Week 2

1.3 Orthonormal bases

1.3.1 Orthonormal bases

Consider the canonical basis in $\mathbb{R}^2$. We learn to use the canonical basis at a young age, whether it is plotting graphs at school or playing Battleships. But why *this* basis, what makes it special? It satisfies two properties that we often take for granted, but are incredibly useful:

- ______________________________,
- ____.

We would like to restrict ourselves to using bases that also have this property. This leads to the notion of an orthonormal basis.

Definition 1.26 (Orthonormal basis) Let $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ be a set of vectors of an n-dimensional vector space V such that

$$\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij} = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases} , \tag{1.8}$$

where δ_{ij} is the Kronecker delta function. We call $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ an *orthonormal basis* of V.

Wikipedia: orthonormal basis

Remark 1.27 Note that Definition 1.26 does not require the vectors to form a basis. In fact, one can check that any set of n vectors satisfying condition (1.8) must form a basis .

Explicitly condition (1.8) implies the n vectors are linearly independent. Suppose that $\lambda_1 \mathbf{e}_1 + \lambda_2 \mathbf{e}_2 + \cdots + \lambda_n \mathbf{e}_n = 0$. We can multiply this relation by the vector $\mathbf{e}_i$, implying that $\lambda_i = 0$. Repeating this for every i, we see the vectors $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ are linearly dependent.

Example 1.28 Recall the Euclidean vector space with canonical basis (1.5) $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ and canonical inner product (1.6). It is quick to check that $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ satisfies condition (1.8) and therefore forms an orthonormal basis.

What nice properties do orthonormal bases have? Once we have fixed an orthonormal basis for our Euclidean vector space, the inner product on that space reduces to the canonical inner product. To see this, let

$(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ be an orthonormal basis for V. Then for any two vectors $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i$, $\mathbf{y} = \sum_{j=1}^n y_j \mathbf{e}_j$, their inner product becomes

$$\langle \mathbf{x}, \mathbf{y} \rangle = \tag{1.9}$$

$$= \sum_{i=1}^n x_i y_i.$$

Furthermore, we will can always find an orthonormal basis for a Euclidean vector space.

Proposition 1.29 *Every (finite-dimensional) Euclidean vector space has an orthonormal basis.*

As a result, we can choose to work only with orthonormal bases from now on. We let $\mathbb{E}^n$ denote an n-dimensional Euclidean vector space with orthonormal basis $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ (and therefore the canonical inner product under this basis).

1.3.2 Orthogonal matrices

Suppose that $\mathcal{B}$ is an orthonormal basis of $\mathbb{E}^n$. What is the condition on the transition matrix that an ordered set of vectors $\mathcal{C}$ also forms an orthonormal basis? The answer to this question is when the transition matrix is *orthogonal*.

Definition 1.30 (Orthogonal matrix) An $n \times n$ matrix M is *orthogonal matrix* if its product with its transpose is equal to the $n \times n$ identity matrix:

Wikipedia: orthogonal matrix

$$\tag{1.10}$$

Remark 1.31 If M is orthogonal then $\det(M) = \pm 1$. This holds as

$$1 = \det(I_n) = \det(M^\top M) = \det(M^\top)\det(M) = \det(M)^2$$
$$\Rightarrow \det(M) = \pm 1$$

(using some properties of the determinant we shall recall in Subsubsection 1.4.2). In particular, the determinant is always non-zero and so it is a transition matrix from one basis to another.

Proposition 1.32 *Let $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ be an orthonormal basis of an n-dimensional Euclidean vector space V. The set of vectors $\mathcal{C} = (\mathbf{f}_1, \ldots, \mathbf{f}_n)$ form an orthonormal basis if and only if the transition matrix ${}_{\mathcal{B}}T_{\mathcal{C}}$ is an orthogonal matrix.*

Proof. Denote the (i,j)-th entry of ${}_{\mathcal{B}}T_{\mathcal{C}}$ as $({}_{\mathcal{B}}T_{\mathcal{C}})_{i,j} = a_{i,j}$. We can write the inner product of $\mathbf{f}_i, \mathbf{f}_j$ as follows:

$$\begin{aligned}
\langle \mathbf{f}_i,\, \mathbf{f}_j \rangle &= \left\langle \sum_{k=1}^{n} a_{k,i}\mathbf{e}_k,\, \sum_{\ell=1}^{n} a_{\ell,j}\mathbf{e}_\ell \right\rangle && \text{(by defn. of transition matrices)}\\
&= \sum_{k,\ell=1}^{n} a_{k,i}a_{\ell,j}\langle \mathbf{e}_k,\, \mathbf{e}_\ell\rangle && \text{(by linearity of inner products)}\\
&= \sum_{k,\ell=1}^{n} a_{k,i}a_{\ell,j}\delta_{k\ell} && \text{(as } \mathcal{B} \text{ is orthonormal)}\\
&= \sum_{k=1}^{n} a_{k,i}a_{k,j} && \text{(nonzero only when } k=\ell)\\
&= \sum_{k=1}^{n} a^{\mathsf{T}}_{i,k}a_{k,j}\\
&= (({}_{\mathcal{B}}T_{\mathcal{C}})^{\mathsf{T}}({}_{\mathcal{B}}T_{\mathcal{C}}))_{ij}
\end{aligned}$$

The vectors $\mathcal{C}$ form an orthonormal basis if and only if $\langle \mathbf{f}_i,\, \mathbf{f}_j \rangle = \delta_{ij}$, therefore if and only if $(({}_{\mathcal{B}}T_{\mathcal{C}})^{\mathsf{T}}({}_{\mathcal{B}}T_{\mathcal{C}}))_{ij} = \delta_{ij}$. This condition is equivalent to $({}_{\mathcal{B}}T_{\mathcal{C}})^{\mathsf{T}}({}_{\mathcal{B}}T_{\mathcal{C}}) = I_n$, *i.e.*, the (i,j)-entry is 1 if $i = j$ and 0 otherwise. Therefore $\langle \mathbf{f}_i,\, \mathbf{f}_j \rangle = \delta_{ij}$ if and only if ${}_{\mathcal{B}}T_{\mathcal{C}}$ is orthogonal. □

Wikipedia: general linear group

Remark 1.33 The set of $n \times n$ transition matrices (with matrix multiplication as the operation) form a group called the *general linear group* $\mathrm{GL}(n,\mathbb{R})$. It can be checked that they form a group: in particular if $M, N \in \mathrm{GL}(n,\mathbb{R})$ then $MN \in \mathrm{GL}(n,\mathbb{R})$ as $\det(MN) = \det(M)\det(N)$ (see Proposition 1.46), both of which have non-zero determinant.

Wikipedia: orthogonal group

The set of $n \times n$ orthogonal matrices form a subgroup of $\mathrm{GL}(n,\mathbb{R})$ called the *orthogonal group* $O(n)$. Checking this forms a group is very similar to $\mathrm{GL}(n,\mathbb{R})$.

1.4 Linear operators

1.4.1 Matrix of a linear operator in a given basis

Wikipedia: linear transformation

Definition 1.34 (Linear transformation/operator) Let V and W be real vector spaces. A map $P\colon V \to W$ is called a *linear transformation* if

$$P(\lambda\mathbf{x} + \mu\mathbf{y}) = \underline{\hspace{6cm}} \tag{1.11}$$

for all $\lambda, \mu \in \mathbb{R}$ and $\mathbf{x}, \mathbf{y} \in V$. This property is commonly referred to as *linearity*.

In the special case that $V = W$ we also refer to P as a *linear operator*: a transformation that operates on a single vector space.

Let $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ be a given (ordered) basis for the vector space V and consider the action of the operator P on the basis vectors. The image $P(\mathbf{e}_i)$ of a basis element is in V and hence there are coefficients $p_{k,i} \in \mathbb{R}$ such that

$$P(\mathbf{e}_i) = \tag{1.12}$$

Representing $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ as standard column vectors as in Subsubsection 1.1.4 this gives

$$P(\mathbf{e}_1) = \begin{bmatrix} p_{1,1} \\ p_{2,1} \\ \vdots \\ p_{n,1} \end{bmatrix} \quad P(\mathbf{e}_2) = \begin{bmatrix} p_{1,2} \\ p_{2,2} \\ \vdots \\ p_{n,2} \end{bmatrix} \quad \cdots \quad P(\mathbf{e}_n) = \begin{bmatrix} p_{1,n} \\ p_{2,n} \\ \vdots \\ p_{n,n} \end{bmatrix}$$

The column vector representing $\mathbf{e}_i$ has a 1 in the i^{th} row and 0s elsewhere. The result of a matrix multiplication by this vector is simply the ______ column of the matrix. Thus, by concatenating these columns, we have a matrix for which multiplication is equivalent to applying the linear operator.

Definition 1.35 (Matrix of a linear operator) Let $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ be a basis for a vector space V, and let P be a linear operator on V. The matrix

$$[P]_{\mathcal{B}} =$$

determined by equation (1.12) is the *matrix of the linear transformation P in the basis $\mathcal{B}$*.

Wikipedia: matrix of the linear transformation P in the basis $\mathcal{B}$

Notice that if the linear operator is invertible then $\mathcal{C} = (P(\mathbf{e}_1), \ldots, P(\mathbf{e}_n))$ is a basis for V and the matrix of the linear operator coincides with the transition matrix from basis $\mathcal{B}$ to a basis $\mathcal{C}$.

Remark 1.36 A linear operator does not depend on a choice of basis, however the matrix of a linear operator does require a choice of basis since we are applying a linear operator to the choice of basis.

Example 1.37 Suppose that we have a vector space V with basis $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2)$. Let P be the following linear operator:

$$P(\mathbf{e}_1) = \mathbf{e}_1$$
$$P(\mathbf{e}_2) = \mathbf{e}_1 + 2\mathbf{e}_2.$$

What is the matrix of P? Does $P(\mathcal{B})$ define a basis?

We wish to consider linear operators, and vectors in general, in different bases. As such we shall introduce the following notation:

Notation 1.38 Let $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ be a basis for a vector space V and let $\mathbf{v}$ be a vector in V. We shall denote the column vector of $\mathbf{v}$ in the basis $\mathcal{B}$ as follows:

$$\text{if} \quad \mathbf{v} = \qquad\qquad \text{then} \quad [\mathbf{v}]_{\mathcal{B}} =$$

Moreover this notation is linear (property (1.11)): that is, if $\mathbf{w} = \sum_{i=1}^{n} w_i \mathbf{e}_i$ is another vector in V and $\lambda, \mu \in \mathbb{R}$ are scalars, then

$$\begin{aligned} \lambda[\mathbf{v}]_{\mathcal{B}} + \mu[\mathbf{w}]_{\mathcal{B}} &= \lambda \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} + \mu \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} \\ &= \begin{bmatrix} \lambda v_1 + \mu w_1 \\ \vdots \\ \lambda v_n + \mu w_n \end{bmatrix} \\ &= [\lambda \mathbf{v} + \mu \mathbf{w}]_{\mathcal{B}}. \end{aligned}$$

Proposition 1.39 *Let $P\colon V \to V$ be a linear operator acting on the vector space V and let $\mathcal{B}$ be a basis for V. Then*

$$[P(\mathbf{v})]_{\mathcal{B}} = \tag{1.13}$$

for all vectors $\mathbf{v} \in V$.

The proof of this proposition is left as an exercise.

We shall consider the effect of multiplication by a transition matrix on column vectors in the given bases.

Lemma 1.40 *Let $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ and $\mathcal{C} = (\mathbf{f}_1, \ldots, \mathbf{f}_n)$ be two bases for a vector space V. Let $T = {}_{\mathcal{B}}T_{\mathcal{C}}$ be the transition matrix from basis $\mathcal{B}$ to basis $\mathcal{C}$ and let $\mathbf{v} \in V$. Then*

$$ \tag{1.14}$$

Proof. Since $T = {}_{\mathcal{B}}T_{\mathcal{C}}$ then (by definition) the i^{th} column of T is $[\mathbf{f}_i]_{\mathcal{B}}$. Thus

$$[\mathbf{f}_i]_{\mathcal{B}} = T \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} {}_{i^{\text{th}}\text{ row}} = T[\mathbf{f}_i]_{\mathcal{C}}$$

We have proven the result whenever $\mathbf{v} \in \mathcal{C}$, but we wish to prove the equality for a general $\mathbf{v} = \sum v_i \mathbf{f}_i \in V$. This follows by linearity (property (1.11)):

$$\begin{aligned} T[\mathbf{v}]_{\mathcal{C}} &= T\left(v_1[\mathbf{f}_1]_{\mathcal{C}} + \cdots + v_n[\mathbf{f}_n]_{\mathcal{C}}\right) \\ &= v_1[\mathbf{f}_1]_{\mathcal{B}} + \cdots + v_n[\mathbf{f}_n]_{\mathcal{B}} \\ &= [\mathbf{v}]_{\mathcal{B}} \end{aligned}$$

□

Lemma 1.41 *Let $\mathcal{B}, \mathcal{C}, \mathcal{D}$ be three bases of a vector space V. Then*

$${}_{\mathcal{B}}T_{\mathcal{D}} =$$

Proof. Lemma 1.40 states that

$$\begin{aligned} [\mathbf{v}]_{\mathcal{B}} &= {}_{\mathcal{B}}T_{\mathcal{C}}\ [\mathbf{v}]_{\mathcal{C}} \\ &= {}_{\mathcal{B}}T_{\mathcal{C}}\ {}_{\mathcal{C}}T_{\mathcal{D}}\ [\mathbf{v}]_{\mathcal{D}} \end{aligned}$$

but this property applied to vectors in the basis $\mathcal{D}$ defines ${}_{\mathcal{B}}T_{\mathcal{D}}$. □

Lemma 1.42 *Let $\mathcal{B}, \mathcal{C}$ be bases of a vector space V. Then ${}_{\mathcal{C}}T_{\mathcal{B}} = ({}_{\mathcal{B}}T_{\mathcal{C}})^{-1}$.*

The proof is left as an exercise, see homework 2.

The lemmas above gives us a way to translate between column vectors in one basis to column vectors in another basis. Using these ideas we can translate from the matrix of a linear operator in one basis to the matrix of the same linear operator in another basis.

Proposition 1.43 *Let P be a linear operator acting on a vector space V. Let $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ and $\mathcal{C} = (\mathbf{f}_1, \ldots, \mathbf{f}_n)$ be two bases for V. Let T be the transition matrix from basis $\mathcal{B}$ to basis $\mathcal{C}$.*

Then $[P]_{\mathcal{C}} = T^{-1}[P]_{\mathcal{B}}T$.

Proof. The i^{th} column of $[P]_{\mathcal{C}}$ is defined to be $[P(\mathbf{f}_i)]_{\mathcal{C}}$. We have that

$$\begin{aligned}[P(\mathbf{f}_i)]_{\mathcal{C}} &= T^{-1}[P(\mathbf{f}_i)]_{\mathcal{B}} && \text{(Lemma 1.40)}\\ &= T^{-1}[P]_{\mathcal{B}}[\mathbf{f}_i]_{\mathcal{B}}\\ &= T^{-1}[P]_{\mathcal{B}}T[\mathbf{f}_i]_{\mathcal{C}}, && \text{the } i^{\text{th}} \text{ column of } T^{-1}[P]_{\mathcal{B}}T.\end{aligned}$$

□

Remark 1.44 This proposition demonstrates that the matrix for a linear operator is not determined by the linear operator alone and a choice of basis must be made.

1.4.2 Determinant and Trace of linear operator

We recall the determinant of a matrix for completeness.

Definition 1.45 (Determinant of a matrix) Let M be an $n \times n$ matrix. The *determinant* of a matrix is defined as

Wikipedia: determinant

$$\det M =$$

where S_n is the group of permutations on the set $\{1, \dots, n\}$.

Wikipedia: Permutation group

However, we shall mostly be concerned with 2×2 and 3×3 matrices and so we also recall the specialised version of this definition to these cases:

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc \tag{1.15}$$

$$\det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = a \det \begin{bmatrix} e & f \\ h & i \end{bmatrix} - b \det \begin{bmatrix} d & f \\ g & i \end{bmatrix} + c \det \begin{bmatrix} d & e \\ g & h \end{bmatrix} \tag{1.16}$$

We also recall (without proof) the following facts about the determinant:

Proposition 1.46 *For all $n \times n$ matrices M and N, the following properties hold:*

(1) M is invertible if and only if ____________________;

(2) $\det(MN) =$ __________________________;

(3) $\det(M^{\mathsf{T}}) =$ _____________;

(4) $\det(\lambda M) =$ __________________ *for all scalars* $\lambda \in \mathbb{R}$.

Notice that because the determinant of the identity is 1, property (2) of the lemma above immediately implies that if M is invertible then $\det(M^{-1}) = (\det M)^{-1}$.

Next we introduce the trace of a matrix.

Definition 1.47 (Trace of a matrix) Let M be an $n \times n$ matrix. We define the *trace* of M to be the sum of the diagonal elements

Wikipedia: trace

$$\operatorname{Tr} M =$$

We shall state without proof several useful properties of the trace map.

Proposition 1.48 *The trace map satisfies the following properties:*

(1) $\operatorname{Tr}(M + N) =$ ____________________ *for all $n \times n$-matrices M and N;*

(2) $\operatorname{Tr}(\lambda M) =$ __________ *for all matrices M and scalars $\lambda \in \mathbb{R}$;*

(3) $\operatorname{Tr}(MN) =$ __________ *for any $m \times n$-matrix M and any $n \times m$-matrix N.*

Remark 1.49

- Notice that together, properties (1) and (2) of the above proposition mean that trace satisfies __________.
- As an aside we also remark that any map that takes matrices to scalars satisfying the properties in Proposition 1.48 must be a scalar multiple of the ____________.

Proposition 1.50 *Let P be a linear operator on the vector space V. Let $\mathcal{B}$ and $\mathcal{C}$ be two bases for V. Then*

$$\det([P]_{\mathcal{B}}) = \_\_\_\_\_\_\_\_\_\_ \text{ and } \operatorname{Tr}([P]_{\mathcal{B}}) = \_\_\_\_\_\_\_\_\_\_.$$

Proof. Let T be the transition matrix from basis $\mathcal{B}$ to basis $\mathcal{C}$.

$$\begin{aligned}
\det([P]_{\mathcal{C}}) &= \det(T^{-1}[P]_{\mathcal{B}}T) \\
&= \det(T^{-1}) \det([P]_{\mathcal{B}}) \det T \\
&= (\det T)^{-1} \det([P]_{\mathcal{B}}) \det T = \det([P]_{\mathcal{B}})
\end{aligned}$$

$$\begin{aligned}
\operatorname{Tr}([P]_{\mathcal{C}}) &= \operatorname{Tr}(T^{-1}[P]_{\mathcal{B}}T) \\
&= \operatorname{Tr}([P]_{\mathcal{B}}TT^{-1}) = \operatorname{Tr}[P]_{\mathcal{B}}
\end{aligned}$$

□

Definition 1.51 (Determinant and Trace of linear operators) Let P be a linear operator on the vector space V. Let $\mathcal{B}$ be any basis for V.

We define the *determinant of the linear operator P* to be the determinant of the matrix $[P]_{\mathcal{B}}$

$$\det P =$$

Similarly we define the *trace of the linear operator P* as the trace of the matrix

$$\operatorname{Tr} P = \underline{\hspace{4cm}}.$$

Remark 1.52 As defined the determinant and trace of a linear operator requires the choice of a basis, however in light of Proposition 1.50 we see that this choice does not actually matter.

1.4.3 Eigenvalues and eigenvectors of linear operators

We can determine some behaviour of a linear operator by considering the vectors whose direction do not change under the action of the operator, only their magnitude.

The word "eigen" comes from German and it means "self" or "own". So an eigenvector or eigenvalue can be seen as a linear operators own proper vector/value.
Wikipedia: eigenvector
Wikipedia: eigenvalue

Definition 1.53 (Eigenvalues and eigenvectors) Let P be a linear operator on the vector space V. An *eigenvector* of P is a nonzero vector $\mathbf{x} \in V$ such that

$$P(\mathbf{x}) = \underline{\hspace{1.5cm}}$$

where $\lambda \in \mathbb{R}$ is the *eigenvalue* associated to $\mathbf{x}$. That is, P scales $\mathbf{x}$ by λ.

Remark 1.54 Of particular note is the case when $\lambda = 1$, as then the associated eigenvectors of λ are completely fixed by the action of P. These vectors will be particularly useful in Subsubsection 1.6.2 when computing rotations in $\mathbb{E}^3$.

The following lemma shows that every eigenvector gives rise to a ____________________________ of eigenvectors.

Lemma 1.55 *Let P be a linear operator on the vector space V. If $\mathbf{x}$ is an eigenvector of P with eigenvalue λ, then all vectors $\mathbf{y} \in \operatorname{span}(\mathbf{x})$ are also eigenvectors with eigenvalue λ.*

Proof. Pick some $\mathbf{y} = \mu\mathbf{x} \in \operatorname{span}(\mathbf{x})$ where $\mu \in \mathbb{R}$. Then by linearity of P

$$P(\mathbf{y}) = P(\mu\mathbf{x}) = \mu P(\mathbf{x}) = \mu\lambda\mathbf{x} = \lambda\mathbf{y}. \qquad \square$$

There is a rich theory behind eigenvalues and eigenvectors, however for our purposes this is all we will need.

Example 1.56 Let $\mathcal{B} = \{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$ be an orthonormal basis for $\mathbb{E}^3$ and consider the linear operator P on $\mathbb{E}^3$ such that

$$P(\mathbf{e}_x) = \mathbf{e}_y, \qquad P(\mathbf{e}_y) = \mathbf{e}_x, \qquad P(\mathbf{e}_z) = -\mathbf{e}_z. \tag{1.17}$$

This operator swaps first two basis vectors and inverts the third one.

The matrix of P in the basis $\mathcal{B}$ is

$$[P]_{\mathcal{B}} =$$

We are told that P has an eigenvalue of $\lambda = 1$, and wish to find its associated eigenvectors. An eigenvector $\mathbf{x} = x\mathbf{e}_x + y\mathbf{e}_y + z\mathbf{e}_z$ with $\lambda = 1$ satisfies

$$P(\mathbf{x}) = \lambda\mathbf{x} = \mathbf{x} \quad \Rightarrow$$

$$\Rightarrow \quad \begin{bmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Therefore, we have the following three equations:

Solving these three simultaneous equations, we find $x = y$, $z = 0$ and so the eigenvectors with eigenvalue 1 are of the form $\mathbf{x} = \alpha(\mathbf{e}_x + \mathbf{e}_y)$ where $\alpha \in \mathbb{R}$.

1.4.4 Orthogonal linear operators

Definition 1.57 (Orthogonal linear operator) A linear operator P acting on a Euclidean vector space V is called an *orthogonal linear operator* if P preserves inner products. That is,

Wikipedia: orthogonal linear operator

(1.18)

The following proposition relates orthogonal operators to orthogonal matrices.

Proposition 1.58 *Let $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ be an orthonormal basis for V and P a linear operator on V. Then the following are equivalent:*

(1) P is an ________________ operator,

(2) $P(\mathcal{B})$ is an ____________________ basis,

(3) $[P]_{\mathcal{B}}$ is an ________________ matrix.

Proof. The equivalence between (2) and (3) is exactly Proposition 1.32 where $[P]_{\mathcal{B}}$ is the transition matrix from $\mathcal{B}$ to $P(\mathcal{B})$. We show that (1) and (2) are equivalent.

Suppose P is an orthogonal operator, then equation (1.18) gives

$$\langle P(\mathbf{e}_i),\, P(\mathbf{e}_j)\rangle = \langle \mathbf{e}_i,\, \mathbf{e}_j\rangle = \delta_{i,j}. \tag{1.19}$$

Hence, $P(\mathcal{B})$ is an orthonormal basis for V.

Conversely, let $P(\mathcal{B})$ be an orthonormal basis. Let $\mathbf{x} = \sum_{i=1}^n x_i\mathbf{e}_i$ and $\mathbf{y} = \sum_{j=1}^n y_j\mathbf{e}_j$ be two arbitrary vectors of V, then by linearity of inner products and of P we see:

$$\begin{aligned}
&\langle P(\mathbf{x}),\, P(\mathbf{y})\rangle \\
&\quad= \left\langle P\left(\sum_{i=1}^n x_i\mathbf{e}_i\right),\, P\left(\sum_{j=1}^n y_j\mathbf{e}_j\right)\right\rangle \\
&\quad= \left\langle \sum_{i=1}^n x_i P(\mathbf{e}_i),\, \sum_{j=1}^n y_j P(\mathbf{e}_j)\right\rangle && \text{(linearity of } P\text{)} \\
&\quad= \sum_{i,j} x_i y_j \langle P(\mathbf{e}_i),\, P(\mathbf{e}_j)\rangle && \text{(linearity of } \langle -,\, -\rangle\text{)} \\
&\quad= \sum_{i=1}^n x_i y_i && (P(\mathcal{B}) \text{ orthornomal}).
\end{aligned}$$

But recall from equation (1.9) that when $\mathbf{x}, \mathbf{y}$ are written in an orthonormal basis, their inner product is $\langle \mathbf{x},\, \mathbf{y}\rangle = \sum_{i=1}^n x_i y_i$. This agrees with $\langle P(\mathbf{x}),\, P(\mathbf{y})\rangle$, and so P is an orthogonal operator. □

Note as a corollary of this proposition that the determinant of an orthogonal operator is ± 1. In particular, orthogonal linear operators are invertible.