Math 4571 (Advanced Linear Algebra)

Lecture #18 of 38 $\sim$ February 23rd, 2026

Orthogonality

- Gram-Schmidt
- Orthogonal Complements
- Orthogonal Projection

This material represents §3.2.1-§3.2.2 from the course notes.

Recall

Recall the notions of orthogonality from last week:

Definition

Two vectors in an inner product space are orthogonal if their inner product is zero. An orthogonal set is one in which every pair of vectors is orthogonal. An orthonormal set is an orthogonal set whose vectors all have norm 1.

The main utility of an orthogonal set is that an orthogonal set of nonzero vectors is linearly independent.

- Thus, if V has dimension n and $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are nonzero and orthogonal, it is a basis for V.
- In that case we can compute the coefficients of any vector easily using the inner product: if $\mathbf{w} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$, then $c_k = \langle \mathbf{w}, \mathbf{v}_k \rangle / \langle \mathbf{v}_k, \mathbf{v}_k \rangle$ for each $1 \le k \le n$.

Gram-Schmidt, I

We now give an algorithm for converting any linearly independent set into an orthogonal set with the same span:

Theorem (Gram-Schmidt Procedure)

Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots\}$ be a linearly independent set in the inner product space V, and set $V_k = \mathrm{span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$. Then there exists an orthogonal set of vectors $\{\mathbf{w}_1, \mathbf{w}_2, \dots\}$ such that, for each $k \geq 1$, $\mathrm{span}(\mathbf{w}_1, \dots, \mathbf{w}_k) = \mathrm{span}(V_k)$ and $\mathbf{w}_k$ is orthogonal to every vector in V_{k-1}. Furthermore, this sequence is unique up to multiplying the elements by nonzero scalars.

The idea is to construct the vectors $\mathbf{w}_k$ recursively. Let's prove the theorem while working out how the construction must go.

Gram-Schmidt, II

Goal: Given independent $\mathbf{v}_1, \dots, \mathbf{v}_k$, construct orthogonal $\mathbf{w}_1, \dots, \mathbf{w}_k$ with $\mathrm{span}(\mathbf{w}_1, \dots, \mathbf{w}_k) = \mathrm{span}(\mathbf{v}_1, \dots, \mathbf{v}_k)$.

- First, we need $\mathrm{span}(\mathbf{w}_1) = \mathrm{span}(\mathbf{v}_1)$, so $\mathbf{w}_1$ is a nonzero multiple of $\mathbf{v}_1$. By rescaling, take $\mathbf{w}_1 = \mathbf{v}_1$.
- Now suppose we have constructed orthogonal $\mathbf{w}_1, \dots, \mathbf{w}_{k-1}$ with $\mathrm{span}(\mathbf{w}_1, \dots, \mathbf{w}_{k-1}) = \mathrm{span}(\mathbf{v}_1, \dots, \mathbf{v}_{k-1})$.

Gram-Schmidt, II

<u>Goal</u>: Given independent $\mathbf{v}_1, \ldots, \mathbf{v}_k$, construct orthogonal $\mathbf{w}_1, \ldots, \mathbf{w}_k$ with $\mathrm{span}(\mathbf{w}_1, \ldots, \mathbf{w}_k) = \mathrm{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$.

- First, we need $\mathrm{span}(\mathbf{w}_1) = \mathrm{span}(\mathbf{v}_1)$, so $\mathbf{w}_1$ is a nonzero multiple of $\mathbf{v}_1$. By rescaling, take $\mathbf{w}_1 = \mathbf{v}_1$.
- Now suppose we have constructed orthogonal $\mathbf{w}_1, \ldots, \mathbf{w}_{k-1}$ with $\mathrm{span}(\mathbf{w}_1, \ldots, \mathbf{w}_{k-1}) = \mathrm{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{k-1})$.
- Then $\mathbf{w}_k \in \mathrm{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}, \mathbf{v}_k) = \mathrm{span}(\mathbf{w}_1, \ldots, \mathbf{w}_{k-1}, \mathbf{v}_k)$ by span properties, so $\mathbf{w}_k = c\mathbf{v_k} - a_1\mathbf{w}_1 - \cdots - a_{k-1}\mathbf{w}_{k-1}$ for some scalars $c, a_1, \ldots, a_{k-1}$.
- Now, if $c = 0$ then $\mathbf{w}_k$ would be linearly dependent with the other $\mathbf{w}_i$, not allowed. So c is nonzero, hence by rescaling we can take $c = 1$.

Gram-Schmidt, III

<u>Goal</u>: Given independent $\mathbf{v}_1, \ldots, \mathbf{v}_k$, construct orthogonal $\mathbf{w}_1, \ldots, \mathbf{w}_k$ with $\mathrm{span}(\mathbf{w}_1, \ldots, \mathbf{w}_k) = \mathrm{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$.

- So, we take $\mathbf{w}_k = \mathbf{v_k} - a_1\mathbf{w}_1 - \cdots - a_{k-1}\mathbf{w}_{k-1}$ for some scalars $a_1, \ldots, a_{k-1}$.
- We need $\mathbf{w}_k$ to be orthogonal to all of the other $\mathbf{w}_i$.
- This is equivalent to $0 = \langle \mathbf{w}_k, \mathbf{w}_i \rangle$
  $= \langle \mathbf{v_k} - a_1\mathbf{w}_1 - \cdots - a_{k-1}\mathbf{w}_{k-1}, \mathbf{w}_i \rangle$
  $= \langle \mathbf{v_k}, \mathbf{w}_i \rangle - a_1 \langle \mathbf{w}_1, \mathbf{w}_i \rangle - \cdots - a_{k-1} \langle \mathbf{w}_{k-1}, \mathbf{w}_i \rangle$
  $= \langle \mathbf{v_k}, \mathbf{w}_i \rangle - a_i \langle \mathbf{w}_i, \mathbf{w}_i \rangle$
  since all of the $\mathbf{w}_*$ are orthogonal to $\mathbf{w}_i$ except $\mathbf{w}_i$.
- So, solving for a_i yields $a_i = \langle \mathbf{v}_k, \mathbf{w}_j \rangle / \langle \mathbf{w}_j, \mathbf{w}_j \rangle$.
- We conclude that (up to rescaling) the vector $\mathbf{w}_k$ is uniquely determined, and is $\mathbf{w}_k = \mathbf{v_k} - a_1\mathbf{w}_1 - \cdots - a_{k-1}\mathbf{w}_{k-1}$ for $a_i = \langle \mathbf{v}_k, \mathbf{w}_j \rangle / \langle \mathbf{w}_j, \mathbf{w}_j \rangle$.

Gram-Schmidt, IV

In the particular situation of a finite-dimensional inner product space, Gram-Schmidt allows us to convert any basis to an orthogonal (or orthonormal) basis:

Corollary

Every finite-dimensional inner product space has an orthonormal basis.

Proof:

Gram-Schmidt, IV

In the particular situation of a finite-dimensional inner product space, Gram-Schmidt allows us to convert any basis to an orthogonal (or orthonormal) basis:

Corollary

Every finite-dimensional inner product space has an orthonormal basis.

Proof:

- Choose any basis $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ for V and apply the Gram-Schmidt procedure: this yields an orthogonal basis $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ for V.
- Now simply normalize each vector in $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ by dividing by its norm: this preserves orthogonality, but rescales each vector to have norm 1, thus yielding an orthonormal basis.

Gram-Schmidt, V

Applying Gram-Schmidt in practice is fairly straightforward, but increasingly tedious as one goes.

- Our algorithm in the proof gives an orthogonal basis, but it is also possible to perform the normalization at each step to construct an orthonormal basis one vector at a time.
- When performing computations by hand, it is generally bad to normalize at each step, because the norm of a vector will often involve square roots, which will then be carried forward.
- When using a computer, however, normalizing at each step can avoid certain numerical instability issues.
- The particular description of the algorithm we have discussed turns out not to be especially numerically stable, but it is possible to modify the algorithm to be more stable.

Gram-Schmidt Examples, I

Example: For $V = \mathbb{R}^3$ with the standard inner product, apply Gram-Schmidt to $\mathbf{v}_1 = (2, 1, 2)$, $\mathbf{v}_2 = (5, 4, 2)$, $\mathbf{v}_3 = (1, 2, 7)$.

Gram-Schmidt Examples, I

Example: For $V = \mathbb{R}^3$ with the standard inner product, apply Gram-Schmidt to $\mathbf{v}_1 = (2, 1, 2)$, $\mathbf{v}_2 = (5, 4, 2)$, $\mathbf{v}_3 = (1, 2, 7)$.

- First, $\mathbf{w}_1 = \mathbf{v}_1 = \boxed{(2, 1, 2)}$.
- Next, $\mathbf{w}_2 = \mathbf{v}_2 - a_1\mathbf{w}_1$, where
  $a_1 = \dfrac{\mathbf{v}_2 \cdot \mathbf{w}_1}{\mathbf{w}_1 \cdot \mathbf{w}_1} = \dfrac{(5, 4, 2) \cdot (2, 1, 2)}{(2, 1, 2) \cdot (2, 1, 2)} = \dfrac{18}{9} = 2$. Thus,
  $\mathbf{w}_2 = \boxed{(1, 2, -2)}$.
- Last, $\mathbf{w}_3 = \mathbf{v}_3 - b_1\mathbf{w}_1 - b_2\mathbf{w}_2$ where
  $b_1 = \dfrac{\mathbf{v}_3 \cdot \mathbf{w}_1}{\mathbf{w}_1 \cdot \mathbf{w}_1} = \dfrac{(1, 2, 7) \cdot (2, 1, 2)}{(2, 1, 2) \cdot (2, 1, 2)} = \dfrac{18}{9} = 2$, and
  $b_2 = \dfrac{\mathbf{v}_3 \cdot \mathbf{w}_2}{\mathbf{w}_2 \cdot \mathbf{w}_2} = \dfrac{(1, 2, 7) \cdot (1, 2, -2)}{(1, 2, -2) \cdot (1, 2, -2)} = \dfrac{-9}{9} = -1$. Thus,
  $\mathbf{w}_3 = \boxed{(-2, 2, 1)}$.

Gram-Schmidt Examples, II

Example: For $V = \mathbb{R}[x]$ with $\langle f, g \rangle = \int_0^1 f(x)g(x)\,dx$, apply Gram-Schmidt to $p_1 = 1$, $p_2 = x$, $p_3 = x^2$, $p_4 = x^3$.

Gram-Schmidt Examples, II

Example: For $V = \mathbb{R}[x]$ with $\langle f, g\rangle = \int_0^1 f(x)g(x)\,dx$, apply Gram-Schmidt to $p_1 = 1$, $p_2 = x$, $p_3 = x^2$, $p_4 = x^3$.

- First, $\mathbf{w}_1 = p_1 = \boxed{1}$.
- Next, $\mathbf{w}_2 = p_2 - a_1\mathbf{w}_1$, where $a_1 = \dfrac{\langle p_2, \mathbf{w}_1\rangle}{\langle \mathbf{w}_1, \mathbf{w}_1\rangle} = \dfrac{\int_0^1 x\,dx}{\int_0^1 1\,dx} = \dfrac{1}{2}$.

  Thus, $\mathbf{w}_2 = \boxed{x - 1/2}$.
- Then, $\mathbf{w}_3 = p_3 - b_1\mathbf{w}_1 - b_2\mathbf{w}_2$ where

  $b_1 = \dfrac{\langle p_3, \mathbf{w}_1\rangle}{\langle \mathbf{w}_1, \mathbf{w}_1\rangle} = \dfrac{\int_0^1 x^2\,dx}{\int_0^1 1\,dx} = \dfrac{1}{3}$, and

  $b_2 = \dfrac{\langle p_3, \mathbf{w}_2\rangle}{\langle \mathbf{w}_2, \mathbf{w}_2\rangle} = \dfrac{\int_0^1 x^2(x-1/2)\,dx}{\int_0^1 (x-1/2)^2\,dx} = \dfrac{1/12}{1/12} = 1$.

  Thus, $\mathbf{w}_3 = \boxed{x^2 - x + 1/6}$.

Gram-Schmidt Examples, III

Example: For $V = \mathbb{R}[x]$ with $\langle f, g \rangle = \int_0^1 f(x)g(x)\,dx$, apply Gram-Schmidt to $p_1 = 1$, $p_2 = x$, $p_3 = x^2$, $p_4 = x^3$.

- We have $\mathbf{w}_1 = \boxed{1}$, $\mathbf{w}_2 = \boxed{x - 1/2}$, and $\mathbf{w}_3 = \boxed{x^2 - x + 1/6}$.

Gram-Schmidt Examples, III

Example: For $V = \mathbb{R}[x]$ with $\langle f, g \rangle = \int_0^1 f(x)g(x)\,dx$, apply Gram-Schmidt to $p_1 = 1$, $p_2 = x$, $p_3 = x^2$, $p_4 = x^3$.

- We have $\mathbf{w}_1 = \boxed{1}$, $\mathbf{w}_2 = \boxed{x - 1/2}$, and $\mathbf{w}_3 = \boxed{x^2 - x + 1/6}$.
- Finally, $\mathbf{w}_4 = p_4 - c_1\mathbf{w}_1 - c_2\mathbf{w}_2 - c_3\mathbf{w}_3$ where

$$b_1 = \frac{\langle p_4, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} = \frac{\int_0^1 x^3\,dx}{\int_0^1 1\,dx} = \frac{1}{4},$$

$$b_2 = \frac{\langle p_4, \mathbf{w}_2 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle} = \frac{\int_0^1 x^3(x - 1/2)\,dx}{\int_0^1 (x - 1/2)^2\,dx} = \frac{3/40}{1/12} = \frac{9}{10}, \text{ and}$$

$$b_3 = \frac{\langle p_4, \mathbf{w}_3 \rangle}{\langle \mathbf{w}_3, \mathbf{w}_3 \rangle} = \frac{\int_0^1 x^3(x^2 - x + 1/6)\,dx}{\int_0^1 (x^2 - x + 1/6)^2\,dx} = \frac{1/120}{1/180} = \frac{3}{2}.$$

Thus, $\mathbf{w}_4 = \boxed{x^3 - (3/2)x^2 + (3/5)x - 1/20}$.

Gram-Schmidt Examples, IV

Although Gram-Schmidt allows us to construct an orthogonal basis for an arbitrary finite-dimensional vector space, there exist infinite-dimensional vector spaces that have no orthogonal basis.

- The precise details are somewhat involved[1], but in fact, the space $\ell^2(\mathbb{R})$ of infinite real sequences $(a_1, a_2, \dots)$ such that $a_1^2 + a_2^2 + \cdots$ is finite, with inner product $\langle (a_1, a_2, \dots), (b_1, b_2, \dots) \rangle = a_1 b_1 + a_2 b_2 + \cdots$, has no orthogonal basis.
- Notice, for example, that the set $\{e_1, e_2, \dots\}$ where e_i has a 1 in the ith coordinate and 0s elsewhere is an orthonormal set but is not a basis. However, since the only other element of the space orthogonal to every vector in the set is the zero vector, it cannot be extended to any larger orthonormal set.

---

[1] As usual with that kind of statement in this class, that means there's a challenge problem where you can explore these ideas in more detail!

Orthogonality and Approximations, I

If V is an inner product space, W is a subspace, and $\mathbf{v}$ is some vector in V, we would like to study the problem of finding a "best approximation" of $\mathbf{v}$ in W.

- For two vectors $\mathbf{v}$ and $\mathbf{w}$, the distance between $\mathbf{v}$ and $\mathbf{w}$ is $||\mathbf{v} - \mathbf{w}||$, so what we are seeking is a vector $\mathbf{w}$ in W that minimizes the quantity $||\mathbf{v} - \mathbf{w}||$.

Orthogonality and Approximations, II

Let's get geometric: suppose we are given a point P in $\mathbb{R}^2$ and wish to find the minimal distance from P to a particular line in $\mathbb{R}^2$.

Orthogonality and Approximations, II

Let's get geometric: suppose we are given a point P in $\mathbb{R}^2$ and wish to find the minimal distance from P to a particular line in $\mathbb{R}^2$.

- Geometrically, the minimal distance is achieved by the segment PQ, where Q is chosen so that PQ is perpendicular to the line.

What about the minimal distance between a point in $\mathbb{R}^3$ and a given plane?

Orthogonality and Approximations, II

Let's get geometric: suppose we are given a point P in $\mathbb{R}^2$ and wish to find the minimal distance from P to a particular line in $\mathbb{R}^2$.

- Geometrically, the minimal distance is achieved by the segment PQ, where Q is chosen so that PQ is perpendicular to the line.

What about the minimal distance between a point in $\mathbb{R}^3$ and a given plane?

- As one can see from a picture, that distance is also minimized by finding the segment perpendicular to the plane.

So it looks like the solution to the optimization problem will involve the notion of "perpendicularity" to the subspace W.

Orthogonal Complements, I

Definition

Let V be an inner product space. If S is a nonempty subset of V, we say a vector $\mathbf{v}$ in V is orthogonal to S if it is orthogonal to every vector in S. The set of all vectors orthogonal to S is denoted $S^{\perp}$ ("S-perpendicular", or often "S-perp" for short).

- We will typically be interested in the case where S is a subspace of V. It is easy to see via the subspace criterion that $S^{\perp}$ is always a subspace of V, even if S itself is not.

Orthogonal Complements, II

Examples:

1. In $\mathbb{R}^3$, if W is the xy-plane consisting of all vectors of the form $(x, y, 0)$, then $W^\perp$ is the z-axis, consisting of the vectors of the form $(0, 0, z)$.

Orthogonal Complements, II

Examples:

1. In $\mathbb{R}^3$, if W is the xy-plane consisting of all vectors of the form $(x, y, 0)$, then $W^\perp$ is the z-axis, consisting of the vectors of the form $(0, 0, z)$.
2. In $\mathbb{R}^3$, if W is the x-axis consisting of all vectors of the form $(x, 0, 0)$, then $W^\perp$ is the yz-plane, consisting of the vectors of the form $(0, y, z)$.

Orthogonal Complements, II

Examples:

1. In $\mathbb{R}^3$, if W is the xy-plane consisting of all vectors of the form $(x, y, 0)$, then $W^\perp$ is the z-axis, consisting of the vectors of the form $(0, 0, z)$.
2. In $\mathbb{R}^3$, if W is the x-axis consisting of all vectors of the form $(x, 0, 0)$, then $W^\perp$ is the yz-plane, consisting of the vectors of the form $(0, y, z)$.
3. In $\mathbb{R}^3$, if W is the span of $(1, 0, 1)$ and $(0, 1, 2)$, then $W^\perp$ is the set of vectors (a, b, c) orthogonal to both $(1, 0, 1)$ and $(0, 1, 2)$, which requires $a + c = 0$ and $b + 2c = 0$, hence $W^\perp$ is the span of $(-1, -2, 1)$.

Orthogonal Complements, II

Examples:

1. In $\mathbb{R}^3$, if W is the xy-plane consisting of all vectors of the form $(x, y, 0)$, then $W^\perp$ is the z-axis, consisting of the vectors of the form $(0, 0, z)$.
2. In $\mathbb{R}^3$, if W is the x-axis consisting of all vectors of the form $(x, 0, 0)$, then $W^\perp$ is the yz-plane, consisting of the vectors of the form $(0, y, z)$.
3. In $\mathbb{R}^3$, if W is the span of $(1, 0, 1)$ and $(0, 1, 2)$, then $W^\perp$ is the set of vectors (a, b, c) orthogonal to both $(1, 0, 1)$ and $(0, 1, 2)$, which requires $a + c = 0$ and $b + 2c = 0$, hence $W^\perp$ is the span of $(-1, -2, 1)$.
4. In any inner product space V, $V^\perp = \{\mathbf{0}\}$ and $\{\mathbf{0}\}^\perp = V$.

We could do more examples like (3), but let's just cut to the chase.

Orthogonal Complements, III

When V is finite-dimensional, we can use Gram-Schmidt to compute an explicit basis of $W^{\perp}$:

Theorem (Basis for Orthogonal Complement)

Suppose W is a subspace of the finite-dimensional inner product space V, and that $S = \{\mathbf{e}_1, \dots, \mathbf{e}_k\}$ is an orthogonal basis for W. If $\{\mathbf{e}_1, \dots, \mathbf{e}_k, \mathbf{e}_{k+1}, \dots, \mathbf{e}_n\}$ is any extension of S to an orthogonal basis for V, the set $\{\mathbf{e}_{k+1}, \dots, \mathbf{e}_n\}$ is an orthogonal basis for $W^{\perp}$. In particular, $V = W \oplus W^{\perp}$ and $\dim(V) = \dim(W) + \dim(W^{\perp})$.

Remark: It is always possible to extend the orthogonal basis $S = \{\mathbf{e}_1, \dots, \mathbf{e}_k\}$ to an orthogonal basis for V: simply extend the linearly independent set S to a basis $\{\mathbf{e}_1, \dots, \mathbf{e}_k, \mathbf{x}_{k+1}, \dots, \mathbf{x}_n\}$ of V, and then apply Gram-Schmidt to obtain an orthogonal basis $\{\mathbf{e}_1, \dots, \mathbf{e}_k, \mathbf{e}_{k+1}, \dots, \mathbf{e}_n\}$ of V. We can also normalize to assume the bases are orthonormal.

Orthogonal Complements, IV

Claim: If $S = \{\mathbf{e}_1, \dots, \mathbf{e}_k\}$ is an orthonormal basis for W and $\{\mathbf{e}_1, \dots, \mathbf{e}_k, \mathbf{e}_{k+1}, \dots, \mathbf{e}_n\}$ is an orthonormal basis for V, then $T = \{\mathbf{e}_{k+1}, \dots, \mathbf{e}_n\}$ is an orthonormal basis for $W^\perp$.

- Proof:

Orthogonal Complements, IV

Claim: If $S = \{\mathbf{e}_1, \dots, \mathbf{e}_k\}$ is an orthonormal basis for W and $\{\mathbf{e}_1, \dots, \mathbf{e}_k, \mathbf{e}_{k+1}, \dots, \mathbf{e}_n\}$ is an orthonormal basis for V, then $T = \{\mathbf{e}_{k+1}, \dots, \mathbf{e}_n\}$ is an orthonormal basis for $W^\perp$.

- Proof: Note that T is orthonormal, hence independent.
- Since each vector in T is orthogonal to every vector in S, each of $\mathbf{e}_{k+1}, \dots, \mathbf{e}_n$ is in $W^\perp$, so we just need to show that T spans $W^\perp$.
- So let $\mathbf{v}$ be any vector in $W^\perp$. Since $\{\mathbf{e}_1, \dots, \mathbf{e}_k, \mathbf{e}_{k+1}, \dots, \mathbf{e}_n\}$ is an orthonormal basis of V, by the orthogonal decomposition theorem we know that $\mathbf{v} = \langle \mathbf{v}, \mathbf{e}_1 \rangle \mathbf{e}_1 + \cdots + \langle \mathbf{v}, \mathbf{e}_k \rangle \mathbf{e}_k + \langle \mathbf{v}, \mathbf{e}_{k+1} \rangle \mathbf{e}_{k+1} + \cdots + \langle \mathbf{v}, \mathbf{e}_n \rangle \mathbf{e}_n$.
- But since $\mathbf{v}$ is in $W^\perp$, $\langle \mathbf{v}, \mathbf{e}_1 \rangle = \cdots = \langle \mathbf{v}, \mathbf{e}_k \rangle = 0$: thus, $\mathbf{v} = \langle \mathbf{v}, \mathbf{e}_{k+1} \rangle \mathbf{e}_{k+1} + \cdots + \langle \mathbf{v}, \mathbf{e}_n \rangle \mathbf{e}_n$, so $\mathbf{v} \in \mathrm{span}(T)$.
- Finally, $V = W \oplus W^\perp$ since $S \cup T$ is a basis of V.

Orthogonal Complements, V

Example: If $W = \text{span}[(2, 1, 2), (1, 2, -2)]$ in $\mathbb{R}^3$ with the standard dot product, find a basis for $W^\perp$.

Orthogonal Complements, V

Example: If $W = \mathrm{span}[(2,1,2),(1,2,-2)]$ in $\mathbb{R}^3$ with the standard dot product, find a basis for $W^\perp$.

- Notice that the vectors $\mathbf{e}_1 = (2,1,2)$ and $\mathbf{e}_2 = (1,2,-2)$ are orthogonal, hence form an orthogonal basis for W.
- It is straightforward to verify that if $\mathbf{v}_3 = (1,2,7)$, then $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{v}_3\}$ is a linearly independent set and therefore a basis for $\mathbb{R}^3$.
- Applying Gram-Schmidt to the set $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{v}_3\}$ yields $\mathbf{w}_1 = \mathbf{e}_1$, $\mathbf{w}_2 = \mathbf{e}_2$, and $\mathbf{w}_3 = (-2,2,1)$, as we worked out earlier.
- By the theorem above, we conclude that $\{\mathbf{e}_3\} = \boxed{(-2,2,1)}$ is an orthogonal basis of $W^\perp$.

Of course, here it would probably have been as easy to solve the linear system $(2,1,2)\cdot(a,b,c) = 0 = (1,2,-2)\cdot(a,b,c)$....

Orthogonal Complements, V

We can give another method for finding a basis for $W^{\perp}$ using matrices:

Theorem (Orthogonal Complements and Matrices)

If A is an $m \times n$ real matrix, then the rowspace of A and the nullspace of A are orthogonal complements of one another in $\mathbb{R}^n$, with respect to the standard dot product.

Remark: There is also a complex version of this theorem: in general, the row space and the complex-conjugate of the nullspace are orthogonal complements, as are the nullspace and the complex-conjugate of the row space.

Orthogonal Complements, VI

Claim: The rowspace and nullspace for a real matrix are orthogonal complements.

- Proof:

Orthogonal Complements, VI

Claim: The rowspace and nullspace for a real matrix are orthogonal complements.

- Proof: Let A be an $m \times n$ real matrix, so that the rowspace and nullspace are both subspaces of $\mathbb{R}^n$.
- By definition, any vector in $\mathrm{rowspace}(A)$ is orthogonal to any vector in $\mathrm{nullspace}(A)$, so $\mathrm{rowspace}(A) \subseteq \mathrm{nullspace}(A)^\perp$ and $\mathrm{nullspace}(A) \subseteq \mathrm{rowspace}(A)^\perp$.
- Furthermore, since $\dim(\mathrm{rowspace}(A)) + \dim(\mathrm{nullspace}(A)) = n$ from our results on the respective dimensions of these spaces, we see that $\dim(\mathrm{rowspace}(A)) = \dim(\mathrm{nullspace}(A)^\perp)$ and $\dim(\mathrm{nullspace}(A)) = \dim(\mathrm{rowspace}(A)^\perp)$.
- Since all these spaces are finite-dimensional, we must therefore have equality everywhere, as claimed.

Orthogonal Complements, VII

Example: If $W = \text{span}[(1,1,-1,1),\ (1,2,0,-2)]$ in $\mathbb{R}^4$ with the standard dot product, find a basis for $W^\perp$.

- Row-reduce the matrix whose rows are the given basis for W:

$$\begin{bmatrix} 1 & 1 & -1 & 1 \\ 1 & 2 & 0 & -2 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & -2 & 4 \\ 0 & 1 & 1 & -3 \end{bmatrix}.$$

Orthogonal Complements, VII

Example: If $W = \mathrm{span}[(1,1,-1,1),\ (1,2,0,-2)]$ in $\mathbb{R}^4$ with the standard dot product, find a basis for $W^\perp$.

- Row-reduce the matrix whose rows are the given basis for W:

$$\begin{bmatrix} 1 & 1 & -1 & 1 \\ 1 & 2 & 0 & -2 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & -2 & 4 \\ 0 & 1 & 1 & -3 \end{bmatrix}.$$

- From the reduced row-echelon form, we see that $\boxed{\{(-4,3,0,1),\ (2,-1,1,0)\}}$ is a basis for the nullspace and hence of $W^\perp$.

Orthogonal Projection, I

As we might expect from geometric intuition, if W is a subspace of the (finite-dimensional) inner product space V, we can decompose any vector uniquely as the sum of a component in W with a component in $W^{\perp}$:

Theorem (Orthogonal Components)

Let V be an inner product space and W be a finite-dimensional subspace. Then every vector $\mathbf{v} \in V$ can be uniquely written in the form $\mathbf{v} = \mathbf{w} + \mathbf{w}^{\perp}$ for some $\mathbf{w} \in W$ and $\mathbf{w}^{\perp} \in W^{\perp}$, and we also have the Pythagorean relation $||\mathbf{v}||^2 = ||\mathbf{w}||^2 + \left|\left|\mathbf{w}^{\perp}\right|\right|^2$.

Draw a picture:

Orthogonal Projection, II

Theorem (Orthogonal Components)

Let W be a finite-dimensional subspace of V. Then every $\mathbf{v} \in V$ can be written as $\mathbf{v} = \mathbf{w} + \mathbf{w}^\perp$ for unique $\mathbf{w} \in W$ and $\mathbf{w}^\perp \in W^\perp$.

<u>Proof</u> (existence):

- Since $\dim W$ is finite, take orthonormal basis $\{\mathbf{e}_1, \dots, \mathbf{e}_k\}$.
- Now set $\mathbf{w} = \langle \mathbf{v}, \mathbf{e}_1 \rangle \mathbf{e}_1 + \langle \mathbf{v}, \mathbf{e}_2 \rangle \mathbf{e}_2 + \cdots + \langle \mathbf{v}, \mathbf{e}_k \rangle \mathbf{e}_k$, and then $\mathbf{w}^\perp = \mathbf{v} - \mathbf{w}$.
- Clearly $\mathbf{w} \in W$ and $\mathbf{v} = \mathbf{w} + \mathbf{w}^\perp$, so we must show $\mathbf{w}^\perp \in W^\perp$.
- To see this, first observe that $\langle \mathbf{w}, \mathbf{e}_i \rangle = \langle \mathbf{v}, \mathbf{e}_i \rangle$ since $\{\mathbf{e}_1, \dots, \mathbf{e}_k\}$ is an orthonormal basis.
- Then $\langle \mathbf{w}^\perp, \mathbf{e}_i \rangle = \langle \mathbf{v} - \mathbf{w}, \mathbf{e}_i \rangle = \langle \mathbf{v}, \mathbf{e}_i \rangle - \langle \mathbf{w}, \mathbf{e}_i \rangle = 0$.
- Thus, $\mathbf{w}^\perp$ is orthogonal to each vector in the orthonormal basis of W, so it is in $W^\perp$.

Orthogonal Projection, III

Theorem (Orthogonal Components)

Let W be a finite-dimensional subspace of V. Then every $\mathbf{v} \in V$ can be written as $\mathbf{v} = \mathbf{w} + \mathbf{w}^{\perp}$ for unique $\mathbf{w} \in W$ and $\mathbf{w}^{\perp} \in W^{\perp}$.

Proof (uniqueness):

- Now suppose we had two decompositions $\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_1^{\perp}$ and $\mathbf{v} = \mathbf{w}_2 + \mathbf{w}_2^{\perp}$.
- By subtracting and rearranging, we see that $\mathbf{w}_1 - \mathbf{w}_2 = \mathbf{w}_2^{\perp} - \mathbf{w}_1^{\perp}$. Let this common vector be $\mathbf{x}$.
- Then $\mathbf{x}$ is in both W and $W^{\perp}$, so $\mathbf{x}$ is orthogonal to itself, but the only such vector is the zero vector.
- Thus, $\mathbf{w}_1 = \mathbf{w}_2$ and $\mathbf{w}_1^{\perp} = \mathbf{w}_2^{\perp}$, so the decomposition is unique.

Orthogonal Projection, IV

Theorem (Orthogonal Components)

If $\mathbf{v} \in V$ *has orthogonal decomposition* $\mathbf{v} = \mathbf{w} + \mathbf{w}^{\perp}$ *for* $\mathbf{w} \in W$ *and* $\mathbf{w}^{\perp} \in W^{\perp}$, *then we have the Pythagorean relation* $||\mathbf{v}||^2 = ||\mathbf{w}||^2 + ||\mathbf{w}^{\perp}||^2$.

<u>Proof</u> (Pythagoreanality):

- Observe that $\langle \mathbf{w}, \mathbf{w}^{\perp} \rangle = 0$ by definition. Then

$$\begin{aligned} ||\mathbf{v}||^2 &= \langle \mathbf{v}, \mathbf{v} \rangle \\ &= \langle \mathbf{w} + \mathbf{w}^{\perp}, \mathbf{w} + \mathbf{w}^{\perp} \rangle \\ &= \langle \mathbf{w}, \mathbf{w} \rangle + \langle \mathbf{w}, \mathbf{w}^{\perp} \rangle + \langle \mathbf{w}^{\perp}, \mathbf{w} \rangle + \langle \mathbf{w}^{\perp}, \mathbf{w}^{\perp} \rangle \\ &= \langle \mathbf{w}, \mathbf{w} \rangle + \langle \mathbf{w}^{\perp}, \mathbf{w}^{\perp} \rangle = ||\mathbf{w}||^2 + ||\mathbf{w}^{\perp}||^2 \end{aligned}$$

  as claimed.

Orthogonal Projection, V

Definition

If V is an inner product space and W is a finite-dimensional subspace with orthonormal basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_k\}$, the *orthogonal projection of $\mathbf{v}$ into W* *is the vector*
$\mathrm{proj}_W(\mathbf{v}) = \langle \mathbf{v}, \mathbf{e}_1 \rangle \, \mathbf{e}_1 + \langle \mathbf{v}, \mathbf{e}_2 \rangle \, \mathbf{e}_2 + \cdots + \langle \mathbf{v}, \mathbf{e}_k \rangle \, \mathbf{e}_k.$

This projection is simply the vector $\mathbf{w}$ from the previous theorem: it is the unique vector with the property that $\mathbf{v} - \mathrm{proj}_W(\mathbf{v}) \in W^{\perp}$.

- If instead we only have an orthogonal basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ of W, the corresponding expression is instead
$$\mathrm{proj}_W(\mathbf{v}) = \frac{\langle \mathbf{v}, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle}\mathbf{u}_1 + \frac{\langle \mathbf{v}, \mathbf{u}_2 \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle}\mathbf{u}_2 + \cdots + \frac{\langle \mathbf{v}, \mathbf{u}_2 \rangle}{\langle \mathbf{u}_k, \mathbf{u}_k \rangle}\mathbf{u}_k.$$

Orthogonal Projection, VI

Example: For $W = \text{span}[(1,0,0), \frac{1}{5}(0,3,4)]$ in $\mathbb{R}^3$ under the standard dot product, compute the orthogonal projection of $\mathbf{v} = (1,2,1)$ into W, and verify $||\mathbf{v}||^2 = ||\mathbf{w}||^2 + ||\mathbf{w}^{\perp}||^2$.

Orthogonal Projection, VI

Example: For $W = \text{span}[(1,0,0), \frac{1}{5}(0,3,4)]$ in $\mathbb{R}^3$ under the standard dot product, compute the orthogonal projection of $\mathbf{v} = (1,2,1)$ into W, and verify $||\mathbf{v}||^2 = ||\mathbf{w}||^2 + ||\mathbf{w}^{\perp}||^2$.

- Notice that the vectors $\mathbf{e}_1 = (1,0,0)$ and $\mathbf{e}_2 = \frac{1}{5}(0,3,4)$ form an orthonormal basis for W.
- Thus, the orthogonal projection is $\mathbf{w} = \text{proj}_W(\mathbf{v}) = \langle \mathbf{v}, \mathbf{e}_1 \rangle \mathbf{e}_1 + \langle \mathbf{v}, \mathbf{e}_2 \rangle \mathbf{e}_2 =$ $1\,(1,0,0) + 2\,(0,3/5,4/5) = \boxed{(1,6/5,8/5)}$.
- We see that $\mathbf{w}^{\perp} = \mathbf{v} - \mathbf{w} = (0, 4/5, -3/5)$ is orthogonal to both $(1,0,0)$ and $(0,3/5,4/5)$, so it is indeed in $W^{\perp}$.
- Furthermore, $||\mathbf{v}||^2 = 6$, while $||\mathbf{w}||^2 = 5$ and $||\mathbf{w}^{\perp}||^2 = 1$, so indeed $||\mathbf{v}||^2 = ||\mathbf{w}||^2 + ||\mathbf{w}^{\perp}||^2$.

Orthogonal Projection, VII

The orthogonal projection gives the answer to the approximation problem we posed earlier:

Corollary (Best Approximations)

If W is a finite-dimensional subspace of the inner product space V, then for any vector $\mathbf{v}$ in V, the orthogonal projection of $\mathbf{v}$ into W is closer to $\mathbf{v}$ than any other vector in W. Explicitly, if $\mathbf{w}$ is the projection, then for any $\mathbf{w}' \in W$, we have $||\mathbf{v} - \mathbf{w}|| \leq ||\mathbf{v} - \mathbf{w}'||$ with equality if and only if $\mathbf{w} = \mathbf{w}'$.

Orthogonal Projection, VIII

Corollary (Best Approximations)

If $\mathbf{w}$ *is the orthogonal projection of* $\mathbf{v}$ *into* W*, then for any* $\mathbf{w}' \in W$*, we have* $||\mathbf{v} - \mathbf{w}|| \le ||\mathbf{v} - \mathbf{w}'||$ *with equality if and only if* $\mathbf{w} = \mathbf{w}'$*.*

Proof:

- Write $\mathbf{v} = \mathbf{w} + \mathbf{w}^{\perp}$ where $\mathbf{w} \in W$ and $\mathbf{w}^{\perp} \in W^{\perp}$.

Orthogonal Projection, VIII

Corollary (Best Approximations)

If $\mathbf{w}$ *is the orthogonal projection of* $\mathbf{v}$ *into* W*, then for any* $\mathbf{w}' \in W$*, we have* $||\mathbf{v}-\mathbf{w}|| \leq ||\mathbf{v}-\mathbf{w}'||$ *with equality if and only if* $\mathbf{w}=\mathbf{w}'$*.*

Proof:

- Write $\mathbf{v}=\mathbf{w}+\mathbf{w}^{\perp}$ where $\mathbf{w} \in W$ and $\mathbf{w}^{\perp} \in W^{\perp}$.
- Now, for any other vector $\mathbf{w}' \in W$, we can write $\mathbf{v}-\mathbf{w}'=(\mathbf{v}-\mathbf{w})+(\mathbf{w}-\mathbf{w}')$, and observe that $\mathbf{v}-\mathbf{w}=\mathbf{w}^{\perp}$ is in $W^{\perp}$, and $\mathbf{w}-\mathbf{w}'$ is in W (since both $\mathbf{w}$ and $\mathbf{w}'$ are, and W is a subspace).
- Thus, $\mathbf{v}-\mathbf{w}'=(\mathbf{v}-\mathbf{w})+(\mathbf{w}-\mathbf{w}')$ is a decomposition of $\mathbf{v}-\mathbf{w}'$ into orthogonal vectors. Taking norms, we see that $||\mathbf{v}-\mathbf{w}'||^2=||\mathbf{v}-\mathbf{w}||^2+||\mathbf{w}-\mathbf{w}'||^2$.
- Then, if $\mathbf{w}' \neq \mathbf{w}$, since the norm of $||\mathbf{w}-\mathbf{w}'||$ is positive, we conclude that $||\mathbf{v}-\mathbf{w}|| < ||\mathbf{v}-\mathbf{w}'||$.

Orthogonal Projection, IX

Example: Find the best approximation to $\mathbf{v} = (3, -3, 3)$ lying in the subspace $W = \mathrm{span}[\frac{1}{3}(1, 2, -2),\ \frac{1}{3}(-2, 2, 1)]$, where distance is measured via the standard dot product.

Orthogonal Projection, IX

<u>Example</u>: Find the best approximation to $\mathbf{v} = (3, -3, 3)$ lying in the subspace $W = \text{span}[\frac{1}{3}(1, 2, -2),\ \frac{1}{3}(-2, 2, 1)]$, where distance is measured via the standard dot product.

- Notice that the vectors $\mathbf{e}_1 = \dfrac{1}{3}(1, 2, -2)$ and $\mathbf{e}_2 = \dfrac{1}{3}(-2, 2, 1)$ form an orthonormal basis for W.
- Thus, the desired vector, the orthogonal projection, is $\text{proj}_W(\mathbf{v}) = \langle \mathbf{v}, \mathbf{e}_1 \rangle \mathbf{e}_1 + \langle \mathbf{v}, \mathbf{e}_2 \rangle \mathbf{e}_2 = -3\mathbf{e}_1 - 3\mathbf{e}_2 = \boxed{(1, -4, 1)}$.

Orthogonal Projection, X

Example: Find the linear polynomial $p(x)$ that minimizes the expression $\int_0^1 (p(x) - e^x)^2 \, dx$.

Orthogonal Projection, X

<u>Example</u>: Find the linear polynomial $p(x)$ that minimizes the expression $\int_0^1 (p(x) - e^x)^2 \, dx$.

- The polynomial minimizing this integral is the orthogonal projection of e^x into $P_1(\mathbb{R})$ under the inner product $\langle f, g \rangle = \int_0^1 f(x)g(x) \, dx$.
- First, by applying Gram-Schmidt to $\{1, x\}$, we can get an orthogonal basis of $P_1(\mathbb{R})$: the result (after rescaling to clear denominators) is $\{1, \, 2x - 1\}$.
- Now, with $p_1 = 1$ and $p_2 = 2x - 1$, we can compute $\langle e^x, p_1 \rangle = e - 1$, $\langle e^x, p_2 \rangle = 3 - e$, $\langle p_1, p_1 \rangle = 1$, and $\langle p_2, p_2 \rangle = 1/3$.
- Then $\mathrm{proj}_{P_2(\mathbb{R})}(e^x) = \dfrac{\langle e^x, p_1 \rangle}{\langle p_1, p_1 \rangle} p_1 + \dfrac{\langle e^x, p_2 \rangle}{\langle p_2, p_2 \rangle} p_2 =$ $\boxed{(10 - 4e) + (18 - 6e)x \approx 0.873 + 1.690x}$.

Summary

We discussed the Gram-Schmidt procedure for constructing orthogonal sets.

We introduced the notion of orthogonal complement and described various ways to compute orthogonal complements.

We discussed orthogonal projection and its primary application to computing best approximations.

Next lecture: More projection, inner products and linear transformations