable* if **S** can be chosen so that $\mathbf{S}^{-1}\mathbf{A}\overline{\mathbf{S}}$ is diagonal.
A is unitarily condiagonalizable $\iff$ **A** is symmetric.

$\mathbf{A}\bar{\mathbf{x}} = \lambda\mathbf{x}$ — $\mathbf{x}$ is an *coneigenvector* of $\mathbf{A}$, corresponding to the *coneigenvalue* λ. A matrix may have infinitely many distinct coneigenvalues or it may have no coneigenvalues at all.

$$\mathbf{A}\overline{\mathbf{A}} = \mathbf{I} \iff \exists\, \mathbf{S}: \ \mathbf{A} = \mathbf{S}\overline{\mathbf{S}}^{-1} \tag{72}$$

$\forall\, \mathbf{A} \in \mathbf{M}_n$: $\mathbf{A}$ is consimilar to $\overline{\mathbf{A}}, \mathbf{A}^\dagger, \mathbf{A}^T$, to a Hermitian matrix, and to a real matrix.
$\forall\, \mathbf{A} \in \mathbf{M}_n$: $\exists\, \mathbf{S}_1, \mathbf{S}_2$ (symmetric), $\mathbf{H}_1, \mathbf{H}_2$ (Hermitian) such that $\mathbf{A} = \mathbf{S}_1\mathbf{H}_1 = \mathbf{H}_2\mathbf{S}_2$

6 Norms for vectors and matrices

Let $\mathbf{V}$ be a vector space over a field $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$).
A function $||\cdot|| : \mathbf{V} \to \mathbb{R}$ is a *vector norm* if $\forall\, \mathbf{x}, \mathbf{y} \in \mathbf{V}$:

$$\left\{ \begin{array}{ll} ||\mathbf{x}|| \geq 0 & \text{Nonnegative} \\ ||\mathbf{x}|| = 0 \iff \mathbf{x} = 0 & \text{Positive} \\ ||c\mathbf{x}|| = |c|\,||\mathbf{x}||\ \forall\, c \in \mathbb{F} & \text{Homogeneous} \\ ||\mathbf{x}+\mathbf{y}|| \leq ||\mathbf{x}|| + ||\mathbf{y}|| & \text{Triangle inequality} \end{array} \right\} \tag{73}$$

The *Euclidean norm* (or l_2 norm) on $\mathbb{C}^n$: $||\mathbf{x}||_2 = \sqrt{|x_1|^2 + \ldots + |x_n|^2}$
The *sum norm* (or l_1 norm) on $\mathbb{C}^n$: $||\mathbf{x}||_1 = |x_1| + \ldots + |x_n|$
The *max norm* (or l_∞ norm) on $\mathbb{C}^n$: $||\mathbf{x}||_\infty = \max\{|x_1| + \ldots + |x_n|\}$
The l_p *norm* on $\mathbb{C}^n$: $||\mathbf{x}||_p = \sqrt[p]{\sum_{i=1}^{n} |x_1|^p}$
A function $< \cdot, \cdot >$: $\mathbf{V} \times \mathbf{V} \to \mathbb{F}$ is an *inner product* if $\forall\, \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbf{V}$:

$$\left\{ \begin{array}{ll} <\mathbf{x},\mathbf{x}> \geq 0 & \text{Nonnegative} \\ <\mathbf{x},\mathbf{x}> = 0 \iff \mathbf{x} = 0 & \text{Positive} \\ <\mathbf{x}+\mathbf{y},\mathbf{z}> = <\mathbf{x},\mathbf{z}> + <\mathbf{y},\mathbf{z}> & \text{Additive} \\ <c\mathbf{x},\mathbf{y}> = c <\mathbf{x},\mathbf{y}>\ \ \forall\, c \in \mathbb{F} & \text{Homogeneous} \\ <\mathbf{x},\mathbf{y}> = \overline{<\mathbf{y},\mathbf{x}>} & \text{Hermitian property} \end{array} \right\} \tag{74}$$

A function $|||\cdot||| : \mathbf{M}_n \to \mathbb{R}$ is a *matrix norm* if $\forall\, \mathbf{A}, \mathbf{B} \in \mathbf{M}_n$:

$$\left\{ \begin{array}{ll} |||\mathbf{A}||| \geq 0 & \text{Nonnegative} \\ |||\mathbf{A}||| = 0 \iff \mathbf{A} = 0 & \text{Positive} \\ |||c\mathbf{A}||| = |c|\,|||\mathbf{A}|||\ \forall\, c & \text{Homogeneous} \\ |||\mathbf{A}+\mathbf{B}||| \leq |||\mathbf{A}||| + |||\mathbf{B}||| & \text{Triangle inequality} \\ |||\mathbf{A}\mathbf{B}||| \leq |||\mathbf{A}|||\,|||\mathbf{B}||| & \text{Submultiplicative} \end{array} \right\} \tag{75}$$

The *maximum column sum matrix norm* $|||\cdot|||_1$ on $\mathbf{M}_n$: $|||\mathbf{A}|||_1 = \max\limits_{1\leq j\leq n} \sum_{i=1}^{n} |a_{ij}|$

The *maximum row sum matrix norm* $|||\cdot|||_\infty$ on $\mathbf{M}_n$: $|||\mathbf{A}|||_\infty = \max\limits_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}|$

The *spectral norm* $|||\cdot|||_2$ on $\mathbf{M}_n$: $|||\mathbf{A}|||_2 = \max\{\sqrt{\lambda} : \lambda \text{ is an eigenvalue of } \mathbf{A}^\dagger\mathbf{A}\}$
General properties:

$$\begin{array}{l} |<\mathbf{x},\mathbf{y}>|^2 \leq <\mathbf{x},\mathbf{x}><\mathbf{y},\mathbf{y}> \\ ||\mathbf{x}|| = \sqrt{<\mathbf{x},\mathbf{x}>} \\ \rho(\mathbf{A}) \leq |||\mathbf{A}||| \end{array} \tag{76}$$

7 Positive Definite Matrices

$$\mathbf{A} \text{ positive definite} \quad \mathbf{x}^\dagger\mathbf{A}\mathbf{x} > 0 \quad \forall \mathbf{x} \in \mathbb{C}^n \neq \mathbf{0} \tag{77}$$

$$\mathbf{A} \text{ positive semidefinite} \quad \mathbf{x}^\dagger\mathbf{A}\mathbf{x} \geq 0 \quad \forall \mathbf{x} \in \mathbb{C}^n \neq \mathbf{0} \tag{78}$$

Similarly are the terms *negative definite* and *negative semidefinite* defined. If a Hermitian matrix is neither positive, nor negative semidefinite, it is called *indefinite*.
Any principal submatrix of a positive definite matrix is positive definite.

$$\mathbf{A} \text{ positive definite} \to \left\{ \begin{array}{l} \sigma(\mathbf{A}) > 0 \\ \operatorname{trace}\mathbf{A} > 0 \\ \det\mathbf{A} > 0 \end{array} \right. \tag{79}$$

$$\begin{aligned}
\mathbf{A}\text{ positive definite} &\iff \sigma(\mathbf{A}) > 0 && (80)\\
\mathbf{A}\text{ positive def. Hermitian} &\iff \det(\mathbf{A}) > 0 \quad \mathbf{A}\text{ Hermitian} && (81)\\
\mathbf{A}\text{ positive definite} &\iff \exists\, \mathbf{C} : \mathbf{A} = \mathbf{C}^\dagger\mathbf{C} \quad \mathbf{C}\text{ nonsingular} && (82)\\
\mathbf{A}\text{ positive definite} &\iff \exists\, \mathbf{L} : \mathbf{A} = \mathbf{L}\mathbf{L}^\dagger \quad \mathbf{L}\text{ nonsingular lower} && (83)\\
& \quad\ \text{triangular with positive diag elements} && (84)\\
\mathbf{A}\text{ positive semidefinite} &\iff \sigma(\mathbf{A}) \geq 0 && (85)\\
\mathbf{A}\text{ positive definite} &\iff \mathbf{A}^{-1}\text{ positive definite} && (86)\\
\mathbf{A}\text{ positive definite} &\iff \mathbf{A}^k\text{ positive definite} \quad \forall\, k \in \mathbb{N} > 0 && (87)\\
& && (88)
\end{aligned}$$

8 Matrices

Block diagonal matrix $\iff$ *diagonal block matrix*, is a square diagonal matrix in which the diagonal elements are square matrices of any size (possibly even 1x1), and the off-diagonal elements are 0. A block diagonal matrix is therefore a block matrix in which the blocks off the diagonal are the zero matrices, and the diagonal matrices are square.
The determinant of the a block diagonal matrix is the product of the determinants of the diagonal elements.
Square root of a matrix. This matrix has to be hermitian, positive semi–definite.
Orthogonal matrix Q : $\mathbf{Q}\mathbf{Q}^T = \mathbf{I}$
Symmetric matrix A : $\mathbf{A} = \mathbf{A}^T$
Skew–Symmetric matrix A : $\mathbf{A} = -\mathbf{A}^T$
Hermitian matrix A : $\mathbf{A} = \mathbf{A}^\dagger$
Skew–Hermitian matrix A : $\mathbf{A} = -\mathbf{A}^\dagger$
Positive semidefinite A $\leftarrow \mathbf{A} = \mathbf{A}^\dagger, \sigma(\mathbf{A}) \in \mathbb{R}^{>=0}$
Scalar matrix A : $\mathbf{A} = a\mathbf{I}, \quad a \in \mathbb{C}$

9 Rank and Range VS. Nullity and Null Space

The rank of a matrix $\mathbf{A}_{n\times m}$ is the smallest value r for which exist $\mathbf{F}_{n\times r}$ and $\mathbf{G}_{r\times m}$ so that $\mathbf{A}_{n\times m} = \mathbf{F}_{n\times r}\mathbf{G}_{r\times m}$:

$$\operatorname{rank}\mathbf{A}_{n\times m} = \operatorname{rank}\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = r \qquad \text{with } \mathbf{A}_{n\times m} = \mathbf{F}_{n\times r}\mathbf{G}_{r\times m} \tag{89}$$

$$\text{Examples:} \quad \begin{array}{lll} \text{rank } \mathbf{T}_6 = & \mathbf{k}\mathbf{k}^\dagger & = 1 \\ \text{rank } \mathbf{T}_6 = & \langle \mathbf{k}\mathbf{k}^\dagger \rangle & \geq 1 \end{array} \qquad \mathbf{T}_6 \to \mathbf{A}_{6\#6}, \quad \mathbf{k} \to \mathbf{F}_{6\#1} \tag{90}$$

$$\operatorname{rank}\mathbf{A}_{n\times m} \begin{cases} = & \text{number of linearly independent vectors in } \mathbf{A} \\ \leq & \min(m,n) \\ = & \operatorname{rank}\mathbf{A}^T = \operatorname{rank}\mathbf{A}^\dagger \\ = & \text{dimension of the range of } \mathbf{A} \\ = & \operatorname{rank}\mathbf{A}\mathbf{X} = \operatorname{rank}\mathbf{Y}\mathbf{A} \qquad \mathbf{X},\mathbf{Y} \text{ non-singular} \\ = & m - \text{nullity } \mathbf{A}_{n\times m} \end{cases} \tag{91}$$

$$\operatorname{rank}\mathbf{A}_{n\times n} \begin{cases} = & n - \text{nullity } \mathbf{A}_{n\times n} \\ < & n \iff \det\mathbf{A}_{n\times n} = 0 \iff |\mathbf{A}| = 0 \end{cases} \tag{92}$$

Range The range (or image) of $\mathbf{A}_{n\times m} \in \mathbf{F}^{n\times m}$ is the subspace of vectors that equal $\mathbf{A}\mathbf{x}$ for all $\mathbf{x} \in \mathbf{F}^m$. The dimension of this subspace is the *rank* of $\mathbf{A}$.

$$\text{range } \mathbf{A}_{n\times m} = \{\mathbf{A}\mathbf{x} : \mathbf{x} \in \mathbf{F}^m\} \tag{93}$$

Null space and Nullity The null space (or kernel) of $\mathbf{A}_{n\times m} \in \mathbf{F}^{n\times m}$ is the subspace of vectors $\mathbf{x} \in \mathbf{F}^m$ for which $\mathbf{A}\mathbf{x} = \mathbf{0}$. The dimension of this subspace is the *nullity* of $\mathbf{A}$.

$$\text{null space } \mathbf{A}_{n\times m} = \{\mathbf{x} \in \mathbf{F}^m : \mathbf{A}\mathbf{x} = \mathbf{0}\}$$

The range of $\mathbf{A}$ is the orthogonal complement of the null space of $\mathbf{A}^\dagger$ and vice versa (the null space of $\mathbf{A}$ is the orthogonal complement of the range of $\mathbf{A}^\dagger$).
The rank of a matrix plus the nullity of the matrix equals the number of columns of the matrix.

10 Diverse

10.1 Basis Transformation

$$\begin{aligned} \text{Vector: } \mathbf{v} &= \mathbf{Mu} && (94) \\ \text{Matrix: } \mathbf{V} &= \mathbf{MUM}^\dagger && (95) \end{aligned}$$

If $\mathbf{M} \in \mathbf{R}^{n\times n}$ is real symmetric, then is $\omega^\dagger \mathbf{M} \omega$ also real ($\omega \in \mathbf{C}^n$).

10.2 Nilpotent Matrix

A matrix $\mathbf{A}$ is *nilpotent* to index k if $\mathbf{A}^k = \mathbf{0}$, but $\mathbf{A}^{k-1} \neq \mathbf{0}$.

$$\mathbf{A} \text{ nilpotent at index } k \rightarrow \begin{cases} \det \mathbf{A} = 0 \\ \sigma(\mathbf{A}) = \{0\} \\ \text{minimal polynomial of } \mathbf{A} \text{ is } t^k \end{cases} \quad (96)$$

11 Groups

A *group* is a set that is closed under a single associative binary operation (multiplication) and such that the identity for and inverses under the operation are contained in the set.

1. The nonsingular matrices in $\mathbf{M}_n(\mathbb{F})$ form the *general linear* group — $GL(n,\mathbb{F})$
2. The set of unitary (respectively real orthogonal) matrices in $\mathbf{M}_n$ form the *n-by-n unitary* group (subgroup of $GL(n,\mathbb{C})$).

12 Diverse

12.1 Joint Diagonalization

The joint diagonalization of a set of square matrices consists in finding the orthonormal change of basis which makes the matrices as diagonal as possible. When all the matrices in the set commute, this can be achieved exactly. When this is not the case, it is always possible to optimize a joint diagonality criterion. This defines an approximate joint diagonalization. When the matrices in the set are ‘almost exactly jointly diagonalizable’, this approach also defines something like the ‘average eigen-spaces’ of the matrix set.

A Example 2×2 Matrix

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{C}^{2\times 2} && (97) \\ p_{\mathbf{A}}(t) &= t^2 - (a+d)t + (ad-bc) && (98) \\ \sigma(\mathbf{A}) &= \left\{ \frac{a+d \pm \sqrt{(a-d)^2 + 4bc}}{2} \right\} && (99) \\ \det \mathbf{A} &= ad - bc && (100) \end{aligned}$$

B IDL–SAR convention: mathematical notation!

Always mathematically correct!!! First index: row number, second: column number. $\mathbf{M}_{row,col} = \mathbf{C}[row,col]$

$$\mathbf{A}_{n\times m} = \begin{bmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{bmatrix} = a[y,x]$$

Always: transpose the given mathematical matrix after generation!!! (no complex conjugate!)

Matrix $\mathbf{M} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad \mathbf{M}_{1,2} = b, \; \mathbf{M}_{2,1} = c$

In IDL: $\mathbf{M} \iff \mathbf{C} = \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \mathbf{M}^T$, so that $\mathbf{C}[1,2] = b,\ \mathbf{C}[2,1] = c$

Examples:

$$\begin{array}{llll}
\mathbf{A}_r & \iff & \text{IDL: } \mathbf{A} \leftarrow \mathbf{A}_r^T & \text{\$ A=transpose(Ar)} \\
\mathbf{C}_r = \mathbf{A}_r\mathbf{B}_r & \iff & \text{IDL: } \mathbf{C} \leftarrow \mathbf{C}_r^T = (\mathbf{A}_r\mathbf{B}_r)^T & \text{\$ C=transpose(Ar \#\# Br)} \\
 & & = (\mathbf{A}^T\mathbf{B}^T)^T = (\mathbf{BA}) & \text{\$ C=B\#\#A=}\mathbf{A\#B} \\
\mathbf{C}_r = \mathbf{A}_r\mathbf{B}_r^\dagger & \iff & \mathbf{C} \leftarrow \mathbf{C}_r^T = (\mathbf{A}_r\mathbf{B}_r^\dagger)^T & \\
 & & = (\mathbf{A}^T\overline{\mathbf{B}})^T = \mathbf{B}^\dagger\mathbf{A} & \text{\$ C=A\#adj(B)} \\
\mathbf{C}_r = \mathbf{A}_r^T\mathbf{B}_r & \iff & \mathbf{C} \leftarrow \mathbf{C}_r^T = (\mathbf{A}_r^T\mathbf{B}_r)^T & \\
 & & = (\mathbf{AB}^T)^T = \mathbf{BA}^T & \text{\$ C=transpose(A)\#B} \\
\mathbf{C}_r = \mathbf{k}_{1r}\mathbf{k}_{2r}^\dagger & \iff & \mathbf{C} = \mathbf{C}_r^T = (\mathbf{k}_{1r}\mathbf{k}_{2r}^\dagger)^T & \\
 & & = (\mathbf{k}_1^T\overline{\mathbf{k}}_2)^T = \mathbf{k}_2^\dagger\mathbf{k}_1 & \text{\$ C=k1\#adj(k2)} \\
= \mathbf{A}_r\mathbf{B}_r\mathbf{C}_r & \iff & (\mathbf{A}_r\mathbf{B}_r\mathbf{C}_r)^T & \\
 & & = (\mathbf{A}^T\mathbf{B}^T\mathbf{C}^T)^T = \mathbf{CBA} & \text{\$ A\#B\#C} \\
\mathbf{A}_r\mathbf{k}_r & \iff & (\mathbf{A}_r\mathbf{k}_r)^T = (\mathbf{A}^T\mathbf{k}^T)^T = \mathbf{kA} & \text{\$ A\#k} \\
\mathbf{A}_r^{-1} & \iff & (\mathbf{A}_r^{-1})^T = (\mathbf{A}_r^T)^{-1}? &
\end{array} \tag{101}$$

Short summary : every matrix and vector should be considered as transposed, in relation between math and IDL.
In mathematics first index is the y–row–index, the second is the x–column-index! In IDL its reversed: first x–column– and then y–row–index.
Using these notations one can use all kinds of matrix multiplications and transpose/conjugate in the mathematical order with **#**. (e.g. in mathematics $\mathbf{w}^\dagger\mathbf{T}\mathbf{w}$ will correspond in IDL to $adj(w)\#T\#w$, but it should be considered that w and T are transposed: $T = \mathbf{T}^T$ and $w = \mathbf{w}^T$. i.e. w is a row–array during $\mathbf{w}$ is a column–vector.)

Matrix input/output in IDL : To give a matrix from math–book or paper into IDL you have to transpose it (before input or right after). To read out the correct values, you have also to transpose the IDL matrix. Since the mathematical matrix indexing and the IDL matrix indexing are transposed and we also transpose the matrix between these systems, *the values have the same indices* in mathematical matrices with mathematical indexing as in IDL matrices with IDL indexes! (i.e. $\mathbf{A}(1,2)_m = A[1,2]_{IDL}$) Consider these (right!) cases: $T6 = \mathbf{T}_6^T$, $\mathbf{T}_{12} = \mathbf{T}_6(0:2,3:5)_m = \mathbf{T}_6[3:5,0:2]_{IDL}$, $T12 = T6[0:2,3:5]_{IDL} = T6(3:5,0:2)_m$, $\mathbf{T}_{12} = transp(T12)$.

Conclusion : Transpose matrices by input/output! And everything else can be used as in mathematical notation: order, matrix multiplication (with #), indices, transpose, conjugate, etc...

C Notation Conventions

Capital bold letters reference to matrices, small bold letter to vectors.
$\mathbf{A}$ is a matrix with dimensions $n \times n$ if no other dimensions are supplied.
^{T}: transpose
†: transpose conjugate complex.
*: conjugate complex.

D Mathematical Glossary

inf: is the infimum of a set (max over a set with bounds)
sup: is the supremum of a set (min over a set with bounds)
$\inf(S) = -\sup(-S)$
Hadamard Product: $\mathbf{A} \circ \mathbf{B}$ is the element-wise multiplication of matrices.

References

[1] G.H. Golub and C.F. Van Loan. *Matrix Computation*, volume 3. The Johns Hopkins University Press, 1996. recommended by Hongwei Zheng.

[2] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1990.

[3] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1991.

[4] G Strang. *Introduction to Linear Algebra*. Wellesley Cambridge Press, 1998.