MIT 18.06 Exam 2, Fall 2018 **Solutions**

Johnson

Problem 1 (33 points):

The matrix A has a nullspace $N(A)$ spanned by

$$\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

and a left nullspace $N(A^T)$ spanned by

$$\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}.$$

(a) What is the **shape** of the matrix A and its **rank**?

(b) If we consider the vector

$$b = \begin{pmatrix} -1 \\ \alpha \\ 0 \\ \beta \end{pmatrix},$$

for **what value(s)** of α and β (if any) is $Ax = b$ solvable? Will the solution (if any) be **unique**?

(c) Give the orthogonal **projections** of

$$y = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}$$

onto **two** of the four fundamental subspaces of A.

Solution:

(a) Since $N(A)$ is a subspace of $\mathbb{R}^3$, the matrix A must have three columns. Since $N(A^T)$ is a subspace of $\mathbb{R}^4$, the matrix A must have four rows. So A is a $\boxed{4 \times 3}$ matrix. The matrix has 3 columns and the null space has dimension 1, and so the rank of the matrix is $r = 3 - 1 = \boxed{2}$.

(b) If $Ax = b$ is solvable, then $b \in C(A)$. Since $C(A)$ is the orthogonal complement of $N(A^T)$, this means that an equivalent condition for $Ax = b$ to be solvable is that b is orthogonal to $N(A^T)$. This gives us two constraints on b:

$$b^T \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = 0 \quad \Longrightarrow \quad -1 + \alpha + \beta = 0,$$

$$b^T \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} = 0 \quad \Longrightarrow \quad -1 + \alpha - \beta = 0.$$

Solving these, we find that $b \in C(A)$ requires $\boxed{\alpha = 1,\ \beta = 0}$. For these values of $\alpha$and β, the solution of $Ax = b$ is **not unique**, since $N(A)$ has dimension 1: given any particuar solution of $Ax = b$, we can add on any multiple of $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ and the resulting vector would still be a solution.

(c) The vector $y = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}$ is in $\mathbb{R}^3$, and so we can **only** project onto $N(A)$ and $C(A^T)$. To project onto $N(A)$, we use the formula to project y onto $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$:

$$p_{N(A)} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \frac{\begin{pmatrix} 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}}{\begin{pmatrix} 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}} = \boxed{\begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix}}.$$

Note, that to compute the projection $\frac{aa^T}{a^Ta}y$, it is *not* necessary to first compute the projection matrix $\frac{aa^T}{a^Ta}$ (although this is certainly allowed). It is a *lot less work* to compute the projection as $a\frac{(a^Ty)}{a^Ta}$ (two dot products) than as $(\frac{aa^T}{a^Ta})y$ (explicitly forming the matrix aa^T).

To compute the projection onto $C(A^T)$, recall that if $p = Py$ is the projection of y onto some subspace, then $(I - P)y$ will project y onto the orthogonal complement of this subspace. Since $C(A^T)$ is orthogonal to

$N(A)$, the projection of y onto $C(A^T)$ is given by:

$$\begin{aligned} p_{C(A^T)} &= \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} - p_{N(A)} \\ &= \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} - \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix} \\ &= \boxed{\begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}}. \end{aligned}$$

Note in particular that it was *not* necessary to explicitly find a basis for $C(A^T)$ or to form its 3×3 projection matrix, and in fact this would be a lot more work!

Problem 2 (34 points):

You have a matrix

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

(a) Give the **ranks** of A, A^T, and A^TA, and also give **bases** for $C(A)$, $N(A)$, and $N(A^TA)$. (**Look carefully at the columns** of A—very little calculation is needed!)

(b) Suppose we are looking for a least-square solution $\hat{x}$ that minimizes $\|b - Ax\|$ for $b = \begin{pmatrix} 0 \\ 2 \\ 1 \\ -1 \end{pmatrix}$. At this minimum, $p = A\hat{x}$ will be the projection of b onto ....? **Find** p. (**Hint**: your answer from (a) should help simplify the calculations.)

Solution:

(a) The first and third columns of A are the same, while the first and second columns are linearly independent. This means that the rank of A is $\boxed{2}$. The rank of $A^T$and the rank of A^TA are **equal** to the rank of A. A basis for $C(A)$ is then just the first two columns $\boxed{\begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \end{pmatrix}}$. The nullspace of

A is one dimensional, and since the first and third columns are the same, a basis for $N(A)$ is given by the vector $\boxed{\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}}$. Finally, $\boxed{N(A) = N(A^TA)}$ (a property we saw several times in class and homework!), and so our basis for $N(A)$ is also a basis for $N(A^TA)$.

(b) First, $p = A\hat{x}$ is the projection of b onto $\boxed{C(A)}$. To find $\hat{x}$ we must solve the normal equations $A^TA\hat{x} = A^Tb$. However, since A only has two linearly independent columns we can simplify our calculations by instead using the basis matrix $B = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix}$ with $C(B) = C(A)$, and solve the normal equations $B^TB\hat{x} = B^Tb$ to find $p = B\hat{x}$. We can calculate

$$B^TB = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 2 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 3 & 6 \end{pmatrix}$$

and

$$B^Tb = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 2 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 2 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \end{pmatrix}.$$

The normal equations are then:

$$\begin{aligned} \begin{pmatrix} 3 & 3 \\ 3 & 6 \end{pmatrix} \hat{x} &= \begin{pmatrix} 0 \\ 3 \end{pmatrix} \\ \implies \hat{x} &= \begin{pmatrix} -1 \\ 1 \end{pmatrix}. \end{aligned}$$

Finally, we can compute

$$p = B\hat{x} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \boxed{\begin{pmatrix} 1 \\ 1 \\ 0 \\ -1 \end{pmatrix}}.$$

Problem 3 (33 points):

Suppose that we apply Gram–Schmidt to the *rows* (in order from top to bottom) of a matrix

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

in order to find three **orthonormal row vectors** q_1^T, q_2^T, q_3^T.

(a) What is q_2?

(b) Suppose that these orthonormal vectors are the **rows** of a matrix $U = \begin{pmatrix} q_1^T \\ q_2^T \\ q_3^T \end{pmatrix}$. Then:

(i) **Circle any** of the following that are **true**: $U^TU = I$ and/or $UU^T = I$?

(ii) **Circle any** of the following that are **true**: $C(A) = C(U)$, $N(A) = N(U)$, $C(A^T) = C(U^T)$, and/or $N(A^T) = N(U^T)$?

(iii) Which is **true**: $A = BU$ or $A = UB$? Is B upper or lower triangular?

Solution:

(a) The rows of A are a^T, $b^T$and $c^T$where

$$a = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, c = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}.$$

We can use Gram-Schmidt to first find three orthogonal vectors u_1, u_2 and u_3. We let $u_1 = a$, then

$$u_2 = b - \frac{u_1^T b}{u_1{}^T u_1} u_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} - \frac{2}{3}\begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 \\ -2 \\ 0 \\ 1 \end{pmatrix}.$$

We can then normalize u_2 to obtain $q_2 = \frac{u_2}{\|u_2\|}$. Now $\|u_2\| = \frac{\sqrt{6}}{3}$ and so

$$q_2 = \boxed{\frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ -2 \\ 0 \\ 1 \end{pmatrix}}.$$

(b) Note that q_1, q_2 and q_3 are the orthonormal columns that you would obtain by applying Gram-Schmidt to the columns of A^T , i.e. $A^T = QR$, where the columns of Q are q_1, q_2 and q_3 . This means that $U = Q^T$.

(i) Since $Q^TQ = I$ for any orthogonal matrix, it follows that $\boxed{UU^T = I}$: this is the dot product of every row of U with every other row. However, since Q is not square, $QQ^T \neq I$ and so $U^TU \neq I$.

(ii) Since A has rank 3, it has full column rank. This means that $C(A) = \mathbb{R}^3$. U also has full column rank, meaning that $C(U) = \mathbb{R}^3$. So $C(A) = C(U)$. Since we obtain U by applying invertible row operations to A, both the null space and row space of A will be unchanged; equivalently, Gram–Schmidt *always* finds an **orthonormal basis for the same space**, in this case the row space. Hence $N(A) = N(U)$ and $C(A^T) = C(U^T)$. Finally, since A and U have full column rank, the dimension of their left nullspaces will both be 0 and so $N(A^T) = N(U^T)$ trivially.

(iii) We know that $A^T = QR$ where R is always upper triangular. Taking the transpose of this equation gives $A = R^T Q^T = BU$, and so $B = R^T$ is **lower triangular**.