Review of Matrix Theory

Notations:

A – capital bold denotes a matrix;

a – lower case bold is a vector;

a – lower case regular is a scalar;

a_{ij} - ij-element of $\mathbf{A}$;

$\det(\mathbf{A})$ - determinant of $\mathbf{A}$;

$tr(\mathbf{A})$ - a trace of $\mathbf{A}$;

Basics

Matrix **A** is defined by its elements a_{ij}:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \quad (1)$$

Sometimes, elements of **A** are denoted as $[\mathbf{A}]_{ij}$.

Sum of 2 matrices is defined element-wise:

$$\mathbf{C} = \mathbf{A} + \mathbf{B} \quad \rightarrow \quad c_{ij} = a_{ij} + b_{ij} \quad (2)$$

Product of matrices is defined as:

$$\mathbf{C} = \mathbf{A} \cdot \mathbf{B} \rightarrow \quad c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \quad (3)$$

Note that the product of **A** and **B** is defined only if the number of columns of **A** is the same as the number of rows of **B**, i.e. **A** and **B** are $m \times n$ and $n \times l$ matrices.

Determinant of a square $n \times n$ matrix det(**A**):

$$\det(\mathbf{A}) = |\mathbf{A}| = \sum_{k+1}^{n} a_{ik} (-1)^{i+k} \mathbf{M}_{ik} \quad (4)$$

where $\mathbf{M}_{ik}$ is the minor of a_{ik}, i.e. the determinant of the submatrix of **A**, which is obtained by deleting *i*-th row and *k*-th column from **A**.

Example:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \rightarrow \quad \det(\mathbf{A}) = a_{11}a_{22} - a_{21}a_{12} \quad (5)$$

The transpose of **A** is defined as

$$\mathbf{B} = \mathbf{A}^T \quad \rightarrow \quad b_{ij} = a_{ji} \quad (6)$$

i.e. row and column indexes are exchanged.

Complex conjugate operation is applied element-wise:

$$\mathbf{B} = \mathbf{A}^* \quad \rightarrow \quad b_{ij} = a_{ij}^* \quad (7)$$

The Hermitian conjugate of A is

$$\mathbf{B} = \mathbf{A}^+ = \left(\mathbf{A}^T\right)^* \quad \rightarrow \quad b_{ij} = a_{ji}^* \quad (8)$$

Product of a matrix **A** and a scalar **c** is defined element-wise:

$$\mathbf{B} = c \cdot \mathbf{A} \rightarrow b_{ij} = c \cdot a_{ij} \quad (9)$$

Some properties of transpose:

$$(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T, (\mathbf{AB})^+ = \mathbf{B}^+\mathbf{A}^+ \tag{10}$$

Properties of det:

$$\begin{aligned} &\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B}); \quad \det(c\cdot\mathbf{A}) = c^n\det(\mathbf{A}) \\ &\det(\mathbf{A}^T) = \det(\mathbf{A}); \quad \det(\mathbf{A}^+) = (\det(\mathbf{A}))^*; \end{aligned} \tag{11}$$

for square **A** and **B**. If det(**A**)=0, **A** is called singular.

Trace of a matrix is the sum of diagonal elements:

$$tr(\mathbf{A}) = \sum_{i=1}^{n} a_{ii} \tag{12}$$

Some properties of trace:

$$\begin{aligned} &tr(\mathbf{A}+\mathbf{B}) = tr(\mathbf{A}) + tr(\mathbf{B}) \\ &tr(\mathbf{AB}) = tr(\mathbf{BA}) \\ &tr(\mathbf{ABC}) = tr(\mathbf{CAB}) = tr(\mathbf{BCA}) \end{aligned} \tag{13}$$

Rank of a matrix is the number of linearly independent columns or rows. Some properties:

$$\begin{aligned} &rank(\mathbf{A}+\mathbf{B}) \le rank(\mathbf{A}) + rank(\mathbf{B}) \\ &rank(\mathbf{AB}) \le \min(rank(\mathbf{A}), rank(\mathbf{B})) \end{aligned} \tag{14}$$

Vector **a** is a $n \times 1$ matrix:

$$\mathbf{a} = [a_1 \quad a_2 \quad .. \quad a_n]^T \tag{15}$$

Sometimes it is called column vector.

Scalar product of two vectors **a** and **b**:

$$\mathbf{a}^+\mathbf{b} = \sum_{i=1}^{n} a_i^* b_i \tag{16}$$

Frobenius or Eucledian norm (length) of a vector is:

$$|\mathbf{a}| = \sqrt{\mathbf{a}^+\mathbf{a}} = \sqrt{\sum_{i=1}^{n} |a_i|^2} \tag{17}$$

Similarly, Frobenius norm of a matrix:

$$\|\mathbf{A}\| = \left(\sum_{i=1}^{n}\sum_{j=1}^{m} |a_{ij}|^2\right)^{1/2} = \sqrt{tr(\mathbf{A}^+\mathbf{A})} \tag{18}$$

Inverse of a $n \times n$ matrix:

$$\mathbf{B} = \mathbf{A}^{-1} \quad \text{if} \quad \mathbf{AB} = \mathbf{BA} = \mathbf{I} \tag{19}$$

I - identity matrix, $[\mathbf{I}]_{ij} = \delta_{ij} = 1$ if i=j, 0 otherwise.

If rank(**A**)<*n*, then det(**A**)=0 and the inverse does not exist → **A** is singular.

Some properties of the inverse:

$$
\begin{aligned}
(\mathbf{AB})^{-1} &= \mathbf{B}^{-1}\mathbf{A}^{-1} \\
\det\left(\mathbf{A}^{-1}\right) &= 1/\det(\mathbf{A}) \\
\left(\mathbf{A}^{T}\right)^{-1} &= \left(\mathbf{A}^{-1}\right)^{T} \\
\left(\mathbf{A}^{+}\right)^{-1} &= \left(\mathbf{A}^{-1}\right)^{+}
\end{aligned} \tag{20}
$$

if all the inverses exist.

The inverse of **A** can be calculated as

$$\mathbf{A}^{-1} = \frac{\mathbf{C}^{T}}{\det(\mathbf{A})}, \quad c_{ij} = (-1)^{i+j}\mathbf{M}_{ij} \tag{21}$$

where **M** is the minor as before.

The matrix inversion lemma (MIL):

$$(\mathbf{A}+\mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}\left(\mathbf{DA}^{-1}\mathbf{B}+\mathbf{C}^{-1}\right)^{-1}\mathbf{DA}^{-1} \tag{22}$$

where **A** is $n \times n$, **B** is $n \times m$, **C** is $m \times m$, **D** is $m \times n$ and all the inverses are assumed to exist.

A special case of (22) is Woodbury's identity:

$$\left(\mathbf{A}+\mathbf{xx}^{+}\right)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{xx}^{+}\mathbf{A}^{-1}}{1+\mathbf{x}^{+}\mathbf{A}^{-1}\mathbf{x}} \tag{23}$$

Note: the product $\mathbf{xx}^{+}$ is defined as

$$\mathbf{B} = \mathbf{xx}^{+} \rightarrow b_{ij} = x_i x_j^{*} \tag{24}$$

i.e. element-wise.

Some special matrices

Symmetric matrix:

$$\mathbf{A} = \mathbf{A}^{T} \rightarrow a_{ij} = a_{ji} \tag{25}$$

Hermitian matrix:

$$\mathbf{A} = \mathbf{A}^{+} \rightarrow a_{ij} = a_{ji}^{*} \tag{26}$$

Unitary matrix:

$$\mathbf{UU}^{+} = \mathbf{I} = \mathbf{U}^{+}\mathbf{U} \rightarrow \mathbf{U}^{-1} = \mathbf{U}^{+} \tag{27}$$

Columns of a unitary matrix are orthogonal, $u_i^{+} u_j = \delta_{ij}$.

Diagonal matrix **A**:

$$a_{ij} = 0 \quad \text{if } i \neq j;\ \mathbf{A} = diag\left(a_{11},\ a_{22}\ ..\ a_{nn}\right) \tag{28}$$

Positive definite matrix:

$$\text{if} \quad \mathbf{x}^{+}\mathbf{Ax} > 0 \quad \forall \mathbf{x} \neq 0 \tag{29}$$

Positive semi-definite matrix:

$$\text{if} \quad \mathbf{x}^{+}\mathbf{Ax} \geq 0 \quad \forall \mathbf{x} \neq 0 \tag{30}$$

If a matrix is (semi)positive-definite, it is also Hermitian. The converse is not true in general.

Projection Matrices

Projection (indempotent) matrix:

$$\mathbf{P}^2 = \mathbf{P} \quad (31)$$

Further, we consider only Hermitian projection matrices, $\mathbf{P}^+ = \mathbf{P}$.

Consider a linear vector space spanned by the columns of $n \times m$ matrix **V**,

$$\mathbf{S}=\text{span}\,(\mathbf{V}) \quad (32)$$

Assume columns of **V** are linearly-independent. Projection of **x** onto **S** is

$$\mathbf{x}_S = \mathbf{Px} \quad , \quad \mathbf{P} = \mathbf{V}\left(\mathbf{V}^+\mathbf{V}\right)^{-1}\mathbf{V}^+ \quad (33)$$

Projection of **x** onto $S_\perp$ is

$$\mathbf{x}_{S\perp} = \mathbf{P}_\perp \mathbf{x} \quad , \quad \mathbf{P}_\perp = \mathbf{I} - \mathbf{P} \quad (34)$$

where $\mathbf{S}_\perp$ is the space orthogonal to **S**.

Eigenvalue Decomposition

Eigenvector of a $n \times n$ matrix:

$$\mathbf{Au} = \lambda\mathbf{u} \;\rightarrow\; (\mathbf{A} - \lambda\mathbf{I})\mathbf{u} = 0 \quad (35)$$

where λ is an eigenvalue. Eigenvectors give "invariant" directions if **A** is considered as linear transformation.

Solution to

$$|\mathbf{A} - \lambda\mathbf{I}| = 0 \quad (36)$$

gives n eigenvalues λ. There are n orthonormal eigenvectors. Define:

$$\mathbf{U} = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \ldots \quad \mathbf{u}_n], \quad \mathbf{UU}^+ = \mathbf{I}$$
$$\mathbf{\Lambda} = diag[\lambda_1 \quad \lambda_2 \quad \ldots \quad \lambda_n]$$

Then,

$$\mathbf{A} = \mathbf{U\Lambda U}^+ = \sum_{i=1}^{n} \lambda_i \mathbf{u}_i \mathbf{u}_i^+ \quad (37)$$

This is eigenvalue decomposition of **A**.

Some properties

$$tr(\mathbf{A}) = \sum_{i=1}^{n} \lambda_i \quad (38)$$

$$\det(\mathbf{A}) = \prod_{i=1}^{n} \lambda_i \quad (39)$$

$$\mathbf{A}^{-1} = \mathbf{U\Lambda}^{-1}\mathbf{V}^+ = \sum_{i=1}^{n} \lambda_i^{-1} \mathbf{u}_i \mathbf{v}_i^+ \quad (40)$$

Let $\lambda(\mathbf{A})$ denotes the eigenvalues of **A**. Then:

$$\lambda(c\mathbf{A}) = c\lambda(\mathbf{A}), \quad c\text{ - scalar} \quad (41)$$
$$\lambda(\mathbf{A}^m) = \lambda^m(\mathbf{A}), \quad \text{m=1, 2, 3....} \quad (42)$$

If **A** is Hermitian, $\mathbf{A} = \mathbf{A}^+$, then $\text{Im}\{\lambda(\mathbf{A})\} = 0$. If **A** is positive definite, then $\lambda_i(\mathbf{A}) > 0$.

$$\lambda_i(\mathbf{I}) = 1, \quad i = 1,2...n. \quad (43)$$
$$\lambda_i(\mathbf{P}) = 1, \quad i = 1...k, \quad \lambda_i(\mathbf{P}) = 0, \quad i = k+1....n \quad (44)$$

where **P** is a projection matrix onto k-dimentional space.

$$\lambda_{\max}(\mathbf{A}) = \max_{\mathbf{x}} \left[\mathbf{x}^{+}\mathbf{A}\mathbf{x} \right], \ |\mathbf{x}| = 1 \qquad (45)$$

If **A** is singular, $\text{rank}(\mathbf{A}) = k < n$, "preudoinverse" can be defined using non-zero eigenvalues only,

$$\mathbf{A}^{\sim 1} = \sum_{i=1}^{K} \lambda_i^{-1} \mathbf{u}_i \mathbf{u}_i^{+}$$

Singular Value Decomposition

Arbitrary $n \times m$ matrix **A** can be decomposed as

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{+} = \sum_{i=1}^{l} \sigma_i \mathbf{u}_i \mathbf{v}_i^{+} \qquad (46)$$

where **U**, **V** are unitary $n \times n$ and $m \times m$ matrices, and Σ is $n \times m$ matrix,

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \qquad \mathbf{\Sigma}_1 = diag(\sigma_1, \sigma_2 ... \sigma_l) \qquad (47)$$

where $\sigma_i \geq 0$ are singular values of **A**, and $\mathbf{u}_i$ are the columns of **U** (the left singular vectors of **A**), $\mathbf{v}_i$ are the columns of **V** (the right singular vectors of **A**).

Note: singular values of **A** are non-negative square roots of the eigenvalues of $\mathbf{A}\mathbf{A}^{+}$. The right singular vectors of **A** are the eigenvectors of $\mathbf{A}^{+}\mathbf{A}$. Note from (46) that

$$\mathbf{A}\mathbf{V}_K = \sigma_K \mathbf{u}_K, \ \mathbf{V}_K^{+}\mathbf{A} = \sigma_K \mathbf{v}_K^{+} \qquad (49)$$

Pseudoinverse $\mathbf{A}^{\sim 1}$ of a $m \times n$ matrix **A** for m>n is defined from the following

$$\mathbf{A}^{\sim 1}\mathbf{A} = \mathbf{I}_{n \times n} \qquad (50)$$

where $\mathbf{I}_{n \times n}$ is $n \times n$ identity matrix. Using the SVD of **A,**

$$\mathbf{A}^{\sim 1} = \sum_{i=1}^{n} \sigma_i^{-1} \mathbf{u}_i \mathbf{v}_i^{+} = \mathbf{U}\mathbf{\Sigma}^{-1}\mathbf{V}^{+} \qquad (51)$$

where

$$\mathbf{\Sigma}^{-1} = \begin{bmatrix} \mathbf{\Sigma}_1^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

$\mathbf{A}^{\sim 1}$ can be expressed as :

$$\mathbf{A}^{\sim 1} = \left(\mathbf{A}^{+}\mathbf{A}\right)^{-1} \mathbf{A}^{+} \qquad (52)$$

The above discussion assumes that **A** has the full column rank, i. e. linearly-independent columns. If n>m and **A** has full row rank, similar expressions hold true.

Pseudoinverse and projection matrix:

$$\mathbf{P}_A = \mathbf{A}\mathbf{A}^{\sim 1}, \quad \mathbf{P}_{\perp A} = \mathbf{I} - \mathbf{A}\mathbf{A}^{\sim 1} \qquad (53)$$

<u>Properties</u> of pseudoinverse:

$$\mathbf{A}\mathbf{A}^{\sim 1}\mathbf{A} = \mathbf{A}, \qquad \mathbf{A}^{\sim 1}\mathbf{A}\mathbf{A}^{\sim 1} = \mathbf{A}^{\sim 1} \qquad (54)$$

$$\left(\mathbf{A}^{+}\right)^{\sim 1} = \left(\mathbf{A}^{\sim 1}\right)^{+}$$

$$\left(\mathbf{A}^{+}\mathbf{A}\right)^{\sim 1} = \mathbf{A}^{\sim 1}\left(\mathbf{A}^{\sim 1}\right)^{+} \qquad (55)$$

$$\left(\mathbf{A}^{+}\mathbf{A}\right)^{\sim 1}\mathbf{A}^{+} = \mathbf{A}^{\sim 1}$$

If **B** is invertible , then

$$\left(\mathbf{BA}\right)^{\sim 1}\mathbf{B} = \mathbf{A}^{\sim 1} \qquad (56)$$

If **a** is a column vector, then

$$\mathbf{a}^{\sim 1} = \frac{\mathbf{a}^{+}}{\left|\mathbf{a}\right|^{2}} \qquad (57)$$

Miscellaneous

Let $a_{(i)}$ be i-th column of **A,** and $b_{(i)}^{T}$ be i-th row of **B**, then

$$\mathbf{AB} = \sum_{i=1}^{n} a_{(i)} b_{(i)}^{T} \qquad (58)$$

Null space of a matrix **A** is a set of vectors **x** that satisfy

$$\mathbf{Ax}=\mathbf{0} \qquad (59)$$

Range of a matrix **A** is a set of vectors **y** that satisfy

$$\mathbf{Ax}=\mathbf{y} \qquad (60)$$

for any **x**. Note that

$$\mathrm{rank}(\mathbf{A})=\dim(\mathbf{y}) \qquad (61)$$

where dim(**y**) is the dimensionality of the **y**. Additionally,

$$\dim(\mathbf{x})+\dim(\mathbf{y})=n \qquad (62)$$

for $n \times n$ matrix.

Important property:

$$\det\left(\mathbf{I}_{n\times n} + \mathbf{AB}\right) = \det\left(\mathbf{I}_{m\times m} + \mathbf{BA}\right) \qquad (63)$$

where **A**, **B** are $n \times m$ and $n \times m$ matrices, and $\mathbf{I}_{n\times n}$ is $n \times n$ identity matrix.

References

Brief reviews of matrices

1. H.L. Van Trees, Optimum Array Processing, Appendix 6.
2. D.H. Johnson, D.E. Dudgeon, Array Signal Processing, Appendix B
3. T.K. Moon, W.C. Stirling, Mathematical Methods and Algorithms for Signal Processing, Appendix C

Books

4. F. Zhang, Matrix theory, Springer, 1999.
5. J.N. Franklin, Matrix Theory, Dover, New York, 1993.
6. M. Marcus, H. Minc, A Survey of Matrix Theory and Matrix Inequalities, Dover, New York, 1992.

Comprehensive modern textbooks[1]

7. R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, 1985.
8. R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, 1994.

[1] strongly recommended to everybody interested in smart antennas, array processing, MIMO systems. Solid knowledge of matrix theory is essential for these fields.