Chapter 2

Operations on matrices

1 Equality

The matrices A and B are equal, that is A = B, if they are of the same order $m \times n$ and

$$a_{ij} = b_{ij}, \text{ for } 1 \leqslant i \leqslant m, \ 1 \leqslant j \leqslant n.$$

Example. The following matrices A and B are equal:

$$A = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}, \ B = \begin{bmatrix} 3 & -2 \\ e^{i\pi} & 1 \end{bmatrix}.$$

2 Transposition

2.1 Definition

The transpose of a matrix A is the matrix denoted by A^T obtained by interchanging the rows and columns of A. The elements of the transpose A^T are a_{ji}. Thus given the $m \times n$ matrix A, its transpose is the $n \times m$ matrix A^T as follows

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \Longrightarrow A^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}.$$

Example: Consider the matrix

$$A = \begin{bmatrix} -1 & 2 & 4 \\ 0 & 5 & 10 \end{bmatrix},$$

then its transpose is

$$A^T = \begin{bmatrix} -1 & 0 \\ 2 & 5 \\ 4 & 10 \end{bmatrix},$$

The transpose of a column vector is a row vector, and vice-versa (conversely).
In the case of block matrices, we have

$$M = \begin{bmatrix} A & D \\ C & B \end{bmatrix} \Longrightarrow M^T = \begin{bmatrix} A^T & C^T \\ D^T & B^T \end{bmatrix}.$$

2.2 Implications

a) Basic properties of transpose

Consider matrices A and B, and a scalar α. Then

(i) $(A^T)^T = A$

(ii) $(A + B)^T = A^T + B^T$

(iii) $(\alpha A)^T = \alpha A^T$

(iv) $(AB)^T = B^T A^T$ (reversed order!)

b) Symmetry

A square matrix A is called a **symmetric** matrix if $A^T = A$. It is called **skew symmetric** if $A^T = -A$.

Example: Consider the matrices

$$A = \begin{bmatrix} 2 & 12 & 20 \\ 12 & 1 & 15 \\ 20 & 15 & 3 \end{bmatrix}, \; B = \begin{bmatrix} 0 & 1 & -3 \\ -1 & 0 & -2 \\ 3 & 2 & 0 \end{bmatrix}.$$

A is symmetric, and B is skew-symmetric.

c) Orthogonality

A matrix A is orthogonal if

(i) A is square,

(ii) $A^TA = AA^T = \mathbf{I} \longleftrightarrow A^T = A^{-1}$.

Hence, square matrices A and B are orthogonal if $AB = \mathbf{I}$.

- **Properties of orthogonal matrices**

  (i) Orthogonal matrices preserve norms. Let Q be an orthogonal matrix and $\vec{v}$ a vector. Then

$$||Q\vec{v}||^2 = (Q\vec{v})^T(Q\vec{v}) = \vec{v}^TQ^TQ\vec{v} = \vec{v}^T\vec{v} = ||\vec{v}||^2.$$

  Consider two matrices A and B. Then we have
  (ii) If A is orthogonal, then A^T and A^{-1} are also orthogonal matrices.
  (iii) If A is orthogonal, then its rows are mutually orthogonal, i.e., the inner product of any two rows is zero.
  (iv) If A is orthogonal, then its columns are mutually orthogonal, i.e., the inner product of any two columns is zero.
  (v) If A is orthogonal, then $\det(A) = \pm 1$. Indeed, $AA^T = \mathbf{I} \implies |A| \cdot |A^T| = 1 \implies |A|^2 = 1$.
  (vi) If A and B are orthogonal, then AB and BA are also orthogonal matrices.

d) Matrix norm

The norm of a matrix A (or the Frobenius norm of a matrix A) is defined as the squareroot of the sum of the squares of its elements:

$$||A|| = \sqrt{\sum_{i,j} a_{ij}^2}.$$

It can readily be shown that

$$||A|| = \sqrt{\sum_i \|\mathbf{a}_i\|^2},$$

where $\mathbf{a}_i$ are the columns vectors of A.

The norm of a matrix can also be defined using the scalar of the matrix and its transpose as the squareroot of the trace of its Gram matrix:

$$||A|| = \sqrt{\mathrm{Tr}(A^TA)}.$$

For a vector, the norm is also called the length (of the vector).

- **Properties of norms**

  Consider two matrices A and B. Then we have
  (i) Norm is nonnegative, $\|A\| \geq 0$.
  (ii) c is number, $\|cA\| = |c| \cdot \|A\|$.
  (iii) Triangle inequality $\|A + B\| \leq \|A\| + \|B\|$.
  (iv) Cauchy-Schwarz inequality $\|AB\| \leq \|A\| \cdot \|B\|$.

Exercise (*determinant of block matrices*).

If A is an $m \times m$ matrix, if B is an $m \times n$ matrix and if $\mathbf{0}$ and $\mathbf{I}$ are zero and identity matrices of appropriate sizes, then given the result $\det \begin{bmatrix} A & B \\ \mathbf{0} & \mathbf{I} \end{bmatrix} = \det(A)$,

1. What are the sizes of $\mathbf{0}$ and $\mathbf{I}$?
2. Compute $M = \begin{bmatrix} \mathbf{0} & A \\ -B & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ B & \mathbf{I} \end{bmatrix}$.
3. Prove that $\det \begin{bmatrix} \mathbf{0} & A \\ -B & \mathbf{I} \end{bmatrix} = \det(AB)$.

3 Addition and subtraction

The matrix $A+B$ is obtained by adding elementwise the two matrices A and B (having the same order). The elements of the resultant matrix, $C \equiv A+B$, are

$$c_{ij} = a_{ij} + b_{ij}.$$

Example. Find the sum of the matrices A and B:

$$A = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 4 & 2 \\ 5 & 0 \end{bmatrix}.$$

Solution. We have

$$A+B = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} + \begin{bmatrix} 4 & 2 \\ 5 & 0 \end{bmatrix} = \begin{bmatrix} 3+4 & -2+2 \\ -1+5 & 1+0 \end{bmatrix} = \begin{bmatrix} 7 & 0 \\ 4 & 1 \end{bmatrix}.$$

Note: The matrix $A-B$ is obtained by subtracting elementwise the two matrices A and B (having the same order). The elements of the resultant matrix, $C \equiv A-B$, are $c_{ij} = a_{ij} - b_{ij}$.

• **Unary and binary operations.** An **unary** operation is an operation with only one operand, i.e. a single input operation.

e.g. the transpose $T : \mathbb{M} \to \mathbb{M}, \; A \mapsto A' = A^T$ is an unary operation on the set $\mathbb{M}$.

A **binary** operation uses two operands. e.g. the multiplication $X : \mathbb{R}\times\mathbb{R} \to \mathbb{R}, \; (x,y) \longmapsto z = x+y$ is a binary operation on the set $\mathbb{M}$.

4 Multiplications

4.1 Multiplication by a scalar

Given any matrix A and any number c (a number is sometimes referred to as a scalar), the matrix cA is obtained from the matrix A by multiplying each element of A by c. For example, consider

$$A = \begin{bmatrix} -1 & 2 & 2 \\ 0 & 0 & 1 \\ -1 & 2 & 0 \end{bmatrix},$$

then for $c = 5$, we have

$$cA = \begin{bmatrix} -5 & 10 & 10 \\ 0 & 0 & 5 \\ -5 & 10 & 0 \end{bmatrix}.$$

4.2 Scalar product of matrices

Let A and B be two matrices of size $n\times m$ and $p\times q$, resp. The matrix product AB is possible/defined only when $m = p$ (i.e., number of columns of A equals the number of rows of B). The resultant matrix, $C \equiv AB$, has the size $n \times q$, and its elements are

$$c_{ij} = \sum_k a_{ik} b_{kj},$$

i.e, the sum of products of elements from the ith row of A and jth column of B.

Example. Find the scalar product of the matrices A and B:

$$A = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 4 & 2 \\ 5 & 0 \end{bmatrix}.$$

Solution. We have

$$AB = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 4 & 2 \\ 5 & 0 \end{bmatrix} = \begin{bmatrix} 3(4)+(-2)(5) & 3(2)+(-2)(0) \\ -1(4)+1(5) & -1(2)+1(0) \end{bmatrix} = \begin{bmatrix} 2 & 6 \\ 1 & -2 \end{bmatrix}.$$

• **Case of block matrices.**

Consider the block matrices

$$M = \left[\begin{array}{c|c} A & D \\ \hline C & B \end{array}\right], \quad M' = \left[\begin{array}{c|c} A' & D' \\ \hline C' & B' \end{array}\right].$$

When the sizes of the constituent matrices allows, the product of the matrices is

$$MM' = \left[\begin{array}{c|c} A & D \\ \hline C & B \end{array}\right] \left[\begin{array}{c|c} A' & D' \\ \hline C' & B' \end{array}\right] = \left[\begin{array}{c|c} AA' + DC' & AD' + DB' \\ \hline CA' + BC' & CD' + BB' \end{array}\right].$$

Example.

$$\left[\begin{array}{cc|c} 1 & -1 & 5 \\ 2 & 3 & -2 \end{array}\right] \left[\begin{array}{cc} 2 & -1 \\ 4 & 3 \\ \hline 3 & -7 \end{array}\right] = \left[\begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 4 & 3 \end{pmatrix} + \begin{pmatrix} 5 \\ -2 \end{pmatrix} \begin{pmatrix} 3 & -7 \end{pmatrix}\right] = \left[\begin{pmatrix} -2 & -4 \\ 16 & 7 \end{pmatrix} + \begin{pmatrix} 15 & -35 \\ -6 & 14 \end{pmatrix}\right] = \begin{bmatrix} 13 & -39 \\ 10 & 21 \end{bmatrix}.$$

4.3 Cayley-Hamilton theorem

Statement: Every square matrix A satifies its own characteristic equation.

Example. Let

$$A = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}.$$

Its characteristic equation is $|A-\lambda I| = 0 \implies (3-\lambda)(1-\lambda)-2 = 0 \implies \lambda^2 - 4\lambda + 1 = 0$. According to Cayley-Hamilton theorem, A satisfies the equation $A^2 - 4A + 1 = 0$. You can check this. That allows simplifying matrix equations.

4.4 Outer product

Let $\mathbf{u}$ and $\mathbf{v}$ be two vectors of length n. Then $\mathbf{u}$ and $\mathbf{v}^T$ are matrices of size $n \times 1$ and $1 \times n$, resp. The scalar product of such matrices, denoted $\mathbf{u}\mathbf{v}^T$, is possible because the number of columns of $\mathbf{u}$ equals the number of rows of $\mathbf{v}^T$. The result is a matrix of size $n \times n$, and its elements are such that

$$A = \mathbf{u}\mathbf{v}^T \implies a_{ij} = u_i v_j.$$

$$A = \begin{bmatrix} u_1v_1 & u_1v_2 & \cdots & u_1v_n \\ u_2v_1 & u_2v_2 & \cdots & u_2v_n \\ \vdots & \vdots & \ddots & \vdots \\ u_nv_1 & u_nv_2 & \cdots & u_nv_n \end{bmatrix}.$$

The scalar product of the column matrix $\mathbf{u}$ and row matrix $\mathbf{v}^T$ is called the **outer product** of (row) vectors $\mathbf{u}$ and $\mathbf{v}$. It is denoted $\mathbf{u} \otimes \mathbf{v} = \mathbf{u}\mathbf{v}^T$. For complex matrices, we have $\mathbf{u} \otimes \mathbf{v} = \mathbf{u}\mathbf{v}^\dagger = \mathbf{u}\left(\mathbf{v}^T\right)^*$.

Example. Find the outer product of the vectors $\mathbf{u}$ and $\mathbf{v}$, $\mathbf{u}$ and $\mathbf{w}$ such that:

$$\mathbf{u} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} 4 \\ 5 \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} 2 \\ i \end{pmatrix}.$$

$$\mathbf{u} \otimes \mathbf{v} = \begin{pmatrix} 3 \\ -2 \end{pmatrix} [4, \quad 5] = \begin{bmatrix} 12 & 15 \\ -8 & -10 \end{bmatrix}; \quad \mathbf{u} \otimes \mathbf{w} = \begin{pmatrix} 3 \\ -2 \end{pmatrix} [2, \quad -i] = \begin{bmatrix} 6 & -3i \\ -4 & 2i \end{bmatrix}.$$

Note that $\mathbf{u} \otimes \mathbf{v} = (\mathbf{v} \otimes \mathbf{u})^T$. Besides, we have $\text{rank}(\mathbf{u} \otimes \mathbf{v}) = 1$, because for any pair (i, j), we have $u_i \text{row}_j = u_j \text{row}_i$.

5 Rotation and permutation matrices

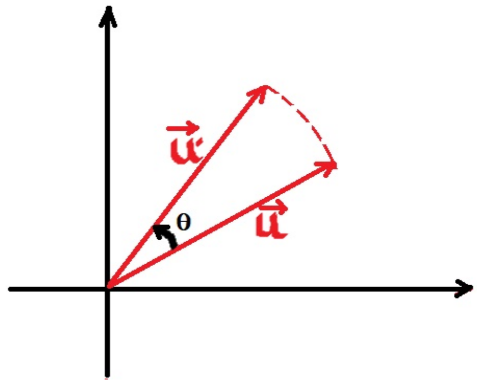
5.1 Rotation matrix

A **rotation matrix** is an orthogonal matrix that rotates a vector in space without changing the vector's length.

Example. The following two-by-two matrix rotates a vector through an angle θ in the x-y plane:

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

The inverse of $R(\theta)$ rotates a vector by $-\theta$, i.e., $R^{-1}(\theta) = R(-\theta)$.



We have $R\,\mathbf{u} = \mathbf{u}'$ and $R^{-1}\,\mathbf{u}' = \mathbf{u}$.

▷ **Homework:** Find the three-by-three matrix that rotates a three-dimensional vector of an angle θ counterclockwise around the z-axis.

5.2 Permutation matrix

An n-by-n **permutation matrix** is a square matrix when multiplying on the left permutes the rows of a matrix, and when multiplying on the right permutes the columns. It is obtained by permutation of rows of the identity matrix of same size.

• Permutation matrix of order 2

For example, let the string $\{1,2\}$ represent the order of the rows or columns of a 2-by-2 matrix. Then the set of all possible permutations of the rows or columns is $\{\{1,2\},\{2,1\}\}$. The first permutation does not achieve a permutation, and the corresponding permutation matrix is simply the identity matrix. The second is a permutation of the rows or columns achieved as follows

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{bmatrix} \text{(row permutation)}, \qquad \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} a_{12} & a_{11} \\ a_{22} & a_{21} \end{bmatrix} \text{(column permutation)}$$

Example:

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 5 & 3 & 4 \\ 2 & -3 & -1 \end{bmatrix} = \begin{bmatrix} 2 & -3 & -1 \\ 5 & 3 & 4 \end{bmatrix} \text{(row permutation)}, \qquad \begin{bmatrix} 5 & 3 \\ -2 & 1 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 5 \\ 1 & -2 \\ 4 & 0 \end{bmatrix} \text{(column permutation)}$$

• Permutation matrix of order 3

The rows or columns of a three-by-three matrix have $A_3^3 = 3! = 6$ possible permutations, namely $\{1,2,3\}$, $\{1,3,2\}$, $\{2,1,3\}$, $\{2,3,1\}$, $\{3,1,2\}$, $\{3,2,1\}$. For example, the row permutation $\{3,1,2\}$ is obtained by

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \text{(row permutation)}.$$

Due to the row permutation $\{3,1,2\}$, we have $Row1 \leftarrow Row3$, $Row2 \leftarrow Row1$, and $Row3 \leftarrow Row2$.

Realize that the string, say $\{3,1,2\}$, also gives the position of the 1 in the corresponding row or column, i.e., in the 1^{st} row, the 1 is located in the 3^{rd} position or column (the remaining entries of the row are zeros); in the 2^{nd} row, the 1 is located in the 1^{st} position, in the 3^{rd} row, the 1 is located in the 2^{nd} position.

The column permutation $\{3,1,2\}$ is obtained by

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \text{(column permutation)}.$$

Obviously, permuting the rows of a column vector will not change its norm.

▷ **Homework:** Write down the six three-by-three permutation matrices corresponding to the permutations $\{1,2,3\}$, $\{1,3,2\}$, $\{2,1,3\}$, $\{2,3,1\}$, $\{3,1,2\}$, $\{3,2,1\}$.

6 Matrix adjoint and inverses

6.1 Adjoint matrix

The **adjoint** of a square matrix A (or **adjunct** or **adjugate**), denoted $A^\dagger$ or adj(A), is the transpose of the matrix of cofactors. So for a 3×3 matrix A, the adjoint reads

$$A^\dagger = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}^T$$

The special case $n = 2$ simply leads to

$$A^\dagger = \begin{bmatrix} a_{22} & -a_{21} \\ -a_{12} & a_{11} \end{bmatrix}^T = \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

• Properties of adjoints

For two matrices A and B, we have

(i) $(AB)^\dagger = B^\dagger A^\dagger$.

(ii) $AA^\dagger = |A|\ \mathbf{I}$.

(iii) $|A^\dagger| = |A|^{n-1}$ where n is the order of A.

• Self-adjoint matrix

A **self-adjoint matrix** (or **Hermitian matrix**) is a complex square matrix that is equal to its own conjugate transpose, i.e. $A^{*T} = A$ (or $a_{ij} = a^*_{ji}$).
Exercise: Consider the following Pauli matrices

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},\ \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix},\ \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Show that $\sigma_{x,y,z}$ are Hermitian matrices.

Note: The conjugate transpose of a matrix is also called its **Hermitian adjoint**.

6.2 Left and right inverses

For a given $n \times m$ matrix A, if rank(A)=min(n, m) then A has a right inverse if $n \leqslant m$ or a left inverse if $m \leqslant n$.

The $m \times n$ matrix B is the right inverse of A if

$$AB = I_n.$$

The $m \times n$ matrix B is the left inverse of A if

$$BA = I_m.$$

Example: Consider the following matrices

$$A = \begin{bmatrix} -3 & -4 \\ 4 & 6 \\ 1 & 1 \end{bmatrix},\ B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

Show that A is left-invertible and B is right-invertible.

6.3 Inverse

• Definition

For a given $n \times n$ matrix A, the $n \times n$ matrix B is the inverse of A if

$$AB = BA = I_n.$$

We denote $B = A^{-1}$.

A^{-1} exists if det$A \neq 0$, i.e., when columns (rows) of A are linearly independent. A^{-1} can be computed with one of the following methods:

— Cofactor method: $[A^{-1}]_{ij} = \frac{1}{\det A} A_{ji}$.

— Gauss-Jordan or Jacobi's method: Assume, by elementary row operations,

$$[AI] = \begin{bmatrix} a_{11} & \cdots & a_{1n} & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \ddots & \\ a_{n1} & \cdots & a_{nn} & 0 & \cdots & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & \cdots & 0 & b_{11} & \cdots & b_{1n} \\ & \ddots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & b_{n1} & \cdots & b_{nn} \end{bmatrix} = [IB]$$

Then $B = A^{-1}$. This will be possible only if rank $A = n$. If rank $A < n$, then A has no inverse.

The special case $n = 2$ simply leads to

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

Example: Compute the inverse of

$$A = \begin{bmatrix} 2 & 0 & -1 \\ 3 & 1 & 2 \\ -1 & 0 & 1 \end{bmatrix}.$$

Verify that

$$A^{-1} = \begin{bmatrix} 1 & 0 & 1 \\ -5 & 1 & -7 \\ 1 & 0 & 2 \end{bmatrix}.$$

• Invertibility condition

We say that a square matrix (of order n) is *invertible*, or has an inverse, if and only if the determinant is not equal to zero. In that case, the matrix is nonsingular or nondegenerate, and has full rank; that is, rank $A = n$.

• Properties of inverses

Consider two invertible matrices A and B, then we have

(i) A is non-singular.

(ii) The inverse matrix of A is unique.

(iii) If c is any non-zero scalar then cA is invertible and $(cA)^{-1} = A^{-1}/c$.

(iv) $(A^{-1})^{-1} = A$.

(v) $(A^{-1})^T = (A^T)^{-1}$.

(vi) $A^{-n} = (A^{-1})^n$.

(vi) $A^{-1}A = AA^{-1} = I$ (left and right invertible).

(vi) $(AB)^{-1} = B^{-1}A^{-1}$.

• Unitary matrix

A complex square matrix A is called a **unitary** matrix if its conjugate transpose (or Hermitian adjoint) is equal to its inverse, i.e., $A^{*T} = A^{-1} \implies AA^{*T} = A^{*T}A = 1$.

Example:

$$U = \frac{1}{2} \begin{bmatrix} 1-i & 1+i \\ 1+i & 1-i \end{bmatrix}, \quad V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}.$$

The rows and columns of a unitary matrix are mutually orthonormal. Unitary matrices are the complex analog of real orthogonal matrices.

- **Unitary matrix**

A complex square matrix A is called an **involutory** matrix if it is equal to its inverse, i.e., $A^{-1} = A \implies AA^{-1} = A^2 = 1$.
Exercise: Consider the following Pauli matrices

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \ \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \ \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Show that $\sigma_{x,y,z}$ are unitary and involutory matrices.

6.4 Singular value decomposition

SVD is an eigenvalue problem and reduction to canonical form for rectangular matrices. It is a generalization of the eigen-decomposition of square normal matrices with an orthonormal basis to non-square matrices. Any $m \times n$ matrix A can be factored into

$$A = QSP^T \text{ or } A = U\Sigma V^T,$$

where

— S is an $m \times n$ "diagonal type" matrix with positive real entries, i.e., $s_{ij} = 0, \ i \neq j; \ s_{ij} \geqslant 0, \ i = j$. It corresponds to a scaling by the singular values μ_j (see definition later) in different directions.

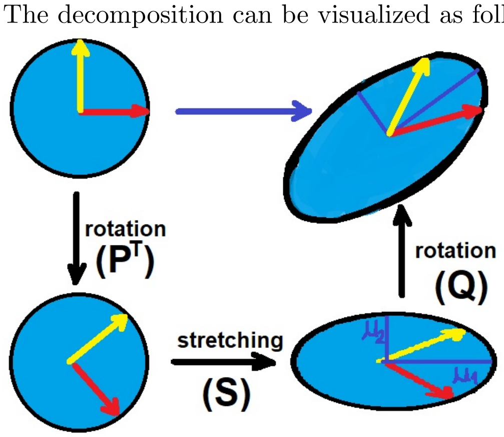
Example of diagonal rectangular matrices:

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix}, \quad \begin{bmatrix} 2 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}.$$

— P is an $n \times n$ orthogonal matrix (P is unitary and corresponds to a rotation).

— Q is an $m \times m$ orthogonal matrix (Q is unitary and corresponds to a rotation).

The decomposition can be visualized as follows:



- **Determination of P, Q and S**

The Gram matrix of A, i.e. A^TA, is a positive semidefinite symmetric $n \times n$-matrix with non-negative eigenvalues λ_i. Suppose that $\lambda_1, \lambda_2, ..., \lambda_r > 0$ and $\lambda_{r+1}, ..., \lambda_n = 0$. Let $\mathbf{g}_1, \mathbf{g}_2, ..., \mathbf{g}_n$ be an orthonormal set of corresponding eigenvectors.
— *Construction of P*: The columns of P are the vectors $\mathbf{g}_1, ..., \mathbf{g}_n$, i.e. $P = [\mathbf{g}_1, ..., \mathbf{g}_n]$.
— *Construction of S:* $s_{ii} = \mu_i = \sqrt{\lambda_i}, \ i = 1, ..., r$. Remaining $s_{ij} = 0$. The μ_i also denoted σ_i are called the **singular values of** A.
— *Construction of Q:* Set $\mathbf{h}_i = \frac{1}{\mu_i} A\mathbf{g}_i, \ i = 1, ..., r$. If $r < m$, complete by $\mathbf{h}_{i+1}, ..., \mathbf{h}_m$ to an orthonormal set of vectors. The columns of Q are the vectors $\mathbf{h}_1, ..., \mathbf{h}_m$, i.e. $Q = [\mathbf{h}_1, ..., \mathbf{h}_m]$.

- **Example**

Find the SVD of the matrix

$$A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}.$$

Step1— First we compute the singular values μ_i by finding the eigenvalues of the Gram matrix of A. $G = A^TA$ is a 3-by-3 matrix and will have 3 non-negative eigenvalues. Alternatively, we can compute the $G' = AA^T$ which is a 2-by-2 matrix and will have 2 positive eigenvalues. Such eigenvalues of G' are also eigenvalues of G. The remaining eigenvalues of G are simply zeros.

$$G \equiv A^T A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} = \begin{bmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{bmatrix},$$

$$G' \equiv AA^T = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} 17 & 8 \\ 8 & -17 \end{bmatrix}.$$

Then the characteristic equation is

$$\det(G' - \lambda I) = 0 \implies \begin{vmatrix} 17-\lambda & 8 \\ 8 & -17-\lambda \end{vmatrix} = 0 \iff (\lambda - 25)(\lambda - 9) = 0 \implies \lambda_1 = 25, \lambda_2 = 9.$$

$$\mu_1 = \sqrt{\lambda_1} = 5, \mu_2 = \sqrt{\lambda_2} = 3 \implies S = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix}.$$

The eigen values of G' are $\lambda_1 = 25, \lambda_2 = 9$. So the eigenvalues of G are $\lambda_1 = 25, \lambda_2 = 9, \lambda_3 = 0$.

Step2— Next we find the right singular vectors (the columns of P) by finding an orthonormal set of eigenvectors of G. It is also possible to proceed by finding the left singular vectors (columns of Q) instead.

For $\lambda_1 = 25$, we have

$$G - 25I = \begin{bmatrix} -12 & 12 & 2 \\ 12 & -12 & -2 \\ 2 & -2 & -17 \end{bmatrix} \overset{ERO}{\sim} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}. \text{ Thus } G\mathbf{x} = 25\mathbf{x} \implies x_1 - x_2 = 0, x_3 = 0.$$

Hence $\mathbf{x}_1 = (x_1, x_1, 0)$. So a unit vector in the kernel of the matrix $G - 25I$ is $\hat{x}_1 = (1/\sqrt{2}, 1/\sqrt{2}, 0)$.

For $\lambda_2 = 9$, we have

$$G - 9I = \begin{bmatrix} 4 & 12 & 2 \\ 12 & 4 & -2 \\ 2 & -2 & -1 \end{bmatrix} \text{ which row-reduces to } \begin{bmatrix} 1 & 0 & -1/4 \\ 0 & 1 & 1/4 \\ 0 & 0 & 0 \end{bmatrix}.$$

A unit vector in the kernel of the matrix $G - 9I$ is $\hat{x}_2 = (1/\sqrt{18}, -1/\sqrt{18}, 4/\sqrt{18})$.

For $\lambda_3 = 0$, we have

$$G = \begin{bmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{bmatrix} \text{ which row-reduces to } \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix}.$$

A vector in the kernel of the matrix G is $\mathbf{x}_3 = (x_1, -x_1, -x_1/2)$. Hence a unit vector in the kernel of G is $\hat{x}_3 = (2/3, -2/3, -1/3)$.

Thus we have $P = [\hat{x}_1 \hat{x}_2 \hat{x}_3]$ or

$$P = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{18} & 2/3 \\ 1/\sqrt{2} & -1/\sqrt{18} & -2/3 \\ 0 & 4/\sqrt{18} & -1/3 \end{bmatrix} \implies P^T = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{18} & -1/\sqrt{18} & 4/\sqrt{18} \\ 2/3 & -2/3 & -1/3 \end{bmatrix}.$$

Step3— We find the left singular vectors (columns of Q). Set $\mathbf{h}_i = \frac{1}{\mu_i} A\hat{x}_i,\ i = 1, 2$. Then $Q = [\mathbf{h}_1, \mathbf{h}_2]$. We have

$$\mathbf{h}_1 = \frac{1}{\mu_1} A\hat{x}_1 = \frac{1}{5} \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

$$\mathbf{h}_2 = \frac{1}{\mu_2} A\hat{x}_2 = \frac{1}{3} \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} \begin{bmatrix} 1/\sqrt{18} \\ -1/\sqrt{18} \\ 4/\sqrt{18} \end{bmatrix} = \begin{bmatrix} 3/\sqrt{18} \\ -3/\sqrt{18} \end{bmatrix}$$

$$\implies Q = \begin{bmatrix} 1/\sqrt{2} & 3/\sqrt{18} \\ 1/\sqrt{2} & -3/\sqrt{18} \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$$

Hence the svd of A is

$$A = QSP^T \implies \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{18} & 2/3 \\ 1/\sqrt{2} & -1/\sqrt{18} & -2/3 \\ 0 & 4/\sqrt{18} & -1/3 \end{bmatrix}^T.$$

6.5 Pseudo-inverse

The **Moore-Penrose inverse** or **pseudo-inverse** (or generalized inverse) of order $n \times m$ of an arbitrary $m \times n$ matrix A with singular value decomposition $A = QSP^T$ is defined by

$$A^+ = PS^+Q^T,$$

where S^+ is the $n \times m$ "diagonal type" matrix with elements $[S^+]_{ii} = 1/\mu_i,\ i = 1, 2, ..., r$ and remaining elements 0. The μ_i are the singular values of A, i.e., $\mu_i = \sqrt{\lambda_i}$ where λ_i are the eigen values of A's Gram matrix.

The pseudo-inverse exists for any matrix A including singular square matrices and non-square matrices. In the particular case when A is a square non-singular matrix, we have $A^+ = A^{-1}$.

- **Properties**

(i) $AA^+A = A$.

(ii) $A^+AA^+ = A^+$.

(iii) $(AA^+)^* = AA^+$ (Hermiticity of AA^+).

(iv) $(A^+A)^* = A^+A$ (Hermiticity of A^+A).

- **Example:**

$$A = \begin{bmatrix} a & 0 \\ \pm b & 0 \end{bmatrix} \Longrightarrow A^+ = \begin{bmatrix} a/(a^2+b^2) & \pm b/(a^2+b^2) \\ 0 & 0 \end{bmatrix}.$$

You can check that $AA^+A = A$ or $A^+AA^+ = A^+$.