Math 2940 Worksheet
Orthogonality, Projections

Week 13
November 21st, 2019

This worksheet covers material from **Sections 6.1 - 6.3**. Please work in collaboration with your classmates to complete the following exercises - this means sharing ideas and asking each other questions.

Question 1. Let $\boldsymbol{c} = \begin{bmatrix} 4/3 \\ -1 \\ 2/3 \end{bmatrix}$ and $\boldsymbol{d} = \begin{bmatrix} 5 \\ 6 \\ -1 \end{bmatrix}$.

(a) Find a unit vector $\boldsymbol{u}$ in the direction of $\boldsymbol{c}$

(b) Show that $\boldsymbol{d}$ is orthogonal to $\boldsymbol{c}$.

(c) Is $\boldsymbol{d}$ orthogonal to $\boldsymbol{u}$? Explain.

Question 2. Let $\boldsymbol{y} = \begin{bmatrix} 7 \\ 6 \end{bmatrix}$ and $\boldsymbol{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Compute the orthogonal projection of $\boldsymbol{y}$ onto $\boldsymbol{u}$, and the component of $\boldsymbol{y}$ orthogonal to $\boldsymbol{u}$.

Question 3. Let U be an $n \times n$ matrix with orthonormal columns. Show that $\det(U) = \pm 1$.

Question 4. Let W be a subspace of $\mathbb{R}^n$. Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be vectors in $\mathbb{R}^n$.

If $\boldsymbol{u}$ is the projection of $\boldsymbol{x}$ onto W and $\boldsymbol{v}$ is the projection of $\boldsymbol{y}$ onto W, show that $\boldsymbol{u} + \boldsymbol{v}$ is the projection of $\boldsymbol{x} + \boldsymbol{y}$ onto W.

Question 5. Let W be the subspace spanned by $\boldsymbol{u}_1 = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}$ and $\boldsymbol{u}_2 = \begin{bmatrix} 5 \\ 1 \\ 4 \end{bmatrix}$.

Write $\boldsymbol{y} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$ as the sum of a vector in W and a vector orthogonal to W.

Question 6.

(a) For vectors in $\mathbb{R}^2$, prove that length squared of a vector is the sum of the squares of its coordinates, with respect to any orthonormal basis.

(b) Repeat for $\mathbb{R}^n$: Show that if $\boldsymbol{v}_1$, ..., $\boldsymbol{v}_n$ and $\boldsymbol{w}_1$, ..., $\boldsymbol{w}_n$ are two orthonormal bases, and

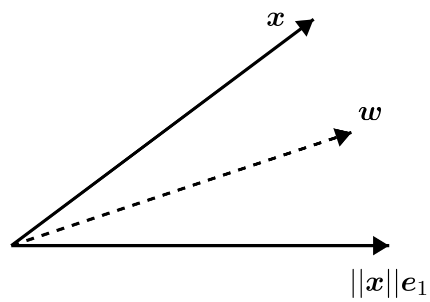
$$a_1\boldsymbol{v}_1 + \cdots + a_n\boldsymbol{v}_n = b_1\boldsymbol{w}_1 + \cdots + b_n\boldsymbol{w}_n$$

then

$$a_1^2 + \cdots + a_n^2 = b_1^2 + \cdots b_n^2$$

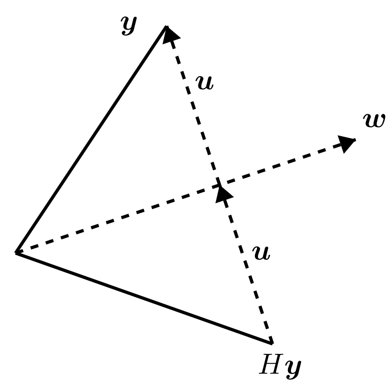
Question 7. In practice, one projection that is vastly used for computations is the Householder transformation, or often referred as the Householder reflection. For example, in scientific computing programs, Householder reflections are used to stably and efficiently calculate the QR decomposition, etc..

(a) Given a specific vector $\boldsymbol{x}$, we want to find a linear transformation that sends $\boldsymbol{x}$ to $||\boldsymbol{x}||\boldsymbol{e}_1$ so that both vectors have the same length. A Householder transformation does so by reflecting across a vector $\boldsymbol{w}$. A graph indicating the transformation on the 2D domain is presented below.



Find a formula for the vector $\boldsymbol{w}$ pictured above in terms of $\boldsymbol{x}$ and $\boldsymbol{e}_1$.

(b) Given a generic vector $\boldsymbol{y}$, find a formula for the vector $\boldsymbol{u}$ pictured below in terms of $\boldsymbol{y}$ and $\boldsymbol{w}$.



(*Hint*: $\boldsymbol{u}$ is the orthogonal complement of $\boldsymbol{y}$ projected onto $\boldsymbol{w}$)

(c) The matrix H that performs the reflection is called the Householder matrix. In other words, we have $||\boldsymbol{x}||\boldsymbol{e}_1 = H\boldsymbol{x}$. Show that the matrix H is given by

$$H = -I + 2\frac{\boldsymbol{w}\boldsymbol{w}^T}{\boldsymbol{w}^T\boldsymbol{w}}$$

(d) Verify that the matrix H has orthogonal columns.

Answer to Question 1.

(a) We first calculate the magnitude of $\boldsymbol{c}$:

$$||\boldsymbol{c}|| = \sqrt{\left(\frac{4}{3}\right)^2 + \left(\frac{-3}{3}\right)^2 + \left(\frac{2}{3}\right)^2} = \sqrt{\frac{4^2+3^2+2^2}{3^2}} = \frac{\sqrt{29}}{3}$$

Then to find a unit vector in the direction of $\boldsymbol{c}$, we take the original vector $\boldsymbol{c}$ and divide it by its magnitude:

$$\boxed{\boldsymbol{u} = \frac{\boldsymbol{c}}{||\boldsymbol{c}||} = \begin{bmatrix} 4/3 \\ -1 \\ 2/3 \end{bmatrix} \bigg/ \left(\frac{\sqrt{29}}{3}\right) = \begin{bmatrix} 4/\sqrt{29} \\ -3/\sqrt{29} \\ 2/\sqrt{29} \end{bmatrix}}$$

(b) To show that $\boldsymbol{d}$ is orthogonal to $\boldsymbol{c}$, we just take the dot product of the two vectors:

$$\boldsymbol{c} \cdot \boldsymbol{d} = \begin{bmatrix} 4/3 \\ -1 \\ 2/3 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 6 \\ -1 \end{bmatrix} = \frac{20}{3} - 6 - \frac{2}{3} = 0$$

$\boxed{\text{Since } \boldsymbol{c} \cdot \boldsymbol{d} = 0, \text{ so } \boldsymbol{d} \text{ is orthogonal to } \boldsymbol{c}.}$

Note: we could also write this dot product as the matrix muliplication $\boldsymbol{c}^T\boldsymbol{d}$.

(c) Again, we can check if the vectors are orthogonal by computing the dot product of the vectors:

$$\boldsymbol{d} \cdot \boldsymbol{u} = \boldsymbol{d} \cdot \frac{\boldsymbol{c}}{||\boldsymbol{c}||} = \frac{\boldsymbol{d} \cdot \boldsymbol{c}}{||\boldsymbol{c}||}$$

Using our answer from part (b),

$$\boxed{\boldsymbol{d} \cdot \boldsymbol{c} = 0, \text{ so } \boldsymbol{d} \text{ is orthogonal to } \boldsymbol{u}}$$

Answer to Question 2. To compute the orthogonal projection of $\boldsymbol{y}$ onto $\boldsymbol{u}$, we want to compute

$$\text{proj}_{\boldsymbol{u}}(\boldsymbol{y}) = \frac{\boldsymbol{y} \cdot \boldsymbol{u}}{||\boldsymbol{u}||^2}\boldsymbol{u}$$

Or in matrix multiplication terms

$$\text{proj}_{\boldsymbol{u}}(\boldsymbol{y}) = \frac{\boldsymbol{y}^T\boldsymbol{u}}{\boldsymbol{u}^T\boldsymbol{u}}\boldsymbol{u}$$

Either way, we can compute

$$\text{proj}_{\boldsymbol{u}}(\boldsymbol{y}) = \frac{7 \cdot 2 + 6 \cdot 1}{2^2 + 1^2} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 4 \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

$$\boxed{\text{proj}_{\boldsymbol{u}}(\boldsymbol{y}) = \begin{bmatrix} 8 \\ 4 \end{bmatrix}}$$

For the component of $\boldsymbol{y}$ orthogonal to $\boldsymbol{u}$, we just compute

$$\boxed{\boldsymbol{y} - \text{proj}_{\boldsymbol{u}}(\boldsymbol{y}) = \begin{bmatrix} 7 \\ 6 \end{bmatrix} - \begin{bmatrix} 8 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}}$$

Answer to Question 3. Let's denote the columns of U as $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, ..., $\boldsymbol{u}_n$.
Since the columns are orthonormal, that means:

$$\boldsymbol{u}_i \cdot \boldsymbol{u}_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

So if we compute $U^T U$ using the above rule, we see that the diagonal entries are 1, and the off-diagonal entries are zero, so

$$U^T U = I$$

Where I is the $n \times n$ identity matrix. Taking the determinant of both sides,

$$\det(U^T U) = \det(I)$$

Using the properties of the determinant,

$$\begin{aligned} \det(U^T)\det(U) &= 1 \\ \det(U)^2 &= 1 \end{aligned}$$

Solving for $\det(U)$,

$$\det(U) = \pm 1$$

Answer to Question 4. First, let us define a matrix A whose columns are the basis vectors of the subspace W. Then we can compute the projections as:

$$\begin{aligned} \boldsymbol{u} = \text{proj}_W(\boldsymbol{x}) = A(A^T A)^{-1} A^T \boldsymbol{x} \\ \boldsymbol{v} = \text{proj}_W(\boldsymbol{y}) = A(A^T A)^{-1} A^T \boldsymbol{y} \end{aligned}$$

Therefore

$$\text{proj}_W(\boldsymbol{x} + \boldsymbol{y}) = A(A^T A)^{-1} A^T (\boldsymbol{x} + \boldsymbol{y}) = A(A^T A)^{-1} A^T \boldsymbol{x} + A(A^T A)^{-1} A^T \boldsymbol{y}$$

And using the definitions of $\boldsymbol{u}$ and $\boldsymbol{v}$ above,

$$\boxed{\text{proj}_W(\boldsymbol{x} + \boldsymbol{y}) = \boldsymbol{u} + \boldsymbol{v}}$$

Answer to Question 5. First, we notice that $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ are orthogonal because:

$$\boldsymbol{u}_1 \cdot \boldsymbol{u}_2 = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 1 \\ 4 \end{bmatrix} = 5 + 3 - 8 = 0$$

Therefore we can project $\boldsymbol{y}$ onto W by just projecting it onto each basis vector and adding them together.
First,

$$\text{proj}_{\boldsymbol{u}_1}(\boldsymbol{y}) = \frac{\boldsymbol{y} \cdot \boldsymbol{u}_1}{||\boldsymbol{u}_1||^2} \boldsymbol{u}_1 = \frac{\begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}}{1^2 + 3^2 + (-2)^2} \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix} = 0$$

For the other component,

$$\text{proj}_{\boldsymbol{u}_2}(\boldsymbol{y}) = \frac{\boldsymbol{y}\cdot\boldsymbol{u}_2}{||\boldsymbol{u}_2||^2}\boldsymbol{u}_2 = \frac{\begin{bmatrix}1\\3\\5\end{bmatrix}\cdot\begin{bmatrix}5\\1\\4\end{bmatrix}}{5^2+1^2+4^2}\begin{bmatrix}5\\1\\4\end{bmatrix} = \frac{28}{42}\begin{bmatrix}5\\1\\4\end{bmatrix} = \begin{bmatrix}10/3\\2/3\\8/3\end{bmatrix}$$

So the orthogonal projection onto W is

$$\text{proj}_W(\boldsymbol{y}) = \text{proj}_{\boldsymbol{u}_1}(\boldsymbol{y}) + \text{proj}_{\boldsymbol{u}_2}(\boldsymbol{y}) = 0 + \begin{bmatrix}10/3\\2/3\\8/3\end{bmatrix} = \begin{bmatrix}10/3\\2/3\\8/3\end{bmatrix}$$

Fot the component orthogonal to W, we compute

$$\boldsymbol{y} - \text{proj}_W(\boldsymbol{y}) = \begin{bmatrix}1\\3\\5\end{bmatrix} = \begin{bmatrix}10/3\\2/3\\8/3\end{bmatrix} = \begin{bmatrix}-7/3\\7/3\\7/3\end{bmatrix}$$

Therefore, we have

$$\boxed{\boldsymbol{y} = \begin{bmatrix}10/3\\2/3\\8/3\end{bmatrix} + \begin{bmatrix}-7/3\\7/3\\7/3\end{bmatrix}}$$

where the first vector is in W and the second is orthogonal to W.

Answer to Question 6.

(a) Let $\boldsymbol{v}$ be a vector in $\mathbb{R}^2$ and $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ be an orthonormal basis in $\mathbb{R}^2$. Then there exist constants a_1 and a_2 such that

$$\boldsymbol{v} = a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2$$

Then we can compute the length squared as

$$\begin{aligned}||\boldsymbol{v}||^2 = \boldsymbol{v}\cdot\boldsymbol{v} &= (a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2)\cdot(a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2)\\ &= a_1\boldsymbol{v}_1\cdot(a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2) + a_2\boldsymbol{v}_2(a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2)\\ &= a_1^2\boldsymbol{v}_1\cdot\boldsymbol{v}_1 + a_1a_2\boldsymbol{v}_1\cdot\boldsymbol{v}_2 + a_2a_1\boldsymbol{v}_2\cdot\boldsymbol{v}_1 + a_2^2\boldsymbol{v}_2\cdot\boldsymbol{v}_2\\ &= a_1^2 + a_2^2\end{aligned}$$

If we repeat the same process for another orthonormal basis $\boldsymbol{w}_1$, $\boldsymbol{w}_2$ we see that

$$\begin{aligned}||\boldsymbol{v}||^2 &= (b_1\boldsymbol{w}_1 + b_2\boldsymbol{w}_2)\cdot(b_1\boldsymbol{w}_1 + b_2\boldsymbol{w}_2)\\ &= b_1^2 + b_2^2\end{aligned}$$

So we see that the length squared of the vector $\boldsymbol{v}$ is the sum of square of its coordinates in any orthonormal basis.

(b) Definining

$$\boldsymbol{v} = a_1\boldsymbol{v}_1 + \cdots a_n\boldsymbol{v}_n = b_1\boldsymbol{w}_1 + \cdots + b_n\boldsymbol{w}_n$$

And repeating part (a) for $\mathbb{R}^n$, we see that

$$\begin{aligned}||\boldsymbol{v}||^2 = \boldsymbol{v}\cdot\boldsymbol{v} &= (a_1\boldsymbol{v}_1 + \cdots + a_n\boldsymbol{v}_n)\cdot(a_1\boldsymbol{v}_1 + \cdots + a_n\boldsymbol{v}_n)\\ &= a_1\boldsymbol{v}_1\cdot(a_1\boldsymbol{v}_1 + \cdots a_n\boldsymbol{v}_n) + \cdots + a_n\boldsymbol{v}_n\cdot(a_1\boldsymbol{v}_1 + \cdots + a_n\boldsymbol{v}_n)\\ &= a_1^2 + a_2^2 + \cdots + a_n^2\end{aligned}$$

And similarly,

$$\begin{aligned}||\boldsymbol{v}||^2 = \boldsymbol{v}\cdot\boldsymbol{v} &= (b_1\boldsymbol{w}_1 + \cdots + b_n\boldsymbol{w}_n)\cdot(b_1\boldsymbol{w}_1 + \cdots + b_n\boldsymbol{w}_n)\\ &= b_1^2 + b_2^2 + \cdots + b_n^2\end{aligned}$$

Putting these together,

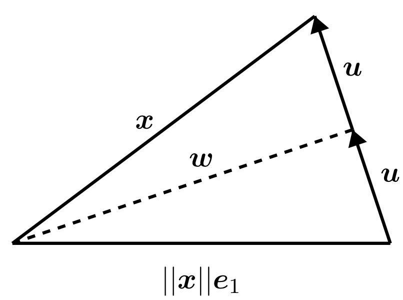
$$\boxed{a_1^2 + \cdots a_n^2 = b_1^2 + \cdots b_n^2}$$

Answer to Question 7.

(a) It turns out that we can compute $\boldsymbol{w}$ as the "average" of $\boldsymbol{x}$ and $||\boldsymbol{x}||\boldsymbol{e}_1$:

$$\boldsymbol{w} = \frac{\boldsymbol{x} + ||\boldsymbol{x}||\boldsymbol{e}_1}{2}$$

To see why, we can define a new vector $\boldsymbol{u}$ according to the figure below:



Then we see that

$$\boldsymbol{w} = ||\boldsymbol{x}||\boldsymbol{e}_1 + \boldsymbol{u}, \qquad \text{and} \qquad \boldsymbol{x} = ||\boldsymbol{x}||\boldsymbol{e}_1 + 2\boldsymbol{u}$$

Rearranging the second equation,

$$\boldsymbol{u} = \frac{\boldsymbol{x} - ||\boldsymbol{x}||\boldsymbol{e}_1}{2}$$

and plugging that into the first equation, we get

$$\boxed{\boldsymbol{w} = \frac{\boldsymbol{x} + ||\boldsymbol{x}||\boldsymbol{e}_1}{2}}$$

(b) Here we see that $\boldsymbol{u}$ is the orthogonal complement of $\boldsymbol{x}$ when projected onto $\boldsymbol{w}$. Therefore

$$\boxed{\boldsymbol{u} = \boldsymbol{y} - \text{proj}_{\boldsymbol{w}}(\boldsymbol{y}) = \boldsymbol{y} - \frac{\boldsymbol{w}\boldsymbol{w}^T}{\boldsymbol{w}^T\boldsymbol{w}}\boldsymbol{y}}$$

(c) Based on part (b), we see that

$$\begin{aligned}H\boldsymbol{y} &= \boldsymbol{y} - 2\boldsymbol{u}\\ &= \boldsymbol{y} - 2\left(\boldsymbol{y} - \frac{\boldsymbol{w}\boldsymbol{w}^T}{\boldsymbol{w}^T\boldsymbol{w}}\boldsymbol{y}\right)\\ &= -\boldsymbol{y} + 2\frac{\boldsymbol{w}\boldsymbol{w}^T}{\boldsymbol{w}^T\boldsymbol{w}}\boldsymbol{y}\\ &= \left(-I + 2\frac{\boldsymbol{w}\boldsymbol{w}^T}{\boldsymbol{w}^T\boldsymbol{w}}\right)\boldsymbol{y}\end{aligned}$$

So we see that

$$\boxed{H = -I = 2\frac{\boldsymbol{w}\boldsymbol{w}^T}{\boldsymbol{w}^T\boldsymbol{w}}}$$

(d) To check that H has orthogonal columns, we can compute $H^T H$ and check whether it is a diagonal matrix. Using our answer from part (c),

$$\begin{aligned}
H^T H &= \left(-I + \frac{2\boldsymbol{w}\boldsymbol{w}^T}{\boldsymbol{w}^T\boldsymbol{w}}\right)^T \left(-I + \frac{2\boldsymbol{w}\boldsymbol{w}^T}{\boldsymbol{w}^T\boldsymbol{w}}\right) \\
&= \left(-I^T + \frac{2\left(\boldsymbol{w}^T\right)^T \boldsymbol{w}^T}{\boldsymbol{w}^T\boldsymbol{w}}\right) \left(-I + \frac{2\boldsymbol{w}\boldsymbol{w}^T}{\boldsymbol{w}^T\boldsymbol{w}}\right) \\
&= \left(-I + \frac{2\boldsymbol{w}\boldsymbol{w}^T}{\boldsymbol{w}^T\boldsymbol{w}}\right) \left(-I + \frac{2\boldsymbol{w}\boldsymbol{w}^T}{\boldsymbol{w}^T\boldsymbol{w}}\right) \\
&= I - 2\frac{\boldsymbol{w}\boldsymbol{w}^T}{\boldsymbol{w}^T\boldsymbol{w}} - 2\frac{\boldsymbol{w}\boldsymbol{w}^T}{\boldsymbol{w}^T\boldsymbol{w}} + 4\frac{\boldsymbol{w}(\boldsymbol{w}^T\boldsymbol{w})\boldsymbol{w}^T}{(\boldsymbol{w}^T\boldsymbol{w})(\boldsymbol{w}^T\boldsymbol{w})} \\
&= I - 4\frac{\boldsymbol{w}\boldsymbol{w}^T}{\boldsymbol{w}^T\boldsymbol{w}} + 4\frac{\boldsymbol{w}\boldsymbol{w}^T}{\boldsymbol{w}^T\boldsymbol{w}} = I
\end{aligned}$$

$$\boxed{\text{Since } H^T H = I, \quad H \text{ has orthogonal columns.}}$$