Introduction to Linear Algebra, Spring 2007
MATH 3130, Section 001

SOLUTIONS TO PROJECT 1

1. Consider a system of m linear equations each in n variables.

   (a) Prove that the 3 row operations preserve the solution set of the linear system. *Note:* This proof is not difficult if you start off correctly. If you need help getting started, ask for some guidance. The discussion of this fact in Section 1.1 of our text book is not rigorous enough, so don't copy that. [**7 points**]

   Solution: Consider any 2 arbitrary rows of the system, say

   $$a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n = b_i$$

   and

   $$a_{j,1}x_1 + a_{j,2}x_2 + \cdots + a_{j,n}x_n = b_j.$$

   Note that these correspond to the ith and jth rows of the augmented matrix. Now, let $\mathbf{s} = (s_1, s_2, \ldots, s_n)$ be any solution to the system. Then $\mathbf{s}$ is a solution to every linear equation in the system. In particular, $\mathbf{s}$ is a solution to the equations above. Next, let c be any nonzero scalar. We need to show that if we apply a single row operation, then $\mathbf{s}$ is still a solution to the system.

   (i) Replacement: Replace the ith linear equation with c times the jth linear equation plus the ith equation. We get the new equation:

   $$c[a_{j,1}x_1 + a_{j,2}x_2 + \cdots + a_{j,n}x_n] + [a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n] = c[b_j] + b_i$$

   This is the new ith row of the augmented matrix. Now, plug the coordinates for $\mathbf{s}$ into the appropriate variables on the left hand side. We get

   $$c[a_{j,1}s_1 + a_{j,2}s_2 + \cdots + a_{j,n}s_n] + [a_{i,1}s_1 + a_{i,2}s_2 + \cdots + a_{i,n}s_n].$$

   But since $\mathbf{s}$ is a solution to the original ith and jth rows of the system, the above expression is equal to

   $$c[b_j] + b_i,$$

   which shows that $\mathbf{s}$ is a solution the new ith equation.

   (ii) Scaling: Replace the ith row with c times the ith row. We get the new equation:

   $$c[a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n] = c[b_i]$$

   This is the new ith row of the augmented matrix. Now, plug the coordinates for $\mathbf{s}$ into the appropriate variables on the left hand side. We get

   $$c[a_{i,1}s_1 + a_{i,2}s_2 + \cdots + a_{i,n}s_n].$$

   But since $\mathbf{s}$ is a solution to the original ith row of the system, the above expression is equal to

   $$c[b_i],$$

   which shows that $\mathbf{s}$ is a solution the new ith equation.

(iii) Interchange: Swap the ith row of the system with the jth row of the system. Since $\mathbf{s}$ is a solution to every equation in the system, $\mathbf{s}$ is certainly still a solution to these rows interchanged.

Therefore, $\mathbf{s}$ is still a solution to the system, regardless which row operations are applied. We have shown that if $\mathbf{s}$ is a solution to the original system, then $\mathbf{s}$ is a solution to the system that results from applying row operations. What we haven't shown is that the new system doesn't have any new solutions that it didn't have before applying row operations. This is also easy to show by reversing the row operations. I didn't take off any points for not showing this part.

(b) What geometric effect does each row operation have on the hyperplanes that correspond to the linear equations in the system? Be as specific as possible and *briefly* justify your answer. [**3 points**]

Solution: Since the solution set of a system of linear equations is equal to the points of intersection of all the hyperplanes and the solution set is preserved under row operations, the points of intersection of all the hyperplanes must be preserved. Here is what happens geometrically.

(i) Replacement: This is the hard one. The points of intersection must stay fixed, but the hyperplane that you are replacing with the sum of itself and the multiple of another hyperplane can move. It gets sheared about the points of intersection (which ends up looking like a rotation).

(ii) Scaling: If $|c| > 1$, then the hyperplane gets "stretched" out. This doesn't change the points of intersection and doesn't look like much because the hyperplanes are extended infinitely far. If $|c| < 1$, then the hyperplane gets "compressed". Again, this doesn't change the points of intersection.

(iii) Interchange: This doesn't do anything except give the hyperplanes new names. The ith hyperplane is now called the jth hyperplane.

2. If $k \geq 1$, then a *k-dimensional linear subspace* of $\mathbb{R}^n$ is equal to the span of some collection of k linearly independent vectors from $\mathbb{R}^n$. Geometrically, these are lines (1-dimensional), planes (2-dimensional), etc. passing through the origin in $\mathbb{R}^n$. The *0-dimensional linear subspace* of $\mathbb{R}^n$ is the zero-vector, $\mathbf{0}$. Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a collection of linearly independent vectors from $\mathbb{R}^n$ and let $W = \text{span}\{\mathbf{v}_1, \ldots \mathbf{v}_k\}$, so that W is a k-dimensional linear subspace of $\mathbb{R}^n$. Note that $1 \leq k \leq n$. Now, let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Define $T(W) = \{T(\mathbf{x}) : \mathbf{x} \in W\}$.

(a) Prove that $T(W)$ is a linear subspace of $\mathbb{R}^m$. [**7 points**]

Solution: Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a collection of linearly independent vectors from $\mathbb{R}^n$ and let $W = \text{span}\{\mathbf{v}_1, \ldots \mathbf{v}_k\}$, so that W is a k-dimensional linear subspace of $\mathbb{R}^n$. Note that $1 \leq k \leq n$. Now, let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Define $T(W) = \{T(\mathbf{x}) : \mathbf{x} \in W\}$. We need to show that $T(W)$ is a linear subspace of $\mathbb{R}^m$. Note that $T(W)$ is the collection of all images of vectors from W. Let $\mathbf{x} \in W$. Then there exists scalars $c_1, \ldots, c_k$ such that

$$x = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k.$$

Since T is a linear transformation, we see that

$$T(x) = T(c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k) = c_1T(\mathbf{v}_1) + \cdots + c_kT(\mathbf{v}_k).$$

This shows that $T(W)$ is equal to the span of $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_k)\}$ in $\mathbb{R}^m$. If $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_k)\}$ in $\mathbb{R}^m$ is linearly independent (which can only happen if $k \leq m$), then $T(W)$ is clearly a k-dimensional linear subspace of $\mathbb{R}^m$ by definition. Assume that $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_k)\}$ in $\mathbb{R}^m$ is linearly dependent. One possibility is that $T(\mathbf{v}_i) = \mathbf{0}$ for all i, in which case, $T(W)$ is only the zero-vector in $\mathbb{R}^m$. If there exists at least one $T(\mathbf{v}_i) \neq \mathbf{0}$, then by Theorem 1.7, we can find the largest collection of linearly independent vectors from $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_k)\}$. This shows that $T(W)$ is a r-dimensional linear subspace where $0 < r < k$.

(b) What are the possible dimensions of $T(W)$? [**3 points**].

Solution: By the discussion above, $T(W)$ is an l-dimensional linear subspace where $0 \leq l \leq k$.

3. Read Section 1.6 and complete Exercise 12. [**10 points**]

Solution:

(a) By looking at the graph, we see that we have the following relationships.

Intersection	Flow in		Flow out
A	x_1	$=$	$40 + x_3 + x_4$
B	200	$=$	$x_1 + x_2$
C	$x_2 + x_3$	$=$	$100 + x_5$
D	$x_4 + x_5$	$=$	60

If we rewrite each equation so that the variables are on the left and the constants are on the right, then we can put the system into a 4×6 augmented matrix. Then row reduce the matrix to reduced echelon form.

$$\begin{bmatrix} 1 & 0 & -1 & -1 & 0 & 40 \\ 1 & 1 & 0 & 0 & 0 & 200 \\ 0 & 1 & 1 & 0 & -1 & 100 \\ 0 & 0 & 0 & 1 & 1 & 60 \end{bmatrix} \to \cdots \to \begin{bmatrix} 1 & 0 & -1 & 0 & 1 & 100 \\ 0 & 1 & 1 & 0 & -1 & 100 \\ 0 & 0 & 0 & 1 & 1 & 60 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Then the general traffic pattern for the network is described by

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 100 + x_3 - x_5 \\ 100 - x_3 + x_5 \\ x_3 \\ 60 - x_5 \\ x_5 \end{bmatrix},$$

where x_3 and x_5 are free.

(b) If x_4 is closed, we can substitute 0 for x_4. Again, put the system in an augmented matrix and row reduce to reduced echelon form. Then the general traffic pattern for the network is described by

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 40 + x_3 \\ 160 - x_3 \\ x_3 \\ 0 \\ 60 \end{bmatrix},$$

where x_3 is free.

(c) If $x_4 = 0$, then the total flow at intersection A is given by $x_1 = x_3 + 40$ (from part (b)). The smallest that the free variable x_3 can be is 0. This implies that the smallest that x_1 can be is 40.

4. Let $T_1 : \mathbb{R}^n \to \mathbb{R}^m$ and $T_2 : \mathbb{R}^m \to \mathbb{R}^k$ be linear transformations. Define $S : \mathbb{R}^n \to \mathbb{R}^k$ via $S(\mathbf{x}) = T_2(T_1(\mathbf{x}))$.

(a) Prove that S is a linear transformation. [**7 points**]

Solution: Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ and let c be a scalar. Then

$$\begin{aligned} S(\mathbf{u} + \mathbf{v}) &= T_2(T_1(\mathbf{u} + \mathbf{v})) \\ &= T_2(T_1(\mathbf{u}) + T_1(\mathbf{v})) \quad \text{(since T}_1\text{ linear)} \\ &= T_2(T_1(\mathbf{u})) + T_2(T_1(\mathbf{v})) \quad \text{(since T}_2\text{ linear)} \\ &= S(\mathbf{u}) + S(\mathbf{v}) \end{aligned}$$

and

$$\begin{aligned} S(c\mathbf{u}) &= T_2(T_1(c\mathbf{u})) \\ &= T_2(cT_1(\mathbf{u})) \quad \text{(since T}_1\text{ linear)} \\ &= cT_2(T_1(\mathbf{u})) \quad \text{(since T}_1\text{ linear)} \\ &= cS(\mathbf{u}). \end{aligned}$$

Therefore, S is a linear transformation.

(b) When will S be one-to-one and onto? Justify your answer. [**3 points**]

Solution: There are many correct answers to this question. Here is one possible answer. If either of T_1 or T_2 is *not* one-to-one, then S is *not* one-to-one (think about the definition or think about Theorem 1.11). So, we need both of T_1 and T_2 to be one-to-one in order for S to be one-to-one. A sufficient condition for S to be onto is if both T_1 and T_2 are onto. But this is stronger than we need for S to be onto. All we need is to be able to find a $\mathbf{y}$ in the range of T_1 for each $\mathbf{b} \in \mathbb{R}^k$ such that $T_2(\mathbf{y}) = \mathbf{b}$. If T_1 and T_2 are one-to-one, then we must have $n \leq m \leq k$, otherwise there isn't enough "room" for $S(\mathbf{e}_1), \ldots, S(\mathbf{e}_n)$. If S is onto, then we must have $n \geq k$. So, a necessary condition for S to be one-to-one and onto is for $n = m$ and $n \leq k$.

5. Suppose $AB = I_n$ (the $n \times n$ identity matrix).

(a) Prove that $A\mathbf{x} = \mathbf{b}$ is consistent for all $b \in \mathbb{R}^n$. [**7 points**]

Solution: We don't know that A and B are square matrices, which implies that we don't know whether A or B is invertible. So, we can't use any facts about invertible matrices. Let $\mathbf{b} \in \mathbb{R}^n$. Multiplying both sides of $AB = I_n$ on the right by $\mathbf{b}$, we obtain

$$(AB)\mathbf{b} = I_n\mathbf{b},$$

which we can rewrite as

$$A(B\mathbf{b}) = \mathbf{b}.$$

Now, notice that $B\mathbf{b}$ is a solution to $A\mathbf{x} = \mathbf{b}$. This shows that $A\mathbf{x} = \mathbf{b}$ is consistent for all $\mathbf{b}$ if we have $AB = I_n$.

(b) Prove that A must have at least as many columns as rows. [**3 points**]

Solution: In order for $A\mathbf{x} = \mathbf{b}$ to be consistent, A must have a pivot in every row (Theorem 1.4). This forces A to have at least as many columns as rows.