Math 54: Worksheet Solutions

March 19

This worksheet is meant to provide some practice with concepts related to orthogonality. Since not all of this material will be reviewed via homework before the midterm, take extra care to make sure that you can solve these kinds of problems! Consider this as your crash course in the orthogonality skills that might show up on the midterm. (Unfortunately, I don't know what's on the midterm, so I can't guarantee that there won't be surprises! What I can say is that this is probably my best worksheet so far.) Detailed solutions are available at `https://math.berkeley.edu/~lindsey/math54`

(1) Let

$$W = \operatorname{span}\left[\begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\end{pmatrix}\right] \subset \mathbb{R}^3.$$

Find an orthogonal basis for W and an orthogonal basis for $W^{\perp}$. [**Hint**: for the second part, apply the fact that $\operatorname{Col}(A)^{\perp} = \operatorname{Null}(A^T)$ to an appropriately chosen matrix A.] Then find an orthonormal basis for each.

Solution: First let's concern ourselves with the first part of the question. Let $\mathbf{v}_1 = \begin{pmatrix}1\\1\\0\end{pmatrix}$ and $\mathbf{v}_2 = \begin{pmatrix}1\\0\\1\end{pmatrix}$, so $W = \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2)$. Evidently $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent, so they form a basis for W. Converting a basis for a vector space into an *orthogonal* basis for the vector space is exactly what Gram-Schmidt is for.

We carry this out as follows. Let $\mathbf{u}_1 = \mathbf{v}_1 = \begin{pmatrix}1\\1\\0\end{pmatrix}$. That's it for the first step.

Since W is only 2-dimensional, there is only one more step to go. Let

$$\begin{aligned}
\mathbf{u}_2 &= \mathbf{v}_2 - \operatorname{Proj}_{\operatorname{span}(\mathbf{u}_1)}(\mathbf{v}_2) = \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 \\
&= \begin{pmatrix}1\\0\\1\end{pmatrix} - \frac{\begin{pmatrix}1\\0\\1\end{pmatrix} \cdot \begin{pmatrix}1\\1\\0\end{pmatrix}}{\begin{pmatrix}1\\1\\0\end{pmatrix} \cdot \begin{pmatrix}1\\1\\0\end{pmatrix}} \begin{pmatrix}1\\1\\0\end{pmatrix} \\
&= \begin{pmatrix}1\\0\\1\end{pmatrix} - \frac{1}{2}\begin{pmatrix}1\\1\\0\end{pmatrix} \\
&= \begin{pmatrix}\frac{1}{2}\\-\frac{1}{2}\\1\end{pmatrix}.
\end{aligned}$$

Therefore $\{\mathbf{u}_1, \mathbf{u}_2\} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix} \right\}$ is an orthogonal basis for W. To make this into an orthonormal basis we take each basis element and divide by its length, so

$$\left\{ \frac{\mathbf{u}_1}{\sqrt{\mathbf{u}_1 \cdot \mathbf{u}_1}}, \frac{\mathbf{u}_2}{\sqrt{\mathbf{u}_2 \cdot \mathbf{u}_2}} \right\} = \left\{ \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{2}\sqrt{\frac{2}{3}} \\ -\frac{1}{2}\sqrt{\frac{2}{3}} \\ \sqrt{\frac{2}{3}} \end{pmatrix} \right\}$$

is an orthonormal basis for W. (One could simplify these radicals differently / better.)

In general, if you want to turn a basis $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ for some vector space V (such as a subspace of $\mathbb{R}^n$) into an *orthogonal* basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ for V, the Gram-Schmidt process looks like this. The first step is to let $\mathbf{u}_1 = \mathbf{v}_1$. Then afterwards, suppose you are at the k-th step, so you have already found $\mathbf{u}_1, \dots, \mathbf{u}_{k-1}$, and you want to find $\mathbf{u}_k$. Then you define

$$\mathbf{u}_k = \mathbf{v}_k - \mathrm{Proj}_{\mathrm{span}(\mathbf{u}_1, \dots, \mathbf{u}_{k-1})}(\mathbf{v}_k) = \mathbf{v}_k - \frac{\mathbf{v}_k \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 - \cdots - \frac{\mathbf{v}_k \cdot \mathbf{u}_{k-1}}{\mathbf{u}_{k-1} \cdot \mathbf{u}_{k-1}}\mathbf{u}_{k-1}.$$

Back to the problem. When finding a basis for $W^\perp$, we usually want to make use of the general fact that $\mathrm{Col}(A)^\perp = \mathrm{Null}(A^T)$.

In this setting, let $A = (\mathbf{v}_1 \ \mathbf{v}_2)$. This means that $W = \mathrm{Col}(A)$, so $W^\perp = \mathrm{Null}(A^T)$, so we want to find a basis for $\mathrm{Null}(A^T)$. Compute A^T:

$$A^T = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}^T = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

(When you transpose a matrix A, you take the first column of A and make it the first row, take the second column of A and make it the second row, etc.)

Then row-reduce:

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}.$$

Then x_3 is free and a general solution is given by

$$\begin{pmatrix} -x_3 \\ x_3 \\ x_3 \end{pmatrix} = x_3 \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix},$$

so $W^\perp = \mathrm{Null}(A^T) = \mathrm{span}\left[\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \right]$, and $\left\{ \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \right\}$ is a basis for $W^\perp$. Since the basis has only one element, it is an orthgonal basis. (If there were more elements in the basis, to find an orthogonal basis we might have to apply Gram-Schmidt!) To find an orthonormal basis, we have to normalize, yielding $\left\{ \begin{pmatrix} -1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix} \right\}$.

(2) With W as in question 1, Compute $\mathrm{Proj}_W \begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix}$ and $\mathrm{Proj}_{W^\perp} \begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix}$.

Solution: For this kind of question, it is often useful to remember the general fact that for any $\mathbf{v}$,

$$\mathbf{v} = \mathrm{Proj}_W(\mathbf{v}) + \mathrm{Proj}_{W^\perp}(\mathbf{v}).$$

Applied to this case, we have that

$$\begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix} = \mathrm{Proj}_W \begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix} + \mathrm{Proj}_{W^\perp} \begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix}. \quad (\star)$$

What this means is that we really only have to compute one of the things we were asked to compute, and then we can recover the other one via this relation.

Now we have an orthogonal basis $\{\mathbf{u}_1, \mathbf{u}_2\}$ for W from the last problem. This means that we can apply the formula

$$\mathrm{Proj}_W(\mathbf{v}) = \frac{\mathbf{v}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{v}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2$$

where $\mathbf{v} = \begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix}$ to reveal

$$\begin{aligned}
\mathrm{Proj}_W \begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix} &= \frac{\begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix}\cdot\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}}{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}\cdot\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \frac{\begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix}\cdot\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix}}{\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix}\cdot\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix}}\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix} \\
&= \frac{5}{2}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \frac{\frac{9}{2}}{\frac{3}{2}}\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} \frac{5}{2} \\ \frac{5}{2} \\ 0 \end{pmatrix} + \begin{pmatrix} \frac{3}{2} \\ -\frac{3}{2} \\ 3 \end{pmatrix} \\
&= \begin{pmatrix} 4 \\ 1 \\ 3 \end{pmatrix}.
\end{aligned}$$

Therefore by the relation $(\star)$ above:

$$\mathrm{Proj}_{W^\perp} \begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix} - \mathrm{Proj}_W \begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix} - \begin{pmatrix} 4 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}.$$

In the last problem we actually computed an orthogonal basis for $W^\perp$, consisting of one element $\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$, so we could have used computed the projection onto $W^\perp$ directly via

$$\mathrm{Proj}_{W^\perp} \begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix} = \frac{\begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix}\cdot\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}}{\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}\cdot\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}}\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} = \frac{3}{3}\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix},$$

and then we could have determined $\mathrm{Proj}_W \begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix}$ via $(\star)$! Note how much easier the second way was in this example given that we already knew an orthogonal basis for $W^\perp$.

(3) With W as in question 1, what is the minimum distance between $\begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix}$ and W?

Solution: This minimum distance from $\mathbf{v}$ to W is the length $\|\mathbf{v} - \mathrm{Proj}_W(\mathbf{v})\|$. Thus we want to compute in this case:

$$\begin{aligned} \left\| \begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix} - \mathrm{Proj}_W \begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix} \right\| &= \left\| \begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix} - \begin{pmatrix} 4 \\ 1 \\ 3 \end{pmatrix} \right\| \\ &= \left\| \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \right\| \\ &= \sqrt{1+1+1} \\ &= \sqrt{3}. \end{aligned}$$

(4) Let $A = \begin{pmatrix} 1 & 2 & 1 \\ 3 & 2 & 3 \\ 2 & 4 & 2 \\ 1 & 1 & 1 \end{pmatrix}$. Find a basis for $\mathrm{Null}(A)^\perp$. [**Hint**: use the fact that $\mathrm{Null}(A)^\perp = \mathrm{Col}(A^T)$.]

Solution: Remember $\mathrm{Null}(A)^\perp = \mathrm{Col}(A^T)$, so we need to find a basis for $\mathrm{Col}(A^T)$.

$$A^T = \begin{pmatrix} 1 & 3 & 2 & 1 \\ 2 & 2 & 4 & 1 \\ 1 & 3 & 2 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 3 & 2 & 1 \\ 0 & -4 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The last matrix is already in row echelon form, and we can see that the first and second columns are pivot columns. Therefore the first and second columns of the **original matrix** A^T (NOT THE FIRST AND SECOND COLUMNS OF THE REDUCED ECHELON FORM) form a basis for $\mathrm{Col}(A^T)$. Thus a suitable basis is

$$\left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 3 \end{pmatrix} \right\}.$$

(5) Let $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3 \in \mathbb{R}^3$. Let $A = (\mathbf{a}_1\ \mathbf{a}_2\ \mathbf{a}_3)$ and $B = (\mathbf{b}_1\ \mathbf{b}_2\ \mathbf{b}_3)$. Convince yourself that

$$A^T B = \begin{pmatrix} \mathbf{a}_1 \cdot \mathbf{b}_1 & \mathbf{a}_1 \cdot \mathbf{b}_2 & \mathbf{a}_1 \cdot \mathbf{b}_3 \\ \mathbf{a}_2 \cdot \mathbf{b}_1 & \mathbf{a}_2 \cdot \mathbf{b}_2 & \mathbf{a}_2 \cdot \mathbf{b}_3 \\ \mathbf{a}_3 \cdot \mathbf{b}_1 & \mathbf{a}_3 \cdot \mathbf{b}_2 & \mathbf{a}_3 \cdot \mathbf{b}_3 \end{pmatrix}.$$

In particular,

$$A^T A = \begin{pmatrix} \mathbf{a}_1 \cdot \mathbf{a}_1 & \mathbf{a}_1 \cdot \mathbf{a}_2 & \mathbf{a}_1 \cdot \mathbf{a}_3 \\ \mathbf{a}_2 \cdot \mathbf{a}_1 & \mathbf{a}_2 \cdot \mathbf{a}_2 & \mathbf{a}_2 \cdot \mathbf{a}_3 \\ \mathbf{a}_3 \cdot \mathbf{a}_1 & \mathbf{a}_3 \cdot \mathbf{a}_2 & \mathbf{a}_3 \cdot \mathbf{a}_3 \end{pmatrix}.$$

Use this to explain why a matrix is orthogonal if and only if the columns of the matrix form an orthonormal set.

By the way, notice that $A^T A$ is always **symmetric**! (*Recall*: a matrix M is by definition symmetric if $M = M^T$.) What is an easier way to prove that $A^T A$ is symmetric?

Solution: By the rules of matrix multiplication, the (i, j) entry (row i, column j) of $A^T B$ should be the dot product of the i-th row of A^T with the j-th column of B (namely $\mathbf{b}_j$).

But the i-th row of A^T is the i-th column of A, namely $\mathbf{a}_i$. Therefore the (i,j) entry of A^T is $\mathbf{a}_i \cdot \mathbf{b}_j$, which is what we were trying to explain.

Remember that a matrix is orthogonal if and only if $A^TA = I$, which is true if and only if the (i,j) entry of A^TA is 0 for $i \neq j$ and 1 for $i = j$. But the (i,j) entry of A^TA is precisely $\mathbf{a}_i \cdot \mathbf{a}_j$, so this last statement is equivalent to saying that $\mathbf{a}_i \cdot \mathbf{a}_j$ are orthogonal when $i \neq j$ and $\mathbf{a}_i \cdot \mathbf{a}_i = 1$ for all i, which is the equivalent to saying that $\mathbf{a}_1, \ldots, \mathbf{a}_n$ form an orthonormal set.

To see that A^TA is symmetric, one can recall the general fact that for matrices X and Y that can be multiplied, $(XY)^T = Y^TX^T$. Therefore $(A^TA)^T = A^T(A^T)^T = A^TA$, so A^TA is symmetric.

(6) Let $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3 \in \mathbb{R}^3$. Let $A = (\mathbf{a}_1\ \mathbf{a}_2\ \mathbf{a}_3)$. Use the result of the last problem to write $\|\mathbf{a}_1 + \mathbf{a}_3\|$ in terms of the entries of A^TA. In other words, suppose

$$A^TA = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{12} & c_{22} & c_{23} \\ c_{13} & c_{23} & c_{33} \end{pmatrix},$$

and come up with a formula for $\|\mathbf{a}_1 + \mathbf{a}_3\|$ in terms of the c_{ij}.

Solution: Compute

$$\begin{aligned} \|\mathbf{a}_1 + \mathbf{a}_3\| &= \sqrt{(\mathbf{a}_1 + \mathbf{a}_3) \cdot (\mathbf{a}_1 + \mathbf{a}_3)} \\ &= \sqrt{(\mathbf{a}_1 + \mathbf{a}_3) \cdot \mathbf{a}_1 + (\mathbf{a}_1 + \mathbf{a}_3) \cdot \mathbf{a}_3} \\ &= \sqrt{\mathbf{a}_1 \cdot \mathbf{a}_1 + \mathbf{a}_3 \cdot \mathbf{a}_1 + \mathbf{a}_1 \cdot \mathbf{a}_3 + \mathbf{a}_3 \cdot \mathbf{a}_3} \\ &= \sqrt{c_{11} + c_{31} + c_{13} + c_{33}}. \end{aligned}$$

(Actually, since A^TA is symmetric, we must have $c_{31} = c_{13}$, so your answer could also be $\sqrt{c_{11} + 2c_{13} + c_{33}}$ or $\sqrt{c_{11} + 2c_{31} + c_{33}}$.

(7) Recall the fact that $\mathrm{Col}(A)^\perp = \mathrm{Null}(A^T)$ for any matrix A. (Try to prove this for a challenge!) Why does this imply the fact that $\mathrm{Null}(A)^\perp = \mathrm{Col}(A^T)$ for any matrix A?

Solution: Apply the fact to the matrix A^T to reveal that $\mathrm{Col}(A^T)^\perp = \mathrm{Null}\left((A^T)^T\right) = \mathrm{Null}(A)$. Take the orthogonal complement of both sides and remember that taking the orthogonal complement twice gives you back the original subspace, yielding $\mathrm{Col}(A^T) = \mathrm{Null}(A)^\perp$.

(8) **Bonus** (for zero points). Suppose that A is a symmetric matrix, and suppose that $\mathbf{v}_1$ and $\mathbf{v}_2$ are eigenvectors of A with corresponding eigenvalues λ_1 and λ_2. Moreover, suppose that $\lambda_1 \neq \lambda_2$. Show that $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthogonal.

Solution: We know that $A\mathbf{v}_1 = \lambda_1\mathbf{v}_1$ and $A\mathbf{v}_2 = \lambda_2\mathbf{v}_2$. We also know that $A = A^T$. Using these facts, compute:

$$\lambda_1(\mathbf{v}_1 \cdot \mathbf{v}_2) = (\lambda_1\mathbf{v}_1)^T\mathbf{v}_2 = (A\mathbf{v}_1)^T\mathbf{v}_2 = \mathbf{v}_1^TA^T\mathbf{v}_2 = \mathbf{v}_1^TA\mathbf{v}_2 = \mathbf{v}_1^T(\lambda_2\mathbf{v}_2) = \lambda_2(\mathbf{v}_1 \cdot \mathbf{v}_2).$$

In summary, $\lambda_1(\mathbf{v}_1 \cdot \mathbf{v}_2) = \lambda_2(\mathbf{v}_1 \cdot \mathbf{v}_2)$, which implies that $(\lambda_1 - \lambda_2)(\mathbf{v}_1 \cdot \mathbf{v}_2) = 0$. But $\lambda_1 \neq \lambda_2$, so $\lambda_1 - \lambda_2 \neq 0$, and we can divide both sides of the last equation by $\lambda_1 - \lambda_2$ to obtain $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$, so $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthogonal.