Day 35 Dot products and geometry in $\mathbb{R}^n$

Vectors in $\mathbb{R}^n$ can be written as linear combinations $\mathbf{x} = \sum x_i\, \mathbf{e}_i$ of the the standard basis elements $\{\mathbf{e}_i\}$. As long as one works only with the standard basis, one can omit the $\mathbf{e}_i$ and record each vector as the row $\mathbf{x} = (x^1, x^2, \ldots, x^n)$ of its coordinates.

Definition. For vectors $\mathbf{x} = (x^1, x^2, \ldots, x^n)$ and $\mathbf{y} = (y^1, y^2, \ldots, y^n)$ in $\mathbb{R}^n$, the **dot product** is

$$\mathbf{x} \cdot \mathbf{y} \;=\; x^1 y^2 + x^2 y^2 + \cdots + x^n y^n$$

and the **norm** of $\mathbf{x}$ is

$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}}.$$

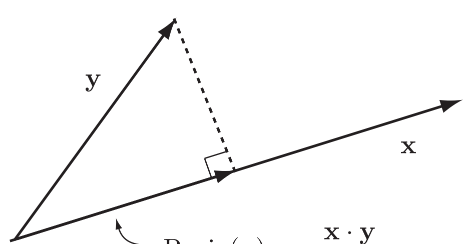
The norm and dot product have simple geometric interpretations:

- $\|\mathbf{x}\|$ is the length of the vector $\mathbf{x}$, regarded as an arrow from the origin to the point $\mathbf{x} \in \mathbb{R}^n$.
- The dot product is related to the angle θ between $\mathbf{x}$ and $\mathbf{y}$ by

  $$\mathbf{x} \cdot \mathbf{y} \;=\; \|\mathbf{x}\|\,\|\mathbf{y}\|\,\cos\theta,$$

  (the span of $\mathbf{x}$ and $\mathbf{y}$ is a plane in $\mathbb{R}^n$, and this formula holds in that plane). Thus one can determine the angle between two vectors by the formula $\cos\theta = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}$.
- Two vectors $\mathbf{x}$ and y are perpendicular (or "orthogonal") if and only if $\mathbf{x} \cdot \mathbf{y} = 0$.
- The **unit vector** in the direction of $\mathbf{x}$ is $\dfrac{\mathbf{x}}{\|\mathbf{x}\|}$.
- Given vectors $\mathbf{x}$ and $\mathbf{y}$ with $\mathbf{x} \neq 0$, we can write $\mathbf{y}$ as the sum of a vector parallel to $\mathbf{x}$ and one orthogonal to $\mathbf{x}$. The parallel one is called the *projection of* $\mathbf{y}$ *onto* $\mathbf{x}$ and is given by

  $$\mathrm{Proj}_{\mathbf{x}}(\mathbf{y}) \;=\; (\|\mathbf{y}\| \cdot \cos\theta)(\text{unit vector in the } \mathbf{x} \text{ direction}) \;=\; \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\|^2}\, \mathbf{x}.$$



Properties. The dot product if a function $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ by $(\mathbf{x}, \mathbf{y}) \mapsto \mathbf{x} \cdot \mathbf{y}$ with three properties. It is:

(a) Bilinear: $(a\mathbf{x} + b\mathbf{y}) \cdot \mathbf{z} = a\mathbf{x} \cdot \mathbf{z} + b\mathbf{y} \cdot \mathbf{z}$ and
$\mathbf{x} \cdot (a\mathbf{y} + b\mathbf{z}) = a\mathbf{x} \cdot \mathbf{y} + b\mathbf{x} \cdot \mathbf{z}$.

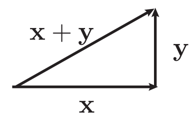
(b) Symmetric: $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$.

(c) Positive Definite: $\mathbf{x} \cdot \mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$ with equality if and only if $\mathbf{x} = 0$.

In fact, many facts about dot products follow using only these three properties. Here are two examples:

Lemma 35.1 (Pythagorean Theorem)**.** *If* $\mathbf{x}$ *and* $\mathbf{y}$ *are orthogonal, then*

$$\|\mathbf{x}+\mathbf{y}\|^2 \;=\; \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$$



Proof. Using the definition of the norm and bilinearity, we have

$$\|\mathbf{x}+\mathbf{y}\|^2 \;=\; (\mathbf{x}+\mathbf{y})\cdot(\mathbf{x}+\mathbf{y}) \;=\; \mathbf{x}\cdot\mathbf{x} \;+\; 2\cancelto{0}{(\mathbf{x}\cdot\mathbf{y})} \;+\; \mathbf{y}\cdot\mathbf{y} \;=\; \|\mathbf{x}\|^2 \;+\; \|\mathbf{y}\|^2.$$

□

Lemma 35.2 (Cauchy-Schwarz Inequality)**.** *For any vectors* $\mathbf{x}, \mathbf{y}$ *in* $\mathbb{R}^n$*,*

$$|\mathbf{x}\cdot\mathbf{y}| \;\le\; \|\mathbf{x}\|\;\|\mathbf{y}\|$$

with equality if and only if one of the vectors is a multiple of the other.

Proof. The inequality follows by taking the absolute value of both sides of the equation $\mathbf{x}\cdot\mathbf{y} \;=\; \|\mathbf{x}\|\,\|\mathbf{y}\|\,\cos\theta$ and noting that $|\cos\theta| \le 1$. Equality holds if and only if $\cos\theta = \pm 1$, which means that one vector is a (positive or negative) multiple of the other. □

Application to statistics

Frequently in statistics, a characteristic of some population is sampled and the results recorded as vector in $\mathbb{R}^n$ where n is the number sampled. For example, if 8 students are sampled and asked how many courses they are currently taking, the resulting data might be recorded as the vector

$$\mathbf{x} \;=\; (4,4,5,4,3,6,4,5) \in \mathbb{R}^8.$$

Definition. Suppose that two characteristics of a population are measured and the resulting data displayed as vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$.

1. The *mean* of the vector $\mathbf{x} = (x^1, x^2, \ldots, x^n)$ is $\bar{x} = \frac{1}{n}\sum_i x^i$.
2. The *deviation vectors* $\mathbf{x}_d$ and $\mathbf{y}_d$ are obtained by subtracting the mean from each entry.

   Example: The vector $\mathbf{x} = (1,2,3,4,5)$ has mean $\bar{x} = \frac{15}{5} = 3$ and deviation vector $\mathbf{x}_d = (-2,-1,0,1,2)$.

3. The **correlation coefficient** r between the characteristics is the cosine of the angle θ between the deviations vectors $\mathbf{x}_d$ and $\mathbf{y}_d$:
$$r \;=\; \cos\theta \;=\; \frac{\mathbf{x}_d\cdot\mathbf{y}_d}{\|\mathbf{x}_d\|\;\|\mathbf{y}_d\|}.$$

One says that the two characteristics are:

- *positively correlated* if $r > 0$,
- *negatively correlated* if $r < 0$, and
- *uncorrelated* if $r = 0$.

If there is a perfect positive correlation ($r = 1$) or a perfect negative correlation ($r = -1$) one can exactly predict the value of one characteristic given the other. When r is close to ± 1 one can almost predict, and one says the two characteristics are "highly correlated".

Homework 35

For the following problems, use a calculator to find the square roots and cosines. Minimize the number of decimal places.

1. Find the length of $\mathbf{u} = (7, 11)$, $\mathbf{v} = (2, 3, 4)$ and $\mathbf{w} = (2, 3, 4, 5)$.
2. Find the angle between (a) $\mathbf{u} = (1, 2, 3)$ and $\mathbf{v} = (2, 3, 4)$, (b) $\mathbf{u} = (1, -1, 2, -2)$ and $\mathbf{v} = (2, 3, 4, 5)$.
3. For each pair of vectors, determine whether the angle between them is acute ($< 90°$), obtuse ($> 90°$) or right.
   (a) $\mathbf{u} = (2, -2)$ and $\mathbf{v} = (5, 4)$.
   (b) $\mathbf{u} = (2, 3, 4)$ and $\mathbf{v} = (2, -8, 5)$.
   (c) $\mathbf{u} = (1, -1, 1, -1)$ and $\mathbf{v} = (3, 4, 5, 3)$.
4. For which choice of k are the vectors $\mathbf{u} = (2, 3, 4)$ and $\mathbf{v} = (1, k, 1)$ orthogonal?
5. Consider the vectors $\mathbf{u} = (1, 1, 1, \cdots, 1)$ and $\mathbf{v} = (1, 0, 0, 0, \cdots, 0)$ in $\mathbb{R}^n$. Determine the angle between them for $n = 2, 3, 4$ and find the limit of this angle as $n \to \infty$.
6. Find the orthogonal projection of $\mathbf{u} = (49, 49, 49)$
   (a) onto the vector $\mathbf{v} = (2, 3, 6)$
   (b) onto the subspace spanned by $\mathbf{v} = (2, 3, 6)$ and $\mathbf{w} = (3, -6, 2)$.
7. Find the orthogonal projection of $\mathbf{u} = (1, 0, 0, 0)$ onto the subspace of $\mathbb{R}^4$ spanned by $\mathbf{v}_1 = (1, 1, 1, 1)$, $\mathbf{v}_2 = (1, 1, -1, -1)$ and $\mathbf{v}_3 = (1, -1, -1, 1)$.
8. Let $\mathbf{v}$ be a vector in $\mathbb{R}^n$. Prove that the set

$$\mathbf{v}^{\perp} = \left\{ \mathbf{w} \in \mathbb{R}^n \;\middle|\; \mathbf{w} \cdot \mathbf{v} = 0 \right\}$$

   is a subspace of $\mathbb{R}^n$ (called the *orthogonal subspace to* $\mathbf{v}$).
9. Consider the vector $\mathbf{v} = (1, 2, 3, 4)$ in $\mathbb{R}^4$. Find a basis of the subspace of $\mathbb{R}^4$ consisting of all vectors perpendicular to $\mathbf{v}$.
10. Five students took aptitude exams in English, mathematics and science. Their scores are shown below.
   (a) Write down the deviation vectors E_d, M_d and S_d.
   (b) Find the correlation coefficients r_{EM}, r_{ES} and r_{MS} between the three pairs of variables. Which pairs of variables are positively/negatively correlated?

Student	E=English	M=Math	S=Science
S1	61	53	53
S2	63	73	78
S3	78	61	82
S4	65	84	96
S5	63	59	71
Mean	66	66	76

Day 36 Inner Product Spaces

On $\mathbb{R}^n$ the standard dot product gives

- lengths $\|\mathbf{v}\|$
- angles $\cos\theta = \dfrac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}$
- projections $\mathrm{Proj}_{\mathbf{x}}(\mathbf{y}) \;=\; \dfrac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|^2}\,\mathbf{x}$.

Inner products are "generalized dot products" that allow us to define lengths, angles and projections in a general vector space.

Definition 36.1. *An* inner product *on a real vector space V is a way of assigning a real number $\langle \mathbf{v}, \mathbf{w}\rangle$ to each pair of vectors $\mathbf{v}, \mathbf{w}$ of V with three properties:*

(a) (Positive definite) $\langle \mathbf{v}, \mathbf{v}\rangle \geq 0$, with equality if and only if $\mathbf{v} = \mathbf{0}$.

(b) (Symmetric) $\langle \mathbf{v}, \mathbf{w}\rangle \;=\; \langle \mathbf{w}, \mathbf{v}\rangle$ for all $\mathbf{v}, \mathbf{w} \in V$.

(c) (Bilinear) (a) $\langle a\mathbf{u} + b\mathbf{v}, \mathbf{w}\rangle = a\langle \mathbf{u}, \mathbf{w}\rangle + b\langle \mathbf{v}, \mathbf{w}\rangle$ and
(b) $\langle \mathbf{u}, c\mathbf{v} + d\mathbf{w}\rangle = c\langle \mathbf{u}, \mathbf{v}\rangle + d\langle \mathbf{u}, \mathbf{w}\rangle$ for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$.

Properties (2) and (3a) imply (3b), so to check the properties one needs only verify (1), (2) and (3a).

An inner product is an addition structure on a vector space; a vector space together with an inner product is called an *inner product space* $(V, \langle\ ,\ \rangle)$. Notice that the definition does not involve any basis.

Given an inner product, one defines the *norm* of a vector $\mathbf{v} \in V$ by

$$\|\mathbf{v}\| \;=\; \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle}. \tag{36.1}$$

Important fact to bear in mind: Each vector space has *many* inner products — we must choose one. Some vector spaces have "standard" inner products.

Examples. (1) On $V = \mathbb{R}^n$, the dot product defines an inner product $\langle \mathbf{v}, \mathbf{w}\rangle = \mathbf{v}\cdot\mathbf{w}$.

(2) One can define many other inner products on $\mathbb{R}^n$. For example, on $V = \mathbb{R}^2$, we can write $\mathbf{v} = (v_1, v_2)$ and $\mathbf{w} = (w_1, w_2)$ and define an inner product by

$$\langle \mathbf{v}, \mathbf{w}\rangle \;=\; 2v_1w_1 + 5v_2w_2.$$

(3) The space $V = M_{\mathbb{R}}(n, m)$ of real $n \times n$ matrices has a standard inner product defined by

$$\langle A, B\rangle \;=\; \mathrm{tr}\,(B^tA)$$

where B^t is the transpose of B. In terms of the entries A_{ij} and B_{ij}, this is

$$\langle A, B\rangle \;=\; \sum_j (B^tA)_{jj} \;=\; \sum_j\sum_i (B^t)_{ji}A_{ij} \;=\; \sum_j\sum_i B_{ij}A_{ij}.$$

This clearly symmetric and satisfies $\langle A, A\rangle = \sum_{ij} |A_{ij}|^2 \geq 0$ with equality if and only if $A = 0$. It is also clear that $\langle aA + b, C\rangle = \langle A, C\rangle + b\langle B, C\rangle$ for any matrices $A, B, C \in V$. Thus this formula defines an inner product on $M_{\mathbb{R}}(n,n)$.

(4) On the vector space $V = \mathcal{C}[a,b]$ of continuous functions on the interval $[a,b]$, the L^2 *inner product* is defined on $f, g \in V$ by

$$\langle f, g\rangle \;=\; \int_0^1 f(x)g(x)\;dx.$$

The following important facts follow algebraically from the definition of inner products, so are true in any inner product space.

Theorem 36.2. *For any two vectors* $\mathbf{v}, \mathbf{w} \in V$ *we have*

(a) $\|\mathbf{v}+\mathbf{w}\|^2 \;=\; \|\mathbf{v}\|^2 + 2\langle \mathbf{v}, \mathbf{w}\rangle + \|\mathbf{w}\|^2.$

(b) The **Cauchy-Schwarz Inequality** $\;|(\mathbf{v},\mathbf{w})| \;\le\; \|\mathbf{v}\|\cdot\|\mathbf{w}\|$ *with equality if and only if one of these vectors is a multiple of the other.*

(c) The **Triangle Inequality** $\;\|\mathbf{v}+\mathbf{w}\| \;\le\; \|\mathbf{v}\| + \|\mathbf{w}\|$, *also with equality if and only if one of these vectors is a multiple of the other.*

Proof. (a) Here's a proof in two-column format using the properties in the definition of inner product:

$$\begin{aligned}
\|\mathbf{v}+\mathbf{w}\|^2 &= \langle \mathbf{v}+\mathbf{w}, \mathbf{v}+\mathbf{w}\rangle && \text{def. of norm}\\
&= \langle \mathbf{v},\mathbf{v}\rangle + \langle \mathbf{v},\mathbf{w}\rangle + \langle \mathbf{w},\mathbf{v}\rangle + \langle \mathbf{w},\mathbf{w}\rangle && \text{bilinearity}\\
&= \langle \mathbf{v},\mathbf{v}\rangle + 2\langle \mathbf{v},\mathbf{w}\rangle + \langle \mathbf{w},\mathbf{w}\rangle && \text{symmetric}\\
&= \|\mathbf{v}\|^2 + 2\langle \mathbf{v},\mathbf{w}\rangle + \|\mathbf{w}\|^2 && \text{def. of norm}
\end{aligned}$$

(b) If $\|\mathbf{w}\| = 0$ then $\mathbf{w} = \mathbf{0}$ and hence the Cauchy-Schwarz inequality holds because both sides are 0. Therefore we can assume that $\|\mathbf{w}\| \neq \mathbf{0}$. Set $t = \dfrac{\langle \mathbf{v},\mathbf{w}\rangle}{\|\mathbf{w}\|^2}$. Then the vector $\mathbf{v} - t\mathbf{w}$ satisfies

$$\begin{aligned}
0 \;\le\; \|\mathbf{v}-t\mathbf{w}\|^2 &= \|\mathbf{v}\|^2 - 2t\,\langle \mathbf{v},\mathbf{w}\rangle + t^2\|\mathbf{w}\|^2\\
&= \|\mathbf{v}\|^2 - 2\frac{|\langle \mathbf{v},\mathbf{w}\rangle|^2}{\|\mathbf{w}\|^2} + \frac{|\langle \mathbf{v},\mathbf{w}\rangle|^2}{\|\mathbf{w}\|^2}\\
&= \|\mathbf{v}\|^2 - \frac{|\langle \mathbf{v},\mathbf{w}\rangle|^2}{\|\mathbf{w}\|^2}.
\end{aligned}$$

After rearranging, this gives $|\langle \mathbf{v},\mathbf{w}\rangle|^2 \;\le\; \|\mathbf{v}\|^2\cdot\|\mathbf{w}\|^2$. The Cauchy-Schwarz inequality follows by taking square roots. Finally, note that equality holds if and only if either $\mathbf{w} = 0$ or $\mathbf{v} = t\mathbf{w}$, which exactly means that one of the vectors is a multiple of the other.

(c) The Triangle Inequality follows from (a) and (b) (HW Problem 5). $\square$

The great thing about inner products is that one can *think geometrically*, even about an abstract vector space. Suppose that V is an inner product space. By analogy with what is done with dot products, we can define the angle' θ between two vectors $\mathbf{v}, \mathbf{w} \in V$ by

$$\cos\theta \;=\; \frac{\langle \mathbf{v},\mathbf{w}\rangle}{\|\mathbf{v}\|\cdot\|\mathbf{w}\|}.$$

(This makes sense because the right-hand side is between -1 and 1 by the Cauchy-Schwarz inequality.) We then say that $\mathbf{v}, \mathbf{w} \in V$ are *orthogonal* if $\langle \mathbf{v},\mathbf{w}\rangle = 0$. In the same spirit, we define the projection of $\mathbf{w}$ onto $\mathbf{v}$ by the same formula we used in $\mathbb{R}^n$:

$$\mathrm{Proj}_{\mathbf{v}}(\mathbf{w}) \;=\; \frac{\langle \mathbf{v},\mathbf{w}\rangle}{\|\mathbf{v}\|^2}\;\mathbf{v}.$$

Examples. (5) In $M_{\mathbb{R}}(2,2)$ with its standard metric $\langle A, B\rangle = \operatorname{tr}(B^t A)$, write $A = \begin{pmatrix} 1 & 5 \\ 2 & 3 \end{pmatrix}$ as the sum $A = cI_2 + B$ for some constant $c \in \mathbb{R}$ and a matrix $B \perp A$.

Solution. Note that $\langle A, I\rangle = \operatorname{tr}(IA) = 4$ and $\|I\|^2 = \operatorname{tr}(I) = 2$. Hence $\operatorname{Proj}_{I_2}(A) = \frac{\langle A, I_2\rangle}{\|I_2\|^2} I = \frac{4}{2} I_2 = 2I_2$, so

$$A = 2I_2 + B \qquad \text{where } B = A - 2I_2 = \begin{pmatrix} -1 & 5 \\ 2 & 1 \end{pmatrix}.$$

(6) In $V = \mathcal{C}[0,1]$ with the L^2 inner product, the angle between the $f(x) = x^3$ and $g(x) = x^5$.

Solution. Using the formula $\cos\theta = \dfrac{\langle f, g\rangle}{\|f\| \cdot \|g\|}$ and computing $\begin{cases} \langle f, g\rangle = \int_0^1 x^8\, dx = \frac{1}{9} \\ \|f\|^2 = \int_0^1 x^6\, dx = \frac{1}{7} \\ \|g\|^2 = \int_0^1 x^{10}\, dx = \frac{1}{11} \end{cases}$

shows that $\cos\theta = \frac{\sqrt{77}}{9}$. Taking arccos gives $\theta = 12.8°$.

Homework 36

1. Verify the statement made after Definition 36.1 that Properties (2) and (3a) imply (3b). *one line!*.
2. Read Example 2 above. Show that the formula $\langle \mathbf{v}, \mathbf{w}\rangle = 2v_1w_1 + 5v_2w_2$ does indeed define an inner product on $\mathbb{R}^2$.
3. Show that the formula given in Example 4 above defines an inner product on $\mathcal{C}[0,1]$.
4. Use the formula given in Example 4 to evaluate the inner products between the following pairs of functions:
   (a) x and x^2 in $\mathcal{C}[0,1]$.
   (b) x and x^2 in $\mathcal{C}[-1,1]$.
   (c) $\sin x$ and $\cos x$ in $\mathcal{C}[0,2\pi]$. *Note that* $\sin 2x = 2\sin x\cos x$
   (d) Find $\|\sin 5x\|$ in $\mathcal{C}[0,2\pi]$. *Note that* $\sin^2 x = \frac{1}{2}(1 - \cos 2x)$
5. Use parts (a) and (b) of Theorem 36.2 to prove part (c) (the Triangle Inequality).
6. Let V be a vector space with an inner product. Prove the *Parallelogram Law*: for any $\mathbf{v}, \mathbf{w} \in V$
   $$\|\mathbf{v} + \mathbf{w}\|^2 + \|\mathbf{v} - \mathbf{w}\|^2 = 2\|\mathbf{v}\|^2 + 2\|\mathbf{w}\|^2.$$
   Use the same format as the proof of Theorem 36.2a) above.
7. Consider the matrix $A = \begin{pmatrix} 1 & 3 & -1 \\ 2 & 0 & 4 \\ 3 & -1 & 1 \end{pmatrix}$ in $M_{\mathbb{R}}(3,3)$ with its standard metric (cf. Examples 3 and 5).
   (a) Find angle between A and I_3.
   (b) Find the projection of A onto I_3
   (c) Write $A = cI_3 + B$ for some matrix $B \perp I_3$.
   (d) What does $B \perp I_3$ mean in terms of the trace of B?
8. Consider the functions $f(x) = \frac{1}{2}x$ and $g(x) = \sqrt{x}$ in $\mathcal{C}[0,4]$ with the L^2 inner product. Write g as the sum of a multiple of f and a function h orthogonal to f.

Day 37 Orthonormal sets and Gram-Schmidt

We start with a topic — distance functions — that gives another way that inner products allow one to think about linear algebra geometrically.

Distance Function. In an inner product space V, we define the distance between two vectors $\mathbf{v}$ and $\mathbf{w}$ by

$$d(\mathbf{v}, \mathbf{w}) \;=\; \|\mathbf{v} - \mathbf{w}\| \tag{36.1}$$

This makes V into a *metric space*, i.e. V is a set with a distance function $V \times V \to \mathbb{R}$ that satisfies, for all $\mathbf{v}, \mathbf{w}$ and $\mathbf{u}$ in V,

1. $d(\mathbf{v}, \mathbf{w}) \geq 0$ with equality if and only if $\mathbf{v} = \mathbf{w}$.
2. $d(\mathbf{v}, \mathbf{w}) = d(\mathbf{w}, \mathbf{v})$.
3. (Triangle inequality) $d(\mathbf{v}, \mathbf{w}) \leq d(\mathbf{v}, \mathbf{u}) + d(\mathbf{u}, \mathbf{w})$

The distance function allows us to talk about perpendicular projections onto subspaces, and about converging sequences of vectors in V; more on these below. Using the distance function, we can think about abstract vector spaces (even infinite-dimensional ones) in rather intuitive geometric terms.

Example. What is the distance between $f(x) = \sqrt{x}$ and $g(x) = x$ in $\mathcal{C}[0,1]$ with the L^2 inner product?

Solution. Computing the square of the distance:

$$(d(f,g))^2 \;=\; \|f - g\|^2 \;=\; \int_0^1 (\sqrt{x} - x)^2\, dx \;=\; \int_0^1 x - 2x^{3/2} + x^2\, dx \;=\; \frac{1}{2} - 2 \cdot \frac{2}{5} + \frac{1}{3} \;=\; \frac{1}{30},$$

so $d(f,g) = \frac{1}{\sqrt{30}}$.

We now come to the main topic of the day:

Definition. Let V be a vector space with an inner product. A set $\{\mathbf{e}_1, \ldots, \}$ of vectors is called

- *orthogonal* if $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = 0$ for all $i \neq j$.
- *orthonormal* if, in addition, each $\mathbf{e}_i$ is a unit vector,this is equivalent to

More concisely, $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is orthonormal if $\langle \mathbf{e}_i, \mathbf{e}_j \rangle \;=\; \delta_{ij} \;=\; \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases}$

Orthonormal sets are *extremely* convenient for computations. This is true because, for vectors that are linear combinations of elements in an orthonormal set, one can compute the coordinates of a vector by taking inner products, and can compute inner products from the coordinates:

Theorem 36.1 (Fourier Expansion Theorem Important!). *If $\{\mathbf{e}_1, \mathbf{e}_2, \ldots\}$ is an orthonormal set and $\mathbf{v}, \mathbf{w} \in V$ are sums*

$$\mathbf{v} = \sum_i a_i\, \mathbf{e}_i \qquad \mathbf{w} = \sum_i b_i\, \mathbf{e}_i$$

then

(a) The coefficients of $\mathbf{v}$ are $a_i = \langle \mathbf{v}, \mathbf{e}_i \rangle$; these are called the **Fourier coefficients** *of $\mathbf{v}$.*

(b) $\langle \mathbf{v}, \mathbf{w} \rangle \;=\; \sum_i a_i b_i$. In particular, $\|\mathbf{v}\|^2 = \sum_i |a_i|^2$.

Proof. (a) Take the inner product of $\mathbf{v}$ with $\mathbf{e}_j$:

$$\langle \mathbf{v}, \mathbf{e}_j \rangle \;=\; \left\langle \sum_i a_i \mathbf{e}_i, \mathbf{e}_j \right\rangle \;=\; \sum a_i \langle \mathbf{e}_i, \mathbf{e}_j \rangle \;=\; \sum_i \delta_{ij} \;=\; a_j.$$

(b) Similarly, using bilinearity,

$$\langle \mathbf{v}, \mathbf{w} \rangle \;=\; \left\langle \sum_i a_i \mathbf{e}_i,\; \sum_j b_j \mathbf{e}_j \right\rangle \;=\; \sum_i \sum_j \; a_i b_j \langle \mathbf{e}_i, \mathbf{e}_j \rangle \;=\; \sum_i \sum_j \; a_i b_j \, \delta_{ij} \;=\; \sum_i a_i b_i.$$

□

Two immediate consequences of Theorem 36.1 are:

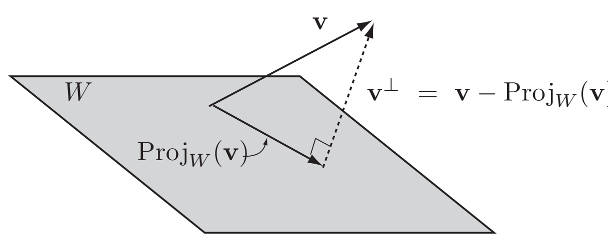
- All orthogonal sets are linearly independent.
- In an n-dimensional vector space, any orthogonal set has at most n vectors, and if it has n vectors then it is a basis.

Proof: To show linear independence, suppose that $\sum_i a_i \mathbf{e}_i = 0$. Then by Theorem 36.1a) with $\mathbf{v} = 0$ we have $a_i = \langle 0, \mathbf{e}_i \rangle = 0$ for all i, as needed. The second fact follows because any set of n linearly independent vectors in an n-dimensional vector space is a basis. □

Orthogonal Projections. In $\mathbb{R}^3$ there is an obvious notion of the perpendicular projection of a vector onto a plane. We now show how to extend the idea to the perpendicular projection onto a subspace in any inner product space. Let V be an inner product space (possibly infinite-dimensional).

Definition. Let $W \subset V$ be a subspace with an orthonormal basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_k\}$ where $k = \dim W$ is finite. The *orthogonal projection of* $\mathbf{v} \in V$ *onto* W is the vector

$$\mathrm{Proj}_W(\mathbf{v}) \;=\; \sum \langle\, \mathbf{v}, \mathbf{e}_j \rangle \; \mathbf{e}_j.$$



Proposition 36.2. *(a)* $\mathbf{v}^{\perp} = \mathbf{v} - \mathit{Proj}_W(\mathbf{v})$ *is perpendicular to* W*, and*

(b) $\mathit{Proj}_W(\mathbf{v})$ *is the vector in* W *closest to* $\mathbf{v}$*.*

Proof. (a) Note that for an arbitrary vector $\mathbf{w} = \sum x^i \mathbf{e}_i$ in W

$$\begin{aligned} \langle \mathbf{v}^{\perp}, \mathbf{w} \rangle \;&=\; \left\langle \mathbf{v} - \mathrm{Proj}_W(\mathbf{v}), \sum x^i \mathbf{e}_i \right\rangle \;=\; \langle \mathbf{v}, \mathbf{w} \rangle - \left\langle \sum_j \langle\, \mathbf{v}, \mathbf{e}_j \rangle \; \mathbf{e}_j, \sum_i x^i \mathbf{e}_i \right\rangle \\ &=\; \langle \mathbf{v}, \mathbf{w} \rangle - \sum_{ij} x^i \langle\, \mathbf{v}, \mathbf{e}_j \rangle \; \delta_{ij}, \end{aligned}$$

which is 0 because $\sum_{ij} x^i \langle\, \mathbf{v}, \mathbf{e}_j \rangle \; \delta_{ij}, = \sum_i \langle \mathbf{v}, x^i \mathbf{e}_i \rangle = \langle \mathbf{v}, \mathbf{w} \rangle$.

(b) The distance from $\mathbf{v}$ to an arbitrary vector $\mathbf{w} = \sum x^i \mathbf{e}_i$ in W, regarded as a function of the x^i, is

$$D(x^1, \cdots x^k) \;=\; \text{dist}(\mathbf{v}, \mathbf{w}) \;=\; \left\| v - \sum x^i \mathbf{e}_i \right\|,$$

Again write $\mathbf{v} = \text{Proj}_W(\mathbf{v}) + \mathbf{v}^\perp$. Then $\text{Proj}_W(\mathbf{v}) - \mathbf{w} = \sum(\langle \mathbf{v}, \mathbf{e}_i - x^i)\ \mathbf{e}_i$ lies in W, so by part (a)

$$\begin{aligned} \|\mathbf{v} - \mathbf{w}\|^2 \;&=\; \|(\text{Proj}_W(\mathbf{v}) - \mathbf{w}) + \mathbf{v}^\perp\|^2 \;=\; \|(\text{Proj}_W(\mathbf{v}) - \mathbf{w})\|^2 \;+\; \|\mathbf{v}^\perp\|^2 \\ &=\; \sum_i \left| \langle \mathbf{v}, \mathbf{e}_i \rangle - x^i \right|^2 \;+\; \|\mathbf{v}^\perp\|^2. \end{aligned}$$

where we have used Theorem 36.1b) in the last step. Thus

$$D(x^1, \cdots x^k) \;=\; \sqrt{\sum_i \;|\langle \mathbf{v}, \mathbf{e}_i \rangle - x^i\big|^2 + C^2}$$

where $C = \|\mathbf{v}^\perp\| \geq 0$ is a constant independent of the numbers x^i. Thus $D(\mathbf{x}) \geq C$ with equality if and only if $x^i = \langle \mathbf{v}, \mathbf{e}_i \rangle$ for all i, that is, if and only if $\mathbf{w} = \sum x^i \mathbf{e}_i$ is $\text{Proj}_W(\mathbf{v})$.

□

Corollary (Bessel's inequality). If $\{\mathbf{e}_i\}$ is orthonormal and $\mathbf{v} = \sum a_i \mathbf{e}_i$, then

$$\sum_i |a_i|^2 \leq \|\mathbf{v}\|^2$$

with equality if and only if $\mathbf{v} \in \text{span}(\mathbf{e}_i)$.

Proof. Writing $\mathbf{v} = \text{Proj}_W(\mathbf{v}) + \mathbf{v}^\perp$ and noting that $\langle \text{Proj}_W(\mathbf{v}),\ \mathbf{v}^\perp \rangle = 0$, we have

$$\|\mathbf{v}\|^2 \;=\; \left\|\text{Proj}_W(\mathbf{v})\right\|^2 + \|\mathbf{v}^\perp\|^2 \;\geq\; \left\|\text{Proj}_W(\mathbf{v})\right\|^2 \;=\; \sum |\langle \mathbf{v}, \mathbf{e}_i \rangle|^2\ \|\mathbf{e}_i\|^2 \;=\; \sum_i |a_i|^2.$$

Equality holds if and only if $\mathbf{v}^\perp = 0$, which means that $\mathbf{v} \in \text{span}(\mathbf{e}_i)$. □

Homework 37

1. Use the Definition 36.1 of inner product to prove that the distance function (36.1) has the three properties stated.
2. Find the distance between each pair of vectors:
   (a) $\mathbf{v} = (1, 2, 3, 4)$ and $\mathbf{w} = (2, 1, 0, 5)$ in $\mathbb{R}^4$ with the standard inner product.
   (b) $\mathbf{v} = \begin{pmatrix} 1 & 2 \\ 0 & -3 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} 2 & -3 \\ 4 & 1 \end{pmatrix}$ in $M_{\mathbb{R}}(2, 2)$ with the standard inner product $\langle A, B \rangle = \text{tr}\,(B^t A)$.
   (c) $f(x) = \sin x$ and $g(x) = \cos x$ in $\mathbf{C}[0, 2\pi]$.
3. In $\mathbb{R}^2$ with the standard dot product, let $\mathbf{e}_1 = \frac{1}{\sqrt{2}}(1, 1)$ and $\mathbf{e}_2 = \frac{1}{\sqrt{2}}(1, -1)$.
   (a) Show that $\{\mathbf{e}_1, \mathbf{e}_2\}$ is an orthonormal set.
   (b) Find the Fourier coefficients of $(3, 4)$ and write $(3, 4)$ as a linear combination of $\mathbf{e}_1$ and $\mathbf{e}_2$.
4. Let W be the subspace of $\mathbf{C}[0, 1]$ spanned by the functions $f(x) = \sqrt{3}x$ and $g(x) = \sqrt{5}(4x^2 - 3x)$.
   (a) Show that $\{f, g\}$ is an orthonormal set.
   (b) Find the projection of $h(x) = x^3$ onto W. Draw a picture with labels.
   (c) How close is h to W?

Day 38 Gram-Schmidt and complex inner products

We have seen that computations are easier and more systematic when one works with an orthonormal basis $\{\mathbf{e}_i\}$. In particular, for vectors

$$\mathbf{v} = \sum_i a_i\, \mathbf{e}_i \qquad \mathbf{w} = \sum_i b_i\, \mathbf{e}_i$$

- $a_i = \langle \mathbf{v}, \mathbf{e}_i \rangle$.
- $\langle \mathbf{v}, \mathbf{w} \rangle \;=\; \sum_i a_i b_i$ and $\|\mathbf{v}\|^2 = \sum_i |a_i|^2$.
- The orthogonal projection of $\mathbf{v} \in V$ onto $W = \text{span}(\mathbf{e}_1, \ldots, \mathbf{e}_k)$ is

$$\text{Proj}_W(\mathbf{v}) \;=\; \sum_{i=1}^{k} \langle\, \mathbf{v}, \mathbf{e}_i \rangle\; \mathbf{e}_i.$$

Of course, to make this work one needs to be able to find an orthonormal basis. There is a simple algorithm for doing so called the Gram-Schmidt Process. It starts with *any* basis, and systematically transmutes it into an orthonormal basis.

The Gram-Schmidt Process. Given a basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ of an inner product space, we can form an orthonormal basis $\{\mathbf{e}_1, \ldots \mathbf{e}_n\}$ by these steps:

STEP 1. Normalize $\mathbf{u}_1$ by setting $\mathbf{e}_1 = \dfrac{\mathbf{u}_1}{\|\mathbf{u}_1\|}$.

STEP 2A. Replace $\mathbf{u}_2$ by the vector $\mathbf{v}_2 = \mathbf{v}_2 - \text{Proj}_{\mathbf{e}_1}(\mathbf{u}_2)$ perpendicular to $\mathbf{e}_1$.

STEP 2B. Normalize $\mathbf{v}_2$ by setting $\mathbf{e}_2 = \dfrac{\mathbf{v}_2}{\|\mathbf{v}_2\|}$.

STEP 3A. Replace $\mathbf{u}_3$ by $\mathbf{v}_3 = \mathbf{u}_3 - \text{Proj}_{W_2}(\mathbf{u}_3)$ perpendicular to $W_2 = \text{span}(\mathbf{e}_1, \mathbf{e}_2)$.

STEP 3B. Normalize $\mathbf{v}_3$, etc.

> At the kth step, calculate
>
> $$\mathbf{v}_k \;=\; \mathbf{u}_k \;-\; \langle \mathbf{u}_k, \mathbf{e}_1 \rangle\; \mathbf{e}_1 \;-\; \cdots \;-\; \langle \mathbf{u}_k, \mathbf{e}_{k-1} \rangle\; \mathbf{e}_{k-1}$$
>
> and then set $\mathbf{e}_k = \dfrac{\mathbf{v}_k}{\|\mathbf{v}_k\|}$.

Tips: When applying the Gram-Schmidt process with specific vectors, it is helpful to:

(i) Pull out common factors. For example, write $\sqrt{2}(1, 1, 1)$ instead of $(\sqrt{2}, \sqrt{2}, \sqrt{2})$.

(ii) Erase the common factor before normalizing. For example, the normalization of $(\sqrt{2}, \sqrt{2}, \sqrt{2})$ is the same as the normalization of $(1, 1, 1)$, which is much easier to compute.

Example 1. Find an orthonormal basis for the subspace W of $\mathbb{R}^4$ generated by the vectors

$$\mathbf{u}_1 = (1, 0, 1, 1) \qquad \mathbf{u}_2 = (1, 0, -2, 0) \qquad \mathbf{u}_3 = (1, -1, 0, 2).$$

Solution. Apply the Gram-Schmidt process to $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$.

STEP 1. Since $\|\mathbf{u}_1\|^2 = (1,0,1,1)\cdot(1,0,1,1) = 3$ the normalization of $\mathbf{u}_1$ is

$$\mathbf{e}_1 \;=\; \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} \;=\; \frac{1}{\sqrt{3}}\,(1,0,1,1).$$

STEP 2.

$$\begin{aligned}\mathbf{v}_2 \;=\; \mathbf{u}_2 - \langle \mathbf{u}_2, \mathbf{e}_1\rangle\, \mathbf{e}_1 \;&=\; (1,0,-2,0) - \tfrac{1}{3}(-1)\,(1,0,1,1)\\ &=\; \tfrac{1}{3}\Big[(3,0,-6,0) + (1,0,1,1)\Big]\\ &=\; \tfrac{1}{3}\,(4,0,-5,1)\end{aligned}$$

To normalize, erase the $\frac{1}{3}$ and compute $(4,0,-5,1)\cdot(4,0,-5,1) = 42$. Hence the normalization is

$$\mathbf{e}_2 \;=\; \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} \;=\; \frac{1}{\sqrt{42}}\,(4,0,-5,1).$$

STEP 3.

$$\begin{aligned}\mathbf{v}_3 \;=\; \mathbf{u}_3 - \langle \mathbf{u}_2, \mathbf{e}_1\rangle\, \mathbf{e}_1 - \langle \mathbf{u}_3, \mathbf{e}_2\rangle\, \mathbf{e}_2 \;&=\; (1,-1,0,2) - \tfrac{1}{3}(3)\,(1,0,1,1) - \tfrac{1}{42}(6)\,(4,0,-5,1)\\ &=\; \frac{1}{7}\Big[(7,-7,0,14) - (7,0,7,7) - (4,0,-5,1)\Big]\\ &=\; (-4,-7,-2,6).\end{aligned}$$

Thus we obtain the orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ for W where

$$\mathbf{e}_1 \;=\; \tfrac{1}{\sqrt{3}}\,(1,0,1,1), \qquad \mathbf{e}_1 \;=\; \tfrac{1}{\sqrt{42}}\,(4,0,-5,1), \qquad \mathbf{e}_1 \;=\; \tfrac{1}{\sqrt{105}}\,(-4,-7,-2,6).$$

Corollary 38.1. *Every vector space with a countable basis has an orthonormal basis* $\{\mathbf{e}_i\}$.

Proof. Start with any basis and apply the Gram-Schmidt process (this works even for infinite dimensional vector spaces). □

Inner products in complex vector spaces. Recall that each complex number $z = a + bi$ has a conjugate $\overline{z} = a - bi$ and a norm

$$|z|^2 \;=\; z\overline{z} \;=\; (a+bi)(a-bi) \;=\; a^2 + b^2,$$

which is real. On the other hand, $z^2 = (a^2 - b^2) + 2abi$ is usually not real. Accordingly, for complex vector spaces we must modify the definition of inner product so that $\langle \mathbf{v}, \mathbf{w}\rangle$ is real.

Definition 38.2. *The standard hermitian inner product on* $\mathbb{C}^n$ *is*

$$\mathbf{v}\cdot\mathbf{w} \;=\; \sum_{i=1}^{n} v_i\overline{w_i}.$$

Taking the conjugate of the second vector ensures that $\|\mathbf{v}\|^2 = \mathbf{v}\cdot\mathbf{v} = \sum v_i\overline{v_i} = \sum |v_i|^2$ is real.

Example 1. $$\begin{aligned}(1, 3-i, 4i)\cdot(1, 1+2i, 2+3i) \;&=\; 1 + (3-i)(1-2i) + 4i(2+3i)\\ &=\; 1 + (3-2-7i) + (-12+8i) \;=\; -10+i.\end{aligned}$$

More generally, the definition of an inner product on a complex vector space is the same as the real case, except that a conjugate appears in the " symmetric" and "bilinear" properties:

Definition 38.3. *Let V be a complex vector space. A (hermitian) inner product on V is a map $V \times V \to \mathbb{C}$ so that*

(a) (Positive definite) $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$, *with equality if and only if* $\mathbf{v} = \mathbf{0}$.

(b) (Conjugate Symmetric) $\langle \mathbf{v}, \mathbf{w} \rangle \;=\; \overline{\langle \mathbf{w}, \mathbf{v} \rangle}$ *for all* $\mathbf{v}, \mathbf{w} \in V$.

(c) (Sesquilinear) *(a)* $\langle a\mathbf{u} + b\mathbf{v}, \mathbf{w} \rangle = a\langle \mathbf{u}, \mathbf{w} \rangle + b\langle \mathbf{v}, \mathbf{w} \rangle$ *and*
(b) $\langle \mathbf{u}, c\mathbf{v} + d\mathbf{w} \rangle = \overline{c}\langle \mathbf{u}, \mathbf{v} \rangle + \overline{d}\langle \mathbf{u}, \mathbf{w} \rangle$ *for all* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$.

"Sesquilinear" means " one-and-one-half linear"; a more descriptive phrase is "conjugate linear in the second variable". *Caution:* Some books – about half – put the conjugate on the first vector.

Again, Properties (2) and (3a) imply (3b), and again we define the norm by

$$\|\mathbf{v}\| \;=\; \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}.$$

See Section 6.1 of the textbook for details on the following examples:

Example 2. The standard hermitian inner product on $V = \mathbb{C}^n$ above.

Example 3. On $M_{\mathbb{C}}(n, m)$, define the inner product of two complex matrices A, B by $\langle A, B \rangle \;=\; \mathrm{tr}\,(B^* A)$ where $B^* = \overline{B}^t$ is the conjugate of the transpose of B.

Example 4. On the vector space $V = \mathcal{C}_{\mathbb{C}}[a, b]$ of continuous complex-valued functions on the interval $[a, b] \subset \mathbb{R}$, define the L^2 inner product by

$$\langle f,\ g \rangle \;=\; \int_a^b f(x)\overline{g(x)}\;dx.$$

Everything we have done with inner products and projections carries over to the complex case, with conjugates appearing in the obvious places. In particular, if $\mathbf{v} = \sum_i a_i\,\mathbf{e}_i$, $\mathbf{w} = \sum_i b_i\,\mathbf{e}_i$ and $W = \mathrm{span}(\mathbf{e}_1, \ldots, \mathbf{e}_k)$, then

$$a_i = \langle \mathbf{v}, \mathbf{e}_i \rangle, \qquad \langle \mathbf{v}, \mathbf{w} \rangle \;=\; \sum_i a_i\overline{b_i}, \qquad \|\mathbf{v}\|^2 = \sum_i |a_i|^2, \qquad \mathrm{Proj}_W(\mathbf{v}) \;=\; \sum_{i=1}^k \langle\ \mathbf{v}, \mathbf{e}_i \rangle\ \mathbf{e}_i.$$

where the first and last of these formulas must have $\langle\ \mathbf{v}, \mathbf{e}_i \rangle$ rather than $\langle\ \mathbf{e}_i, \mathbf{v} \rangle$.

Homework 38

1. Normalize the vectors $(2, 1, -2) \in \mathbb{R}^3$ and $(1, 7, 1, 7) \in \mathbb{R}^4$.
2. Show that $\mathbf{v}_1 = \frac{1}{7}(3, 6, -2)$, $\mathbf{v}_2 = \frac{1}{7}(-2, 3, 6)$ and $\mathbf{v}_3 = \frac{1}{7}(6, -2, 3)$ are an orthonormal set in $\mathbb{R}^3$. Without further computations, explain why they are a basis of $\mathbb{R}^3$.
3. Perform the Gram-Schmidt process on the following basis of $\mathbb{R}^2$:

$$\mathbf{v}_1 = \begin{pmatrix} -3 \\ 4 \end{pmatrix} \qquad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 7 \end{pmatrix}$$

Illustrate your work with sketches showing the perpendicular projection you are using.

4. Perform the Gram-Schmidt process on the following basis of $\mathbb{R}^3$:

$$\mathbf{v}_1 = \begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix} \qquad \mathbf{v}_2 = \begin{pmatrix} b \\ c \\ 0 \end{pmatrix} \qquad \mathbf{v}_3 = \begin{pmatrix} d \\ e \\ f \end{pmatrix}$$

   Here a, c, f are positive numbers and the other constants are arbitrary. Illustrate your work with a sketch as in the previous problem.

5. Consider the following vectors in $\mathbb{R}^4$:

$$\mathbf{v}_1 = \begin{pmatrix} 4 \\ 2 \\ 2 \\ 1 \end{pmatrix} \qquad \mathbf{v}_2 = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 2 \end{pmatrix} \qquad \mathbf{v}_3 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix} \qquad\qquad \mathbf{v} = \begin{pmatrix} 0 \\ 5 \\ 1 \\ -2 \end{pmatrix}$$

   (a) Use the Gram-Schmidt process to find an orthonormal basis of the subspace $W = \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$.

   (b) Find the orthogonal projection of $\mathbf{v}$ onto W.

6. Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$ be an orthonormal basis for an inner product space V. Let W be the subspace spanned by $\mathbf{v}_1, \ldots, \mathbf{v}_k$ and let W' be the subspace spanned by $\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n$. Prove that $W \perp W'$, that is, $\langle w, w' \rangle = 0$ for every $w \in W$ and $w' \in W'$. *This can be done in one line!*

7. Consider the vector space of $\mathcal{C}[-1, 1]$ with its L^2 inner product

$$\langle f, g \rangle \;=\; \int_{-1}^{1} f(x)g(x)\, dx.$$

   Apply the Gram-Schmidt process to $\mathbf{u}_1 = 1, \mathbf{u}_2 = x, \mathbf{u}_3 = x^2$ and $\mathbf{u}_4 = x^3$ to obtain an orthonormal basis $\{P_0(x), P_1(x), P_2(x), P_3(x)\}$ of the subspace $\mathbf{P}_3(\mathbb{R})$ of $\mathcal{C}[-1, 1]$.

   One could continue, obtaining a sequence of polynomials $P_1, P_2, P_3, \ldots$ that are orthonormal with $\deg P_n = n$. These are (up to normalization) the *Legendre polynomials* that occur in physics and differential equations. The next two are (again up to normalization)

$$P_4(x) = \tfrac{1}{8}(35x^4 - 30x^2 + 3). \qquad P_5(x) = \tfrac{1}{8}(63x^5 - 70x^3 + 15x).$$

8. For the vectors $\mathbf{x} = (2, 1+i, i)$ and $\mathbf{y} = (2-i, 2, 1+i)$ in $\mathbb{C}^3$ with the standard hermitian inner product (Definition 38.2), compute $\langle \mathbf{x}, \mathbf{y} \rangle$, $\|\mathbf{x}\|$, and $\|\mathbf{y}\|$, and use these to find $\|\mathbf{x} + \mathbf{y}\|$.

9. Use the inner product on $M_{\mathbb{C}}(2, 2)$ given in Example 3 above to find $\langle A, B \rangle$ and $\|A\|$ for

$$A \;=\; \begin{pmatrix} 1 & 2+i \\ 3 & i \end{pmatrix} \qquad \text{and} \qquad B \;=\; \begin{pmatrix} 1+i & 0 \\ i & -i \end{pmatrix}$$