Proceeding as before, we 'complete the square' with respect to x_2 (we don't need to complete the square for x_1): we have

$$\begin{aligned} & -x_1^2 + 2x_2x_3 \\ = & -x_1^2 + \frac{1}{2}(x_2+x_3)^2 - \frac{1}{2}(x_2-x_3)^2 \end{aligned}$$

Hence, if we let

$$\begin{aligned} y_1 &= x_1 \\ y_2 &= \tfrac{1}{\sqrt{2}}(x_2+x_3) \\ y_3 &= \tfrac{1}{\sqrt{2}}(x_2-x_3) \end{aligned}$$

then we have

$$\underline{x}^t A \underline{x} = -y_1^2 + y_2^2 - y_3^2.$$

Furthermore, if we let

$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix},$$

and defined $P = Q^{-1}$, then

$$P^t A P = \begin{bmatrix} -1 & & \\ & 1 & \\ & & -1 \end{bmatrix}.$$

Hence, $p = 1, q = 2$ and

$$\operatorname{sig}(B_A) = -1.$$

3.3 Euclidean spaces

Throughout this section V will be a finite dimensional $\mathbb{R}$-vector space and $\mathbb{K} = \mathbb{R}$.

Definition 3.3.1. Let $B \in \mathrm{Bil}_{\mathbb{R}}(V)$ be a symmetric bilinear form. We say that B is an *inner product on V* if B satisfies the following property:

$$B(v,v) \geq 0, \quad \text{for every } v \in V, \text{ and } \; B(v,v) = 0 \Leftrightarrow v = 0_V.$$

If $B \in \mathrm{Bil}_{\mathbb{K}}(V)$ is an inner product on V then we will write

$$\langle u, v \rangle \stackrel{def}{=} B(u,v).$$

Remark 3.3.2. Suppose that $\langle , \rangle$ is an inner product on V. Then, we have the following properties:

i) $\langle \lambda u + v, w \rangle = \lambda \langle u, w \rangle + \langle v, w \rangle$, for every $u, v, w \in V, \lambda \in \mathbb{K}$,

ii) $\langle u, \lambda v + w \rangle = \lambda \langle u, v \rangle + \langle u, w \rangle$, for every $u, v, w \in V, \lambda \in \mathbb{K}$,

iii) $\langle u, v \rangle = \langle v, u \rangle$, for every $u, v \in V$.

iv) $\langle v, v \rangle \geq 0$, for every $v \in V$, with equality precisely when $v = 0_V$.

Property *iv*) is often referred to as the **positive-definite** property of an inner product.

Definition 3.3.3. An *Euclidean space, or inner product space,* is a pair $(V, \langle , \rangle)$, where V is a finite dimensional $\mathbb{R}$-vector space and $\langle , \rangle$ is an inner product on V.

Given an inner product space $(V, \langle , \rangle)$ we define the *norm function on V (with respect to $\langle , \rangle$)* to be the function

$$||.|| : V \rightarrow \mathbb{R}_{\geq 0} \; ; \; v \mapsto ||v|| = \sqrt{\langle v, v \rangle}.$$

For any $v \in V$ we define the *length of v (with respect to $\langle , \rangle$)* to be $||v|| \in \mathbb{R}_{\geq 0}$.

Let $(V_1, \langle , \rangle_1), (V_2, \langle , \rangle_2)$ be inner product spaces. Then, we say that a linear morphism

$$f : V_1 \to V_2,$$

is an *Euclidean morphism* if, for every $u, v \in V_1$ we have

$$\langle u, v \rangle_1 = \langle f(u), f(v) \rangle_2.$$

An Euclidean morphism whose underlying linear morphism is an isomorphism is called a *Euclidean isomorphism*.

If $f : (V, \langle , \rangle) \to (V, \langle , \rangle)$ is a Euclidean morphism such that the domain and codomain are the same Euclidean space, then we say that f is an *orthogonal morphism, or an orthogonal transformation*. We denote the set of all orthgonal transformations of $(V, \langle , \rangle)$ by $O(V, \langle , \rangle)$, or simply $O(V)$ when there is no confusion.

Example 3.3.4. 1. We define *n-dimensional Euclidean space,* denoted $\mathbb{E}^n$, to be the Euclidean space $(\mathbb{R}^n, \cdot)$, where $\cdot$ is the usual 'dot product' from analytic geometry: that is, for $\underline{x}, \underline{y} \in \mathbb{R}^n$ we have

$$\underline{x} \cdot \underline{y} \overset{def}{=} \underline{x}^t \underline{y} = x_1 y_1 + ... + x_n y_n.$$

It easy to check that $\cdot$ is bilinear and symmetric and, moreover, we have

$$\underline{x} \cdot \underline{x} = \underline{x}^t \underline{x} = x_1^2 + ... + x_n^2 \geq 0,$$

with equality precisely when $\underline{x} = \underline{0}$.

Given $\underline{x} \in \mathbb{E}^n$, the length of $\underline{x}$ is

$$||\underline{x}|| = \sqrt{x_1^2 + ... + x_n^2}.$$

2. Consider the symmetric bilinear form $B_A \in \mathrm{Bil}_{\mathbb{R}}(\mathbb{R}^3)$ where

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

Then, you can check that

$$\underline{x} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \in \mathbb{R}^3,$$

has the property that

$$B_A(\underline{x}, \underline{x}) = -2 < 0,$$

so that B_A is not an inner product on $\mathbb{R}^3$.

3. Let $B_A \in \mathrm{Bil}_{\mathbb{R}}(\mathbb{R}^4)$ be the symmetric bilinear form defined by

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

Then, B_A is an inner product: indeed, let $\underline{x} \in \mathbb{R}^3$. Then, we have

$$B_A(\underline{x}, \underline{x}) = x_1^2 + 2x_1x_2 + 2x_2^2 + 2x_3^2 + 2x_3x_4 + x_4^2 = (x_1 + x_2)^2 + x_2^2 + x_3^2 + (x_3 + x_4)^2 \geq 0,$$

and we have $B_A(\underline{x}, \underline{x}) = 0$ precisely when

$$x_1 + x_2 = 0,\ x_2 = 0,\ x_3 = 0,\ x_3 + x_4 = 0,$$

so that $x_1 = x_2 = x_3 = x_4 = 0$ and $\underline{x} = 0$.

With respect to this inner product, the vector

$$\underline{x} = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 1 \end{bmatrix} \in \mathbb{R}^4,$$

has length

$$||\underline{x}|| = \sqrt{\langle \underline{x}, \underline{x} \rangle} = \sqrt{2}.$$

4. In fact, a symmetric bilinear form B on an n-dimensional $\mathbb{R}$-vector space V is an inner product precisely when $\mathrm{sig}(B) = n$.[64]

5. Consider the linear morphism $T_A \in \mathrm{End}_{\mathbb{R}}(\mathbb{R}^2)$, where

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

Then, T_A is an orthogonal transformation of $\mathbb{E}^2$: indeed, for any $\underline{x}, \underline{y} \in \mathbb{R}^2$, we have

$$T_A(\underline{x}) \cdot T_A(\underline{y}) = (A\underline{x})^t (A\underline{y}) = \underline{x}^t A^t A \underline{y} = \underline{x}^t \underline{y} = \underline{x} \cdot \underline{y},$$

since $A^{-1} = A^t$.

This example highlights a more general property of orthogonal transformations of $\mathbb{E}^n$ to be discussed later:

$$A \in O(\mathbb{E}^n) \textbf{ if and only if } A^{-1} = A^t.$$[65]

6. If $(V, \langle , \rangle)$ is an Euclidean space then id_V is always an orthogonal transformation.

Remark 3.3.5. 1. A Euclidean space is simply a $\mathbb{R}$-vector space V equipped with an inner product. This means that it is possible for the <u>same</u> $\mathbb{R}$-vector space V to have two distinct Euclidean space structures (ie, we can equip the same $\mathbb{R}$-vector space with two distinct inner products). However, as we will see shortly, given a $\mathbb{R}$-vector space V there is *essentially* only one Euclidean space structure on V: this means that we can find a Euclidean isomorphism between the two distinct Euclidean space structures on V.

2. It is important to remember that the norm function $||.||$ is **not linear**. In fact, the norm function is **not additive**: indeed, let $v \in V$ be nonzero. Then,

$$0 = ||0_v|| = ||v + (-v)||,$$

so that if $||.||$ were additive then we would have $||v|| + || - v|| = 0$, for every $v \in V$. As $||v||, || - v|| \geq 0$ then we would have that

$$||v|| = || - v|| = 0, \text{ for every } v \in V.$$

That is, every $v \in V$ would have length 0. However, the only $v \in V$ that can have length 0 is $v = 0_V$.

Moreover, for any $v \in V, \lambda \in \mathbb{K}$, we have

$$||\lambda v|| = |\lambda| ||v||.$$

[64] This is shown in a few paragraphs.

Theorem 3.3.6. *Let $(V, \langle , \rangle)$ be an Euclidean space. Then,*

a) for any $u, v \in V$ we have

$$||u + v|| \leq ||u|| + ||v||. \qquad \textit{(triangle inequality)}$$

b) $||v|| = 0$ if and only if $v = 0_V$.

c) if $\langle u, v\rangle = 0$ then

$$||u||^2 + ||v||^2 = ||u + v||^2. \qquad \textit{(Pythagoras' theorem)}$$

d) for any $u, v \in V$ we have

$$|\langle u, v\rangle| \leq ||u|| ||v||. \qquad \textit{(Cauchy-Schwarz inequality)}$$

Proof: Left as an exercise for the reader. □

We will now show that there is essentially only one Euclidean space structure that we can give an arbitrary finite dimensional $\mathbb{R}$-vector space. Moreover, this Euclidean space structure is well-known to us all.

Lemma 3.3.7. *Suppose that $\langle , \rangle$ is an inner product on V. Then, $\langle , \rangle \in \mathrm{Bil}_{\mathbb{R}}(V)$ is nondegenerate.*

Proof: We need to show the following property of $\langle , \rangle$:

$$\text{if } v \in V \text{ is such that } \langle u, v\rangle = 0, \text{ for every } u \in V, \text{ then } v = 0_V.$$

So, suppose that $v \in V$ is such that $\langle u, v\rangle = 0$, for every $u \in V$. In particular, we must have

$$\langle v, v\rangle = 0 \implies v = 0_V,$$

by the defining property of an inner product (Remark 3.3.2, iv)). Hence, $\langle , \rangle$ is nondegenerate. □

Hence, using Sylvester's law of inertia (Theorem 3.2.6), we know that for an Euclidean space $(V, \langle , \rangle)$ there is an ordered basis $\mathcal{B} \subset V$ such that

$$[\langle , \rangle]_{\mathcal{B}} = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix}, \quad \text{where } d_i \in \{1, -1\}, \ n = \dim V.$$

Moreover, since $\langle , \rangle$ is an inner product we must have that $\mathrm{sig}(\langle , \rangle) = n$: indeed, we have

$$\mathrm{sig}(\langle , \rangle) = p - q \in \{-n, -(n-1), \ldots, n-1, n\},$$

so that $\mathrm{sig}(\langle , \rangle) = n$ if and only if $q = 0$, so that there are no -1s appearing on the diagonal of $[\langle , \rangle]_{\mathcal{B}}$. If some $d_i = -1$ then we would have

$$0 \leq \langle b_i, b_i\rangle = -1,$$

which is impossible. Hence, we must have $d_1 = d_2 = \ldots = d_n = 1$, so that

$$[\langle , \rangle]_{\mathcal{B}} = I_n.$$

Theorem 3.3.8 (Classification of Euclidean spaces)**.** *Let $(V, \langle , \rangle)$ be an Euclidean space, $n = \dim V$. Then, there is a Euclidean isomorphism*

$$f : (V, \langle , \rangle) \to \mathbb{E}^n.$$

Proof: Let $\mathcal{B} \subset V$ be an ordered basis such that

$$[\langle , \rangle]_{\mathcal{B}} = I_n.$$

Then, let

$$f = [-]_{\mathcal{B}} : V \to \mathbb{R}^n,$$

be the $\mathcal{B}$-coordinate morphism. Then, this is an isomorphism of $\mathbb{R}$-vector spaces so that we need only show that

$$\langle u, v \rangle = [u]_{\mathcal{B}} \cdot [v]_{\mathcal{B}},$$

for every $u, v \in V$. Now, let $u, v \in V$ and suppose that

$$u = \sum_{i=1}^{n} \lambda_i b_i, \ \ u = \sum_{j=1}^{n} \mu_j b_j.$$

Then,

$$\langle u, v \rangle = \langle \sum_{i=1}^{n} \lambda_i b_i, \sum_{j=1}^{n} \mu_j b_j \rangle = \sum_{i,j} \lambda_i \mu_j \langle b_i, b_j \rangle = \sum_{i=1}^{n} \lambda_i \mu_i,$$

where we have used bilinearity of $\langle , \rangle$ and that

$$\langle b_i, b_j \rangle = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

Now, we also have

$$[u]_{\mathcal{B}} \cdot [v]_{\mathcal{B}} = [u]_{\mathcal{B}}^t [v]_{\mathcal{B}} = [\lambda_1 \ \cdots \ \lambda_n] \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix} = \sum_{i=1}^{n} \lambda_i \mu_i = \langle u, v \rangle,$$

and the result follows. □

Corollary 3.3.9. *Let $(V_1, \langle , \rangle_1), (V_2, \langle , \rangle_2)$ be Euclidean spaces. Then, if $\dim V_1 = \dim V_2$ then $(V_1, \langle , \rangle_1)$ and $(V_2, \langle , \rangle_2)$ are Euclidean-isomorphic.*

Proof: By Theorem 3.3.8 we have Euclidean isomorphisms

$$f_1 : (V_1, \langle , \rangle_1) \to \mathbb{E}^n, \ \ f_2 : (V_2, \langle , \rangle_2) \to \mathbb{E}^n.$$

Then, as the composition of two Euclidean isomorphisms is again a Euclidean isomorphism[66] then we obtain an isomorphism

$$f_2^{-1} \circ f_1 : (V_1, \langle , \rangle_1) \to (V_2, \langle , \rangle_2).$$

□

In fact, the condition defining an Euclidean morphism (not necessarily an isomorphism) is extremely strong: if $(V_1, \langle , \rangle_1)$ and $(V_2, \langle , \rangle_2)$ are Euclidean spaces and $f : V_1 \to V_2$ is a Euclidean morphism, then it is easy to check that we must have

$$||v|| = ||f(v)||, \text{ for every } v \in V,$$

so that f is *length preserving*. If you think about what this means geometrically then we obtain that

'Euclidean morphisms are always injective'

since no nonzero vector can be mapped to 0_{V_2} by f. As a consequence, we obtain

[66] Check this.

Proposition 3.3.10. *Let $(V_1, \langle, \rangle_1), (V_2, \langle, \rangle_2)$ be Euclidean spaces of the same dimension. Then, if there exists an Euclidean morphism*

$$f : V_1 \to V_2,$$

it must automatically be an Euclidean isomorphism.

Corollary 3.3.11. *Let $(V, \langle, \rangle)$ be an Euclidean space. Then, every Euclidean endomorphism*

$$f : V \to V$$

is an orthogonal transformation (= Euclidean isomorphism). Hence, we have

$$O(V) = \{f \in \mathrm{End}_{\mathbb{R}}(V) \mid f \text{ is Euclidean}\}.$$

Definition 3.3.12. The set of orthogonal transformations of $\mathbb{E}^n$ is called the *orthogonal group of size* n and is denoted $O(n)$.

Suppose that $g \in O(n)$ is an orthogonal transformation of $\mathbb{E}^n$ and identify g with its standard matrix $[g]_{\mathcal{S}^{(n)}}$. Then, we must have, for every $\underline{x}, \underline{y} \in \mathbb{R}^n$, that

$$\underline{x} \cdot \underline{y} = (g\underline{x}) \cdot (g\underline{y}) = (g\underline{x})^t (g\underline{y}) = \underline{x}^t g^t g \underline{y},$$

so that

$$\underline{x}^t \underline{y} = \underline{x}^t g^t g \underline{y},$$

for every $\underline{x}, \underline{y} \in \mathbb{R}^n$. Hence, by Lemma 3.1.6, we must have that

$$g^t g = I_n.$$

Hence, we see that we can identify

$$[-]_{\mathcal{S}^{(n)}} : O(n) \to \{X \in Mat_n(\mathbb{R}) \mid X^t X = I_n\}.$$

Moreover, this identification satisfies the following properties:

- $[\mathrm{id}_{\mathbb{E}^n}]_{\mathcal{S}^{(n)}} = I_n$,
- for every $f, g \in O(n)$, $[f \circ g]_{\mathcal{S}^{(n)}} = [f]_{\mathcal{S}^{(n)}} [g]_{\mathcal{S}^{(n)}}$.

Hence, the correspondence

$$[-]_{\mathcal{S}^{(n)}} : O(n) \to \{X \in Mat_n(\mathbb{R}) \mid X^t X = I_n\},$$

is an *isomorphism of groups*.

From now on, when we consider orthogonal transformations $g \in O(n)$ we will identify g with its standard matrix. Then, the previous discussion shows that $g \in \mathrm{GL}_n(\mathbb{R})$ and $g^t g = I_n$.

Let's think a little bit more about the condition

$$A^t A = I_n,$$

for $A \in Mat_n(\mathbb{R})$.

i) If A is such that $A^t A = I_n$ then we must have that $\det(A)^2 = 1$, since $\det(A) = \det(A^t)$. In particular, $\det(A) \in \{1, -1\}$[67] so that $A \in \mathrm{GL}_n(\mathbb{R})$: the inverse of A is $A^{-1} = A^t$. Furthermore, this implies that we must have

$$AA^t = AA^{-1} = I_n,$$

so that

[67] It is NOT true that if $A \in \mathrm{GL}_n(\mathbb{R})$ such that $\det A = 1$ then $A \in O(n)$. For example, consider

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

Then, it is not the case that $A^t A = I_2$ so that $A \notin O(2)$.

$$\boxed{A^tA = I_n \textbf{ if and only if } AA^t = I_n.}$$

ii) Let us write

$$A = [a_1 \ \cdots \ a_n],$$

so that the i^{th} column of A is a_i. Then, as $A \in \mathrm{GL}_n(\mathbb{R})$ we have that $\{a_1, \dots, a_n\}$ is linearly independent and defines a basis of $\mathbb{R}^n$. Moreover, as the i^{th} row of A^t is a_i^t, then the condition $A^tA = I_n$ implies that

$$a_i \cdot a_j = a_i^t a_j = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases} \quad .$$

In particular, we see that **each column of A has length** 1[68] (with respect to the inner product $\cdot$), and that the $\cdot$-complement of a_i is precisely

$$\mathrm{span}_{\mathbb{R}}\{a_j \mid j \neq i\}.$$

iii) A matrix $A \in Mat_n(\mathbb{R})$ such that

$$A^tA = I_n,$$

will be called an *orthogonal matrix*.

iv) A matrix $A \in Mat_n(\mathbb{R})$ is an orthogonal matrix if and only if for every $\underline{x}, \underline{y} \in \mathbb{R}^n$ we have

$$(A\underline{x}) \cdot (A\underline{y}) = \underline{x} \cdot \underline{y}.$$

We can interpret this result using the slogan

'orthogonal transformations are the 'rigid' transformations'

Example 3.3.13. 1. Let $\theta \in \mathbb{R}$ and consider the matrix

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \in Mat_2(\mathbb{R}).$$

Then, you may know already that R_θ corresponds to the 'rotate by θ counterclockwise' morphism of $\mathbb{R}^2$. If not, then this is easily seen: since R_θ defines a linear transformation of $\mathbb{R}^2$ we need only determine what happens to the standard basis of $\mathbb{R}^2$. We have

$$R_\theta e_1 = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}, \ R_\theta e_2 = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix},$$

and by considering triangles and the unit circle the result follows.

You can check easily that

$$R_\theta^t R_\theta = I_2,$$

so that $\mathbb{R}_\theta \in O(2)$.

In fact, it can be shown that every orthogonal transformation of $\mathbb{R}^2$ that has determinant 1 is of the form R_θ, for some θ. Moreover, every orthogonal transformation of $\mathbb{R}^2$ is of one of the following forms:

$$R_\theta, \ \text{ or } \ \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} R_\theta.$$

[68]Similarly, we obtain that each row must have legnth 1

3.3.1 Orthogonal complements, bases and the Gram-Schmidt process

Definition 3.3.14. Let $(V, \langle , \rangle)$ be an Euclidean space, $S \subset V$ a nonempty subset. We define the *orthogonal complement of* S, denoted $S^\perp$, to be the $\langle , \rangle$-complement of S defined in Definition 3.1.15. Hence,

$$S^\perp = \{v \in V \mid \langle v, s\rangle = 0, \text{ for every } s \in S\} = \{v \in V \mid \langle s, v\rangle = 0, \text{ for every } s \in S\}.$$

$S^\perp$ is a subspace of V, for any subset $S \subset V$.[69]

Proposition 3.3.15. *Let $(V, \langle , \rangle)$ be an Euclidean space and $U \subset V$ a subspace. Then,*

$$V = U \oplus U^\perp.$$

Proof: We know that $\dim V = \dim U + \dim U^\perp$ by Proposition 3.1.17. Hence, if we show that $U \cap U^\perp = \{0_V\}$ then we must have

$$V = U + U^\perp = U \oplus U^\perp.^{70}$$

Assume that $v \in U \cap U^\perp$. Then, $v \in U$ and $v \in U^\perp$ so that

$$0 = \langle v, v\rangle \implies v = 0_V,$$

since $\langle , \rangle$ is an inner product. The result follows. □

Remark 3.3.16. 1. Just as we have shown before, we have

$$S^\perp = (\text{span}_{\mathbb{R}} S)^\perp.$$

2. If we are thinking geometrically (as we should do whenever we are given any Euclidean space V) then we see that the orthogonal complement $U^\perp$ of a subspace U is the subspace of V which is 'perpendicular' to U. For example, consider the Euclidean space $\mathbb{E}^3$, U is the 'x-axis', which we'll denote L. Then, the subspace that is perpendicular to the x-axis is the $x = 0$-plane Π. Indeed, we have

$$L = \left\{ \begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix} \in \mathbb{R}^3 \right\}, \text{ and } \Pi = \left\{ \begin{bmatrix} 0 \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 \right\}.$$

It is easy to check that $\Pi = L^\perp$.[71]

Definition 3.3.17. Let $(V, \langle , \rangle)$ be an Euclidean space, $U \subset V$ a subspace and $v \in V$. Then, we define the *projection of v onto U* to be the vector $\text{proj}_U v$ defined as follows: using Proposition 3.3.15 we know that $V = U \oplus U^\perp$ so that there exists (unique!) $u \in U, z \in U^\perp$ such that $v = u + z$. Then, we define

$$\text{proj}_U v \stackrel{def}{=} u \in U.$$

Remark 3.3.18. In fact, the assignment

$$\text{proj}_U : V \to U \ ; \ v \mapsto \text{proj}_U v,$$

is precisely the 'projection onto U' morphism defined earlier. As a consequence we see that

$$\text{proj}_U(v + v') = \text{proj}_U v + \text{proj}_U v', \text{ and } \text{proj}_U \lambda v = \lambda \, \text{proj}_U v.$$

We can think of $\text{proj}_U v$ in more geometric terms.

[69] Check this.
[70] This follows from the dimension formula.
[71] Do this!

Proposition 3.3.19. *Let $(V, \langle , \rangle)$ be an Euclidean space, $U \subset V$ a subspace and $v \in V$. Then, $\text{proj}_U v \in U$ is the unique vector in U such that*

$$|| \text{proj}_U v - v|| \leq ||u - v||, \ u \in U.$$

Hence, we can say that $\text{proj}_U v$ is the closest vector to v in U.

Proof: Let $u \in U$. Then, we have

$$(\text{proj}_U v - v) + (u - \text{proj}_U v) = (u - v),$$

and, since $\text{proj}_U v - v \in U^\perp$ (Definition 3.3.17) and $u - \text{proj}_U v \in U$, then

$$||u - v||^2 = || \text{proj}_U v - v||^2 + ||u - \text{proj}_U v||^2 \geq || \text{proj}_U v - v||^2,$$

where we have used Pythagoras' theorem (Theorem 3.3.6). Hence, we have

$$||u - v|| \geq || \text{proj}_U v - v||, \quad \text{for any } u \in U.$$

Suppose that $w \in U$ is such that

$$||w - v|| \leq ||u - v||, \quad \text{for any } u \in U.$$

This implies that we must have

$$||w - v|| = || \text{proj}_U v - v||,$$

by what we have just shown.

Now, using Pythagoras' theorem, and that $v - \text{proj}_U v \in U^\perp, \text{proj}_U v - w \in U$, we obtain

$$||v-w||^2 = ||v-\text{proj}_U v+\text{proj}_U v-w||^2 = ||v-\text{proj}_U v||^2+|| \text{proj}_U v-w||^2 \implies || \text{proj}_U v-w||^2 = 0,$$

and $\text{proj}_U v = w$. Hence, $\text{proj}_U v$ is the unique element of U satisfying the above inequality. □

Example 3.3.20. Consider the Euclidean space $\mathbb{E}^2$ and let $L \subset \mathbb{R}^2$ be a line through the origin. Suppose that $v \in \mathbb{R}^2$ is an arbitrary vector. What does $\text{proj}_L v$ look like geometrically?

Using Proposition 3.3.19 we know that $w = \text{proj}_L v \in L$ is the unique vector in L that is closest to v.

- if $v \in L$ then $\text{proj}_L v = v$, as $v \in L$ is the closest vector v (trivially).
- if $v \notin L$ then consider the line L' perpendicular to L and for which the endpoint of the vector v lies on L' (so it might not be the case that L' is a line through the origin). The point of intersection $L \cap L'$ defines the vector $\text{proj}_L v$.

In fact, it is precisely this geometric intuition that guides the definition of $\text{proj}_L v$: we have defined $\text{proj}_L v \in L$ as the unique vector such that

$$v = \text{proj}_L v + z, \ z \in L^\perp.$$

Definition 3.3.21. Let $(V, \langle , \rangle)$ be an Euclidean space. We say that a subset $S \subset V$ is an *orthogonal set* if, for every $s, t \in S$, $s \neq t$, we have

$$\langle s, t \rangle = 0.$$

Lemma 3.3.22. *Let $S \subset V$ be an orthogonal set of nonzero vectors. Then, S is linearly independent.*

Proof: Left as en exercise for the reader. □

Lemma 3.3.23. *Let $S = \{s_1, \dots, s_k\} \subset V$ be an orthogonal set and such that S contains only nonzero vectors. Then, for any $v \in V$, we have*

$$\text{proj}_{\text{span}_\mathbb{R} S} v = \frac{\langle v, s_1 \rangle}{\langle s_1, s_1 \rangle} s_1 + \dots + \frac{\langle v, s_k \rangle}{\langle s_k, s_k \rangle} s_k.$$

Proof: Since S is linearly independent we have that S forms a basis of $\text{span}_{\mathbb{R}} S$. Hence, for any $v \in V$, we can write

$$\text{proj}_{\text{span}_{\mathbb{R}} S} v = \lambda_1 s_1 + ... + \lambda_k s_k,$$

for unique $\lambda_1, ..., \lambda_k \in \mathbb{R}$. Hence, for each $i = 1, ..., k$, we have

$$\langle \text{proj}_{\text{span}_{\mathbb{R}} S} v, s_i \rangle = \lambda_i \langle s_i, s_i \rangle,$$

using that S is orthogonal. Hence, we have that

$$\lambda_i = \frac{\langle \text{proj}_{\text{span}_{\mathbb{R}} S} v, s_i \rangle}{\langle s_i, s_i \rangle}.$$

Now, since $v - \text{proj}_{\text{span}_{\mathbb{R}} S} v \in (\text{span}_{\mathbb{R}} S)^{\perp}$ we see that, for each i,

$$0 = \langle v - \text{proj}_{\text{span}_{\mathbb{R}} S} v, s_i \rangle = \langle v, s_i \rangle - \langle \text{proj}_{\text{span}_{\mathbb{R}} S} v, s_i \rangle \implies \langle v, s_i \rangle = \langle \text{proj}_{\text{span}_{\mathbb{R}} S}, s_i \rangle.$$

The result follows. □

Definition 3.3.24. Let $(V, \langle , \rangle)$ be an Euclidean space. A basis $\mathcal{B} \subset V$ is called an *orthogonal basis* if it is an orthogonal set.

An orthogonal basis $\mathcal{B}$ is called *orthonormal* if, for every $b \in \mathcal{B}$, we have $||b|| = 1$.

Remark 3.3.25. 1. Recall that we defined an orthogonal matrix $A \in Mat_n(\mathbb{R})$ to be a matrix such that

$$A^t A = I_n.$$

The remarks at the end of the previous section imply that **the columns of an orthogonal matrix define an orthonormal basis**.

2. Not every basis in an Euclidean space is an orthogonal basis: for example, consider the Euclidean space $\mathbb{E}^2$. Then,

$$\mathcal{B} = \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) = (b_1, b_2),$$

is a basis of $\mathbb{R}^2$ but we have

$$b_1 \cdot b_2 = 1 \neq 0.$$

3. It is not true that any orthogonal set $E \subset V$ defines an orthogonal basis of $\text{span}_{\mathbb{R}} E$: for example, let $v \in V$ be nonzero and consider the subset $E = \{0_V, v\}$. Then, E is orthogonal[72] but E is not a basis, as E is a linearly dependent set. However, if E contains nonzero vectors and is orthogonal then E is an orthogonal basis of $\text{span}_{\mathbb{R}} E$, by Lemma 3.3.22.

At first glance it would appear to be quite difficult to determine an orthogonal (or orthonormal) basis of V. This is essentially the same problem as coming up with an orthogonal matrix. Moreover, it is hard to determine whether orthogonal bases even exist!

It is a quite remarkable result that given **ANY** basis $\mathcal{B}$ of an Euclidean space $(V, \langle , \rangle)$ we can determine an **orthonormal** basis $\mathcal{B}'$ of V. This is the **Gram-Schmidt process**.

Theorem 3.3.26 (Gram-Schmidt process)**.** *Let $(V, \langle , \rangle)$ be an Euclidean space, $\mathcal{B} = (b_1, ..., b_n) \subset V$ an arbitrary ordered basis of V. Then, there exists an orthonormal basis $\mathcal{B}' = (b'_1, ..., b'_n) \subset V$.*

Proof: Consider the following algorithm: define

$$c_1 = b_1.$$

We inductively define c_i: for $2 \leq i \leq n$ define

$$c_i = b_i - \text{proj}_{E_{i-1}} b_i,$$

where $E_{i-1} \stackrel{def}{=} \text{span}_{\mathbb{R}}\{c_1, ..., c_{i-1}\}$.

If $i < j$ then

$$\langle c_i, c_j \rangle = 0,$$

since $c_j \in E_{j-1}^{\perp}$ by construction[73], and $c_i \in E_{j-1}$.

[72] Check this.

[73] Think about why this is true. What is the definition of c_j?

Hence, $\mathcal{C} = (c_1, \dots, c_n)$ is an orthogonal basis. To obtain an orthonormal basis $\mathcal{B}' = (b'_1, \dots, b'_n)$ given an orthogonal basis $\mathcal{C}$, we simply set

$$b'_i = \frac{c_i}{||c_i||}.$$

Then, we have

$$||b'_i|| = 1,$$

and $\mathcal{B}'$ is an orthonormal basis. $\square$

Corollary 3.3.27. *Let $(V, \langle , \rangle)$ be an Euclidean space, $E \subset V$ an orthogonal set consisting of nonzero vectors. Then, E can be extended to an orthogonal basis of V.*

Proof: Left as an exercise for the reader. $\square$

Remark 3.3.28. 1. Let's illuminate exactly what we have done in the proof of Theorem 3.3.26, making use of Lemma 3.3.23.

Let $\mathcal{B} = (b_1, \dots, b_n)$ be **any** basis. We can organise the algorithm from Theorem 3.3.26 into a table

$$\begin{aligned}
c_1 = &\quad b_1 \\
c_2 = &\quad b_2 - \tfrac{\langle b_2, c_1\rangle}{\langle c_1, c_1\rangle} c_1 \\
c_3 = &\quad b_3 - \tfrac{\langle b_3, c_1\rangle}{\langle c_1, c_1\rangle} c_1 - \tfrac{\langle b_3, c_2\rangle}{\langle c_2, c_2\rangle} c_2 \\
\vdots & \\
c_n = &\quad b_n - \tfrac{\langle b_n, c_1\rangle}{\langle c_1, c_1\rangle} c_1 - \dots - \tfrac{\langle b_n, c_{n-1}\rangle}{\langle c_{n-1}, c_{n-1}\rangle} c_{n-1}
\end{aligned}$$

Then $\mathcal{C} = (c_1, \dots, c_n)$ is an orthogonal basis of V. To obtain an orthonormal basis of V we set

$$b'_i = \frac{c_i}{||c_i||}, \text{ for each } i.$$

Then, $\mathcal{B}' = (b'_1, \dots, b'_n)$ is orthonormal.

In practice in can be quite painful to actually perform the Gram-Schmidt process (if $\dim V$ is large). However, it is important to know that the Gram-Schmidt process allows us to show that **orthonormal bases exist**.

2. If $\mathcal{B}$ is orthogonal to start with then the basis $\mathcal{C}$ we obtain after performing the Gram-Schmidt process is just $\mathcal{C} = \mathcal{B}$.

3. It is important to remember that **the Gram-Schmidt process depends on the inner product $\langle , \rangle$ used to define the Euclidean space $(V, \langle , \rangle)$.**

Example 3.3.29. Let $V = \mathbb{E}^2$ and consider the basis

$$\mathcal{B} = \left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 5 \end{bmatrix} \right).$$

Let's perform the Gram-Schmidt process to obtain an orthogonal basis $\mathcal{C} = (c_1, c_2)$ of $\mathbb{E}^2$. We have

$$\begin{aligned}
c_1 = &\quad \begin{bmatrix} 1 \\ -1 \end{bmatrix} \\
c_2 = &\quad \begin{bmatrix} 2 \\ 5 \end{bmatrix} - \frac{\begin{bmatrix} 2 \\ 5 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix}}{\begin{bmatrix} 1 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \end{bmatrix} - \tfrac{2.1+5.(-1)}{1^2+(-1)^2} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \end{bmatrix} + \tfrac{3}{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 7/2 \\ 7/2 \end{bmatrix}
\end{aligned}$$

Then, you can check that

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 7/2 \\ 7/2 \end{bmatrix} = 7/2 - 7/2 = 0.$$

If we define

$$b'_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \ b'_2 = \frac{2}{7\sqrt{2}} \begin{bmatrix} 7/2 \\ 7/2 \end{bmatrix},$$

we have that $\mathcal{B}' = (b'_1, b'_2)$ is orthonormal.

Corollary 3.3.30 (QR factorisation)**.** *Let $A \in \mathrm{GL}_n(\mathbb{R})$. Then, there exists an orthogonal matrix $Q \in O(n)$ and a upper-triangular matrix R such that*

$$A = QR.$$

Proof: This is a simple restatement of the Gram-Schmidt process. Suppose that

$$A = [a_1 \ \cdots \ a_n].$$

Then $\mathcal{B} = (a_1, \dots, a_n)$ is an ordered basis of $\mathbb{R}^n$. Apply the Gram-Schmidt process (with respect to the dot product) to obtain an orthonormal basis $\mathcal{B}' = (b_1, \dots, b_n)$ as above. Then, we have

$$\begin{aligned} b_1 = & \ \tfrac{1}{r_1} a_1 \\ b_2 = & \ \tfrac{1}{r_2} (a_2 - (a_2 \cdot b_1) b_1) \\ \vdots & \\ b_n = & \ \tfrac{1}{r_n} (a_n - (a_n \cdot b_1) b_1 - \dots - (a_n \cdot b_{n-1}) b_{n-1}) \end{aligned}$$

where $r_i \in \mathbb{R}_{>0}$ is the length of the c_i vectors from the Gram-Schmidt process. We have also slightly modified the Gram-Schmidt process (in what way?) but you can check that $(b_1, \dots, b_n)$ is an orthonormal basis.[74]

By moving all b_i terms to the left hand side of the above equations we obtain the table

$$\begin{aligned} r_1 b_1 \quad &= a_1 \\ (a_2 \cdot b_1) b_1 + r_2 b_2 \quad &= a_2 \\ &\vdots \\ (a_n \cdot b_1) b_1 + \dots + (a_n \cdot b_{n-1}) b_{n-1} + r_n b_n \quad &= a_n \end{aligned}$$

and we can rewrite these equations using matrices: if

$$Q = [b_1 \ \cdots \ b_n] \in O(n), \quad R = \begin{bmatrix} r_1 & a_2 \cdot b_1 & a_3 \cdot b_1 & \cdots & a_n \cdot b_1 \\ 0 & r_2 & a_3 \cdot b_2 & \cdots & a_n \cdot b_2 \\ 0 & 0 & r_3 & \cdots & a_n \cdot b_3 \\ \vdots & & & \ddots & \vdots \\ 0 & & \cdots & & r_n \end{bmatrix},$$

then we see that the above equations correspond to

$$QR = A.$$

□

3.4 Hermitian spaces

[74] Do this!