Midterm Solutions

November 8, 2013

Problem 1

We are given the LU decomposition of a matrix A as

$$A = LU = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 \\ 3 & 2 & 4 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 & 1 & 2 & 1 \\ 0 & 0 & 2 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

(a) What is the rank $r(A)$ of matrix A? Explain

(b) Find a basis for the null space of A. Show your work.

(c) Find a basis for the column space of A. Show your work

(d) For $\vec{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ do solution(s) to $A\vec{x} = \vec{b}$ exist? Clearly motivate your answer. Note that it is not necessary to compute the solution(s).

(a) First, recall that $r(A) = \dim(R(A)) = \dim(\text{rowsp}(A))$. Now, since we are already given the LU decomposition of A, we can tell the rank of matrix A by counting the number of non-zero rows in U, which corresponds to the number of linearly independent rows in A, i.e. $\dim(\text{rowsp}(A))$. There are 3 non-zero rows in U, hence $r(A) = 3$.

(b) There are a couple of ways to get the basis for the null space of A. (The method presented here was not the only way to get full credit.)
We recall that for $A \in \mathsf{R}^{m \times n}$, the null space $N(A)$ is defined as,

$$N(A) = \{\vec{x} \in \mathsf{R}^n | A\vec{x} = \vec{0}\}.$$

We know from the lectures that $N(U) = N(A)$. For completeness, this can be proven as follows (but not necessary to do this on the midterm solutions): First, we know that $N(U) \subseteq N(A)$ since if $\vec{x} \in N(U)$, then $U\vec{x} = \vec{0}$ and $A\vec{x} = LU\vec{x} = L\vec{0} = \vec{0}$, so $\vec{x} \in N(A)$. Then, we know that L is nonsingular as it is a triangular matrix with non-zero entries on the diagonal. In other words, L has full rank. Because $A = LU$, we then know that $N(A) \subseteq N(U)$. We can also convince ourselves of the latter in the following way: suppose that we have a $\vec{x} \in N(A)$ and $\vec{x} \notin N(U)$, then $U\vec{x} = \vec{y} \neq \vec{0}$ and we get a contradiction $A\vec{x} = LU\vec{x} = L\vec{y} \neq \vec{0}$, since L has full column rank.

Using that $N(A) = N(U)$, we can find a basis for the null space of A by working with U directly.

Suppose, $\vec{x} \in N(U)$, $\vec{x} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \end{bmatrix}^T$. Then,

$$U\vec{x} = \begin{bmatrix} 1 & 2 & 0 & 1 & 2 & 1 \\ 0 & 0 & 2 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

We get the following set of equations

$$\begin{aligned} x_1 + 2x_2 + x_4 + 2x_5 + x_6 = 0 &\implies x_1 = -2x_2 - x_4 - 2x_5 - x_6, \\ 2x_3 + 2x_4 = 0 &\implies x_3 = -x_4, \\ x_6 = 0. \end{aligned}$$

Choosing the free parameters, $x_2 = r,\ x_4 = s,\ x_5 = t$, we plug in to the equations above to get

$$\begin{aligned} x_1 &= -2r - s - 2t \\ x_2 &= r \\ x_3 &= -s \\ x_4 &= s \\ x_5 &= t \\ x_6 &= 0. \end{aligned}$$

Hence, every vector in the null space of A can written as,

$$\vec{x} = \begin{bmatrix} -2r - s - 2t \\ r \\ -s \\ s \\ t \\ 0 \end{bmatrix} = r \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -1 \\ 0 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -2 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} = r\vec{a} + s\vec{b} + t\vec{c}.$$

Since, $r,\ s,$ and t are free parameters, the entire null space of A is comprised of linear combinations of the three vectors $\vec{a},\ \vec{b},\ \vec{c}$, i.e.

$$N(A) = \left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\}.$$

[Note that depending on the choice of the free parameters, the resultant basis vectors could differ. However, we only need to make sure that the vectors are linearly independent and they indeed span the null space.] You can easily, verify that the vectors $\vec{a},\ \vec{b},\ \vec{c}$ are in the null space of U by multiplying each one of them by U. Also, by observation it is easy to see that these three vectors are linearly independent. Finally, the dimension of the null space is $\dim(N(A)) = n - r(A) = 6 - 3 = 3$ by the subtle theorem. Hence, we indeed found a basis for the null space of A.

(c) The idea here is to multiply L and U to get A, and then the columns of A corresponding to the pivot positions of U will form the column space of A. Therefore

$$A = LU = \begin{bmatrix} 1 & 2 & 0 & 1 & 2 & 1 \\ 2 & 4 & 2 & 4 & 4 & 2 \\ 2 & 4 & 2 & 4 & 4 & 3 \\ 3 & 6 & 4 & 7 & 6 & 7 \end{bmatrix}$$

The column space can therefore be taken as columns $\{1, 3, 6\}$ or $\{1, 4, 6\}$ of A, and column 3 may be scaled by $\frac{1}{2}$. Your answer here affects the next part so this was taken into account. (Note - many students took the columns from U and not A and this is incorrect.)

(d) There are a few ways to do this. The easiest way is if you notice that

$$\vec{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \\ 3 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 0 \\ 2 \\ 2 \\ 4 \end{bmatrix}$$

which means that it is indeed spanned by the columns of A. A more systematic way to arrive at this is by actually solving for the coefficients of the linear combinations of the columns in the column space of A which will give $\vec{b}$

$$\alpha_1 \begin{bmatrix} 1 \\ 2 \\ 2 \\ 3 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 \\ 2 \\ 2 \\ 4 \end{bmatrix} + \alpha_3 \begin{bmatrix} 1 \\ 2 \\ 3 \\ 7 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

There is no need to fully solve this system, only need to row reduce the augmented matrix and show that there is no inconsistent row.

$$\left(\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 2 & 2 & 2 & 1 \\ 2 & 2 & 3 & 1 \\ 3 & 4 & 7 & 1 \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 0 & 2 & 0 & -1 \\ 0 & 2 & 1 & -1 \\ 0 & 4 & 4 & -2 \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 0 & 2 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 4 & 0 \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 0 & 2 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Yet another way to do this is by looking at

$$\begin{aligned} Ax &= b \\ LUx &= b \\ Ly &= b \\ Ux &= y \end{aligned}$$

You can solve for y by performing forward substitution or computing the inverse of L and multiplying this to b. You should get

$$y = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}$$

Unfortunately, many people made mistakes in computing the inverse of L. The most common mistake was to take L and simply multiply all off-diagonal elements by -1. Careful though! This could be done for the elementary building blocks of L (the C_i's introduced in class) but it can *not* by done on L itself. You can easily convince yourself by multiplying the matrix you would get in this case with L. You will not get the identity.

At this point you can explicitly solve $Ux = y$ to get the solution or simply observe that wherever U has all zero rows (in this case only the last row) y also has zero component (the last two components). Therefore the system has at least one solution, in fact infinitely many solutions.

Problem 2

Indicate whether the following statements are TRUE or FALSE. Motivate your answers clearly.

(a) If the vectors $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m$ in R^n span a subspace S in R^n, the dimension of S is m.

(b) If for the $m \times n$ matrix A, the equation $A\vec{x} = \vec{b}$ always has at least one solution for every choice of $\vec{b}$, then the only solution to $A^T\vec{y} = \vec{0}$ is $\vec{y} = \vec{0}$.

(c) If the m vectors $\vec{x}_i, i = 1, \ldots, m$ are orthogonal, they are also independent. Note here that $\vec{x}_i \in \mathsf{R}^n$ and $n > m$.

(a) FALSE.

Take for example $\vec{x}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\vec{x}_2 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$. Then $m = 2$ but the span of $\{\vec{x}_1, \vec{x}_2\}$ only has dimension 1.

The crux here is to recognize that is just says the vectors *span* then subspace. It does not say the vectors form a basis. The vectors form a spanning set and this spanning set can be redundant.

(b) TRUE.

Since $A\vec{x} = \vec{b}$ has a solution for any $\vec{b}$, $rank(A) = m$. Then we know that $rank(A) + nullity(A^T) = m$ or $nullity(A^T) = 0$ which implies the only solution of $A^T\vec{y} = \vec{0}$ is $\vec{y} = \vec{0}$

(c) TRUE.

If $c_1, \ldots, c_m$ are scalars satisfying $c_1\vec{x}_1 + \ldots + c_m\vec{x}_m = \vec{0}$, then we may write

$$(c_1\vec{x}_1 + \ldots + c_m\vec{x}_m)^T\vec{x}_i = \vec{0}^T\vec{x}_i = \vec{0},$$

or, equivalently,

$$c_1(\vec{x}_1^T\vec{x}_i) + \ldots + c_i(\vec{x}_i^T\vec{x}_i) + \ldots + c_m(\vec{x}_m^T\vec{x}_i) = \vec{0}.$$

Since the vectors $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\}$ are orthogonal, all the inner products in the previous line vanish with the exception of $c_i(\vec{x}_i^T\vec{x}_i)$. Thus, we have

$$c_i(\vec{x}_i^T\vec{x}_i) = \vec{0}$$

Now, since $\vec{x}_i^T\vec{x}_i \neq 0$ (we implicitly know that the $\vec{x}_i$ are nonzero for the conclusion to hold), we must have $c_i = 0$. Since this argument holds for all $i = 1, \ldots, m$, we deduce that the set of vectors $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\}$ is linearly independent.

Note that we here implicitly assumed that none of the vectors are $\vec{0}$.

Problem 3

(a) (i) Using Gram-Schmidt orthogonalization, create an orthonormal basis for R^2 from the vectors

$$\vec{a}_1 = \begin{bmatrix} 3 \\ -4 \end{bmatrix}, \quad \text{and } \vec{a}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

(ii) Find the QR decomposition of the matrix

$$A = \begin{bmatrix} 3 & 1 & 1 \\ -4 & 1 & -1 \end{bmatrix},$$

where Q is a 2×2 matrix and R is 2×3.

(b) The system $A\vec{x} = \vec{b}$, with $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \vec{b} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$ has the solution $\vec{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. The vector $\vec{b}$ is now perturbed. The new vector is equal to $\vec{b} + \delta\vec{b}$ with $\delta\vec{b} = \begin{bmatrix} 0.001 \\ 0 \end{bmatrix}$. Estimate the maximum perturbation we can expect in th solution $\vec{x}$, measured in the vector 2-norm. For matrix-norms, you may use the Frobenius norm.

(a) (i) $\vec{q}_1 = \frac{\vec{a}_1}{\|\vec{a}_1\|_2} = \frac{1}{\sqrt{3^2+(-4)^2}} \begin{bmatrix} 3 \\ -4 \end{bmatrix} = \begin{bmatrix} 3/5 \\ -4/5 \end{bmatrix}$.

$\vec{w}_2 = a_2 - (\vec{q}_1^T \vec{a}_2)\vec{q}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \left(\begin{bmatrix} 3/5 & -4/5 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) \begin{bmatrix} 3/5 \\ -4/5 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} -3/25 \\ 4/25 \end{bmatrix} = \begin{bmatrix} 28/25 \\ 21/25 \end{bmatrix}$

Consider $\|\vec{w}_2\|_2 = \sqrt{\left(\frac{28}{25}\right)^2 + \left(\frac{21}{25}\right)^2} = \frac{7}{5}$. So $\vec{q}_2 = \frac{\vec{w}_2}{\|\vec{w}_2\|} = \begin{bmatrix} 4/5 \\ 3/5 \end{bmatrix}$.

Therefore, the orthogonal basis is $\left\{ \begin{bmatrix} 3/5 \\ -4/5 \end{bmatrix}, \begin{bmatrix} 4/5 \\ 3/5 \end{bmatrix} \right\}$.

(ii) The first two columns of matrix A are vectors $\vec{a}_1$ and $\vec{a}_2$ from part (i). We know that $\vec{a}_1, \vec{a}_2$ form a basis of R^2. So $Q = \begin{bmatrix} \vec{q}_1 & \vec{q}_2 \end{bmatrix} = \begin{bmatrix} 3/5 & 4/5 \\ -4/5 & 3/5 \end{bmatrix}$

We want to find R such that $A = QR$. So $R = Q^T A = \begin{bmatrix} 5 & -1/5 & 7/5 \\ 0 & 7/5 & 1/5 \end{bmatrix}$.

(b) Here, we use the perturbation result we derived in class:

$$\frac{\|\vec{\delta x}\|}{\|x\|} \leq \kappa(A) \frac{\|\delta\vec{b}\|}{\|\vec{b}\|},$$

We know that $\kappa(A) = \|A\| \|A^{-1}\|$. We will use the Frobenius norm in the evaluation. Clearly $\|A\|_F = \sqrt{2^2 + 1^2 + 2^2 + 1^2} = \sqrt{10}$. We also see that

$$A^{-1} = \frac{1}{3} \left[\begin{array}{cc} 2 & -1 \\ -1 & 2 \end{array} \right],$$

giving $\|A^{-1}\|_F = \frac{\sqrt{10}}{3}$, and so $\kappa(A) = 10/3$. Further we have that $\|\vec{x}\|_2 = \sqrt{2}, \|\vec{b}\|_2 = 3\sqrt{2}, \|\delta\vec{b}\|_2 = 0.001$. This gives $\|\vec{\delta x}\|_2 \leq \frac{10}{3}\sqrt{2}\frac{0.001}{3\sqrt{2}} = 0.01/9$.

Not everybody remembered this formula. A common mistake was that the formula was used not including norms, or with only some of the components including norms.