1. Question: Inner product and orthonormal basis of $\mathbb{R}^n$ (*elementary*)

1.1. Show that the dot product for vectors, i.e. $\mathbf{x}^\top \mathbf{y}$ with $\mathbf{x}, \mathbf{y} \in V$ for some vector space V, is an inner product. Which norm and metric does it induce?

1.2. Can you think of a definition of an inner product $\langle \cdot, \cdot \rangle_{\text{ex}}$ that isn't the dot product $\langle \cdot, \cdot \rangle_{\text{dot}}$?
If two vectors are orthogonal in $(V, \langle \cdot, \cdot \rangle_{\text{dot}})$, are they also orthogonal in $(V, \langle \cdot, \cdot \rangle_{\text{ex}})$?

1.3. Show that, for some inner product space $(V, \langle \cdot, \cdot \rangle)$ with zero element $\mathbf{0}$, the following holds: *All vectors* $\mathbf{v} \in V$ *are orthogonal to* $\mathbf{0}$*, and* $\mathbf{0}$ *is the only vector in* V *that is orthogonal to itself.*

Solution:

1.1. Let us assume that all vectors in V have n entries. We can easily verify that the dot product fulfills all properties of an inner product:

1. Clearly, $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^\top \mathbf{y} = \sum_{i=1}^n x_i y_i = \sum_{i=1}^n y_i x_i = \mathbf{y}^\top \mathbf{x} = \langle \mathbf{y}, \mathbf{x} \rangle$.
2. $$\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = (\mathbf{x} + \mathbf{y})^T \mathbf{z} = \left(\mathbf{x}^T + \mathbf{y}^T\right) \mathbf{z} = \mathbf{x}^T \mathbf{z} + \mathbf{y}^T \mathbf{z} = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle,$$
   where we have used that matrix multiplication distributes over addition.
3. Clearly, for a scalar c, $\langle c\mathbf{x}, \mathbf{y} \rangle = (c\mathbf{x})^\top \mathbf{y} = \sum_{i=1}^n (c x_i) y_i = \sum_{i=1}^n c(x_i y_i) = c(\mathbf{x}^\top \mathbf{y}) = c\langle \mathbf{x}, \mathbf{y} \rangle$.
4. $$\langle \mathbf{x}, \mathbf{x} \rangle = \mathbf{x}^T \mathbf{x} = x_1^2 + x_2^2 + \cdots + x_n^2 \geq 0$$
   with equality if and only if $x_i = 0 \ \forall i \in \{1, \ldots, n\}$, that is, $\mathbf{x} = \mathbf{0}$. Furthermore, the dot product induces the euclidean norm and distance, which we can also easily verify:
   $$\sqrt{\mathbf{x}^\top \mathbf{x}} = \sqrt{\textstyle\sum_{i=1}^n x_i^2} \text{ and } \sqrt{(\mathbf{x} - \mathbf{y})^\top (\mathbf{x} - \mathbf{y})} = \sqrt{\textstyle\sum_{i=1}^n (x_i - y_i)^2}.$$

1.2. One example would be $\langle \boldsymbol{x}, \boldsymbol{y} \rangle := \boldsymbol{x}^\top \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \boldsymbol{y}$.
No, the angles between vectors can generally vary across different inner products. See question **2.2** for an example.

1.3. First, we note that the fact that $\mathbf{0}$ is the only vector orthogonal to itself follows immediately from the definition of inner products, which states that $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ if and only if $\mathbf{u} = \mathbf{0}$.
Furthermore, combining this fact with the third and second properties of inner products, we get

$$0 = \langle \mathbf{v} - \mathbf{v}, \mathbf{0} \rangle = \langle \mathbf{v}, \mathbf{0} \rangle + \langle -\mathbf{v}, \mathbf{0} \rangle,$$

So both $\langle \mathbf{v}, \mathbf{0} \rangle$ and $\langle -\mathbf{v}, \mathbf{0} \rangle$ must be equal to 0 for all vectors $\mathbf{v} \in V$, which proves that all vectors $\mathbf{v} \in V$ are orthogonal to $\mathbf{0}$.

2. Question: Angles between Vectors and Projection onto a Line (*elementary*)

2.1. Find the angle in between vectors $\mathbf{a} = (8, -2, 16)^\top$ and $\mathbf{b} = (-9, 8, 12)^\top$ in radian and degrees.

2.2. Calculate the angle between the vectors $\boldsymbol{x} = [1, 1]^\top, \boldsymbol{y} = [-1, 1]^\top \in \mathbb{R}^2$ with regards to *both the dot product and the inner product defined as*

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle := \boldsymbol{x}^\top \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \boldsymbol{y}\,.$$

2.3. Let $\mathbf{y} = \begin{bmatrix} 7 \\ 6 \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$. Find the orthogonal projection of $\mathbf{y}$ onto u. Then write $\mathbf{y}$ as the sum of two orthogonal vectors, one in span$\{\mathbf{u}\}$ and one orthogonal to $\mathbf{u}$.

2.4. Project

(i) the vector $\begin{pmatrix} 2 \\ -1 \\ 4 \end{pmatrix}$ orthogonally onto the line $\left\{ c \begin{pmatrix} -3 \\ 1 \\ -3 \end{pmatrix} \middle| \, c \in \mathbb{R} \right\}$

(ii) $\begin{pmatrix} -1 \\ -1 \end{pmatrix}$ orthogonally onto the line $y = 3x$.

Solution:

2.1.

$$\begin{aligned} \cos\omega &= \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|} = \frac{8 \times (-9) + (-2) \times 8 + 16 \times 12}{\sqrt{8^2 + (-2)^2 + 16^2} \cdot \sqrt{(-9)^2 + 8^2 + 12^2}} \\ &= \frac{104}{18 \cdot 17} = \frac{104}{306} \\ \Longrightarrow \omega &= \arccos\left(\frac{104}{306}\right) \approx 1.22. \end{aligned}$$

So the angle is $\approx 1.22\,$radian $\approx 68°$

2.2. For the dot product, we can immediately infer from

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 1 \end{bmatrix} = -1 + 1 = 0$$

that $\mathbf{x}$ and $\mathbf{y}$ are orthogonal, so the angle between them equals 90°.

Meanwhile, for $\langle \boldsymbol{x}, \boldsymbol{y} \rangle := \boldsymbol{x}^\top \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \boldsymbol{y}$ we get that the angle ω between $\boldsymbol{x}$ and $\boldsymbol{y}$ is given by

$$\cos\omega = \frac{\langle \boldsymbol{x}, \boldsymbol{y} \rangle}{\|\boldsymbol{x}\|\|\boldsymbol{y}\|} = -\frac{1}{3} \Longrightarrow \omega \approx 1.91\text{rad} \approx 109.5°,$$

and $\boldsymbol{x}$ and $\boldsymbol{y}$ are not orthogonal.

2.3. Compute

$$\begin{aligned} \mathbf{y} \cdot \mathbf{u} &= \begin{bmatrix} 7 \\ 6 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 40 \\ \mathbf{u} \cdot \mathbf{u} &= \begin{bmatrix} 4 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 20 \end{aligned}$$

The orthogonal projection of $\mathbf{y}$ onto $\mathbf{u}$ is

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}} \mathbf{u} = \frac{40}{20} \mathbf{u} = 2 \begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

and the component of $\mathbf{y}$ orthogonal to $\mathbf{u}$ is

$$\mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} 7 \\ 6 \end{bmatrix} - \begin{bmatrix} 8 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

The sum of these two vectors is $\mathbf{y}$. That is,

$$\underset{\substack{\uparrow \\ y}}{\begin{bmatrix} 7 \\ 6 \end{bmatrix}} = \underset{\substack{\uparrow \\ \hat{y}}}{\begin{bmatrix} 8 \\ 4 \end{bmatrix}} + \underset{\substack{\uparrow \\ (y - \hat{y})}}{\begin{bmatrix} -1 \\ 2 \end{bmatrix}}$$

This decomposition of $\mathbf{y}$ is illustrated in Figure 3. Note: If the calculations above are correct, then $\{\hat{\mathbf{y}}, \mathbf{y} - \hat{\mathbf{y}}\}$ will be an orthogonal set. As a check, compute

$$\hat{\mathbf{y}} \cdot (\mathbf{y} - \hat{\mathbf{y}}) = \begin{bmatrix} 8 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 2 \end{bmatrix} = -8 + 8 = 0$$

2.4. (i) $$\frac{\begin{pmatrix} 2 \\ -1 \\ 4 \end{pmatrix} \cdot \begin{pmatrix} -3 \\ 1 \\ -3 \end{pmatrix}}{\begin{pmatrix} -3 \\ 1 \\ -3 \end{pmatrix} \cdot \begin{pmatrix} -3 \\ 1 \\ -3 \end{pmatrix}} \cdot \begin{pmatrix} -3 \\ 1 \\ -3 \end{pmatrix} = \frac{-19}{19} \cdot \begin{pmatrix} -3 \\ 1 \\ -3 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \\ 3 \end{pmatrix}$$

(ii) Writing the line as $\left\{ c \cdot \binom{1}{3} \middle| \, c \in \mathbb{R} \right\}$ gives this projection:

$$\frac{\binom{-1}{-1} \cdot \binom{1}{3}}{\binom{1}{3} \cdot \binom{1}{3}} \cdot \binom{1}{3} = \frac{-4}{10} \cdot \binom{1}{3} = \binom{-2/5}{-6/5}$$

3. Question: Gram-Schmidt Process *(elementary)*

3.1. Carry out the Gram-Schmidt orthonormalization process on the following pair of vectors in $\mathbb{R}^2$ to obtain an orthonormal basis:

$$\begin{bmatrix} 2 \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} -1 \\ 3 \end{bmatrix}.$$

3.2. Carry out the Gram-Schmidt orthonormalization process on the following three vectors in $\mathbb{R}^3$ to obtain an orthonormal basis:

$$\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 8 \\ 1 \\ -6 \end{bmatrix}, \text{ and } \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

3.3. Perform an Eigendecomposition of the following matrix. How do the spectral theorem and the Gram-Schmidt process help you here?

$$\boldsymbol{A} = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & 2 \\ 2 & 2 & 3 \end{bmatrix}.$$

(You may use the fact that $\det(\mathbf{A} - \lambda \mathbf{I}) = -(\lambda - 1)^2(\lambda - 7)$ to save time.)

Solution:

3.1. We apply the Gram-Schmidt algorithm with $\mathbf{b}_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\mathbf{b}_2 \begin{bmatrix} -1 \\ 3 \end{bmatrix}$.

First, set $u_1 := \mathbf{b}_1$ and normalize it:

$$\|\mathbf{b}_1\| = \sqrt{2^2 + 1^2} = \sqrt{5}$$
$$\implies w_1 = \frac{1}{\sqrt{5}}(2, 1) = \left(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}} \right).$$

Second, find w_2 :

$$\begin{aligned}
\mathbf{b}_2 - \frac{\langle \mathbf{b}_1, u_1\rangle}{\langle u_1, u_1\rangle}\mathbf{b}_1 &= (-1,3) - \frac{(-1,3)\cdot(2,1)}{(2,1)\cdot(2,1)}(2,1)\\
&= (-1,3) - \frac{1}{5}(2,1)\\
&= \left(-\frac{7}{5}, \frac{14}{5}\right)\\
\text{Since } \left|\left(-\frac{7}{5}, \frac{14}{5}\right)\right| &= \frac{7}{\sqrt{5}}\\
\text{it follows that } w_2 &= \frac{\sqrt{5}}{7}\left(-\frac{7}{5}, \frac{14}{5}\right)\\
&= \left(-\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)
\end{aligned}$$

Now w_1, w_2 constitute an orthonormal basis for $\mathbb{R}^2$.

3.2. We apply the Gram-Schmidt algorithm with $\mathbf{b}_1 = \begin{bmatrix}1\\2\\0\end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix}8\\1\\-6\end{bmatrix}$, $\mathbf{b}_3 = \begin{bmatrix}0\\0\\1\end{bmatrix}$.

$$u_1 = \mathbf{b}_1 = \begin{bmatrix}1\\2\\0\end{bmatrix} \text{ and } w_1 = \frac{\mathbf{b}_1}{\|\mathbf{b}_1\|} = \frac{1}{\sqrt{5}}\begin{bmatrix}1\\2\\0\end{bmatrix}$$

$$u_2 = \mathbf{b}_2 - \frac{\langle \mathbf{b}_2, u_1\rangle}{\langle u_1, u_1\rangle}u_1 = \begin{bmatrix}8\\1\\-6\end{bmatrix} - \frac{\left\langle \begin{bmatrix}8\\1\\-6\end{bmatrix}, \begin{bmatrix}1\\2\\0\end{bmatrix}\right\rangle}{\left\langle \begin{bmatrix}1\\2\\0\end{bmatrix}, \begin{bmatrix}1\\2\\0\end{bmatrix}\right\rangle}\begin{bmatrix}1\\2\\0\end{bmatrix} = \begin{bmatrix}8\\1\\-6\end{bmatrix} - \frac{10}{5}\begin{bmatrix}1\\2\\0\end{bmatrix} = \begin{bmatrix}6\\-3\\-6\end{bmatrix}$$

$$\Longrightarrow w_2 = \frac{u_2}{\|u_2\|} = \frac{1}{9}\begin{bmatrix}6\\-3\\-6\end{bmatrix} = \frac{1}{3}\begin{bmatrix}2\\-1\\-2\end{bmatrix}$$

$$u_3 = \mathbf{b}_3 - \frac{\langle \mathbf{b}_3, u_1\rangle}{\langle u_1, u_1\rangle}u_1 - \frac{\langle \mathbf{b}_3, u_2\rangle}{\langle u_2, u_2\rangle}u_2 = \begin{bmatrix}0\\0\\1\end{bmatrix} - \frac{\left\langle \begin{bmatrix}0\\0\\1\end{bmatrix}, \begin{bmatrix}1\\2\\0\end{bmatrix}\right\rangle}{\left\langle \begin{bmatrix}1\\2\\0\end{bmatrix}, \begin{bmatrix}1\\2\\0\end{bmatrix}\right\rangle}\begin{bmatrix}1\\2\\0\end{bmatrix} - \frac{\left\langle \begin{bmatrix}0\\0\\1\end{bmatrix}, \begin{bmatrix}6\\-3\\-6\end{bmatrix}\right\rangle}{\left\langle \begin{bmatrix}6\\-3\\-6\end{bmatrix}, \begin{bmatrix}6\\-3\\-6\end{bmatrix}\right\rangle}\begin{bmatrix}6\\-3\\-6\end{bmatrix}$$

$$= \begin{bmatrix}0\\0\\1\end{bmatrix} - \frac{0}{5}\begin{bmatrix}1\\2\\0\end{bmatrix} - \frac{-6}{81}\begin{bmatrix}6\\-3\\-6\end{bmatrix} = \frac{1}{9}\begin{bmatrix}4\\-2\\5\end{bmatrix}$$

$$\Longrightarrow w_3 = \frac{u_3}{\|u_3\|} = \frac{1}{3\sqrt{5}}\begin{bmatrix}4\\-2\\5\end{bmatrix}$$

3.3. The characteristic polynomial of $\boldsymbol{A}$ is

$$p_{\boldsymbol{A}}(\lambda) = -(\lambda-1)^2(\lambda-7),$$

so that we obtain the eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 7$, where λ_1 is a repeated eigenvalue. Following our standard procedure for computing eigenvectors, we obtain the eigenspaces

$$E_1 = \text{span}\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \underbrace{\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}}_{=:\boldsymbol{x}_1} \right\}, \quad E_7 = \text{span}\left\{ \underbrace{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}_{=:\boldsymbol{x}_3} \right\}.$$

We see that $\boldsymbol{x}_3$ is orthogonal to both $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$. However, since $\boldsymbol{x}_1^\top \boldsymbol{x}_2 = 1 \neq 0$, they are not orthogonal. The **spectral theorem** states that there exists an orthogonal basis, but the one we have is not orthogonal. However, we can construct one.

To construct such a basis, we exploit the fact that $\boldsymbol{x}_1, \boldsymbol{x}_2$ are eigenvectors associated with the same eigenvalue λ. Therefore, for any $\alpha, \beta \in \mathbb{R}$ it holds that

$$\boldsymbol{A}\left(\alpha \boldsymbol{x}_1 + \beta \boldsymbol{x}_2\right) = \boldsymbol{A}\boldsymbol{x}_1 \alpha + \boldsymbol{A}\boldsymbol{x}_2 \beta = \lambda\left(\alpha \boldsymbol{x}_1 + \beta \boldsymbol{x}_2\right),$$

i.e., any linear combination of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ is also an eigenvector of $\boldsymbol{A}$ associated with λ. The **Gram-Schmidt algorithm** is a method for iteratively constructing an orthogonal/orthonormal basis from a set of basis vectors using such linear combinations. Therefore, even if $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are not orthogonal, we can apply the Gram-Schmidt algorithm and find eigenvectors associated with $\lambda_1 = 1$ that are orthogonal to each other (and to $\boldsymbol{x}_3$). In our example, *without normalization*, we will obtain

$$\boldsymbol{x}_1' = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{x}_2' = \frac{1}{2}\begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$$

which are orthogonal to each other, orthogonal to $\boldsymbol{x}_3$, and eigenvectors of $\boldsymbol{A}$ associated with $\lambda_1 = 1$.

Lastly, we can normalize , $\boldsymbol{x}_1'$, $\boldsymbol{x}_2'$ and $\boldsymbol{x}_3$ to obtain an orthogonal *and thereby easily invertible* matrix with eigenvectors as columns:

$$\mathbf{P} := \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

and we get the following eigendecomposition:

$$\mathbf{A} = \mathbf{P}\,\text{diag}(1,1,7)\mathbf{P}^\top = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 7 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} & 1\frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}.$$

4. Question: Projection onto general Subspaces *(elementary)*

4.1. How does the formula for orthogonal projection onto subspaces from definition 5.7 simplify if the given basis is not only orthogonal but orthonormal?

4.2. In $\mathbb{R}^3$, let

$$W = \text{span}\left\{ \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \right\}$$

be the subspace spanned by the vectors $(1,1,2)^\top$ and $(1,1,-1)^\top$. What point of W is closest to the vector $(4,5,-2)^\top$?

4.3. In $\mathbb{R}^3$, find the orthogonal projection of $(2,2,5)^\top$ on the subspace

$$W = \operatorname{span}\left\{ \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} \right\}.$$

4.4. Let $\{\mathbf{w_1}, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ be an orthonormal basis of the subspace $W \subset V$. Prove that the vectors $(v - \mathrm{pr}_W(v))$ and $\mathbf{w}_k$ ar orthogonal $\forall v \in V$, $k = 1, \dots, m$; and hence $v - \mathrm{pr}_W(v)$ is orthogonal to every vector in W.

Solution:

4.1. All vectors $\mathbf{w}_j$ in an orthonormal basis have length one, i.e. $\langle \mathbf{w}_j, \mathbf{w}_j \rangle = 1$, so the formula simplifies to

$$\mathrm{proj}_W(\mathbf{v}) = \pi_W(\mathbf{v}) = \sum_{j=1}^{m} \langle \mathbf{v}, \mathbf{w}_j \rangle \mathbf{w}_j .$$

4.2. Clearly, $(1,1,2)^\top$ and $(1,1,-1)^\top$ are already orthogonal, so we can immediately use the formula from from definition 5.7.
Let

$$\begin{aligned} w_1 &= (1,1,2) \\ w_2 &= (1,1,-1) \\ W &= \mathrm{Span}\,(w_1, w_2) \\ v &= (4,5,-2) \end{aligned}$$

Since $w_1 \perp w_2$, the projection of v onto W is as follows.

$$\begin{aligned} \mathrm{proj}_{w_1} v &= \frac{w_1 \cdot v}{w_1 \cdot w_1} w_1 = \frac{5}{6}(1,1,2) = \left(\frac{5}{6}, \frac{5}{6}, \frac{10}{6}\right) \\ \mathrm{proj}_{w_2} v &= \frac{w_2 \cdot v}{w_2 \cdot w_2} w_2 = \frac{11}{3}(1,1,-1) = \left(\frac{11}{3}, \frac{11}{3}, -\frac{11}{3}\right) \\ \mathrm{proj}_W v &= \mathrm{proj}_{w_1} v + \mathrm{proj}_{w_2} v \\ &= \left(\frac{27}{6}, \frac{27}{6}, -\frac{6}{3}\right) = \left(\frac{9}{2}, \frac{9}{2}, -2\right) \end{aligned}$$

Hence $\left(\frac{9}{2}, \frac{9}{2}, -2\right)^\top$ is the closest vector.

4.3. Let

$$\begin{aligned} v_1 &= (2,1,1), \\ v_2 &= (0,2,1), \\ v &= (2,2,5), \\ W &= \mathrm{Span}\,(v_1, v_2). \end{aligned}$$

Since v_1 and v_2 are not orthogonal to each other, we have to find an orthogonal basis for W. The GramSchmidt orthogonalization gives us the following:

$$\begin{aligned} \|v_1\| &= \sqrt{6} \\ u_1 &= \frac{1}{\sqrt{6}}(2,1,1) \\ v_2 - \mathrm{proj}_{v_1} v_2 &= (0,2,1) - \left(1, \frac{1}{2}, \frac{1}{2}\right) = \left(-1, \frac{3}{2}, \frac{1}{2}\right) \\ \left\| \left(-1, \frac{3}{2}, \frac{1}{2}\right) \right\| &= \sqrt{1 + \frac{9}{4} + \frac{1}{4}} = \sqrt{\frac{14}{4}} = \frac{\sqrt{14}}{2} \\ u_2 &= \frac{2}{\sqrt{14}}\left(-1, \frac{3}{2}, \frac{1}{2}\right) = \left(-\frac{2}{\sqrt{14}}, \frac{3}{\sqrt{14}}, \frac{1}{\sqrt{14}}\right) \end{aligned}$$

Hence u_1, u_2 constitutes an orthonormal basis for W. Now we find the projection of v onto W :

$$\begin{aligned}
\operatorname{proj}_{u_1} v &= (u_1 \cdot v)\, u_1 \\
&= \frac{11}{6}(2,1,1) = \left(\frac{11}{3}, \frac{11}{6}, \frac{11}{6}\right) \\
\operatorname{proj}_{u_2} v &= (u_2 \cdot v)\, u_2 \\
&= 1\left(-1, \frac{3}{2}, \frac{1}{2}\right) = \left(-1, \frac{3}{2}, \frac{1}{2}\right), \\
\operatorname{proj}_W v &= \operatorname{proj}_{u_1} v + \operatorname{proj}_{u_2} v \\
&= \left(\frac{11}{3} - 1, \frac{11}{6} + \frac{3}{2}, \frac{11}{6} + \frac{1}{2}\right) \\
&= \left(\frac{8}{3}, \frac{10}{3}, \frac{7}{3}\right)
\end{aligned}$$

Therefore $\left(\frac{8}{3}, \frac{10}{3}, \frac{7}{3}\right)^\top$ is the closest vector.

4.4. This is a straightforward calculation:

$$\begin{aligned}
\langle v - \operatorname{pr}_W(v), u_k\rangle &= \langle v, u_k\rangle - \langle \mathbf{pr}_W(v), u_k\rangle \\
&= \langle v, u_k\rangle - \left\langle \sum_{j=1}^m \langle v, u_j\rangle\, u_j, u_k \right\rangle \\
&= \langle v, u_k\rangle - \sum_{j=1}^m \langle v, u_j\rangle \langle u_j, u_k\rangle \\
&= \langle v, u_k\rangle - \langle v, u_k\rangle \langle u_k, u_k\rangle = 0
\end{aligned}$$

The transition from line 3 to line 4 is justified by the orthogonality of the u's; the last step requires that the u's be unit vectors.

The second assertion of the lemma follows from the fact that every element of W can be written as a linear combination of the orthonormal basis vectors.

If you have any questions or feedback, please feel free to contact me via E-mail at hannah.kuempel@stat.uni-muenchen.de!!

Also, thank you to the authors of the books *Linear Algebra and Its Applications* & *Mathematics for Machine Learning* as well as I Seul Bee whose exercises this sheet was inspired by.