Jared Andreatta

MATH 531 **Homework 2** **Spring 2025**

For background, assume that a set of linearly independent vectors $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ form a vector subspace of $\mathbb{R}^n$. Call this $\mathcal{W}_X$.

Using Gram-Schmidt, we can find $\{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_k\}$, a set of k orthonormal vectors in $\mathbb{R}^n$ that also spans $\mathcal{W}_X$. This means:
1. For all $\mathbf{y} \in \mathcal{W}_X$, there exist $\alpha_1, \ldots, \alpha_k$ such that

$$\mathbf{y} = \sum_i \mathbf{z}_i \alpha_i.$$

2. For all $\alpha \in \mathbb{R}^k$, if

$$\mathbf{y} = \sum_i \mathbf{z}_i \alpha_i,$$

then $\mathbf{y} \in \mathcal{W}_X$.

1. Let Z be a matrix formed by taking these orthonormal vectors as columns, so

$$Z = [\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_k].$$

(Note: in `R`, this would be the `cbind` function to combine the vectors into a matrix.)

Show that

$$Z^T Z = I_m,$$

where I_m is the $m \times m$ identity matrix.

What is the dimension of this identity matrix?

Proof. We have $Z = [z_1, \ldots, z_k]$ with the set $\{z_k\}$ being the column vectors of Z. So we use this to compute the product $Z^T Z$.

$$Z^T Z = [z_1^T, z_2^T, \ldots, z_k^T][z_1, z_2, \ldots, z_k] = \begin{bmatrix} z_1^T z_1 & \cdots & z_1^T z_k \\ \vdots & \ddots & \vdots \\ z_k^T z_1 & \cdots & z_k^T z_k \end{bmatrix}$$

We know that these vectors are orthogonal, so $z_i^T z_j = 0$ when $i \neq j$. We can simplify the matrix.

$$\begin{bmatrix} z_1^T z_1 & \cdots & z_1^T z_k \\ \vdots & \ddots & \vdots \\ z_k^T z_1 & \cdots & z_k^T z_k \end{bmatrix} = \begin{bmatrix} z_1^T z_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & z_k^T z_k \end{bmatrix}$$

Note that the only nonzero entries are on the diagonal of the matrix. Also, these vectors are normal, so we know that $\|z_i\| = 1$ for $i = 1, \ldots, k$. Given that $z_i^T z_i = \|z_i\|^2$, we can simplify further.

$$\begin{bmatrix} \|z_1\|^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \|z_k\|^2 \end{bmatrix} = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} = I_m$$

Hence, $Z^T Z = I_m$. $\square$

For this particular problem, we note that $m = 1$, meaning that the identity matrix is I_k, or it has dimension $k \times k$. This is because $Z^T \in \mathbb{R}^{k \times n}$ and $Z \in \mathbb{R}^{n \times k}$.

2. Show that the trace of the matrix ZZ^T is k. Under what circumstances will $ZZ^T = I_n$?

Proof. From the previous question, we know that $Z^T Z = I_k$. We can use the property that $\text{trace}(AB) = \text{trace}(BA)$. Therefore, we have

$$\text{trace}(ZZ^T) = \text{trace}(Z^T Z) = \text{trace}(I_k) = \sum_{j=1}^{k} (I_k)_{jj}.$$

We know that the diagonal elements of the identity matrix are 1, so we can simplify.

$$\sum_{j=1}^{k} (I_k)_{jj} = (I_k)_{11} + (I_k)_{22} + \ldots + (I_k)_{kk} = k(1) = k \quad \square$$

To address the next part of the question, $ZZ^T = I_n$ if $k = n$ and the set of vectors $\{z_1, \ldots, z_n\}$ forms an orthonormal basis for $\mathbb{R}^n$.

3. Consider the decomposition of any $\mathbf{y} \in \mathbb{R}^n$ and for the subspace $\mathcal{W}_X$ as

$$\mathbf{y} = \mathbf{u} + \mathbf{v},$$

where $\mathbf{u} \in \mathcal{W}_X$ and $\mathbf{v} \in \mathcal{W}_X^{\perp}$.

Show that if $\mathbf{u} = ZZ^T\mathbf{y}$ and $\mathbf{v} = \mathbf{y} - \mathbf{u}$, then this decomposition holds.

Proof. First, we examine to see if $u = ZZ^T y \in \mathcal{W}_X$. We know that the column vectors of Z form an orthonormal basis for $\mathcal{W}_X$. Since this is true, we can say that any column vector of Z, denoted by $z_i \in \mathbb{R}^n$, can be written as a linear combination of the product $ZZ^T y$ for some $y \in \mathbb{R}^n$. This means that we can say that $u = ZZ^T y \in \mathcal{W}_X$ since it is in the space spanned by $\{z_k\}$.

Next, we check that $v \in \mathcal{W}_X^{\perp}$. We can easily do this by taking the inner product $\langle v, u \rangle$. By definition, these two vectors are supposedly orthogonal, which we can verify by seeing if the inner product is equal to zero. We can substitute for u and v using the definition from above.

$$\langle v, u \rangle = \langle y + (-ZZ^T y), ZZ^T y \rangle = 0$$

We can use linearity in the first argument for inner products to expand this inner product term.

$$\begin{aligned}\langle y + (-ZZ^Ty), ZZ^Ty\rangle &= \langle y, ZZ^Ty\rangle - \langle ZZ^Ty, ZZ^Ty\rangle \\ &= y^TZZ^Ty - (ZZ^Ty)^TZZ^TY \\ &= y^TZZ^Ty - y^TZ(Z^TZ)Z^Ty\end{aligned}$$

Recall from above that we showed $Z^TZ = I_m$. Therefore,

$$y^TZZ^Ty - y^TZ(Z^TZ)Z^Ty = y^TZZ^Ty - y^TZZ^Ty = 0$$

Thus, for every $y \in \mathbb{R}^n$, v is orthogonal to the space spanned by the column space of Z, so $v \in \mathcal{W}^\perp$. With this, we can conclude that $y = u + v$ with $u = ZZ^Ty$ and $v = y - u$ is indeed an orthogonal decomposition of y in $\mathbb{R}^n$. $\square$

4. Now suppose that there are two other vectors $\mathbf{u}'$ and $\mathbf{v}'$ that also satisfy this decomposition for a given $\mathbf{y}$ and $\mathcal{W}_X$.

   Show that $\mathbf{u} = \mathbf{u}'$ and $\mathbf{v} = \mathbf{v}'$.

   Proof. Suppose that there exists $u' \in \mathcal{W}_X$ and $v' \in \mathcal{W}_X^\perp$ that satisfies the decomposition $y = u' + v' = u + v$. We can rearrange this equality into $(u - u') = (v' - v)$. We can rearrange the RHS to find $(u - u') = -(v - v')$. We can denote $w = u - u'$ and $w = -(v - v')$.

   We know that $u, u' \in \mathcal{W}_X$, and since $\mathcal{W}_X$ is a linear subspace, we know that it is closed under addition. Therefore, the sum $w = u + (-1)u' \in \mathcal{W}_X$. The same can be said for $w = -(v - v')$, since it is the orthogonal complement to W_X. Thus, $w \in \mathcal{W}_X^\perp$.

   We now know that w belongs to both subspaces, but we also know that, since they are orthogonal complements, the intersection of these sets is $\mathcal{W}_X \cap \mathcal{W}_X^\perp = \{0\}$. Therefore, $w = 0$, since it is the only vector that is orthogonal to either subspace. Therefore, we can substitute that $0 = u - u'$ and $0 = -(v - v')$, or $u = u'$ and $v = v'$. $\square$

Note: We have defined $ZZ^T\mathbf{y}$ as the *projection* of $\mathbf{y}$ on $\mathcal{W}_X$. The problems above show that the projection vector is *unique* for fixed $\mathbf{y}$ and $\mathcal{W}_X$. This eliminates the concern that a different choice for the orthonormal vectors may give a different projection. Also note that the uniqueness only depends on the *subspace* not on the vectors that generate it. So if $\mathcal{W}$ is the same when using a different set of column vectors, then the projection will be the same. This is an abstract way of describing the practical result that if you reparametrize a linear model, the least squares predictions will not change!