Solutions to Assignment 10

Math 217, Fall 2002

5.4.18 Define $T : \mathbb{R}^3 \to \mathbb{R}^3$ by $T(\mathbf{x}) = A\mathbf{x}$ where A is a 3×3 matrix with eigenvalues 5 and -2. Does there exist a basis $\mathcal{B}$ for $\mathbb{R}^3$ such that the $\mathcal{B}$-matrix for T is a diagonal matrix?

We know that if C is the matrix giving the $\mathcal{B}$-matrix for T, then A is similar to C. So this question can be restated as follows: is A similar to a diagonal matrix? We know that this can happen if and only if A has n linearly independent eigenvectors (theorem 5, page 320). We also know that eigenvectors corresponding to different eigenvalues are linearly independent (theorem 2, pg 307). The characteristic equation of this matrix has degree 3, so one of the two given eigenvalues must occur with multiplicity two (an aside, complex roots come in pairs, so it can't be the case that the remaining root of the characteristic equation is complex). We conclude that A will be diagonalizable if and only if the eigenspace corresponding to the eigenvalue with multiplicity two has dimension two.

5.4.20 Show that if A is similar to B, then A^2 is similar to B^2 (where A and B are square).

Because A is similar to B, there is a P such that $A = PBP^{-1}$. Squaring both sides of this equation, we find that $A^2 = (PBP^{-1})^2 = (PBP^{-1})(PBP^{-1}) = PBP^{-1}PBP^{-1} = PBBP^{-1} = PB^2P^{-1}$, so A^2 is similar to B^2.

5.4.24 Show that if A and B are similar, then they have the same rank.

The proof is not difficult, but the chain of reasoning one has to follow is somewhat long. We will need to prove the following lemma:

Lemma 1. *If D and C are an $n \times n$ matrices such that C is invertible, then* $\text{rank}(CD) = \text{rank}(D)$.

Before giving the proof of this lemma, let's see how we will use it.

Because A is similar to B, there is an invertible matrix P such that $A = PBP^{-1}$, and thus $AP = PB$. The lemma will show us that $\text{rank}(B) = \text{rank}(PB)$. Of course, $\text{rank}(PB) = \text{rank}(AP)$, and by the Rank Theorem $\text{rank}(AP) = \text{rank}((AP)^{\text{T}})$. Then the lemma will also show that $\text{rank}((AP)^{\text{T}}) = \text{rank}(P^{\text{T}}A^{\text{T}}) = \text{rank}(A^{\text{T}})$ (recall that by the IMT, P^{t} is invertible if and only if P is invertible). Again applying the Rank Theorem, we will conclude that $\text{rank}(B) = \text{rank}(PB) = \text{rank}(AP) = \text{rank}((AP)^{\text{T}}) = \text{rank}(P^{\text{T}}A^{\text{T}}) = \text{rank}(A^{\text{T}}) = \text{rank}(A)$.

So all we have to do is to prove the lemma.

Let C and D be as given in the statement of the lemma. Then the Rank Theorem also tells us that $\text{rank}(D)+\dim(\text{Nul}(D)) = n = \text{rank}(CD)+\dim(\text{Nul}(CD))$ so it is enough to show that $\dim(\text{Nul}(D)) = \dim(\text{Nul}(CD))$. To do this we will show that $\text{Nul}(D) = \text{Nul}(CD)$. If $\mathbf{x} \in \text{Nul}(D)$, then $D\mathbf{x} = \mathbf{0}$, and thus $CD\mathbf{x} = \mathbf{0}$ so $\mathbf{x} \in \text{Nul}(CD)$. On the other hand, if $\mathbf{x} \in \text{Nul}(CD)$, then $CD\mathbf{x} = \mathbf{0}$ and because C is invertible, $C^{-1}CD\mathbf{x} = C^{-1}\mathbf{0}$, or $D\mathbf{x} = \mathbf{0}$. Thus $\text{Nul}(D) = \text{Nul}(CD)$, so $\dim(\text{Nul}(D)) = \dim(\text{Nul}(CD))$, and we conclude that $\text{rank}(D) = \text{rank}(CD)$ as required.

6.1.10 Find a unit vector in the direction of $\mathbf{u} = \begin{bmatrix} -6 \\ 4 \\ -3 \end{bmatrix}$.

The unit vector in the direction of $\mathbf{u}$ is $\frac{\mathbf{u}}{\|\mathbf{u}\|}$. Here $\| \mathbf{u} \|$ is

$$\sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{(-6)^2 + 4^2 + (-3)^2} = \sqrt{61}$$

and thus $\frac{\mathbf{u}}{\|\mathbf{u}\|} = \begin{bmatrix} \frac{-6}{\sqrt{61}} \\ \frac{4}{\sqrt{61}} \\ \frac{-3}{\sqrt{61}} \end{bmatrix}$.

6.1.24 Verify the parallelogram law for vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$: $\|\mathbf{u}+\mathbf{v}\|^2 + \|\mathbf{u}-\mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2$.

We know that the dot product distributes, so,

$$\|\mathbf{u}+\mathbf{v}\|^2 + \|\mathbf{u}-\mathbf{v}\|^2 = \left(\sqrt{(\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v})}\right)^2 + \left(\sqrt{(\mathbf{u}-\mathbf{v})\cdot(\mathbf{u}-\mathbf{v})}\right)^2 =$$

$$(\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v})+(\mathbf{u}-\mathbf{v})\cdot(\mathbf{u}-\mathbf{v}) = \mathbf{u}\cdot\mathbf{u}+\mathbf{u}\cdot\mathbf{v}+\mathbf{v}\cdot\mathbf{u}+\mathbf{v}\cdot\mathbf{v}+\mathbf{u}\cdot\mathbf{u}-\mathbf{u}\cdot\mathbf{v}-\mathbf{v}\cdot\mathbf{u}+\mathbf{v}\cdot\mathbf{v} =$$

$$2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2.$$

6.1.26 Let $\mathbf{u} = \begin{bmatrix} 5 \\ -6 \\ 7 \end{bmatrix}$, and let W be the set of all $\mathbf{x} \in \mathbb{R}^3$ such that $\mathbf{u} \cdot \mathbf{x} = 0$. What theorem in Chapter 4 can be used to show that W is a subspace of $\mathbb{R}^3$. Describe W in geometric language.

Note that $\mathbf{x} \in W$ if and only if $\mathbf{u} \cdot \mathbf{x} = 0$ or rather, if $\mathbf{u}^{\mathrm{T}}\mathbf{x} = 0$. Thus W is the Null space of the matrix $\mathbf{u}^{\mathrm{T}}$. We know from theorem 2, page 227, that a Nul space is a vector space. In geometric language, W consists of all vectors perpendicular to $\mathbf{u}$, that is, W is the plane going through the origin which is perpendicular to $\mathbf{u}$.

6.2.14 Let $\mathbf{y} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} 7 \\ 1 \end{bmatrix}$. Write $\mathbf{y}$ as the sum of a vector in Span$\{\mathbf{u}\}$ and a vector orthogonal to $\mathbf{u}$.

First we calculate $\mathrm{proj}_{\mathbf{u}}(y) = \hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u} = \frac{20}{50}\begin{bmatrix} 7 \\ 1 \end{bmatrix} = \begin{bmatrix} 14/5 \\ 2/5 \end{bmatrix}$. Let $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} 2 \\ 6 \end{bmatrix} - \begin{bmatrix} 14/5 \\ 2/5 \end{bmatrix} = \begin{bmatrix} -4/5 \\ 28/5 \end{bmatrix}$. Now by construction it is clear that $\mathbf{y} = \mathbf{z} + \hat{\mathbf{y}}$, that $\hat{\mathbf{y}} \in$ Span$\{\mathbf{u}\}$, and, as was argued on pg 386, that $\mathbf{z}$ is orthogonal to $\mathbf{u}$.

6.2.32 Let $\{\mathbf{v}_1, \mathbf{v}_2\}$ be an orthogonal set of nonzero vectors, and c_1, c_2 be any nonzero scalars. Show that $\{c_1\mathbf{v}_1, c_2\mathbf{v}_2\}$ is also an orthogonal set.

We have that $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$, and thus $(c_1\mathbf{v}_1) \cdot (c_2\mathbf{v}_2) = c_1c_2(\mathbf{v}_1 \cdot \mathbf{v}_2) = c_1c_2 0 = 0$. We conclude that $c_1\mathbf{v}_1$ and $c_2\mathbf{v}_2$ are orthogonal.

6.2.34 Given $\mathbf{u} \neq \mathbf{0}$ in $\mathbb{R}^n$, let $L = $ Span$\{\mathbf{u}\}$. For $\mathbf{y} \in \mathbb{R}^n$, the reflection of $\mathbf{y}$ in L is the point $\mathrm{refl}_L(\mathbf{y})$ defined by $\mathrm{refl}_L(\mathbf{y}) = 2\mathrm{proj}_L(\mathbf{y}) - \mathbf{y}$. Show that the mapping $\mathbf{y} \to \mathrm{refl}_L(\mathbf{y})$ is a linear transformation.

Well, there are some things we have to check. First, for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$, note that

$$T(\mathbf{x}+\mathbf{y}) = \mathrm{refl}_L(\mathbf{x}+\mathbf{y}) = 2\mathrm{proj}_L(\mathbf{x}+\mathbf{y}) - \mathbf{x} - \mathbf{y} = 2\frac{(\mathbf{x}+\mathbf{y})\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u} - \mathbf{x} - \mathbf{y} =$$

$$2\frac{\mathbf{x}\cdot\mathbf{u}+\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u} - \mathbf{x} - \mathbf{y} = 2\frac{\mathbf{x}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u} - \mathbf{x} + 2\frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u} - \mathbf{y} =$$

$$2\text{proj}_L(\mathbf{x}) - \mathbf{x} + 2\text{proj}_L(\mathbf{y}) - \mathbf{y} = \text{refl}_L(\mathbf{x}) + \text{refl}_L(\mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y}).$$

For all $c \in \mathbb{R}$,

$$T(c\mathbf{x}) = \text{refl}_L(c\mathbf{x}) = 2\text{proj}_L(c\mathbf{x}) - c\mathbf{x} = 2\frac{c\mathbf{x}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u} - c\mathbf{x} =$$

$$2c\frac{\mathbf{x}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u} - c\mathbf{x} = c\left(2\frac{\mathbf{x}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u} - \mathbf{x}\right) = c(2\text{proj}_L(\mathbf{x}) - \mathbf{x}) = c\text{refl}_L(\mathbf{x}) = cT(\mathbf{x}).$$

We conclude that T is a linear transformation.

6.3.16 Let $\mathbf{y} = \begin{bmatrix} 3 \\ -1 \\ 1 \\ 13 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \\ -1 \\ 2 \end{bmatrix}$, and $\mathbf{v}_2 = \begin{bmatrix} -4 \\ 1 \\ 0 \\ 3 \end{bmatrix}$. Find the distance from $\mathbf{y}$ to the subspace of $\mathbb{R}^4$ spanned by $\mathbf{v}_1$ and $\mathbf{v}_2$.

Call W the subspace spanned by $\mathbf{v}_1$ and $\mathbf{v}_2$. Then they are asking us to calculate

$$\|\mathbf{y} - \text{proj}_W(\mathbf{y})\| = \|\mathbf{y} - \frac{\mathbf{v}_1\cdot\mathbf{y}}{\mathbf{v}_1\cdot\mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{v}_2\cdot\mathbf{y}}{\mathbf{v}_2\cdot\mathbf{v}_2}\mathbf{v}_2\| =$$

$$\| \begin{bmatrix} 3 \\ -1 \\ 1 \\ 13 \end{bmatrix} - \frac{30}{10}\begin{bmatrix} 1 \\ -2 \\ -1 \\ 2 \end{bmatrix} - \frac{26}{26}\begin{bmatrix} -4 \\ 1 \\ 0 \\ 3 \end{bmatrix} \| = \| \begin{bmatrix} 4 \\ 4 \\ 4 \\ 4 \end{bmatrix} \| = \sqrt{4^2 + 4^2 + 4^2 + 4^2} = 8.$$

Done.

6.3.20 Let $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix}$, and let $\mathbf{u}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$. It can be shown that $\mathbf{u}_4$ is not in the subspace W spanned by $\mathbf{u}_1$ and $\mathbf{u}_2$. Use this fact to construct a nonzero vector $\mathbf{v}$ in $\mathbb{R}^3$ that is orthogonal to $\mathbf{u}_1$ and $\mathbf{u}_2$.

We know by theorem 8, page 395, that $\mathbf{v} = \mathbf{u}_4 - \hat{\mathbf{u}}_4 = \mathbf{u}_4 - \text{proj}_W(\mathbf{u}_4)$ is orthogonal to every vector in W. It is also true that $\mathbf{u}_1 \cdot \mathbf{u}_2 = 0$, so

$$\mathbf{v} = \mathbf{u}_4 - \hat{\mathbf{u}}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - (\frac{\mathbf{u}_1\cdot\mathbf{u}_4}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{u}_2\cdot\mathbf{u}_4}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2)$$

$$= \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - \left(\frac{1}{6}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} + \frac{-1}{30}\begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 4/5 \\ 2/5 \end{bmatrix}.$$

6.3.24 Let W be a subspace of $\mathbb{R}^n$ with an orthogonal basis $\{\mathbf{w}_1, \dots, \mathbf{w}_p\}$, and let $\{\mathbf{v}_1, \dots, \mathbf{v}_q\}$ be an orthogonal basis for $W^\perp$.

(a) Explain why $\{\mathbf{w}_1, \dots, \mathbf{w}_p \mathbf{v}_1, \dots, \mathbf{v}_q\}$ is an orthogonal set.

We know that $\mathbf{w}_i \cdot \mathbf{w}_j = 0$ for all $i \neq j$, $1 \leq i, j \leq p$, and that $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ for all $i \neq j$, $1 \leq i, j \leq q$. So it remains to check that $\mathbf{w}_i \cdot \mathbf{v}_j = 0$ for all $1 \leq i \leq p$ and $1 \leq j \leq q$. We know this is true, however, because $\mathbf{v}_j \in W^\perp$, which is by definition the set of all vectors whose dot product with elements of W is zero.

(b) Explain why the set in part (a) spans $\mathbb{R}^n$

By the Orthogonal Decomposition Theorem, any $\mathbf{y} \in \mathbb{R}^n$ can be written as $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$ where $\hat{\mathbf{y}} \in W$ and $\mathbf{z} \in W^\perp$. This means we can write any vector in $\mathbb{R}^n$ using the vectors from the set $\{\mathbf{w}_1, \dots, \mathbf{w}_p \mathbf{v}_1, \dots, \mathbf{v}_q\}$, and thus this set spans.

(c) Show that $\dim(W) + \dim(W^\perp) = n$.

We know that $\{\mathbf{w}_1, \dots, \mathbf{w}_p \mathbf{v}_1, \dots, \mathbf{v}_q\}$ spans $\mathbb{R}^n$, so if we can show that this set is linearly independent, then $n = p + q = \dim(W) + \dim(W^\perp)$ as required (I am using here the fact that a linearly independent spanning set is a basis).

Suppose that we have $c_i, d_j \in \mathbb{R}$ for $i = 1, \dots, p$ and $j = 1, \dots, q$ such that $\sum_{i=1}^p c_i \mathbf{w}_i + \sum_{j=1}^q d_i \mathbf{v}_i = \mathbf{0}$ and not all the c_i, d_j are zero. Because the $\mathbf{w}_i$ are linearly independent, it can not be the case that all the d_i are zero (otherwise we could write $\sum_{i=1}^p c_i \mathbf{w}_i + \sum_{j=1}^q d_i \mathbf{v}_i = \sum_{i=1}^p c_i \mathbf{w}_i = \mathbf{0}$ where not all the c_i are zero, a contradiction). A similar argument for the $\mathbf{v}_j$ shows that all the c_i are not zero. Thus we can write $\mathbf{u} = \sum_{i=1}^p c_i \mathbf{w}_i = -\sum_{j=1}^q d_i \mathbf{v}_i$, and thus there is a nonzero element of $\mathbf{u} \in W^\perp$ which is also in W ($\sum_{i=1}^p c_i \mathbf{w}_i$ cannot be identically zero because not all the c_i are zero and the $\mathbf{w}_i$ are linearly independent). Because $\mathbf{u} \in W^\perp$, we know that the dot product of $\mathbf{u}$ with anything in W is zero. But $\mathbf{u}$ is also in W, so $\mathbf{u} \cdot \mathbf{u} = 0$, and hence that $\mathbf{u} = \mathbf{0}$. This is a contradiction ($\mathbf{u}$ was supposed to be nonzero), and thus completes the proof.