MATH 217 SPRING 2014
WRITTEN HOMEWORK 10
SOLUTIONS

Section 5.2

Problem 40. Denote the column vectors of A by $\mathbf{v}_1, \dots \mathbf{v}_n$, so that

$$A = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ | & | & \cdots & | \end{bmatrix}.$$

Since the vectors $\mathbf{v}_1, \dots \mathbf{v}_n$ are orthogonal to one another, they are linearly independent and thus form a basis for $\mathbb{R}^n$. We can easily obtain an orthonormal basis by divining each $\mathbf{v}_j$ by its length. That is, for each $j = 1, 2, \dots, n$, let

$$\mathbf{u}_j = \frac{1}{||\mathbf{v}_j||}\mathbf{v}_j, \quad \text{or equivalently } \mathbf{v}_j = ||\mathbf{v}_j||\mathbf{u}_j.$$

Then the vectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ are orthonormal. Notice that

$$A = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ | & | & \cdots & | \end{bmatrix} = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \\ | & | & \cdots & | \end{bmatrix} \begin{bmatrix} ||\mathbf{v}_1|| & 0 & 0 & 0 & \dots & 0 \\ 0 & ||\mathbf{v}_2|| & 0 & 0 & \dots & 0 \\ 0 & 0 & ||\mathbf{v}_3|| & 0 & \dots & 0 \\ 0 & 0 & 0 & ||\mathbf{v}_4|| & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & ||\mathbf{v}_n|| \end{bmatrix},$$

which is the QR-factorization. That is, the upper-diagonal R-matrix is in fact diagonal.

Section 5.3

Problem 30. Let $L : \mathbb{R}^m \to \mathbb{R}^n$ be length preserving. Let $\mathbf{x} \in \ker(L)$. Then $L\mathbf{x} = \mathbf{0}$. Since L is length preserving, $||\mathbf{x}|| = ||L\mathbf{x}|| = ||\mathbf{0}||$, and so $\mathbf{x} = \mathbf{0}$. Thus $\ker(L) = \{\mathbf{0}\}$.

Since $\ker(L) = \mathbf{0}$, the rank-nullity theorem implies that the dimension of the image of L is m. Since $\operatorname{im}(L) \subset \mathbb{R}^n$, we must have $m \leq n$.

Denoting by $\mathbf{e}_j$ the j-th standard vector in $\mathbb{R}^m$, and letting A be the matrix for the transformation L, we see that

$$A = \begin{bmatrix} | & | & \cdots & | \\ L\mathbf{e}_1 & L\mathbf{e}_2 & \cdots & L\mathbf{e}_m \\ | & | & \cdots & | \end{bmatrix}.$$

Since the vectors $L\mathbf{e}_1, L\mathbf{e}_2, \dots, L\mathbf{e}_m \in \mathbb{R}^n$ span $\operatorname{im}(L)$, and the dimension of $\operatorname{im}(L)$ is m, the vectors $L\mathbf{e}_1, L\mathbf{e}_2, \dots, L\mathbf{e}_m \in \mathbb{R}^n$ must in fact form a basis for $\operatorname{im}(L)$. In fact, these vectors are orthogonal to eachother. To see this, consider that $||\mathbf{e}_i + \mathbf{e}_j||^2 = ||L(\mathbf{e}_i + \mathbf{e}_j)||^2 = ||L\mathbf{e}_i + L\mathbf{e}_j||^2$. Assume that $i \neq j$ so that $\mathbf{e}_i \cdot \mathbf{e}_j = 0$. In terms of the dot product, this gives

$$(\mathbf{e}_i + \mathbf{e}_j) \cdot (\mathbf{e}_i + \mathbf{e}_j) = (L\mathbf{e}_i + L\mathbf{e}_j) \cdot (L\mathbf{e}_i + L\mathbf{e}_j).$$

Expanding both sides of this dot product gives

$$||\mathbf{e}_i||^2 + ||\mathbf{e}_j||^2 = ||L\mathbf{e}_i||^2 + ||L\mathbf{e}_j||^2 + 2(L\mathbf{e}_i) \cdot (L\mathbf{e}_j).$$

Since L is length preserving, the above equation simplifies to $(L\mathbf{e}_i) \cdot (L\mathbf{e}_j) = 0$. We can therefore see that the colums of A form an orthonormal basis for $\text{im}(L) \subset \mathbb{R}^n$.

The matrix $A^T A$ is

$$A^T A = \begin{bmatrix} - & L\mathbf{e}_1 & - \\ - & L\mathbf{e}_2 & - \\ \vdots & \vdots & \vdots \\ - & L\mathbf{e}_m & - \end{bmatrix} \begin{bmatrix} | & | & \cdots & | \\ L\mathbf{e}_1 & L\mathbf{e}_2 & \cdots & L\mathbf{e}_m \\ | & | & \cdots & | \end{bmatrix}.$$

Since the vectors $L\mathbf{e}_1, L\mathbf{e}_2, \ldots L\mathbf{e}_m$ are orthonormal, the above matrix multiplication give $A^T A = I_m$.

Let $\mathbf{x} \in \mathbb{R}^n$. Then

$$\begin{aligned} AA^T\mathbf{x} = A(A^T\mathbf{x}) &= \begin{bmatrix} | & | & \cdots & | \\ L\mathbf{e}_1 & L\mathbf{e}_2 & \cdots & L\mathbf{e}_m \\ | & | & \cdots & | \end{bmatrix} \begin{bmatrix} (L\mathbf{e}_1) \cdot \mathbf{x} \\ (L\mathbf{e}_2) \cdot \mathbf{x} \\ (L\mathbf{e}_3) \cdot \mathbf{x} \\ \vdots \\ (L\mathbf{e}_m) \cdot \mathbf{x} \end{bmatrix} \\ &= [(L\mathbf{e}_1) \cdot \mathbf{x}]L\mathbf{e}_1 + [(L\mathbf{e}_2) \cdot \mathbf{x}]L\mathbf{e}_2 + \cdots + [(L\mathbf{e}_m) \cdot \mathbf{x}]L\mathbf{e}_m, \end{aligned}$$

which is the projection of $\mathbf{x}$ onto $\text{im}(A)$. Thus AA^T is the matrix for the the projection onto $\text{im}(L)$ in $\mathbb{R}^n$.

As a simple example, consider $L : \mathbb{R}^2 \to \mathbb{R}^3$ given by $L \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}$. The matrix for this transformation is

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix},$$

which clearly has rank 2. It is easy to see

$$A^T A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad AA^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

The matrix AA^T shown above is the matrix for projection onto the xy-plane in $\mathbb{R}^3$:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}.$$

Problem 31. Let A be an orthogonal matrix, i.e. it has orthonormal columns. Then A is invertible and A^{-1} is orthogonal as well. In fact, by Theorem 5.3.7, $A^{-1} = A^T$, so the columns of A^T are orthonormal. Since the columns of A^T are the rows of A, it follows that the rows of A are orthonormal as well.

Problem 46. We start with $M = QR$, where $M \in \mathbb{R}^{n\times m}$ has linearly independent columns, $Q \in \mathbb{R}^{n\times m}$ has orthonormal columns, and $R \in \mathbb{R}^{m\times m}$ is upper triangular with positive entries on the diagonal. Since Q has orthonormal columns, $Q^TQ = I_m$. Multiplying on the left by Q^T, the equation $M = QR$ becomes $Q^TM = Q^TQR = I_mR = R$. Equivalently, $R = Q^TM$.

Problem 48. Since A is invertible, it has rank n. Since rank(A) =rank(A^T), we see that rank$(A^T) = n$, so A^T is invertible as well, and so it has a unique QR-factorization. That is, we can write $A^T = QR$ where Q is orthogonal and R is upper-triangular with positive entries on the diagonal. Taking the transpose of both sides of this equation gives $(A^T)^T = (QR)^T$, or equivalently $A = R^TQ^T$. Notice that R^T is lower triangular with positive entries on the diagonal, and Q^T is orthogonal.

Problem 72. Let $V = \text{span}(\mathbf{v})$. An orthonormal basis for V is $\mathbf{u} = \frac{1}{||\mathbf{v}||}\mathbf{v}$. According to Theorem 5.3.10, the matrix for the orthogonal projection onto V is $P = QQ^T$, where

$$Q = \begin{bmatrix} | \\ \mathbf{u} \\ | \end{bmatrix} = \frac{1}{||\mathbf{v}||}\begin{bmatrix} | \\ \mathbf{v} \\ | \end{bmatrix} = \frac{1}{||\mathbf{v}||}\begin{bmatrix} 1 \\ a \\ a^2 \\ \vdots \\ a_{n-1} \end{bmatrix}.$$

Then P is

$$P = QQ^T = \frac{1}{||\mathbf{v}||^2}\begin{bmatrix} 1 \\ a \\ a^2 \\ \vdots \\ a_{n-1} \end{bmatrix}\begin{bmatrix} 1 & a & a^2 & \cdots & a_{n-1} \end{bmatrix} = \frac{1}{||\mathbf{v}||^2}\begin{bmatrix} 1 & a & a^2 & \cdots & a^{n-1} \\ a & a^2 & a^3 & \cdots & a_n \\ a^2 & a^3 & a^4 & \cdots & a^{n+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a^{n-1} & a^n & a^{n+1} & \cdots & a^{2n-2} \end{bmatrix},$$

which is a Hankel matrix according to the definition in Exercise 71.

In the case $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$, we have $||\mathbf{v}||^2 = 21$, so

$$P = \frac{1}{21}\begin{bmatrix} 1 & 2 & 4 \\ 2 & 4 & 8 \\ 4 & 8 & 16 \end{bmatrix}.$$

Section 5.4

Problem 10.

(a) Since the system is consistent, it has some solution $\mathbf{x}$. Since $\ker(A)$ and $(\ker(A))^\perp$ are complementary spaces, we can write $\mathbf{x}$ in the form $\mathbf{x} = \mathbf{x}_h + \mathbf{x}_0$ where $\mathbf{x}_h \in \ker A$ and $\mathbf{x}_o \in (\ker A)^\perp$. We thus have $\mathbf{b} = A\mathbf{x} = A(\mathbf{x}_h + \mathbf{x}_0) = A\mathbf{x}_h + A\mathbf{x}_0 = A\mathbf{x}_0$, where we have used the fact that $A\mathbf{x}_h = \mathbf{0}$ since $\mathbf{x}_h \in \ker A$. Thus $A\mathbf{x}_0 = \mathbf{b}$, and so x_0 solves the system as well.

(b) Suppose $\mathbf{x}_0$ and $\mathbf{x}_1$ are two solutions to the system $A\mathbf{x} = \mathbf{b}$, and $\mathbf{x}_0$ and $\mathbf{x}_1$ are both in $(\ker A)^\perp$. Then $A(\mathbf{x}_0 - \mathbf{x}_1) = A\mathbf{x}_0 - A\mathbf{x}_1 = \mathbf{b} - \mathbf{b} = \mathbf{0}$, and so $\mathbf{x}_0 - \mathbf{x}_1 \in \ker A$. On the other hand $\mathbf{x}_0 - \mathbf{x}_1 \in (\ker A)^\perp$ since it is a linear combination of vectors in $(\ker A)^\perp$. Since $\ker A$ and $(\ker A)^\perp$ are complementary subspaces, their intersection consists of only the zero vector. This implies $\mathbf{x}_0 - \mathbf{x}_1 = \mathbf{0}$ and so $\mathbf{x}_0 = \mathbf{x}_1$, thus there is only one solution to $A\mathbf{x} = \mathbf{b}$ which lies in $(\ker A)^\perp$.

(c) Let $\mathbf{x}_0 \in (\ker A)^\perp$ and $\mathbf{x}_1 \notin (\ker A)^\perp$ be two solutions to the linear system $A\mathbf{x} = \mathbf{b}$. The vector $\mathbf{x}_1$ can be written as the sum $\mathbf{x}_1 = \mathbf{x}_1^{||} + \mathbf{x}_1^\perp$, where $\mathbf{x}_1^{||} \in \ker A$ and $\mathbf{x}_1^\perp \in (\ker A)^\perp$. Notice then that $\mathbf{b} = Ax_1 = A\mathbf{x}_1^{||} + A\mathbf{x}_1^\perp = A\mathbf{x}_1^\perp$, and so $\mathbf{x}_1^\perp$ also solves the linear system $A\mathbf{x} = \mathbf{b}$. According to part (b), we must have $\mathbf{x}_1^\perp = \mathbf{x}_0$. Thus $\mathbf{x}_1 = \mathbf{x}_0 + \mathbf{x}_1^{||}$, where $\mathbf{x}_1^{||}$ is a nonzero vector in $\ker A$. Notice

$$||\mathbf{x}_1||^2 = \mathbf{x}_1 \cdot \mathbf{x}_1 = (\mathbf{x}_0 + \mathbf{x}_1^{||}) \cdot (\mathbf{x}_0 + \mathbf{x}_1^{||}) = ||\mathbf{x}_0||^2 + ||\mathbf{x}_1^{||}||^2 + 2(\mathbf{x}_0) \cdot (\mathbf{x}_1^{||}).$$

Since $\mathbf{x}_1^{||} \in \ker A$ and $\mathbf{x}_0 \in (\ker A)^\perp$, $(\mathbf{x}_0) \cdot (\mathbf{x}_1^{||}) = 0$, so the above equation is

$$||\mathbf{x}_1||^2 = ||\mathbf{x}_0||^2 + ||\mathbf{x}_1^{||}||^2 > ||\mathbf{x}_0||^2,$$

since $\mathbf{x}_1^{||} \neq \mathbf{0}$.

Problem 11.

(a) Let $\mathbf{y}_1, \mathbf{y}_2 \in \mathbb{R}^m$, and denote $L^+(\mathbf{y}_1) = \mathbf{x}_1$ and $L^+(\mathbf{y}_2) = \mathbf{x}_2$. This means that $A\mathbf{x}_1 = \mathbf{y}_1$, $A\mathbf{x}_2 = \mathbf{y}_2$, and that $\mathbf{x}_1, \mathbf{x}_2 \in (\ker A)^\perp$. Since $(\ker A)^\perp$ is a subspace of $\mathbb{R}^n$, we have $k\mathbf{x}_1 + \mathbf{x}_2 \in (\ker A)^\perp$ for any scalar $k \in \mathbb{R}$. Consider the linear combination $k\mathbf{y}_1 + \mathbf{y}_2$. Applying L^+ to this vector, we have by definition

$$L^+(k\mathbf{y}_1 + \mathbf{y}_2) = \left(\text{the solution to the system } \; A\mathbf{x} = k\mathbf{y}_1 + \mathbf{y}_2 \text{ which lies in } \; (\ker A)^\perp\right).$$

Notice that $A(k\mathbf{x}_1 + \mathbf{x}_2) = kA\mathbf{x}_1 + A\mathbf{x}_2 = k\mathbf{y}_1 + \mathbf{y}_2$, so $k\mathbf{x}_1 + \mathbf{x}_2$ clearly solves the system $A\mathbf{x} = k\mathbf{y}_1 + \mathbf{y}_2$. As mentioned above, $k\mathbf{x}_1 + \mathbf{x}_2 \in (\ker A)^\perp$, thus $L^+(k\mathbf{y}_1 + \mathbf{y}_2) = k\mathbf{x}_1 + \mathbf{x}_2 = kL^+(\mathbf{y}_1) + L^+(\mathbf{y}_2)$. This proves that L^+ is a linear transformation.

(b) Denote $L^+(\mathbf{y}) = \mathbf{x}$. By definition, $L\mathbf{x} = \mathbf{y}$. This which immediately implies $L(L^+(\mathbf{y})) = L(\mathbf{x}) = \mathbf{y}$.

(c) Let $\mathbf{x} \in \mathbb{R}^n$. $\mathbf{x}$ can be written in the form $\mathbf{x} = \mathbf{x}^{||} + \mathbf{x}^\perp$, where $\mathbf{x}^{||} \in \ker L$ and $\mathbf{x}^\perp \in (\ker L)^\perp$. Then $L\mathbf{x} = L(\mathbf{x}^{||} + \mathbf{x}^\perp) = L\mathbf{x}^{||} + L\mathbf{x}^\perp = L\mathbf{x}^\perp$. Alpying L^+ we find that

$$L^+(L\mathbf{x}) = L^+(L\mathbf{x}^\perp) = \left(\text{the solution to the system } \; A\mathbf{x} = A\mathbf{x}^\perp \text{ which lies in } \; (\ker A)^\perp\right),$$

which is clearly $\mathbf{x}^\perp$. Thus $L^+(L(\mathbf{x})) = \mathbf{x}^\perp$. In other words, L^+L is the orthogonal projection onto $(\ker L)^\perp$.

(d) Let $\mathbf{y} \in \ker(L^+)$, so that $L^+(\mathbf{y}) = \mathbf{0}$. By definition, this means that $L\mathbf{0} = \mathbf{y}$, so we must have $\ker(L^+) = \mathbf{0}$. By the rank-nullity theorem, $\dim(\operatorname{im} L^+) = m$.

Notice that since $\operatorname{rank}(A) = m$, the rank-nullity theorem implies that $\dim(\ker A) = n - m$, and so $\dim((\ker A)^\perp) = n - (n-m) = n$. The image of L^+ is clearly contained in $(\ker A)^\perp$. Since $\dim(\operatorname{im} L^+) = m = \dim((\ker A)^\perp)$, we must have $\operatorname{im}(L^+) = (\ker A)^\perp$.

(e) Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. Then $\ker A = \operatorname{span}\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right)$, and so $(\ker A)^\perp = \operatorname{span}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right)$.

Let $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \in \mathbb{R}^2$. Then $A\mathbf{x} = \mathbf{y}$ for any $\mathbf{x} \in \mathbb{R}^3$ of the form $\mathbf{x} = \begin{bmatrix} y_1 \\ y_2 \\ t \end{bmatrix}$ for any $t \in \mathbb{R}$.

If $\mathbf{x}$ is to be in $(\ker A)^\perp$, we must choose $t = 0$. Thus we have

$$L^+ \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ 0 \end{bmatrix}.$$

Section 5.5

Problem 18. Let $\mathcal{B} = (g_1, g_2, \ldots, g_n)$ be an orthonormal basis for V. Let $f \in V$ have the coordinates $[f]_\mathcal{B} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$ for some real numbers $c_1, c_2, \ldots, c_n$. Cleatly then $||[f]_\mathcal{B}||^2 = c_1^2 + c_2^2 + \cdots + c_n^2$.

Since $f = c_1g_1 + c_2g_2 + \cdots + c_ng_n$, we have

$$||f||^2 = \langle c_1g_1 + c_2g_2 + \cdots + c_ng_n, c_1g_1 + c_2g_2 + \cdots + c_ng_n \rangle.$$

By the bilinear property of the inner product we can expand the inner product:

$$||f||^2 = \sum_{i,j=1}^{n} \langle c_ig_i, c_jg_j \rangle.$$

Since the basis elements of $\mathcal{B}$ are orthogonal to one another, many terms in the expansion are zero. The nonzero terms are the ones with $i = j$:

$$||f||^2 = \langle c_1g_1, c_1g_1 \rangle + \langle c_2g_2, c_2g_2 \rangle + \cdots + \langle c_ng_n, c_ng_n \rangle = c_1^2\langle g_1, g_1 \rangle + c_2^2\langle g_2, g_2 \rangle + \cdots + c_n^2\langle g_n, g_n \rangle.$$

Since the basis elements have norm 1, this simplifies to

$$||f||^2 = c_1^2 + c_2^2 + \cdots + c_n^2 = ||[f]_\mathcal{B}||^2.$$