Jared Andreatta

MATH 531 **Homework 1** **Spring 2025**
(Due: 19th January, Tuesday)

1. Let $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ be a set of k orthogonal vectors in $\mathbb{R}^n$. Show that this set is linearly independent when $k \leq n$. Also explain why this problem does not make sense when $k > n$.

Proof. We will show that the set $\{x_k\}$ is linearly independent by contradiction. So assume that $\{x_k\}$ is linearly dependent, that is, we can write any vector $x_a \in \{x_k\}$ can be written as a linear combination of all of the other vectors in the set.

$$x_a = \sum_{i=1}^{k} \alpha_i x_i$$

We also know that all of the vectors in the set are orthogonal to each other, that is, for any vector $x_b \in \{x_k\}$, we have

$$\langle x_a, x_b \rangle = 0$$

when $b \neq a$. We can expand this inner product by using our definition of x_a.

$$\langle \sum_{i=1}^{k} \alpha_i x_i, x_b \rangle = 0$$

Since the vectors in the set are orthogonal, the inner product is only nonzero when $a = b$. Thus, we can simplify the expression to

$$\begin{aligned} \langle \alpha_b x_b, x_b \rangle &= 0 \\ a_b \langle x_b, x_b \rangle &= 0 \end{aligned}$$

If x_b is not the nontrivial zero-vector, then the only value of a_b that satisfies the equation is $a_b = 0$, which entails that $a_b = 0$ for $b = 1, ..., k$. This contradicts the original assumption that the set of vectors $\{x_k\}$ is linearly dependent. Therefore, we can conclude that $\{x_k\}$ is linearly independent. $\square$

In regards to the second part of the question, it would not hold that the set $\{x_k\}$ is linearly independent. That's because, in the space $\mathbb{R}^n$, there can only be n independent vectors. If $k > n$, then it is inevitable that at least one of the vectors in the space can be written as a linear combination of the other vectors in the set.

2. Suppose $\mathcal{V}_n \in \mathbb{R}^n$ is a vector space. Prove the following results.

   (a) $\mathbf{0} \in \mathcal{V}_n$.
   Proof. Any vector space is closed under scalar multiplication. This means for any vector $x \in \mathcal{V}_n$, we have that $\alpha x \in \mathcal{V}_n$ for any $\alpha \in \mathbb{R}$. Let $\alpha = 0$. We have $\alpha x = 0 \cdot x = 0$. Since $\alpha x \in \mathcal{V}_n$, then $0 \in \mathcal{V}_n$. □

   (b) If $\boldsymbol{x} \in \mathcal{V}_n$ and $\boldsymbol{x} \perp \mathcal{V}_n$ then $\boldsymbol{x} = \mathbf{0}$.
   Proof. Let $y \in \mathcal{V}_n$ be any vector in $\mathcal{V}_n$. If $x \perp \mathcal{V}_n$, then that means that $x^T y = 0$ for every y. Given that $x \in \mathcal{V}_n$, for every nonzero y, this implies that x is the zero vector. Hence, $x = 0$. □

   (c) $\mathcal{V}_n^{\perp} = \{\boldsymbol{x} : \boldsymbol{x} \perp \mathcal{V}_n\}$ is a vector space.
   Proof. We can redefine $\mathcal{V}_n^{\perp}$ using the notion of an inner product as $\mathcal{V}_n^{\perp} = \{x : \langle x, y\rangle = 0 \text{ for every } y \in \mathcal{V}_n\}$. We must now verify that $\mathcal{V}_n^{\perp}$ is closed under addition and scal multiplication, and that $0 \in \mathcal{V}_n^{\perp}$.

   First, we check closure under addition. Let $x_1, x_2 \in \mathcal{V}_n^{\perp}$. We have that $\langle x_1 + x_2, y\rangle = \langle x_1, y\rangle + \langle x_2, y\rangle = 0 + 0 = 0$. Therefore, $x_1 + x_2 \in \mathcal{V}_n^{\perp}$.

   Next, we check that it is closed under scalar multiplication. Let $\alpha \in \mathbb{R}$ and $x \in \mathcal{V}_n^{\perp}$. We have $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle = \alpha \cdot 0 = 0$. Therefore, $\alpha x \in \mathcal{V}_n^{\perp}$.

   Lastly, we verify that $0 \in \mathcal{V}_n^{\perp}$. Let x be any vector in $\mathcal{V}_n$. We take the dot product $\langle 0, x\rangle = 0 \cdot x = 0$. Clearly, these are orthogonal, which entails that $0 \in \mathcal{V}_n^{\perp}$.

   We have proven that $\mathcal{V}_n^{\perp}$ is a subspace of $\mathbb{R}^n$. Therefore, we can conclude that $\mathcal{V}_n^{\perp}$ is indeed a vector space. □

3. Let $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}$ be a basis of a vector space $\mathcal{W}$. Then show that $\boldsymbol{y} \in \mathcal{W}^{\perp}$ if and only if (aka iff) $\boldsymbol{y} \perp \boldsymbol{x}_i,\ i = 1, 2, \ldots, k$.

   Proof. First, we define $\mathcal{W}^{\perp}$, which is the orthogonal complement to $\mathcal{W}$, as
   $\mathcal{W}^{\perp} = \{y : y^T x = 0 \text{ for every } x \in \mathcal{W}\}$. Now, assume that $y \in \mathcal{W}^{\perp}$. Since the set $\{x_1, ..., x_k\}$ forms a basis for $\mathcal{W}$, then we know that $x_i \in \mathcal{W}$ for $i = 1, ..., k$. Since $y \in \mathcal{W}^{\perp}$, we know that it is orthogonal to every vector in $\mathcal{W}$, so by definition, $y^T x_i = 0$ for $i = 1, ..., k$. Thus, $y \perp x_i$ for $i = 1, ..., k$.

   Next, assume $y \perp x_i$ for $i = 1, ..., k$. We note that y is orthogonal to all of the basis vectors, which implies that y is orthogonal to the space spanned by $\{x_1, ..., x_k\}$. This means that, by definition, y is in the orthogonal complement of $\mathcal{W}$, or $y \in \mathcal{W}^{\perp}$. Thus, we conclude that $y \in \mathcal{W}^{\perp}$ iff $y \perp x_i$ for $i = 1, ..., k$. □

4. Recall the Cauchy-Schwartz inequality for two vectors , $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$

$$|\mathbf{x}^T\mathbf{y}| \leq \|\mathbf{x}\|\|\mathbf{y}\|$$

use this inequality to show that the sample correlation coefficient between two data vectors is always in the range $[-1, 1]$

Proof. Let $\tilde{x}$ and $\tilde{y}$ be the mean-centered vectors of their corresponding data vector, that is, $\tilde{x} = (x_1 - \bar{x}, ..., x_n - \bar{x}$ and $\tilde{y} = (y_1 - \bar{y}, ..., y_n - \bar{y})$ where $\bar{x}$ and $\bar{y}$ are their respective means. The sample correlation coefficient r can be written as such:

$$r = \frac{\tilde{x}^T\tilde{y}}{\|\tilde{x}\|\|\tilde{y}\|}$$

We can apply the Cauchy-Schwarz inequality to the numerator term.

$$|\tilde{x}^T\tilde{y}| \leq \|\tilde{x}\|\|\tilde{y}\|.$$

We divide both sides by $\|\tilde{x}\|\|\tilde{y}\|$, so as to match the correlation coefficient. Note that the denominator term can be written as an absolute value since $\|\cdot\| \geq 0$ if the vector inside the norm is nonzero.

$$\left|\frac{\tilde{x}^T\tilde{y}}{\|\tilde{x}\|\|\tilde{y}\|}\right| = \frac{\|\tilde{x}\|\|\tilde{y}\|}{\|\tilde{x}\|\|\tilde{y}\|} \to |r| \leq 1$$

Since the LHS term is an absolute value, this implies that $r \in [-1, 1]$. $\square$

5. Recall the projection of a vector $\mathbf{y}$ onto $\mathbf{x}$ is given by $\hat{\mathbf{y}} = \beta\mathbf{x}$ with $\beta = (\mathbf{x}^T\mathbf{y})/(\mathbf{x}^T x)$ Using the fact that $\|\mathbf{y} - \hat{\mathbf{y}}\| \geq 0$ for all $\mathbf{x}$ and $\mathbf{y}$ prove the Cauchy-Schwartz inequality

Proof. We have that $\hat{y} = \beta x$ with $\beta = \frac{x^T y}{x^T x}$, and $\|y - \hat{y}\| \geq 0$. First, we square both sides of the inequality. This operation yields $\|y - \hat{y}\|^2 \geq 0$. We can expand the LHS of the inequality.

$$\begin{aligned} \|y - \hat{y}\|^2 &= \langle y - \hat{y}, y - \hat{y}\rangle \\ &= (y - \hat{y})(y - \hat{y}) \\ &= \|y\|^2 - 2\beta x^T y + \|\beta x\|^2 \end{aligned}$$

Next, we substitute for β.

$$\begin{aligned} \|y\|^2 - 2\beta x^T y + \|\beta x\|^2 &= \|y\|^2 - 2\frac{x^T y}{\|x\|^2}(x^T y) + \frac{(x^T y)^2}{(\|x\|^2)^2}\|x\|^2 \\ &= \|y\|^2 - \frac{(x^T y)^2}{\|x\|^2} \end{aligned}$$

Now, we go back to the original inequality and rearrange it.

$$
\begin{aligned}
\|y\|^2 - \frac{(x^T y)^2}{\|x\|^2} &\geq 0 \\
\|y\|^2 &\geq \frac{(x^T y)^2}{\|x\|^2} \\
\|y\|^2 \|x\|^2 &\geq (x^T y)^2
\end{aligned}
$$

We take the square root of both sides then we flip the inequality to find

$$
\begin{aligned}
\sqrt{\|y\|^2 \|x\|^2} &\geq \sqrt{(x^T y)^2} \\
|x^T y| &\leq \|x\| \|y\|
\end{aligned}
$$

Thus, we have proven the Cauchy-Schwarz inequality. □

6. Extra Credit Who were Cauchy and Schwarz in the equality given in the previous problem? If you could go to dinner with either one who would you choose?

Augustin-Louis Cauchy was a French mathematician born in the 18th century. Cauchy was a pioneer in the field of mathematical analysis and has several mathematical tools/concepts named after him, such as Cauchy sequences and Cauchy-Schwarz inequality. He is also well known for his contributions to complex analysis and physics. Hermann Schwarz was a German mathematician that was born later in the 19th century. He was best known for his pioneering contributions to the field of complex analysis. If I had to choose, I would have dinner with Cauchy. I think it'd be interesting to pick the brain of a pioneer of a field that fascinates, but also troubles, so many mathematics students to this day.