Projections

Zdeněk Dvořák

March 10, 2015

Theorem 1 (Properties of orthonormal bases)**.** *Let* $\mathbf{V}$ *be an inner product space and let* $B = v_1, \ldots, v_n$ *be an orthonormal basis of* $\mathbf{V}$*.*

1. *The coordinates of a vector* v *with respect to* B *are* $(\langle v, v_1\rangle, \langle v, v_2\rangle, \ldots, \langle v, v_n\rangle)$*.*
2. *If the coordinates of* $u, v \in \mathbf{V}$ *with respect to* B *are* $(\alpha_1, \ldots, \alpha_n)$ *and* $(\beta_1, \ldots, \beta_n)$*, respectively, then* $\langle u, v\rangle = \alpha_1\overline{\beta_1} + \ldots + \alpha_n\overline{\beta_n}$*.*
3. *If the coordinates of* $v \in \mathbf{V}$ *with respect to* B *are* $(\beta_1, \ldots, \beta_n)$*, then* $\|v\| = \sqrt{|\beta_1|^2 + \ldots + |\beta_n|^2}$*.*

Example 1. *Consider the space* $\mathcal{P}_2$ *of real polynomials of degree at most two, with inner product defined by*

$$\langle p, q\rangle = \int_0^1 p(x)q(x)\,dx.$$

Find an orthonormal basis of $\mathcal{P}_2$*.*

We apply the Gram-Schmidt process to the standard basis $1, x, x^2$ *of the space* $\mathcal{P}_2$*.*

- $v_1' = 1$, $\|1\| = 1$, $u_1 = 1$.
- $v_2' = x - \langle x, 1\rangle 1 = x - 1/2$, $\|x - 1/2\| = \sqrt{1/12}$, $u_2 = \sqrt{3}(2x-1)$.
- $v_3' = x^2 - \langle x^2, 1\rangle 1 - \left\langle x^2, \sqrt{3}(2x-1)\right\rangle \sqrt{3}(2x-1) = x^2 - 1/3 - (2x-1)/2 = x^2 - x + 1/6$, $\|x^2 - x + 1/6\| = \sqrt{1/180}$, $u_3 = \sqrt{5}(6x^2 - 6x + 1)$

Hence, an orthonormal basis is $1, \sqrt{3}(2x-1), \sqrt{5}(6x^2-6x+1)$*.*

1 Orthogonal complement and projection

Definition 1. *Let $\mathbf{V}$ be an inner product space and let $S \subseteq \mathbf{V}$. The orthogonal complement of S is*

$$S^{\perp} = \{u : u \perp s \textit{ for all } s \in S\}.$$

Lemma 2. *Let $\mathbf{V}$ be an inner product space over the field $\mathbf{F}$ and let $S \subseteq \mathbf{V}$.*

- *$S^{\perp}$ is a subspace of $\mathbf{V}$.*
- *If $T \subseteq S$, then $S^{\perp} \subseteq T^{\perp}$.*
- *$S^{\perp} = span(S)^{\perp}$.*
- *If $x \in S \cap S^{\perp}$, then $x = o$.*

Proof.

- Suppose that $u, v \in S^{\perp}$ and $\alpha \in \mathbf{F}$. For every $s \in S$, we have

$$\langle u + v, s\rangle = \langle u, s\rangle + \langle v, s\rangle = 0$$
$$\langle \alpha v, s\rangle = \alpha \langle v, s\rangle = 0,$$

  and thus $u + v, \alpha v \in S^{\perp}$.

- If $u \in S^{\perp}$, then $u \perp t$ for every $t \in T \subseteq S$, and thus $u \in T^{\perp}$.

- Suppose that $x \in S^{\perp}$, and consider any $v \in \mathrm{span}(S)$, $v = \alpha_1 s_1 + \ldots + \alpha_n s_n$ for some $s_1, \ldots, s_n \in S$ and $\alpha_1, \ldots, \alpha_n \in \mathbf{F}$. We have

$$\langle v, x\rangle = \alpha_1 \langle s_1, x\rangle + \ldots + \alpha_n \langle s_n, x\rangle = 0,$$

  and thus $x \perp v$. It follows that $x \in \mathrm{span}(S)^{\perp}$, and thus $S^{\perp} \subseteq \mathrm{span}(S)^{\perp}$. By the previous claim, $\mathrm{span}(S)^{\perp} \subseteq S^{\perp}$, since $S \subseteq \mathrm{span}(S)$.

- If $x \in S \cap S^{\perp}$, then $x \perp x$, and thus $0 = \langle x, x\rangle$ and $x = o$.

□

Lemma 3. *Let $\mathbf{V}$ be an inner product space and let $\mathbf{U}$ be its subspace. If $v_1, \ldots, v_n$ is an orthonormal basis of $\mathbf{V}$ and $\mathbf{U} = span(v_1, \ldots, v_m)$, then $\mathbf{U}^{\perp} = span(v_{m+1}, \ldots, v_n)$.*

Proof. Since the basis is orthonormal, we have $v_{m+1}, \ldots, v_n \perp v_1, \ldots, v_m$, and thus $v_{m+1}, \ldots, v_n \in \{v_1, \ldots, v_m\}^{\perp} = \mathbf{U}^{\perp}$. Since $\mathbf{U}^{\perp}$ is a subspace,

$\text{span}(v_{m+1}, \ldots, v_n)$ is a subspace of $\mathbf{U}^\perp$. By Lemma 2, we have $\mathbf{U} \cap \mathbf{U}^\perp = \{o\}$, and thus

$$\begin{aligned} n &= \dim(\mathbf{U}) + \dim(\text{span}(v_{m+1}, \ldots, v_n)) \\ &\leq \dim(\mathbf{U}) + \dim(\mathbf{U}^\perp) \\ &= \dim(\mathbf{U} \cap \mathbf{U}^\perp) + \dim(\mathbf{U} + \mathbf{U}^\perp) \\ &\leq 0 + n. \end{aligned}$$

It follows that $\dim(\mathbf{U}^\perp) = \dim(\text{span}(v_{m+1}, \ldots, v_n))$, and $\mathbf{U}^\perp = \text{span}(v_{m+1}, \ldots, v_n)$. □

Therefore, we can determine the basis of $\mathbf{U}^\perp$ as follows.

Algorithm 1. *Let* $\mathbf{V}$ *be an inner product space of finite dimension.*

Input: *A subspace* $\mathbf{U}$ *of* $\mathbf{V}$.

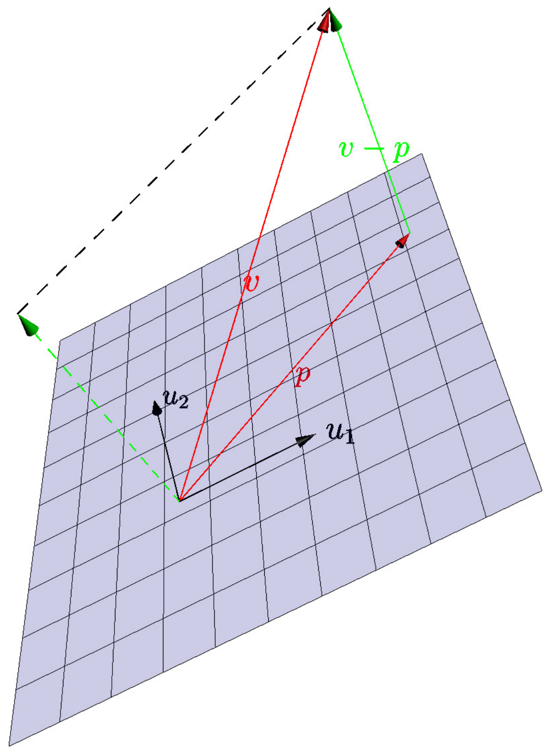
Output: *A basis* $w_1, \ldots, w_k$ *of* $\mathbf{U}^\perp$.

- *Let* $v_1, \ldots, v_n$ *be a basis of* $\mathbf{V}$, *and* $u_1, \ldots, u_m$ *a basis of* $\mathbf{U}$.
- *Apply the Gram-Schmidt process on* $u_1, \ldots, u_m, v_1, \ldots, v_n$, *giving an orthonormal basis* $z_1, \ldots, z_m, w_1, \ldots, w_k$ *of* $\mathbf{V}$.

Then $z_1, \ldots, z_m$ *is an orthonormal basis of* $\mathbf{U}$, *and* $w_1, \ldots, w_k$ *is an orthonormal basis of* $\mathbf{U}^\perp$.

Example 2. *Let* $\mathbf{U} = \text{span}((1,1,1),(1,2,3))$ *be a plane in* $\mathbf{R}^3$. *Find the coefficients of the equation* $ax + by + cz = 0$ *of this plane.*

We are looking for a non-zero vector (a,b,c) *such that* $(a,b,c)\cdot(x,y,z) = 0$ *for every* $(x,y,z) \in \mathbf{U}$, *i.e.,* $(a,b,c) \in \mathbf{U}^\perp$. *The Gram-Schmidt process on* $(1,1,1),(1,2,3),(1,0,0),(0,1,0),(0,0,1)$ *returns* $\frac{\sqrt{3}}{3}(1,1,1)$, $\frac{\sqrt{2}}{2}(-1,0,1)$, $\frac{\sqrt{6}}{6}(1,-2,1)$, *and thus* $\mathbf{U}^\perp = \text{span}\left(\frac{\sqrt{6}}{6}(1,-2,1)\right) = \text{span}((1,-2,1))$. *The equation of the plane* $\mathbf{U}$ *is* $x - 2y + z = 0$.



Theorem 4. *Let* $\mathbf{V}$ *be an inner product space and let* $\mathbf{U}$ *be its subspace of finite dimension.*

- *For every* $v \in \mathbf{V}$*, there exist unique* $p \in \mathbf{U}$ *and* $q \in \mathbf{U}^{\perp}$ *such that* $v = p + q$*.*
  - *If* $B = u_1, \ldots, u_k$ *is an orthonormal basis of* $\mathbf{U}$*, then the coordinates of* p *with respect to* B *are* $(\langle v, u_1 \rangle, \ldots, \langle v, u_k \rangle)$*, and thus* $p = \langle v, u_1 \rangle u_1 + \ldots + \langle v, u_k \rangle u_k$*.*
- $\mathbf{V} = \mathbf{U} + \mathbf{U}^{\perp}$*, and if* $\mathbf{V}$ *has a finite dimension, then* $\dim(\mathbf{V}) = \dim(\mathbf{U}) + \dim(\mathbf{U}^{\perp})$*.*
- $\left(\mathbf{U}^{\perp}\right)^{\perp} = \mathbf{U}$*.*

Proof.

- Consider any $x \in \mathbf{U}$, and let $(\alpha_1, \ldots, \alpha_k)$ be its coordinates with respect to B. Now, $v - x \in \mathbf{U}^{\perp} = \{u_1, \ldots, u_k\}^{\perp}$ if and only if $v - x \perp u_i$ for $i = 1, \ldots, k$, that is,

$$0 = \langle v - x, u_i \rangle = \langle v, u_i \rangle - \langle x, u_i \rangle = \langle v, u_i \rangle - \alpha_i.$$

Therefore, the vector p with coordinates $(\langle v, u_1 \rangle, \ldots, \langle v, u_k \rangle)$ is the only element of $\mathbf{U}$ such that $q = v - p \in \mathbf{U}^{\perp}$.

- By the first claim, every element of $\mathbf{V}$ belongs to $\mathbf{U} + \mathbf{U}^{\perp}$. Since $\mathbf{U} \cap \mathbf{U}^{\perp} = \{o\}$ has dimension 0, it follows that $\dim(\mathbf{V}) = \dim(\mathbf{U}) + \dim(\mathbf{U}^{\perp})$.

- Note that each $u \in \mathbf{U}$ satisfies $u \perp x$ for every $x \in \mathbf{U}^{\perp}$, and thus $u \in \left(\mathbf{U}^{\perp}\right)^{\perp}$.

  Conversely, consider any $v \in \left(\mathbf{U}^{\perp}\right)^{\perp}$. By the first claim, there exist $p \in \mathbf{U}$ and $q \in \mathbf{U}^{\perp}$ such that $v = p + q$. Note that $v \perp q$ and $p \perp q$, and thus $0 = \langle v, q\rangle = \langle p + q, q\rangle = \langle p, q\rangle + \langle q, q\rangle = \langle q, q\rangle$. Therefore, $q = o$ and $p = v$, and thus $v \in \mathbf{U}$.

$\square$

Warning: Theorem 4 is not necessarily true if $\mathbf{U}$ has infinite dimension.

Example 3. *Consider the space $\mathcal{P}$ of all real polynomials in variable x, and its subspace $\mathbf{U} = span(x - 1, x^2 - 1, x^3 - 1, \ldots)$. Note that a polynomial p belongs to U if and only if the sum of its coefficients is 0, and thus $\mathbf{U} \neq \mathcal{P}$. Let us define the inner product of two polynomials by $\langle \sum_{i=0}^{n} \alpha_i x^i, \sum_{i=0}^{n} \beta_i x^i\rangle = \sum_{i=0}^{n} \alpha_i \beta_i$.*

Then for a polynomial $p = \sum_{i=0}^{n} \alpha_i x^i$, we have $\left\langle p, x^k - 1\right\rangle = 0$ if and only if $\alpha_k = \alpha_0$. Consequently, $p \in \mathbf{U}^{\perp}$ if and only if $\alpha_0 = \alpha_1 = \alpha_2 = \ldots$. Since p has only finitely many non-zero coefficients, this is only possible if $p = 0$, and thus $\mathbf{U}^{\perp} = \{0\}$. Consequently, $\mathbf{U} + \mathbf{U}^{\perp} = \mathbf{U} \neq \mathcal{P}$. Also, $\left(\mathbf{U}^{\perp}\right)^{\perp} = \{0\}^{\perp} = \mathcal{P} \neq \mathbf{U}$.

Definition 2. *Let $\mathbf{V}$ be an inner product space and let $\mathbf{U}$ be its subspace of finite dimension. For $v \in \mathbf{V}$, the <u>orthogonal projection of v on $\mathbf{U}$</u> is the vector $p \in \mathbf{U}$ such that $v - p \in \mathbf{U}^{\perp}$.*

Lemma 5 (Basic properties of the projection)**.** *Let $\mathbf{V}$ be an inner product space and let $\mathbf{U}$ be its subspace of finite dimension. Let $P : \mathbf{V} \to \mathbf{U}$ be the function mapping each vector to its projection on $\mathbf{U}$. Then*

1. *P is a linear function,*

2. *if $u_1, \ldots, u_k$ is an orthonormal basis of $\mathbf{U}$, then $P(v) = \langle v, u_1\rangle u_1 + \ldots + \langle v, u_k\rangle u_k$ for every $v \in \mathbf{V}$,*

3. *$P(u) = u$ for every $u \in \mathbf{U}$, and*

4. *$P(P(v)) = P(v)$ for every $v \in \mathbf{V}$.*

Proof. 1. We have $v_1 - P(v_1), v_2 - P(v_2) \in \mathbf{U}^\perp$, and thus $(v_1+v_2)-(P(v_1)+P(v_2)) \in \mathbf{U}^\perp$ and $\alpha v_1 - \alpha P(v_1) \in \mathbf{U}^\perp$. Consequently, $P(v_1+v_2) = P(v_1)+P(v_2)$ and $P(\alpha v_1) = \alpha P(v_1)$.

2. This holds by Theorem 4.

3. This holds since $u - u = o \in \mathbf{U}^\perp$.

4. This holds by the previous item, since $P(v) \in \mathbf{U}$.

$\square$

Lemma 6 (Bessel's inequality, Parseval's theorem)**.** *Let* $\mathbf{V}$ *be an inner product space and let* $S = \{v_1, \ldots, v_m\}$ *be a finite orthonormal set in* $\mathbf{V}$*. For every* $v \in \mathbf{V}$*,*

$$\|v\| \geq \sqrt{|\langle v, v_1\rangle|^2 + \ldots + |\langle v, v_m\rangle|^2},$$

and the equality holds if and only if $v \in span(S)$*.*

Equivalently, for every $v \in \mathbf{V}$*, if* p *is the projection of* v *on* $span(S)$*, then* $\|v\| \geq \|p\|$*.*

Proof. By Theorem 4, the coordinates of p with respect to the orthonormal basis $v_1, \ldots, v_m$ of $\mathrm{span}(S)$ are $(\langle v, v_1\rangle, \ldots, \langle v, v_m\rangle)$, and by Theorem 1, we have

$$\|p\| = \sqrt{|\langle v, v_1\rangle|^2 + \ldots + |\langle v, v_m\rangle|^2}.$$

However, by the definition of the projection, we have $v - p \perp p$, and by the Pythagoras theorem,

$$\|v\|^2 = \|p\|^2 + \|v-p\|^2 \geq \|p\|^2,$$

with equality if and only if $v - p = o$, i.e., $v = p \in \mathrm{span}(S)$. $\square$

Lemma 7. *Let* $\mathbf{V}$ *be an inner product space and let* $\mathbf{U}$ *be its subspace of finite dimension. Let* $p \in \mathbf{U}$ *be the projection of* $v \in \mathbf{V}$*. Then* p *is the vector of* $\mathbf{U}$ *closest to* v*, that is,*

$$\|v - x\| > \|v - p\|$$

for every $x \in \mathbf{U} \setminus \{p\}$*.*

Proof. Note that $p - x \in \mathbf{U}$ and $v - p \in \mathbf{U}^\perp$, and thus $p - x \perp v - p$. By Pythagoras theorem, we have

$$\|v-p\|^2 + \|p-x\|^2 = \|v-x\|^2,$$

and since $p \neq x$, $\|p - x\| > 0$ and $\|v - x\| > \|v - p\|$. $\square$

Example 4. *Let* $\mathbf{U} = span((1,1,1),(1,2,3))$ *be a plane in* $\mathbf{R}^3$*. Determine the distance of the point* $v = (3,5,1)$ *from* $\mathbf{U}$*.*

In Example 2, we determined that $u_1, u_2 = \frac{\sqrt{3}}{3}(1,1,1), \frac{\sqrt{2}}{2}(-1,0,1)$ *is an orthonormal basis of* $\mathbf{U}$*, and thus the projection* p *of* v *on* $\mathbf{U}$ *is*

$$p = (v \cdot u_1)u_1 + (v \cdot u_2)u_2 = 3(1,1,1) - (-1,0,1) = (4,3,2).$$

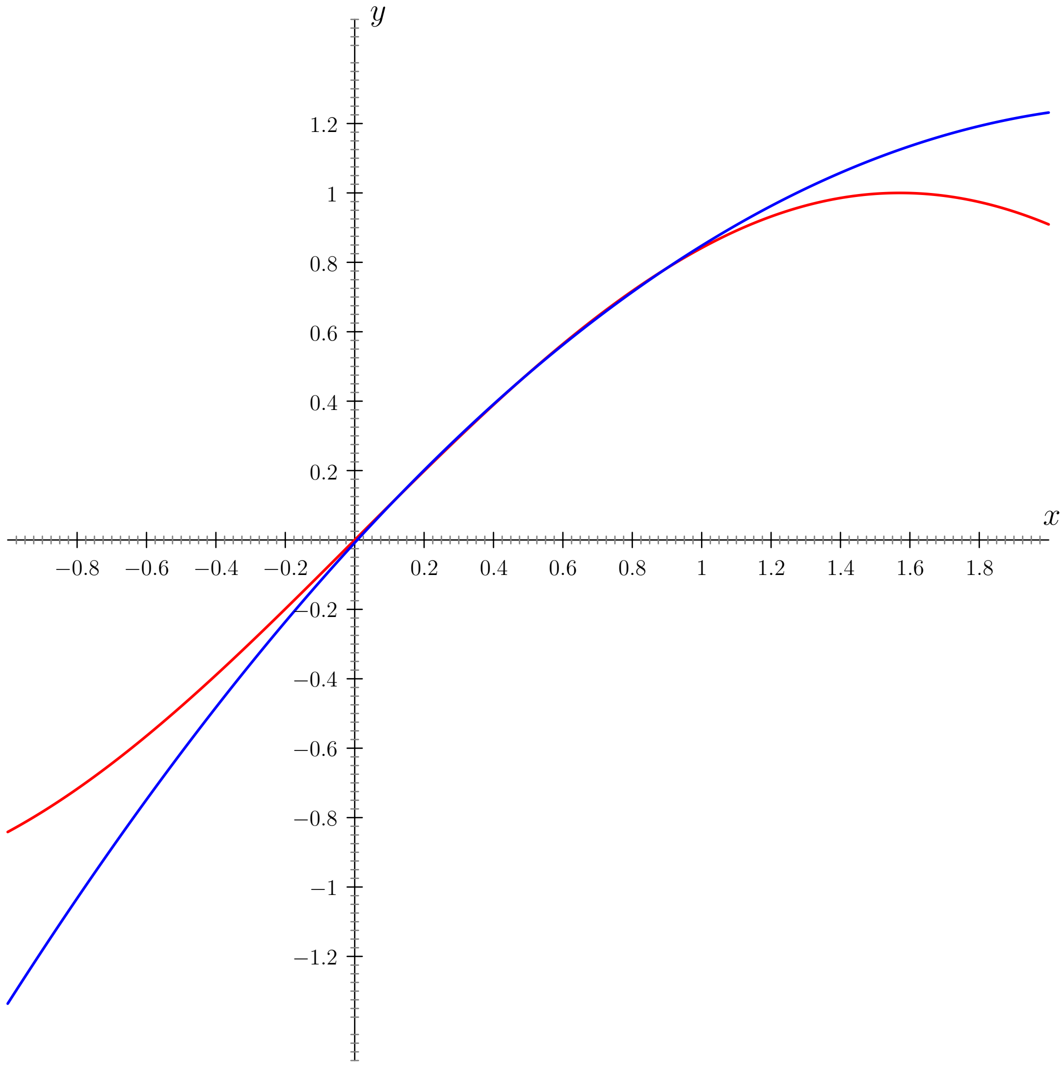
Hence, the distance is $|v - p| = |(-1,2,-1)| = \sqrt{6}$*.*

Example 5. *Find the polynomial* p *of degree at most two that approximates* $\sin x$ *on the interval* $[0,1]$ *the best, i.e., such that* $\int_0^1 (p(x) - \sin(x))^2\,dx$ *is minimum.*

Consider $\sin x$ *as an element of the vector space* $\mathbf{V}$ *of continuous functions from* $[0,1]$ *to* $\mathbf{R}$*, and let* $\mathbf{U} = \mathcal{P}_2$ *be its subspace. By Lemma 7,* p *is the projection of* $\sin x$ *on* $\mathbf{U}$*. Let* $B = u_1, u_2, u_3 = 1, \sqrt{3}(2x-1), \sqrt{5}(6x^2 - 6x + 1)$ *be the orthonormal basis of* $\mathcal{P}_2$ *that we determined in Example 1. By Theorem 4,* $p = \langle \sin x, u_1\rangle\, u_1 + \langle \sin x, u_2\rangle\, u_2 + \langle \sin x, u_3\rangle\, u_3$*.*

$$\begin{aligned}
\langle \sin x, u_1\rangle &= \int_0^1 \sin x\,dx \approx 0.4597 \\
\langle \sin x, u_2\rangle &= \sqrt{3}\int_0^1 \sin x(2x-1)\,dx \approx 0.2471 \\
\langle \sin x, u_3\rangle &= \sqrt{5}\int_0^1 \sin x(6x^2 - 6x + 1)\,dx \approx 0.0176
\end{aligned}$$

Hence, $p \approx -0.2361x^2 + 1.092x - 0.008$*.*


y
1.2
1
0.8
0.6
0.4
0.2
x
−0.8
−0.6
−0.4
−0.2
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
−0.2
−0.4
−0.6
−0.8
−1
−1.2