Unitary, orthogonal, Hermitian matrices

1 Unitary matrices

Take N orthonormal basis vectors in $\mathbb{C}^N$: $\mathbf{u}_1, \ldots, \mathbf{u}_N$.

$$\begin{aligned}
\mathbf{u}_1 &= (u_{11}, u_{21}, \ldots, u_{N1}) \\
&\vdots \\
\mathbf{u}_j &= (u_{1j}, u_{2j}, \ldots, u_{Nj}) \\
&\vdots \\
\mathbf{u}_N &= (u_{1N}, u_{2N}, \ldots, u_{NN}).
\end{aligned}$$

Orthonormal means

$$(\mathbf{u}_j, \mathbf{u}_k) = \sum_{l=1}^{N} \overline{u}_{lj}\, u_{lk} = \delta_{jk} \tag{1}$$

Construct a $N \times N$ matrix U by filling the j-th column with $\mathbf{u}_j$:

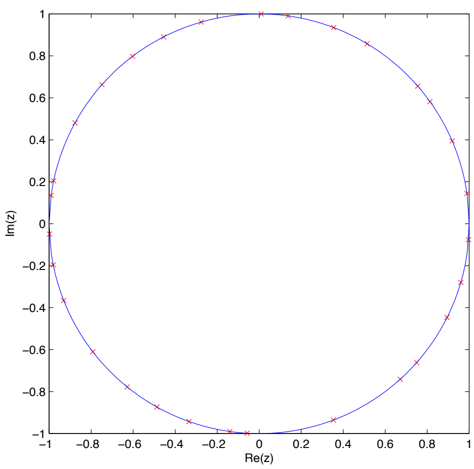
$$U = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1N} \\ u_{21} & u_{22} & \cdots & u_{2N} \\ \multicolumn{4}{c}{\dotfill} \\ u_{N1} & u_{N2} & \cdots & u_{NN} \end{pmatrix}$$

The matrix U is called unitary

The eigenvalues of U are complex numbers of modulo one, i.e.

$$e^{i\theta_1}, \ldots, e^{i\theta_N}.$$

The eigenvalues of a unitary matrix lie on the unit circle in the complex plane.



A 2×2 example of unitary matrix:

$$\begin{pmatrix} e^{i\phi_1}\cos\theta & -e^{-i(\chi-\phi_2)}\sin\theta \\ e^{i(\chi+\phi_1)}\sin\theta & e^{i\phi_2}\cos\theta \end{pmatrix}$$

You can check that this is constructed of orthonormal vectors as described above.

2 Hermitian inner product

The Hermitian product is a map $\mathbb{C}^N \times \mathbb{C}^N \to \mathbb{C}$, denoted by $(\mathbf{v}, \mathbf{w})$, with the following properties:

1. $(\mathbf{v}, \mathbf{w}) = \overline{(\mathbf{w}, \mathbf{v})}$;
2. $(\mathbf{v}, \mathbf{w} + \mathbf{z}) = (\mathbf{v}, \mathbf{w}) + (\mathbf{v}, \mathbf{z})$;
3. $(\mathbf{v}, \alpha\mathbf{w}) = \alpha(\mathbf{v}, \mathbf{w})$, from 1. $(\alpha\mathbf{v}, \mathbf{w}) = \overline{\alpha}(\mathbf{v}, \mathbf{w}) \quad \alpha \in \mathbb{C}$
4. $(\mathbf{v}, \mathbf{v}) \geq 0$, where the equality holds if and only if $\mathbf{v} = \mathbf{0}$.

- Two vectors are said to be orthogonal if $(\mathbf{v}, \mathbf{w}) = 0$.
- We define the *norm* of a vector to be $|\mathbf{v}| = \sqrt{(\mathbf{v}, \mathbf{v})}$.
  Example:
  $$(\mathbf{v}, \mathbf{w}) = \sum_{j=1}^{N} \overline{v}_j w_j.$$
  This is the inner product that we will always consider.
- The *Hermitian conjugate* or *adjoint* $M^\dagger$ of a matrix M is defined by
  $$(\mathbf{v}, M\mathbf{w}) = (M^\dagger\mathbf{v}, \mathbf{w}) \quad \forall \mathbf{v}, \mathbf{w} \in \mathbb{C}^N.$$
- Written in terms of components this formula becomes
  $$\sum_{j,k=1}^{N} \overline{v}_j m_{jk} w_k = \sum_{j,k=1}^{N} \overline{m^\dagger}_{kj} \overline{v}_j \; w_k, \tag{2}$$
  which implies $m_{jk} = \overline{m^\dagger}_{kj}$ or $m^\dagger_{jk} = \overline{m}_{kj}$.
- Sometimes $M^\dagger$ is called *conjugate transpose* of M and is denoted by M^*.

  In other words, if
  $$M = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix}$$

then
$$M^\dagger = \begin{pmatrix} \overline{m}_{11} & \overline{m}_{21} & \overline{m}_{31} \\ \overline{m}_{12} & \overline{m}_{22} & \overline{m}_{32} \\ \overline{m}_{13} & \overline{m}_{23} & \overline{m}_{33} \end{pmatrix}$$

For complex matrices the adjoint (conjugate transpose) takes the place of the transpose of a real matrix.

3 Hermitian inner product and orthogonal matrices:

Example: $\mathbb{R}^3$.

Let $\mathbf{a} = (a_1, a_2, a_3)$ and $\mathbf{b} = (b_1, b_2, b_3)$ two vectors in $\mathbb{R}^3$.

We have

$$\mathbf{a} \cdot \mathbf{b} = (\mathbf{a}, \mathbf{b}) = |\mathbf{a}|\,|\mathbf{b}| \cos\phi$$

Here ϕ is the angle between $\mathbf{a}$ and $\mathbf{b}$: see the illustration in Figure 1. You can check that this operation is linear in both arguments

Let $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ be the orthogonal unit vectors along the three axes x_1, x_2 and x_3 respectively.

We have

$$\begin{aligned} (\mathbf{i}, \mathbf{j}) = (\mathbf{j}, \mathbf{k}) = (\mathbf{i}, \mathbf{k}) = 0 \\ (\mathbf{i}, \mathbf{i}) = (\mathbf{j}, \mathbf{j}) = (\mathbf{k}, \mathbf{k}) = 1. \end{aligned}$$

If we write $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ and $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$ we obtain

$$(\mathbf{a}, \mathbf{b}) = a_1 b_1 + a_2 b_2 + a_3 b_3. \tag{3}$$

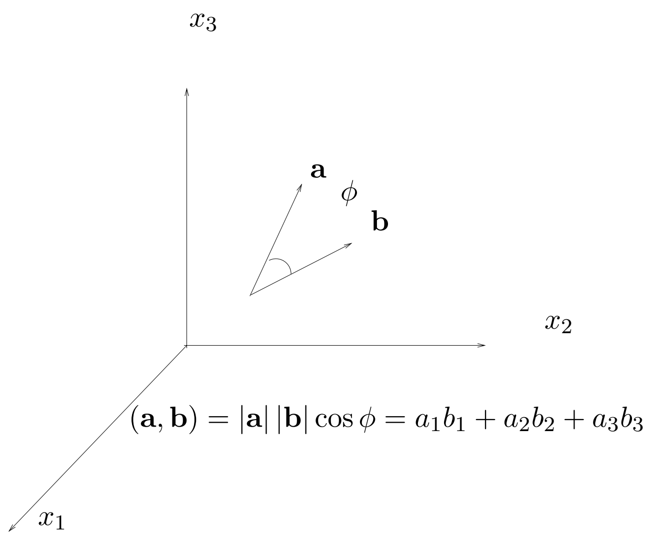


Figure 1:

Rigid rotations in space are achieved by 3×3 ***orthogonal matrices, i.e.***

$$\mathbf{a}' = O\mathbf{a} \quad \text{and} \quad \mathbf{b}' = O\mathbf{b}.$$

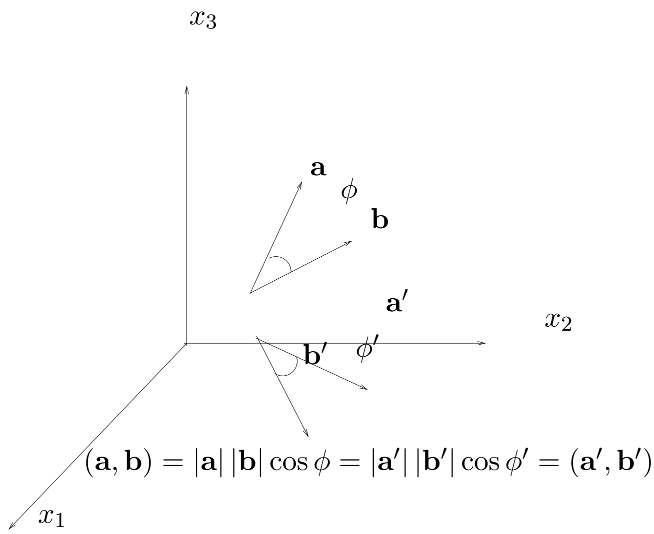


Figure 2:

Rigid rotations in space preserve length and angles: see Figure 2. This means that

$$(\mathbf{a},\mathbf{b}) = (O\mathbf{a}, O\mathbf{b}) = (\mathbf{a}',\mathbf{b}').$$

Let $\mathbf{x} = (x_1, \ldots, x_N)$ and $\mathbf{y} = (y_1, \ldots, y_N)$ be two vectors in $\mathbb{R}^N$. The scalar product in $\mathbb{R}^N$ is defined

$$(\mathbf{x},\mathbf{y}) = \sum_{j=1}^{N} x_j y_j$$

Definition: *An $N \times N$ real matrix is called orthogonal if it preserves the scalar product in $\mathbb{R}^N$*

$$(O\mathbf{x}, O\mathbf{y}) = (\mathbf{x},\mathbf{y}).$$

In particular O preserves the length of a vector, i.e.

$$|O\mathbf{x}|^2 = (O\mathbf{x}, O\mathbf{x}) = (\mathbf{x},\mathbf{x}) = |\mathbf{x}|^2 = \sum_{j=1}^{N} x_i^2.$$

Let $O = (o_{ij})$ be an $N \times N$ orthogonal matrix.
The i-th element $(O\mathbf{x})_i$ of the vector $O\mathbf{x}$ is

$$(O\mathbf{x})_i = \sum_{j=1}^{N} o_{ij} x_j$$

The definition in terms of matrix elements reads

$$(O\mathbf{x}, O\mathbf{y}) = \sum_{i,j,k=1}^{N} o_{ij}x_j o_{ik}y_k = \sum_{i,j,k=1}^{N} x_j o_{ij} o_{ik} y_k = \sum_{j=1} x_j y_j = (\mathbf{x}, \mathbf{y}).$$

This means

$$\sum_{i=1}^{N} o_{ij} o_{ik} = \delta_{jk}$$

Thus, $\sum_{j,k=1}^{N} x_j \delta_{jk} y_k = \sum_{j=1}^{N} x_j y_j$

In other words, take the j-th and k-th column, multiply element by element and add along the rows.

The columns of an orthogonal matrix are a set of orthonormal vectors.

Now, let $O^t = (o^t_{ij})$ be the ***transpose matrix*** of $O = (o_{ij})$, i.e. $o^t_{ij} = o_{ji}$

But because O preserves the scalar product

$$\sum_{i=1}^{N} o_{ij} o_{ik} = \delta_{jk} = \sum_{i=1}^{N} o^t_{ji} o_{ik}.$$

Or equivalently $O^t O = I$. In other words $O^t = O^{-1}$

Note also that

$$O^t O = I$$
$$O O^t O = O$$

Therefore $OO^t = I$.

An equivalent definition of an orthogonal matrix is

$$OO^t = O^t O = I. \tag{4}$$

In terms of matrix elements equation (4) reads

$$\sum_{i=1}^{N} o_{ij} o_{ik} = \sum_{i=1}^{N} o_{ji} o_{ki} = \delta_{jk}.$$

Both the sets of rows and columns of an orthogonal matrix are orthonormal bases.

To construct the 2×2 orthogonal that corresponds to a *anti-clockwise* rotation of an angle θ in $\mathbb{R}^2$ (the two-dimensional real plane), let $\mathbf{i}$ and $\mathbf{j}$ be the unit vectors along the two axes x_1, x_2 of a right-handed reference frame. Then

$$\mathbf{i} \mapsto \cos\theta \mathbf{i} + \sin\theta \mathbf{j}$$
$$\mathbf{j} \mapsto -\sin\theta \mathbf{i} + \cos\theta \mathbf{j}$$

In components the above equations read

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$$

Therefore the matrix is

$$O = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

4 Hermitian matrices

- $H = (h_{jk})$ is called *Hermitian* or *self-adjoint* if

$$H^\dagger = H \text{ or } h_{jk} = \overline{h}_{kj}. \tag{5}$$

- If the elements of H are real, then trivially $H = H^t$ or $h_{jk} = h_{kj}$, i.e. H is *symmetric*.

1. **The eigenvalues of a Hermitian matrix are real.**

   Suppose $\mathbf{x}$ is an eigenvector of H with eigenvalue $\lambda \neq 0$, then

   $$(\mathbf{x}, H\mathbf{x}) = (\mathbf{x}, \lambda\mathbf{x}) = \lambda(\mathbf{x}, \mathbf{x}) = (H\mathbf{x}, \mathbf{x}) = (\lambda\mathbf{x}, \mathbf{x}) = \overline{\lambda}(\mathbf{x}, \mathbf{x}).$$

   Since $(\mathbf{x}, \mathbf{x})$ is real and positive $\lambda = \overline{\lambda}$. Note: this proof does not use the fact that the scalar product is defined in a complex linear space. Therefore, it holds for symmetric matrices too. That is, the eigenvalues of a symmetric matrix are real.

2. **Eigenvectors of a Hermitian matrix corresponding to different eigenvalues are orthogonal.**

   Let $H\mathbf{x} = \lambda\mathbf{x}$ and $H\mathbf{y} = \mu\mathbf{y}$. Then

   $$(\mathbf{x}, H\mathbf{y}) = \mu(\mathbf{x}, \mathbf{y}) \quad \text{and} \quad (H\mathbf{x}, \mathbf{y}) = \lambda(\mathbf{x}, \mathbf{y}). \tag{6}$$

   Since $(\mathbf{x}, H\mathbf{y}) = (H\mathbf{x}, \mathbf{y})$, by subtracting both sides of these equations we obtain

   $$0 = (\mu - \lambda)(\mathbf{x}, \mathbf{y}). \tag{7}$$

   Since $\lambda \neq \mu$ it must be that $(\mathbf{x}, \mathbf{y}) = 0$.

3. **A $N \times N$ Hermitian matrix can be diagonalized and admits a set of N orthogonal eigenvectors**.

Note: Facts 2 and 3 hold for symmetric matrices too.

A matrix H is called anti-Hermitian if $H = -H^\dagger$. If H is Hermitian iH is anti-Hermitian.

5 More on unitary matrices

Unitary matrices are the generalization to complex matrices of orthogonal matrices.

Definition: A $N \times N$ complex matrix $U = (u_{jk})$ is called *unitary* if it preserves the Hermitian inner product:

$$(U\mathbf{v}, U\mathbf{w}) = (\mathbf{v}, \mathbf{w}) \tag{8}$$

Immediate consequences:

1. $(U\mathbf{v}, U\mathbf{w}) = (\mathbf{v}, U^\dagger U\mathbf{w}) = (\mathbf{v}, \mathbf{w}) \Rightarrow UU^\dagger = U^\dagger U = I$. In components

   $$\sum_{k=1}^{N} \overline{u}_{jk} u_{lk} = \sum_{k=1}^{N} \overline{u}_{kj} u_{kl} = \delta_{jl}.$$

   Equivalently, the rows (columns) of a unitary matrix are sets of orthonormal vectors, ie. they are bases in $\mathbb{C}^N$.

2. Hermitian (symmetric) matrices are diagonalized by unitary (orthogonal) matrices.

   The eigenvectors $\{\mathbf{u}_j\}_{j=1,\ldots,N}$ of a Hermitian (symmetric) matrix are orthogonal. If they are normalized so that $|\mathbf{x}_j| = 1$, then they are an orthonormal basis in $\mathbb{C}^N$ ($\mathbb{R}^N$). By definition, the matrix U whose columns are the vectors $\mathbf{u}_j$ is unitary (orthogonal).

3. **The eigenvalues of a unitary matrix are complex numbers with absolute value one.**

   If $\mathbf{x}$ is an eigenvector with eigenvalue λ, then $(U\mathbf{x}, U\mathbf{x}) = (\lambda\mathbf{x}, \lambda\mathbf{x}) = |\lambda|^2 (\mathbf{x}, \mathbf{x}) = (\mathbf{x}, \mathbf{x})$, therefore $|\lambda|^2 = 1$ (This is true for orthogonal matrices too, but note that the eigenvalues of an orthogonal matrix are not real.)

4. **The product of two unitary matrices is a unitary matrix.**

   Let U and V be unitary matrices, then $(VU\mathbf{v}, VU\mathbf{w}) = (U\mathbf{v}, U\mathbf{w}) = (\mathbf{v}, \mathbf{w})$. Therefore, by definition VU is unitary. Note that $(VU)^\dagger = U^\dagger V^\dagger$.

5. **A unitary matrix maps orthonormal bases into orthonormal bases.**

   $(\mathbf{v}, \mathbf{w}) = 0 \Rightarrow (U\mathbf{v}, U\mathbf{w}) = 0$

6. **Eigenvectors of a unitary matrix corresponding to different eigenvalues are orthogonal.**

7. **A $N \times N$ unitary matrix can be diagonalized and admits a set of N orthogonal eigenvectors.**

The statements 4, 5, 6, 7 are true for orthogonal matrices too.

8. **Unitary matrices are diagonalized by unitary matrices, i.e.** $B^\dagger U B = \Lambda$**, with** $\Lambda = \mathrm{diag}(e^{i\theta_1}, \ldots, e^{i\theta_N})$**, i.e.**

$$\Lambda = \begin{pmatrix} e^{i\theta_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{i\theta_N} \end{pmatrix}, \tag{9}$$

   Note that B and Λ are unitary matrices too.

   Important: This statement is not true for orthogonal matrices, i.e. orthogonal matrices are not diagonalized by orthogonal matrices.

6 Groups

Definition: A set G equipped with a multiplication law, is said to be a group if

1. $g_1, g_2 \in G$ implies that $g_3 = g_1 g_2$ belongs to G too.
2. $(g_1 g_2) g_3 = g_1 (g_2 g_3)$
3. There exists an element $e \in G$ such that $eg = ge = g$ for every element $g \in G$.
4. For every element $g \in G$ there must be an inverse $g^{-1} \in G$ such that $gg^{-1} = g^{-1}g = e$.

Examples:

- The permutations of N letters form a group.
- Let $\mathrm{GL}(N, \mathbb{C})$ the set of all complex matrices M such that $\det M \neq 0$. $\mathrm{GL}(N, \mathbb{C})$ equipped with the usual matrix multiplication is a group.(Remember that M^{-1} exists if and only if $\det(M) \neq 0$.)

6.1 The Unitary Group

The set of $N \times N$ unitary matrices form a group denoted by $\mathrm{U}(N)$. It is a subgroup of $\mathrm{GL}(N, \mathbb{C})$.

Checks:

1. If U and V are unitary so is UV.
2. Associative: if U, V, W are unitary, then $(UV)W = U(VW)$
3. The identity matrix $\begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}$ is trivially unitary.
4. If U is unitary, so is $U^\dagger$ and $U^{-1} = U^\dagger$.