

11.2 Orthonormal Bases and Gram–Schmidt

A basis lets us write every vector in coordinates, but finding those coordinates costs work: in Definition 1.5.2 the scalars $a_1, \dots, a_n$ come from solving a linear system. In an inner product space one kind of basis makes that cost disappear. If the basis vectors are mutually orthogonal unit vectors, then the i -th coordinate of v is just $\langle v, e_i \rangle$, one inner product, computed without reference to the others. This section defines such bases, shows that every finite-dimensional inner product space has one, and gives the algorithm that produces them.

11.2.1 Orthonormal lists

The standard basis of $\mathbb{R}^n$ has two features beyond being a basis: distinct vectors are orthogonal, and each has length 1. The two conditions can be packed into a single formula using the Kronecker delta δ_{ij} , which is 1 when $i = j$ and 0 otherwise.

An orthonormal list is a list of unit vectors that are pairwise perpendicular.

Definition 11.2.1. Orthonormal List

Let V be an inner product space. A list $(e_1, \dots, e_k)$ of vectors of V is **orthonormal** if

$$\langle e_i, e_j \rangle = \delta_{ij} \quad \text{for all } i, j \in \{1, \dots, k\},$$

that is, if $\langle e_i, e_j \rangle = 0$ whenever $i \neq j$ (**orthogonal**) and $\|e_i\| = 1$ for every i (**normalized**).

In words: the single equation $\langle e_i, e_j \rangle = \delta_{ij}$ carries two conditions at once, because the case $i \neq j$ says the vectors are orthogonal and the case $i = j$ says $\|e_i\|^2 = 1$. The list must be **ordered**, as every list in this book is, but nothing in the definition depends on the order. Note that no vector of an orthonormal list is 0, since $\|0\| = 0 \neq 1$; so by Theorem 11.1.4, **every orthonormal list is linearly independent**. That single sentence is why orthonormality is so useful: independence, which normally takes an argument, comes for free.

Definition 11.2.2. Orthonormal Basis

An **orthonormal basis** of an inner product space V is an orthonormal list that is a basis of V .

Examples.

- **The standard basis.** In F^n with the standard inner product, $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij}$ by inspection. This is the model, and the notation $\mathbf{e}_i$ is deliberately reused for orthonormal vectors in a general space.
- **A rotated basis of $\mathbb{R}^2$.** The list $(\frac{1}{\sqrt{2}}(1, 1), \frac{1}{\sqrt{2}}(1, -1))$ is orthonormal: the inner product of the two vectors is $\frac{1}{2}(1 - 1) = 0$, and each has norm $\frac{1}{\sqrt{2}}\sqrt{2} = 1$. Geometrically it is the standard basis turned through $\pi/4$.
- **A basis of $\mathbb{R}^3$.** The columns of $\frac{1}{3} \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{pmatrix}$ form an orthonormal list: each has norm

$\frac{1}{3}\sqrt{1+4+4} = 1$, and the three pairwise inner products are $\frac{1}{9}(2+2-4)$, $\frac{1}{9}(2-4+2)$ and $\frac{1}{9}(4-2-2)$, all 0. Being independent and of length $3 = \dim \mathbb{R}^3$, it is a basis (Theorem 1.6.4 (a)).

- **Matrix units.** In $M_{m \times n}(F)$ with the Frobenius inner product, the matrix units $\mathbf{E}_{ij}$ satisfy $\langle \mathbf{E}_{ij}, \mathbf{E}_{pq} \rangle = 1$ if $(i, j) = (p, q)$ and 0 otherwise, since only one entry of each is nonzero. So the standard basis of $M_{m \times n}(F)$ is orthonormal.
- **The Lagrange basis.** On $F[x]_{\leq n}$ with the node inner product $\langle p, q \rangle = \sum_{i=0}^n p(c_i) \overline{q(c_i)}$ of the previous section, the Lagrange polynomials $\ell_0, \dots, \ell_n$ of Definition 6.6.1 vanish at every node but their own, where they take the value 1; that is, $\ell_i(c_j) = \delta_{ij}$. Hence $\langle \ell_i, \ell_j \rangle = \sum_k \ell_i(c_k) \overline{\ell_j(c_k)} = \delta_{ij}$: the Lagrange basis is an orthonormal basis for this inner product. An interpolation formula and an orthonormal expansion turn out to be the same statement.
- **The degenerate case.** The empty list is orthonormal, vacuously, and it is an orthonormal basis of $V = \{0\}$. This is not a joke: it is the base case of every induction below.

Nonexample by minimal change. In $\mathbb{R}^2$, the list $((1, 1), (1, -1))$ is orthogonal, and it is a basis. It is **not** orthonormal, because $\|(1, 1)\| = \sqrt{2} \neq 1$. The failure is only in the normalization, and it is repaired by dividing each vector by its norm, which changes neither the spans nor the orthogonality (Theorem 11.1.2 (b)). Scaling is always available; orthogonality is the real content.

11.2.2 Coordinates in an orthonormal basis

Here is the payoff promised in the opening.

Theorem 11.2.1. Coordinates, Inner Products, and Norms in an Orthonormal Basis

Let $(e_1, \dots, e_n)$ be an orthonormal basis of an inner product space V , and let $u, v \in V$. Then:

- (a) $v = \sum_{i=1}^n \langle v, e_i \rangle e_i$;
- (b) $\langle u, v \rangle = \sum_{i=1}^n \langle u, e_i \rangle \overline{\langle v, e_i \rangle}$;
- (c) $\|v\|^2 = \sum_{i=1}^n |\langle v, e_i \rangle|^2$ (**Parseval's identity** for a finite orthonormal basis).

Idea. For (a) we already know that $v = \sum a_i e_i$ for unique scalars a_i ; the only question is what they are. Pair both sides with e_j : orthonormality kills every term except the j -th, and that one survives with coefficient 1. Parts (b) and (c) then follow by expanding, since the coordinates are now known.

Proof.

- (a) Since $(e_1, \dots, e_n)$ is a basis, Theorem 1.5.1 gives scalars $a_1, \dots, a_n \in F$ with $v = \sum_i a_i e_i$. Fix j . Pairing with e_j and using (IP1),

$$\langle v, e_j \rangle = \sum_{i=1}^n a_i \langle e_i, e_j \rangle = \sum_{i=1}^n a_i \delta_{ij} = a_j,$$

by Definition 11.2.1. So $a_j = \langle v, e_j \rangle$ for every j , which is (a).

- (b) Write $u = \sum_i \langle u, e_i \rangle e_i$ and $v = \sum_j \langle v, e_j \rangle e_j$ by (a). Expanding the first slot with (IP1) and the second with Theorem 11.1.1 (b),

$$\langle u, v \rangle = \sum_i \sum_j \langle u, e_i \rangle \overline{\langle v, e_j \rangle} \langle e_i, e_j \rangle.$$

Only the terms with $i = j$ survive, and there $\langle e_i, e_i \rangle = 1$. This is (b).

- (c) Take $u = v$ in (b): the i -th term becomes $\langle v, e_i \rangle \overline{\langle v, e_i \rangle} = |\langle v, e_i \rangle|^2$ by Theorem 0.6.2 (c), and the left side is $\|v\|^2$. ■

Read (b) again: in the coordinates supplied by an orthonormal basis, **every** inner product becomes the standard inner product of F^n . So an orthonormal basis does not merely simplify coordinates; it turns an abstract inner product space into F^n with the dot product. We record that as a corollary below, and use it repeatedly: **to make an abstract statement concrete, fix an orthonormal basis.**

Example. Reading Off Coordinates

Let $\mathbf{e}_1 = \frac{1}{3}(1, 2, 2)$, $\mathbf{e}_2 = \frac{1}{3}(2, 1, -2)$, $\mathbf{e}_3 = \frac{1}{3}(2, -2, 1)$, the orthonormal basis of $\mathbb{R}^3$ found above. Write $\mathbf{v} = (3, 0, 3)$ in this basis, and check Parseval's identity.

Solution. By Theorem 11.2.1 (a), the coordinates are

$$\begin{aligned}\langle \mathbf{v}, \mathbf{e}_1 \rangle &= \frac{1}{3}(3 + 0 + 6) = 3, \\ \langle \mathbf{v}, \mathbf{e}_2 \rangle &= \frac{1}{3}(6 + 0 - 6) = 0, \\ \langle \mathbf{v}, \mathbf{e}_3 \rangle &= \frac{1}{3}(6 + 0 + 3) = 3,\end{aligned}$$

so $\mathbf{v} = 3\mathbf{e}_1 + 0\mathbf{e}_2 + 3\mathbf{e}_3$. No system was solved. Checking directly, $3\mathbf{e}_1 + 3\mathbf{e}_3 = (1, 2, 2) + (2, -2, 1) = (3, 0, 3)$. For Parseval, $\|\mathbf{v}\|^2 = 9 + 0 + 9 = 18$, and indeed $3^2 + 0^2 + 3^2 = 18$.

Quick check. With the same $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, find the coordinates of $\mathbf{v} = (1, 1, 1)$ and verify Parseval's identity.

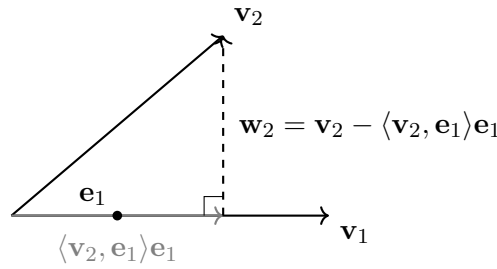
Solution. $\langle \mathbf{v}, \mathbf{e}_1 \rangle = \frac{1}{3}(1 + 2 + 2) = \frac{5}{3}$, $\langle \mathbf{v}, \mathbf{e}_2 \rangle = \frac{1}{3}(2 + 1 - 2) = \frac{1}{3}$, $\langle \mathbf{v}, \mathbf{e}_3 \rangle = \frac{1}{3}(2 - 2 + 1) = \frac{1}{3}$. Then $\frac{25}{9} + \frac{1}{9} + \frac{1}{9} = \frac{27}{9} = 3 = \|\mathbf{v}\|^2$, as Parseval requires.

11.2.3 The Gram–Schmidt process

Everything so far assumed that an orthonormal basis was handed to us. Does one always exist? The answer is yes in finite dimension, and the proof is an algorithm that we will use constantly.

The idea comes from a picture in $\mathbb{R}^3$. Given independent $\mathbf{v}_1, \mathbf{v}_2$, normalize $\mathbf{v}_1$ to get $\mathbf{e}_1$. The vector $\mathbf{v}_2$ is generally not perpendicular to $\mathbf{e}_1$; but its “shadow” on the line through $\mathbf{e}_1$ is

the multiple $\langle \mathbf{v}_2, \mathbf{e}_1 \rangle \mathbf{e}_1$, and subtracting the shadow leaves a vector $\mathbf{w}_2$ perpendicular to $\mathbf{e}_1$. Normalize it. Then repeat with $\mathbf{v}_3$, subtracting its shadows on both $\mathbf{e}_1$ and $\mathbf{e}_2$.



Project, subtract, normalize. The right angle at the foot is the only thing that has to be checked, and one line of algebra checks it.

Theorem 11.2.2. Gram–Schmidt Process

Let V be an inner product space and let $(v_1, \dots, v_m)$ be a **linearly independent** list in V . Define recursively

$$w_k = v_k - \sum_{i=1}^{k-1} \langle v_k, e_i \rangle e_i, \quad e_k = \frac{w_k}{\|w_k\|} \quad (k = 1, \dots, m).$$

Then every w_k is nonzero, so the recursion is defined, and:

- (a) $(e_1, \dots, e_m)$ is an orthonormal list;
- (b) $\text{span}(e_1, \dots, e_k) = \text{span}(v_1, \dots, v_k)$ for every $k = 1, \dots, m$;
- (c) if $(v_1, \dots, v_r)$ is already orthonormal for some $r \leq m$, then $e_i = v_i$ for $i \leq r$.

In particular, every finite-dimensional inner product space has an orthonormal basis.

Idea. Induction on k . Hypothesis: $(e_1, \dots, e_k)$ is orthonormal and spans $\text{span}(v_1, \dots, v_k)$. Two things need care. First, $w_k \neq 0$: if it were 0, then v_k would lie in $\text{span}(e_1, \dots, e_{k-1}) = \text{span}(v_1, \dots, v_{k-1})$, contradicting independence. That is the only place independence is used, and it is used exactly there. Second, $w_k \perp e_j$ for $j < k$: pair and watch the sum collapse.

Proof. We prove by induction on k that $w_k \neq 0$, that $(e_1, \dots, e_k)$ is orthonormal, and that $\text{span}(e_1, \dots, e_k) = \text{span}(v_1, \dots, v_k)$.

Base case $k = 1$. The empty sum is 0, so $w_1 = v_1$, which is nonzero because a list containing 0 is dependent. Then $e_1 = v_1/\|v_1\|$ has norm 1 by Theorem 11.1.2 (b), and $\text{span}(e_1) = \text{span}(v_1)$ since each vector is a nonzero multiple of the other.

Induction step. Let $2 \leq k \leq m$ and suppose the claim holds for $k - 1$. Put $U = \text{span}(e_1, \dots, e_{k-1}) = \text{span}(v_1, \dots, v_{k-1})$.

First, $w_k \neq 0$. If $w_k = 0$, then $v_k = \sum_{i < k} \langle v_k, e_i \rangle e_i \in U = \text{span}(v_1, \dots, v_{k-1})$, so v_k would be a linear combination of the earlier v_i , making $(v_1, \dots, v_m)$ dependent, contrary to hypothesis. Hence $\|w_k\| > 0$ and e_k is defined, with $\|e_k\| = 1$.

Next, $e_k \perp e_j$ for every $j < k$. By (IP1) and the induction hypothesis $\langle e_i, e_j \rangle = \delta_{ij}$,

$$\langle w_k, e_j \rangle = \langle v_k, e_j \rangle - \sum_{i=1}^{k-1} \langle v_k, e_i \rangle \delta_{ij} = \langle v_k, e_j \rangle - \langle v_k, e_j \rangle = 0,$$

and e_k is a scalar multiple of w_k , so $\langle e_k, e_j \rangle = 0$ as well. Together with the induction hypothesis, $(e_1, \dots, e_k)$ is orthonormal, which is (a).

For (b), e_k is a linear combination of v_k and $e_1, \dots, e_{k-1}$, hence lies in $\text{span}(v_1, \dots, v_k)$ by the induction hypothesis; so $\text{span}(e_1, \dots, e_k) \subseteq \text{span}(v_1, \dots, v_k)$. Conversely $v_k = \|w_k\|e_k + \sum_{i < k} \langle v_k, e_i \rangle e_i \in \text{span}(e_1, \dots, e_k)$, and the earlier v_i lie there by the induction hypothesis. The two spans are equal, completing the induction.

For (c), suppose $(v_1, \dots, v_r)$ is orthonormal and that $e_i = v_i$ for all $i < k$, where $k \leq r$. Then $\langle v_k, e_i \rangle = \langle v_k, v_i \rangle = 0$ for $i < k$, so $w_k = v_k$ and $\|w_k\| = 1$, giving $e_k = v_k$. Induction on k proves (c).

Finally, let V be finite-dimensional. If $V = \{0\}$, the empty list is an orthonormal basis. Otherwise take any basis $(v_1, \dots, v_n)$ of V and apply the process: by (a) the output is orthonormal, and by (b) it spans $\text{span}(v_1, \dots, v_n) = V$, so it is an orthonormal basis. ■

Two remarks on the formulas. The vector subtracted from v_k is exactly the expansion of v_k in the orthonormal basis $(e_1, \dots, e_{k-1})$ of U , which by [Theorem 11.2.1](#) (a) is the best available “copy of v_k inside U ”; the next section will call it the orthogonal projection of v_k onto U . And the subtraction is the same move that produced the vector z in the proof of [Theorem 11.1.5](#). **Project, then subtract** is one move used twice.

Corollary 11.2.3. Extending an Orthonormal List

Let V be a finite-dimensional inner product space and let $(e_1, \dots, e_r)$ be an orthonormal list in V . Then there are vectors $e_{r+1}, \dots, e_n \in V$ such that $(e_1, \dots, e_n)$ is an orthonormal basis of V .

Proof. The list $(e_1, \dots, e_r)$ is linearly independent by [Theorem 11.1.4](#), since its vectors are orthogonal and nonzero. By [Theorem 1.6.5](#) there are $v_{r+1}, \dots, v_n \in V$ making $(e_1, \dots, e_r, v_{r+1}, \dots, v_n)$ a basis of V . Apply [Theorem 11.2.2](#) to that basis. By part (c) the first r output vectors are $e_1, \dots, e_r$ unchanged, and by (a) and (b) the whole output is an orthonormal basis of V . ■

11.2.4 Two worked examples

Example. Gram–Schmidt in $\mathbb{R}^3$

Apply the Gram–Schmidt process of [Theorem 11.2.2](#) to $\mathbf{v}_1 = (1, 1, 1)$, $\mathbf{v}_2 = (0, 1, 1)$, $\mathbf{v}_3 = (0, 0, 1)$ in $\mathbb{R}^3$ with the dot product.

Solution. Step 1. $\|\mathbf{v}_1\|^2 = 3$, so $\mathbf{e}_1 = \frac{1}{\sqrt{3}}(1, 1, 1)$.

Step 2. $\langle \mathbf{v}_2, \mathbf{e}_1 \rangle = \frac{1}{\sqrt{3}}(0 + 1 + 1) = \frac{2}{\sqrt{3}}$, so

$$\mathbf{w}_2 = (0, 1, 1) - \frac{2}{3}(1, 1, 1) = \frac{1}{3}(-2, 1, 1).$$

Then $\|\mathbf{w}_2\|^2 = \frac{1}{9}(4 + 1 + 1) = \frac{2}{3}$, and $\mathbf{e}_2 = \frac{1}{\sqrt{6}}(-2, 1, 1)$.

Step 3. $\langle \mathbf{v}_3, \mathbf{e}_1 \rangle = \frac{1}{\sqrt{3}}$ and $\langle \mathbf{v}_3, \mathbf{e}_2 \rangle = \frac{1}{\sqrt{6}}$, so

$$\mathbf{w}_3 = (0, 0, 1) - \frac{1}{3}(1, 1, 1) - \frac{1}{6}(-2, 1, 1) = (0, -\frac{1}{2}, \frac{1}{2}).$$

Then $\|\mathbf{w}_3\|^2 = \frac{1}{2}$ and $\mathbf{e}_3 = \frac{1}{\sqrt{2}}(0, -1, 1)$.

Check. The three vectors have norm 1, and $\langle \mathbf{e}_1, \mathbf{e}_2 \rangle = \frac{1}{\sqrt{18}}(-2 + 1 + 1) = 0$, $\langle \mathbf{e}_1, \mathbf{e}_3 \rangle = \frac{1}{\sqrt{6}}(0 - 1 + 1) = 0$, $\langle \mathbf{e}_2, \mathbf{e}_3 \rangle = \frac{1}{\sqrt{12}}(0 - 1 + 1) = 0$. The span condition is visible too: $\mathbf{e}_1$ is a multiple of $\mathbf{v}_1$, and $\mathbf{e}_2$ is a combination of $\mathbf{v}_1$ and $\mathbf{v}_2$.

The next example runs the same three steps in a function space, where no picture is available and only the algebra is left.

Example. Gram-Schmidt on $1, x, x^2$

Apply the Gram-Schmidt process of [Theorem 11.2.2](#) to $(1, x, x^2)$ in $\mathbb{R}[x]_{\leq 2}$ with

$$\langle p, q \rangle = \int_{-1}^1 p(t)q(t) dt.$$

Solution. Step 1. $\|1\|^2 = \int_{-1}^1 1 dt = 2$, so $e_1 = 1/\sqrt{2}$.

Step 2. $\langle x, e_1 \rangle = \frac{1}{\sqrt{2}} \int_{-1}^1 t dt = 0$, since the integrand is odd. So $w_2 = x$, and $\|x\|^2 = \int_{-1}^1 t^2 dt = \frac{2}{3}$, giving $e_2 = \sqrt{3/2}x$.

Step 3. $\langle x^2, e_1 \rangle = \frac{1}{\sqrt{2}} \cdot \frac{2}{3}$ and $\langle x^2, e_2 \rangle = \sqrt{3/2} \int_{-1}^1 t^3 dt = 0$, again by oddness. Hence

$$w_3 = x^2 - \frac{1}{\sqrt{2}} \cdot \frac{2}{3} \cdot \frac{1}{\sqrt{2}} = x^2 - \frac{1}{3}.$$

Its squared norm is

$$\int_{-1}^1 \left(t^2 - \frac{1}{3}\right)^2 dt = \frac{2}{5} - \frac{4}{9} + \frac{2}{9} = \frac{8}{45},$$

so $e_3 = \frac{\sqrt{10}}{4}(3x^2 - 1)$.

The monic polynomials produced along the way are 1 , x , and $x^2 - \frac{1}{3}$. These are, up to scaling, the first three **Legendre polynomials**; the usual normalization takes $P_2 = \frac{1}{2}(3x^2 - 1) = \frac{3}{2}(x^2 - \frac{1}{3})$, fixed by the convention $P_k(1) = 1$ rather than by the norm. Section 10 of this chapter takes up such families in general.

Notice how much work the oddness of the integrand did: two of the three inner products were

0 before any computation. Symmetry of the weight is worth looking for.

Warning. **Gram–Schmidt depends on the order of the list.** Run the process on $((0, 1), (1, 1))$ in $\mathbb{R}^2$: it returns $((0, 1), (1, 0))$. Run it on $((1, 1), (0, 1))$, the same two vectors in the other order: it returns $(\frac{1}{\sqrt{2}}(1, 1), \frac{1}{\sqrt{2}}(-1, 1))$. Both outputs are orthonormal bases of $\mathbb{R}^2$, and they are different. An orthonormal basis is never unique (for $\dim V \geq 1$ one can always replace e_1 by $-e_1$); the process returns **an** orthonormal basis adapted to the given order, not **the** orthonormal basis.

Quick check. Run Gram–Schmidt on $((1, 0), (1, 1))$ in $\mathbb{R}^2$, and then on $((1, 1), (1, 0))$. Do you get the same list?

Solution. First order: $\mathbf{e}_1 = (1, 0)$; $\langle (1, 1), \mathbf{e}_1 \rangle = 1$, so $\mathbf{w}_2 = (1, 1) - (1, 0) = (0, 1)$ and $\mathbf{e}_2 = (0, 1)$. Output $((1, 0), (0, 1))$.

Second order: $\mathbf{e}_1 = \frac{1}{\sqrt{2}}(1, 1)$; $\langle (1, 0), \mathbf{e}_1 \rangle = \frac{1}{\sqrt{2}}$, so $\mathbf{w}_2 = (1, 0) - \frac{1}{2}(1, 1) = \frac{1}{2}(1, -1)$ and $\mathbf{e}_2 = \frac{1}{\sqrt{2}}(1, -1)$. Output $(\frac{1}{\sqrt{2}}(1, 1), \frac{1}{\sqrt{2}}(1, -1))$. Not the same list, and not even the same set of lines.

Warning. **The classical process is numerically fragile.** Computed with rounding, the vectors $\mathbf{e}_1, \mathbf{e}_2, \dots$ drift out of orthogonality, and the drift grows with k and with how nearly dependent the input list is. Nothing above is wrong — the fault is in the arithmetic, not the mathematics — but in practice one uses a rearranged version or reflections instead. Section 8 of this chapter gives the reflection-based method, and Chapter 24 explains why it is the stable one.

11.2.5 All inner product spaces of the same dimension look alike

Corollary 11.2.4. Classification of Finite-Dimensional Inner Product Spaces

Let V be an inner product space over F with $\dim V = n$. Then there is an isomorphism $T: V \rightarrow F^n$ with

$$\langle Tu, Tv \rangle = \langle u, v \rangle \quad \text{for all } u, v \in V,$$

where the right-hand inner product is that of V and the left-hand one is the standard inner product of F^n . An isomorphism with this property is an **isometric isomorphism**, so every n -dimensional inner product space over F is isometrically isomorphic to F^n .

Proof. By Theorem 11.2.2, V has an orthonormal basis $\mathcal{E} = (e_1, \dots, e_n)$. Let $T: V \rightarrow F^n$ be the coordinate map $Tv = [v]_{\mathcal{E}}$, an isomorphism by Corollary 2.5.5. By Theorem 11.2.1 (a), the i -th entry of Tv is $\langle v, e_i \rangle$. Hence the standard inner product of Tu and Tv in F^n is $\sum_i \langle u, e_i \rangle \overline{\langle v, e_i \rangle}$, which equals $\langle u, v \rangle$ by Theorem 11.2.1 (b).

So dimension is a complete invariant of finite-dimensional inner product spaces over a fixed field: there is nothing to classify. This should be compared with Theorem 2.5.4, which said the same for vector spaces; adding an inner product adds no new invariant. What becomes interesting is the interaction of the inner product with a given operator, and that is the subject of Chapter 12. Isometries in general are the subject of Section 7 of this chapter.

11.2.6 Gram matrices

Given a list of vectors, all of its inner products can be collected into one matrix, and that matrix turns out to answer questions about the list.

Definition 11.2.3. Gram Matrix

Let $(v_1, \dots, v_k)$ be a list in an inner product space V over F . Its **Gram matrix** is $\mathbf{G} \in M_k(F)$ with entries

$$(\mathbf{G})_{ij} = \langle v_j, v_i \rangle \quad (1 \leq i, j \leq k).$$

The order of the subscripts looks backwards and is chosen on purpose: it makes $\mathbf{G}$ act on coefficient vectors the right way round, as part (b) below shows. For a real list the distinction disappears, since then $\langle v_j, v_i \rangle = \langle v_i, v_j \rangle$.

For example, the Gram matrix of $((1, 1, 0), (1, 0, 1), (0, 1, 1))$ in $\mathbb{R}^3$ is $\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$, with the squared norms on the diagonal.

Over $\mathbb{C}$ the order of the subscripts becomes visible. For $\mathbf{v}_1 = (1, 0)$ and $\mathbf{v}_2 = (i, 1)$ in $\mathbb{C}^2$, the off-diagonal entries are

$$(\mathbf{G})_{12} = \langle \mathbf{v}_2, \mathbf{v}_1 \rangle = i, \quad (\mathbf{G})_{21} = \langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \bar{i} = -i,$$

so $\mathbf{G} = \begin{pmatrix} 1 & i \\ -i & 2 \end{pmatrix}$, which is Hermitian but not symmetric. Part (b) below says $\mathbf{x}^* \mathbf{G} \mathbf{x} = \|x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2\|^2$; with $\mathbf{x} = (1, 1)$ both sides are 3, since $\mathbf{v}_1 + \mathbf{v}_2 = (1 + i, 1)$. Writing $(\mathbf{G})_{ij} = \langle \mathbf{v}_i, \mathbf{v}_j \rangle$ instead would transpose $\mathbf{G}$ and leave that identity reading $\mathbf{x}^* \mathbf{G}^\top \mathbf{x}$.

Theorem 11.2.5. Properties of the Gram Matrix

Let $(v_1, \dots, v_k)$ be a list in an inner product space V over F , with Gram matrix $\mathbf{G}$. Then:

- (a) $\mathbf{G}$ is **Hermitian**, $\mathbf{G}^* = \mathbf{G}$, and its diagonal entries are the real numbers $\|v_i\|^2 \geq 0$;
- (b) $\mathbf{x}^* \mathbf{G} \mathbf{x} = \|\sum_j x_j v_j\|^2$ for every $\mathbf{x} \in F^k$;
- (c) $\mathbf{G}$ is invertible if and only if $(v_1, \dots, v_k)$ is linearly independent;
- (d) $\det \mathbf{G}$ is a real number with $\det \mathbf{G} \geq 0$, and $\det \mathbf{G} > 0$ exactly when the list is linearly independent.

Idea. Part (b) is the whole theorem: the quadratic expression $\mathbf{x}^* \mathbf{G} \mathbf{x}$ is a squared length, so it vanishes only for the coefficient vectors of genuine dependence relations. That identifies $\text{nul}(\mathbf{G})$ with the set of dependence relations and gives (c) at once. For (d), fix an orthonormal basis of the span and write $\mathbf{G} = \mathbf{A}^* \mathbf{A}$ for a coordinate matrix $\mathbf{A}$; a determinant of that shape is a squared modulus.

Proof.

(a) By (IP2), $(\mathbf{G}^*)_{ij} = \overline{(\mathbf{G})_{ji}} = \overline{\langle v_i, v_j \rangle} = \langle v_j, v_i \rangle = (\mathbf{G})_{ij}$. The diagonal entries are $\langle v_i, v_i \rangle = \|v_i\|^2$, real and ≥ 0 by Theorem 11.1.1 (a).

(b) Write $u = \sum_j x_j v_j$. Expanding by (IP1) and Theorem 11.1.1 (b),

$$\|u\|^2 = \sum_i \sum_j x_j \overline{x_i} \langle v_j, v_i \rangle = \sum_i \overline{x_i} \sum_j (\mathbf{G})_{ij} x_j,$$

and the right-hand side is exactly the entry of the 1×1 matrix $\mathbf{x}^* \mathbf{G} \mathbf{x}$.

(c) We show $\text{nul}(\mathbf{G}) = \{\mathbf{x} \in F^k : \sum_j x_j v_j = 0\}$. If $\sum_j x_j v_j = 0$, then for each i , $(\mathbf{G}\mathbf{x})_i = \sum_j \langle v_j, v_i \rangle x_j = \langle \sum_j x_j v_j, v_i \rangle = 0$ by (IP1), so $\mathbf{G}\mathbf{x} = \mathbf{0}$. Conversely, if $\mathbf{G}\mathbf{x} = \mathbf{0}$, then $\mathbf{x}^* \mathbf{G} \mathbf{x} = 0$, so $\|\sum_j x_j v_j\|^2 = 0$ by (b) and hence $\sum_j x_j v_j = 0$ by Theorem 11.1.2 (a). Now $\mathbf{G}$ is invertible if and only if $\mathbf{G}\mathbf{x} = \mathbf{0}$ forces $\mathbf{x} = \mathbf{0}$ (Theorem 3.4.4, (a) $\Leftrightarrow$ (b)), which by the displayed equality of sets says exactly that the only dependence relation among the v_j is the trivial one.

(d) Let $U = \text{span}(v_1, \dots, v_k)$, let $r = \dim U$, and let $(e_1, \dots, e_r)$ be an orthonormal basis of U , which exists by Theorem 11.2.2. Let $\mathbf{A} \in M_{r \times k}(F)$ have as its j -th column the coordinate vector of v_j , so that $v_j = \sum_p a_{pj} e_p$. By Theorem 11.2.1 (b),

$$(\mathbf{G})_{ij} = \langle v_j, v_i \rangle = \sum_{p=1}^r a_{pj} \overline{a_{pi}} = (\mathbf{A}^* \mathbf{A})_{ij},$$

so $\mathbf{G} = \mathbf{A}^* \mathbf{A}$.

If the list is dependent, then $\mathbf{G}$ is singular by (c) and $\det \mathbf{G} = 0$. If it is independent, then $r = k$ and $\mathbf{A}$ is square. By Theorem 7.4.5, $\det(\mathbf{A}^*) = \det(\overline{\mathbf{A}})$, and $\det(\overline{\mathbf{A}}) = \overline{\det \mathbf{A}}$, because by the Leibniz formula (Theorem 7.4.3) the determinant is a sum of products of entries and conjugation preserves sums and products (Theorem 0.6.2). Hence, by Theorem 7.5.4,

$$\det \mathbf{G} = \det(\mathbf{A}^*) \det(\mathbf{A}) = \overline{\det \mathbf{A}} \det \mathbf{A} = |\det \mathbf{A}|^2.$$

This is real and ≥ 0 , and it is > 0 because $\mathbf{A}$ is the matrix of an isomorphism onto its coordinates, hence invertible with $\det \mathbf{A} \neq 0$ (Theorem 7.5.2). ■

Over $\mathbb{R}$, with $V = \mathbb{R}^n$ and the dot product, part (d) is the statement Chapter 7 reached from Cauchy–Binet: Corollary 7.7.3 says that $\det(\mathbf{A}\mathbf{A}^\top) \geq 0$ for a real $\mathbf{A}$, with equality exactly when the rows of $\mathbf{A}$ are dependent. Taking the rows of $\mathbf{A}$ to be $\mathbf{v}_1, \dots, \mathbf{v}_k$ makes $\mathbf{A}\mathbf{A}^\top$ the Gram matrix, so that corollary is the special case of (d) in which the ambient space is $\mathbb{R}^n$. The pointer left there is now paid: a Gram matrix is exactly a matrix of pairwise inner products, and its determinant is nonnegative for that reason.

Example. Deciding Independence with a Gram Matrix

Decide whether $((1, 1, 0), (1, 0, 1), (0, 1, -1))$ is independent in $\mathbb{R}^3$, using its Gram matrix.

Solution. The pairwise dot products give

$$\mathbf{G} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix},$$

and expanding along the first row, $\det \mathbf{G} = 2(4 - 1) - 1(2 + 1) + 1(-1 - 2) = 6 - 3 - 3 = 0$. By [Theorem 11.2.5](#) (d) the list is dependent. Indeed $\mathbf{G}\mathbf{x} = \mathbf{0}$ for $\mathbf{x} = (1, -1, -1)$, and [Theorem 11.2.5](#) (c) says this is a dependence relation: $(1, 1, 0) - (1, 0, 1) - (0, 1, -1) = \mathbf{0}$.

Two features of this method are worth noting. It never needs coordinates in the ambient space, so it works verbatim for polynomials and functions, where writing the vectors as columns is not an option. And the answer comes with a witness: a null vector of $\mathbf{G}$ is a dependence relation, ready to read off.

11.2.7 Exercises

11.2.7.1 A. Check your understanding

Exercise. A1

- Define an orthonormal list, using the Kronecker delta, and say what an orthonormal basis is.
- Why is every orthonormal list linearly independent? Name the result used.
- Write down the coordinates of v in an orthonormal basis $(e_1, \dots, e_n)$, and state Parseval's identity.
- True or false: the Gram–Schmidt process ([Theorem 11.2.2](#)) applied to a list and to the same list in reverse order produces the same orthonormal list. Justify your answer.
- State the criterion for linear independence in terms of the Gram matrix.
- True or false: every 3-dimensional real inner product space is isometrically isomorphic to $\mathbb{R}^3$ with the dot product. Justify your answer.

Solution.

- A list $(e_1, \dots, e_k)$ with $\langle e_i, e_j \rangle = \delta_{ij}$ for all i, j ([Definition 11.2.1](#)); an orthonormal basis is such a list that is also a basis ([Definition 11.2.2](#)).
- Its vectors are pairwise orthogonal and nonzero (each has norm 1), so [Theorem 11.1.4](#) applies.
- $v = \sum_i \langle v, e_i \rangle e_i$, and $\|v\|^2 = \sum_i |\langle v, e_i \rangle|^2$ ([Theorem 11.2.1](#)).
- False. On $((0, 1), (1, 1))$ the process returns $((0, 1), (1, 0))$, while on $((1, 1), (0, 1))$ it returns $(\frac{1}{\sqrt{2}}(1, 1), \frac{1}{\sqrt{2}}(-1, 1))$.
- The list is independent if and only if its Gram matrix is invertible, if and only if

$\det \mathbf{G} > 0$ (Theorem 11.2.5 (c), (d)).

- (f) True. By Corollary 11.2.4, the coordinate map with respect to any orthonormal basis is an isometric isomorphism onto $\mathbb{R}^3$.

11.2.7.2 B. Practice

Exercise. B1: Gram–Schmidt with numbers

Apply the Gram–Schmidt process of Theorem 11.2.2 to each list, with the dot product.

- (a) $((1, 0, 1), (1, 1, 0), (0, 1, 1))$ in $\mathbb{R}^3$.
 (b) $((1, 1, 0, 0), (1, 0, 1, 0), (0, 0, 1, 1))$ in $\mathbb{R}^4$.

Solution.

- (a) $\|(1, 0, 1)\|^2 = 2$, so $\mathbf{e}_1 = \frac{1}{\sqrt{2}}(1, 0, 1)$. Next $\langle(1, 1, 0), \mathbf{e}_1\rangle = \frac{1}{\sqrt{2}}$, so

$$\mathbf{w}_2 = (1, 1, 0) - \frac{1}{2}(1, 0, 1) = \frac{1}{2}(1, 2, -1),$$

with $\|\mathbf{w}_2\|^2 = \frac{1}{4} \cdot 6 = \frac{3}{2}$, giving $\mathbf{e}_2 = \frac{1}{\sqrt{6}}(1, 2, -1)$. Finally $\langle(0, 1, 1), \mathbf{e}_1\rangle = \frac{1}{\sqrt{2}}$ and $\langle(0, 1, 1), \mathbf{e}_2\rangle = \frac{1}{\sqrt{6}}$, so

$$\mathbf{w}_3 = (0, 1, 1) - \frac{1}{2}(1, 0, 1) - \frac{1}{6}(1, 2, -1) = \frac{2}{3}(-1, 1, 1),$$

with $\|\mathbf{w}_3\|^2 = \frac{4}{9} \cdot 3 = \frac{4}{3}$, giving $\mathbf{e}_3 = \frac{1}{\sqrt{3}}(-1, 1, 1)$.

- (b) $\mathbf{e}_1 = \frac{1}{\sqrt{2}}(1, 1, 0, 0)$. Then $\langle(1, 0, 1, 0), \mathbf{e}_1\rangle = \frac{1}{\sqrt{2}}$, so

$$\mathbf{w}_2 = (1, 0, 1, 0) - \frac{1}{2}(1, 1, 0, 0) = \frac{1}{2}(1, -1, 2, 0),$$

with $\|\mathbf{w}_2\|^2 = \frac{1}{4} \cdot 6 = \frac{3}{2}$ and $\mathbf{e}_2 = \frac{1}{\sqrt{6}}(1, -1, 2, 0)$. Finally $\langle(0, 0, 1, 1), \mathbf{e}_1\rangle = 0$ and $\langle(0, 0, 1, 1), \mathbf{e}_2\rangle = \frac{2}{\sqrt{6}}$, so

$$\mathbf{w}_3 = (0, 0, 1, 1) - \frac{1}{3}(1, -1, 2, 0) = \frac{1}{3}(-1, 1, 1, 3),$$

with $\|\mathbf{w}_3\|^2 = \frac{1}{9} \cdot 12 = \frac{4}{3}$ and $\mathbf{e}_3 = \frac{1}{2\sqrt{3}}(-1, 1, 1, 3)$.

Exercise. B2: Gram–Schmidt on polynomials

In $\mathbb{R}[x]_{\leq 2}$ with $\langle p, q \rangle = \int_0^1 p(t)q(t) dt$, apply the Gram–Schmidt process of Theorem 11.2.2 to $(1, x, x^2)$. Compare the monic outputs with those of Example (Gram–Schmidt on $1, x, x^2$) and explain the difference.

Solution. $\|1\|^2 = 1$, so $e_1 = 1$. Then $\langle x, 1 \rangle = \int_0^1 t dt = \frac{1}{2}$, so $w_2 = x - \frac{1}{2}$ and

$$\|w_2\|^2 = \int_0^1 \left(t - \frac{1}{2}\right)^2 dt = \frac{1}{3} - \frac{1}{2} + \frac{1}{4} = \frac{1}{12},$$

giving $e_2 = \sqrt{3}(2x - 1)$. Next $\langle x^2, e_1 \rangle = \frac{1}{3}$ and

$$\langle x^2, e_2 \rangle = \sqrt{3} \int_0^1 (2t^3 - t^2) dt = \sqrt{3} \left(\frac{1}{2} - \frac{1}{3} \right) = \frac{\sqrt{3}}{6},$$

so $w_3 = x^2 - \frac{1}{3} - \frac{\sqrt{3}}{6} \cdot \sqrt{3}(2x - 1) = x^2 - x + \frac{1}{6}$. Its squared norm is

$$\int_0^1 \left(t^2 - t + \frac{1}{6} \right)^2 dt = \frac{1}{180},$$

so $e_3 = \sqrt{5}(6x^2 - 6x + 1)$.

The monic outputs are 1 , $x - \frac{1}{2}$, $x^2 - x + \frac{1}{6}$, whereas on $[-1, 1]$ they were 1 , x , $x^2 - \frac{1}{3}$. Orthogonality is a property of the inner product, not of the polynomials: changing the interval changes which polynomials are orthogonal. These are the shifted Legendre polynomials, and the substitution $t \mapsto 2t - 1$ carries one family to the other.

Exercise. B3: Coordinates without solving a system

Let $\mathbf{e}_1 = \frac{1}{3}(1, 2, 2)$, $\mathbf{e}_2 = \frac{1}{3}(2, 1, -2)$, $\mathbf{e}_3 = \frac{1}{3}(2, -2, 1)$ in $\mathbb{R}^3$.

- Write $\mathbf{v} = (1, -1, 4)$ in this basis.
- Compute $\|\mathbf{v}\|^2$ in two ways, and state which theorem says the answers agree.
- Compute $\langle \mathbf{v}, \mathbf{u} \rangle$ for $\mathbf{u} = (3, 0, 3)$ from the coordinates of $\mathbf{v}$ and $\mathbf{u}$, using [Example \(Reading Off Coordinates\)](#) for $\mathbf{u}$.

Solution.

- $\langle \mathbf{v}, \mathbf{e}_1 \rangle = \frac{1}{3}(1 - 2 + 8) = \frac{7}{3}$, $\langle \mathbf{v}, \mathbf{e}_2 \rangle = \frac{1}{3}(2 - 1 - 8) = -\frac{7}{3}$, $\langle \mathbf{v}, \mathbf{e}_3 \rangle = \frac{1}{3}(2 + 2 + 4) = \frac{8}{3}$. So $\mathbf{v} = \frac{7}{3}\mathbf{e}_1 - \frac{7}{3}\mathbf{e}_2 + \frac{8}{3}\mathbf{e}_3$.
- Directly, $\|\mathbf{v}\|^2 = 1 + 1 + 16 = 18$. From the coordinates, $\frac{49}{9} + \frac{49}{9} + \frac{64}{9} = \frac{162}{9} = 18$. They agree by [Theorem 11.2.1](#) (c).
- The coordinates of $\mathbf{u}$ are $(3, 0, 3)$ by [Example \(Reading Off Coordinates\)](#), so by [Theorem 11.2.1](#) (b),

$$\langle \mathbf{v}, \mathbf{u} \rangle = \frac{7}{3} \cdot 3 + \left(-\frac{7}{3}\right) \cdot 0 + \frac{8}{3} \cdot 3 = 7 + 8 = 15.$$

Directly, $\langle \mathbf{v}, \mathbf{u} \rangle = 3 + 0 + 12 = 15$.

Exercise. B4: Gram matrices

For each list, write down the Gram matrix and use it to decide whether the list is independent.

- $((1, 1, 0), (1, 0, 1), (0, 1, 1))$ in $\mathbb{R}^3$ with the dot product.
- $((1, i), (i, 1))$ in $\mathbb{C}^2$ with the standard inner product.
- $(1, x, x^2 - \frac{1}{3})$ in $\mathbb{R}[x]_{\leq 2}$ with $\langle p, q \rangle = \int_{-1}^1 pq$.

Solution.

- (a) $\mathbf{G} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$, and expanding along the first row, $\det \mathbf{G} = 2(4 - 1) - 1(2 - 1) + 1(1 - 2) = 6 - 1 - 1 = 4 > 0$. By [Theorem 11.2.5](#) (d) the list is independent.
- (b) $\langle (1, i), (1, i) \rangle = 1 + 1 = 2$ and $\langle (i, 1), (1, i) \rangle = i\bar{1} + 1\bar{i} = i - i = 0$, so $\mathbf{G} = 2\mathbf{I}_2$ by [Theorem 11.2.5](#) (a). Then $\det \mathbf{G} = 4 > 0$ and the list is independent; in fact it is orthogonal, so Theorem 11.1.4 already gives the answer.
- (c) These are the monic outputs of [Example \(Gram-Schmidt on \$1, x, x^2\$ \)](#), so they are pairwise orthogonal, and $\mathbf{G} = \text{diag}(2, \frac{2}{3}, \frac{8}{45})$. Its determinant is $2 \cdot \frac{2}{3} \cdot \frac{8}{45} = \frac{32}{135} > 0$, so the list is independent. A Gram matrix is diagonal exactly when the list is orthogonal.

11.2.7.3 C. Going deeper

Exercise. C1: The Gram matrix determines the list

Let $(v_1, \dots, v_k)$ be a list in an inner product space V and $(w_1, \dots, w_k)$ a list in an inner product space W , both over F , with the **same** Gram matrix: $\langle v_j, v_i \rangle = \langle w_j, w_i \rangle$ for all i, j .

- (a) Prove that $\sum_j x_j v_j = 0$ if and only if $\sum_j x_j w_j = 0$, for every $\mathbf{x} \in F^k$.
- (b) Deduce that there is an isomorphism $T: \text{span}(v_1, \dots, v_k) \rightarrow \text{span}(w_1, \dots, w_k)$, well defined by $Tv_j = w_j$ for every j .
- (c) Prove that T is an isometric isomorphism, that is, $\langle Tu, Tu' \rangle = \langle u, u' \rangle$ for all u, u' in the domain.

Solution.

- (a) Let $\mathbf{G}$ be the common Gram matrix. By [Theorem 11.2.5](#) (b), $\|\sum_j x_j v_j\|^2 = \mathbf{x}^* \mathbf{G} \mathbf{x} = \|\sum_j x_j w_j\|^2$. By Theorem 11.1.2 (a), one side is 0 if and only if the other is, and a norm is 0 only at the zero vector.
- (b) Every element of $U = \text{span}(v_1, \dots, v_k)$ is $\sum_j x_j v_j$ for some $\mathbf{x}$. Define $T(\sum_j x_j v_j) = \sum_j x_j w_j$. This is well defined: if $\sum_j x_j v_j = \sum_j y_j v_j$, then $\sum_j (x_j - y_j) v_j = 0$, so $\sum_j (x_j - y_j) w_j = 0$ by (a), that is $\sum_j x_j w_j = \sum_j y_j w_j$. It is linear because the formula is linear in $\mathbf{x}$, it is surjective onto $\text{span}(w_1, \dots, w_k)$ by construction, and it is injective: if $T(\sum x_j v_j) = 0$, then $\sum x_j w_j = 0$, so $\sum x_j v_j = 0$ by (a).
- (c) Let $u = \sum_j x_j v_j$ and $u' = \sum_i y_i v_i$. Expanding by (IP1) and Theorem 11.1.1 (b),

$$\langle u, u' \rangle = \sum_i \sum_j x_j \bar{y}_i \langle v_j, v_i \rangle,$$

and the same computation in W gives $\langle Tu, Tu' \rangle = \sum_i \sum_j x_j \bar{y}_i \langle w_j, w_i \rangle$. The two sums are equal term by term, since the Gram matrices agree.

Exercise. C2: Gram-Schmidt on a dependent list

Let $(v_1, \dots, v_m)$ be a list in an inner product space, **not** assumed independent, and run the recursion of [Theorem 11.2.2](#).

- (a) Let k be the smallest index with $v_k \in \text{span}(v_1, \dots, v_{k-1})$. Prove that the recursion is defined up to e_{k-1} and that $w_k = 0$, so the step $e_k = w_k / \|w_k\|$ fails.
- (b) Repair the process: show that if, whenever $w_k = 0$, we discard v_k and continue with v_{k+1} , the output is an orthonormal basis of $\text{span}(v_1, \dots, v_m)$.
- (c) Run the repaired process on $((1, 1, 0), (2, 2, 0), (1, 0, 1))$ in $\mathbb{R}^3$.

Solution.

- (a) The list $(v_1, \dots, v_{k-1})$ is independent, since a dependence among its members would, by Theorem 1.4.1, put some v_j with $j < k$ in the span of its predecessors, contradicting the minimality of k . So Theorem 11.2.2 applies to it and produces $e_1, \dots, e_{k-1}$ with $\text{span}(e_1, \dots, e_{k-1}) = \text{span}(v_1, \dots, v_{k-1}) =: U$. Since $v_k \in U$, Theorem 11.2.1 (a) applied inside U gives $v_k = \sum_{i < k} \langle v_k, e_i \rangle e_i$, which is precisely the vector subtracted, so $w_k = 0$. Dividing by $\|w_k\| = 0$ is impossible.
- (b) Let $e_1, \dots, e_s$ be the vectors that the repaired process outputs. The proof of Theorem 11.2.2 shows, unchanged, that after processing $v_1, \dots, v_k$ the accumulated list is orthonormal and spans $\text{span}(v_1, \dots, v_k)$: the only step that used independence was the claim $w_k \neq 0$, and in the repaired process a step with $w_k = 0$ discards a v_k that lies in the span already built, so the span does not change. Taking $k = m$, the output is an orthonormal list spanning $\text{span}(v_1, \dots, v_m)$, hence an orthonormal basis of it.
- (c) $\mathbf{e}_1 = \frac{1}{\sqrt{2}}(1, 1, 0)$. For $\mathbf{v}_2 = (2, 2, 0)$: $\langle \mathbf{v}_2, \mathbf{e}_1 \rangle = \frac{4}{\sqrt{2}} = 2\sqrt{2}$, so $\mathbf{w}_2 = (2, 2, 0) - 2\sqrt{2} \cdot \frac{1}{\sqrt{2}}(1, 1, 0) = \mathbf{0}$, and $\mathbf{v}_2$ is discarded. For $\mathbf{v}_3 = (1, 0, 1)$: $\langle \mathbf{v}_3, \mathbf{e}_1 \rangle = \frac{1}{\sqrt{2}}$, so $\mathbf{w}_3 = (1, 0, 1) - \frac{1}{2}(1, 1, 0) = \frac{1}{2}(1, -1, 2)$, with $\|\mathbf{w}_3\|^2 = \frac{6}{4} = \frac{3}{2}$ and $\mathbf{e}_2 = \frac{1}{\sqrt{6}}(1, -1, 2)$. The output $(\frac{1}{\sqrt{2}}(1, 1, 0), \frac{1}{\sqrt{6}}(1, -1, 2))$ is an orthonormal basis of the plane spanned by the three given vectors.