MIT 18.06 Exam 1 Solutions, Fall 2022
Johnson

Problem 1 (6+6+6+6+6+6=36 points):

Fill in the blanks:

(a) *Any* solution x to $Ax = b$ (if it exists) is always a sum of a vector in the ______ space of A plus a vector in the ______ space of A.

(b) $Ax = b$ is solvable if (and only if) b is orthogonal to every vector in the ______ space of A.

(c) If A is a 4×3 matrix and $Ax = b$ is *not* solvable for some b and the solutions are *not* unique when they exist, possible values for the rank of A are ______ (list all possibilities).

(d) $C(AB)$ *must* ______ (**contain** $\supseteq$ / **be contained in** $\subseteq$ / **equal** $=$) the column space of ______ (A **or** B) for *all* 4×4 matrices A and B.

(e) If $x, y, z \in \mathbb{R}^n$ are n-component vectors, then the number of operations to compute $xy^T z$ scales proportional to ______ (n, n^2, or n^3) for large n if you compute it in the order $(xy^T)z$, or proportional to ______ (n, n^2, or n^3) if you compute it in the order $x(y^T z)$.

(f) If x_1 and x_2 are *both* solutions to $Ax = b$, then the vector $x_1 - x_2$ must be in the ______ space of A.

Solutions:

(a) Any solution is a sum of a vector in the **null space** of A plus a vector in the **row space** of A. The reason for this is that $N(A)$ and $C(A^T)$ are orthogonal complements, so together they span the whole space of possible "inputs" x to A (i.e. all of $\mathbb{R}^n$ if A is $m \times n$).

(b) b must be orthogonal to every vector in the **left nullspace** of A—since $N(A^T)$ is the orthogonal complement of $C(A)$, being orthogonal to $N(A^T)$ is equivalent to being in $C(A)$, and the condition for $Ax = b$ to be solvable is for $b \in C(A)$.

(c) A must be **rank deficient** for the solutions to not be unique *and* not necessarily exist, so the rank must be **0, 1, or 2**.

(d) $C(AB)$ must **be contained in** the column space of A, i.e. $C(AB) \subseteq C(A)$. Intuitively, the "output" of the linear operator AB comes from A, so its column space must be related to that of A (not B, which affects the "inputs" to A). More precisely, $C(AB)$ consists of vectors ABx for any x, but $ABx = A(Bx)$ is A times some vector, so it must be in $C(A)$, and hence $C(AB) \subseteq C(A)$.

The converse is not true unless B is invertible, however: $C(A) = C(AB)$ only if B is an invertible matrix, since that's the only way you can go from any vector Ay in $C(A)$ to a vector $Ay = AB(\text{some vector})$ in $C(AB)$, by setting $(\text{some vector}) = B^{-1}y$.

(e) $(xy^T)z$ requires $\boxed{\sim n^2}$ arithmetic operations (forming the $n \times n$ matrix xy^T requires n^2 multiplications, and then multiplying a matrix times the vector z is also $\sim n^2$), while $x(y^Tz)$ requires $\boxed{\sim n}$ arithmetic operations (the dot product y^Tz costs $\sim n$, and then multiplying the resulting scalar by the vector x costs another n multiplications).

(f) $x_1 - x_2$ must be in the **null space** of A. The only way $Ax = b$ can have multiple solutions is for them to differ by something in $N(A)$. We can see this explicitly from $A(x_1 - x_2) = Ax_1 - Ax_2 = b - b = \vec{0}$. (Many people answered "column space" here, which is a nonsensical answer: the column space contains vectors like b, which is not even the right "shape" for x if A is non-square.)

Hot tip: remember that for an $m \times n$ matrix takes inputs ("x") in $\mathbb{R}^n$ to outputs ($b = Ax$) in $\mathbb{R}^m$. The **"input" subspaces** of $\mathbb{R}^n$ are $N(A)$ and $C(A^T)$, while the **"output" subspaces** of $\mathbb{R}^m$ are $C(A)$ and $N(A^T)$. So if any problem is asking you about "inputs" (or things you add or dot-product with inputs), like parts (a) or (f), it must involve $N(A)$ and/or $C(A^T)$. And if any problem is asking you about "outputs" (or things you add or dot-product with outputs), like part (b), it must be asking about $C(A)$ and $N(A^T)$.

Another common mistake: many people also write something like "solution space" for some of the answers, which I guess means the set of solutions x to $Ax = b$. First, this makes no sense because the solutions depend on *both* A and b, so how could there be a "solution space of A" by itself? Second, the set of solutions x is *not* generally a subspace, because it does not include $x = 0$ (except in the special case $b = 0$, of course).

Problem 2 (6+11+6+11=34 points):

If

$$\underbrace{\begin{pmatrix} 1 & 2 & 4 & 2 & 5 \\ & 2 & 3 & 5 & 6 \\ & & 3 & 4 & 3 \\ & & & 4 & 3 \\ & & & & 5 \end{pmatrix}}_{A} B \underbrace{\begin{pmatrix} 4 & 1 & 1 \\ & 1 & 1 \\ & & 2 \end{pmatrix}}_{C} x = b$$

has the complete solution

$$x = \begin{pmatrix} 7 \\ 1 \\ 2 \end{pmatrix} + \alpha_1 \begin{pmatrix} 2 \\ 3 \\ -4 \end{pmatrix},$$

for any scalar α_1, then:

(a) What is the size and rank of B?

(b) The ________ space of B must be spanned by the basis ________.

(c) In part (b), you could **alternatively** have found a basis for the ________ space of B, which is also fully determined by the information given because it is ________ to your answer from (b).

(d) Give a possible matrix B.

Solutions:

(a) B must be 5×3 in order for ABC to make sense. From the complete solution, we have a 1d nullspace, so the 5×3 matrix ABC must have rank $r = 2$. A and C are clearly invertible matrices (upper triangular with all of their pivots), so they can't add any new vectors to the nullspace: the nullspace must come from B, so B must also be **rank 2**.

(b) As in the previous part, we know a vector in the nullspace from the α_1 term of the complete solution: we must have

$$ABC \begin{pmatrix} 2 \\ 3 \\ -4 \end{pmatrix} = \vec{0} \implies B \left[C \begin{pmatrix} 2 \\ 3 \\ -4 \end{pmatrix} \right] = A^{-1}\vec{0} = \vec{0},$$

but that means that

$$C \begin{pmatrix} 2 \\ 3 \\ -4 \end{pmatrix} = \begin{pmatrix} 4 \cdot 2 + 3 - 4 \\ 3 - 4 \\ 2 \cdot (-4) \end{pmatrix} = \boxed{\begin{pmatrix} 7 \\ -1 \\ -8 \end{pmatrix}}$$

must be in $N(B)$. Hence $\boxed{[7, -1, -8]}$ is a basis for the 1d **null space** of B.

(c) Once we know $N(B)$, we also know the **row space** of B, because the row space $C(B^T)$ is **orthogonal** to $N(B)$. That is, we just need the 2d subspace (plane) of vectors orthogonal to $[7, -1, -8]$, which we could find by computing the nullspace of the 1-row matrix $\begin{pmatrix} 7 & -1 & 8 \end{pmatrix}$.

(d) We don't know anything other than the size, rank, and nullspace / row space of B. So, we just need any 5×3 matrix of rank 2 (**two independent columns**) with $[7, -1, -8]$ in its nullspace. That is, if its columns are $B = \begin{pmatrix} c_1 & c_2 & c_3 \end{pmatrix}$, then we need $7c_1 - c_2 - 8c_3 = \vec{0}$, or equivalently $\boxed{c_2 = 7c_1 - 8c_3}$. So, we can just pick the first and third columns to be any independent (non-parallel) vectors we want—say, for example, $[1, 0, 0, 0, 0]$ and $[0, 1, 0, 0, 0]$—and compute the second column by this formula:

$$\boxed{B = \begin{pmatrix} 1 & 7 & 0 \\ 0 & -8 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}.$$

Of course, there are infinitely many other possible choices for B, but they should all share these basic features.

Problem 3 (5+6+13+6=30 points):

Consider the matrix $A = BCD$ given by:

$$A = \underbrace{\begin{pmatrix} 1 & 0 & 2 & 0 \\ & 1 & 0 & 3 \\ & & -1 & 0 \\ & & & 1 \end{pmatrix}}_{B} \underbrace{\begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & 0 & 1 \\ 0 & 2 & -1 & 0 \\ 1 & 4 & 0 & 1 \end{pmatrix}}_{C} \underbrace{\begin{pmatrix} 2 & & & \\ & 1 & & \\ & & -2 & \\ & & & 3 \end{pmatrix}}_{D}.$$

(a) Write A^{-1} in terms of B^{-1}, C^{-1}, and D^{-1} (without computing any numbers).

(b) To compute the **sum** x of the four columns of A^{-1}, you could solve $Ax = b$ for x using what right-hand-side vector b?

(c) Compute the sum of the columns of A^{-1}.

(d) A basis for the column space $C(A)$ is ______.

Solutions:

(a) $A^{-1} = \boxed{D^{-1}C^{-1}B^{-1}}$.

(b) The sum of the columns of A^{-1} is $x = A^{-1}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$, or equivalently the solution to $Ax = \boxed{\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}}$.

(c) We want to compute

$$x = A^{-1}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = D^{-1}\underbrace{C^{-1}\underbrace{B^{-1}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}}_{c}}_{d}.$$

As usual, we want to do this without explicitly multiplying or inverting matrices, since that is a lot more effort than we usually need. In particular, we can break it into the following steps. (You can get a lot of partial credit here just by outlining the steps below: backsubstitution, elimination, diagonal solve.)

(Note that explicitly multiplying $A = BCD$ and then solving $Ax = [1, 1, 1, 1]$ by elimination is also possible, but is substantially more work. Multiplying BC is more work than handling B separately via backsubstitution, and multiplying CD is also more work than solving D separately. Think of the arithmetic counts: backsubstitution takes $\sim n^2$ operations, while multiplying two matrices is $\sim n^3$. Similarly, computing $D^{-1}d$ requires only n divisions for a diagonal D, whereas multiplying C by a diagonal matrix takes n^2 multiplications. If the matrix is already factored for you into "nice" matrices, you waste a lot of effort if you throw that factorization away by multiplying the factors together!)

(i) Compute $c = B^{-1}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$. Since $B$is upper-triangular, we can do this by back-substitution, from bottom to top: $c_4 = 1$, $c_3 = -1$, $c_2 = 1 - 3c_4 = -2$, $c_1 = 1 - 2c_3 = 3$, so $c = \begin{pmatrix} 3 \\ -2 \\ -1 \\ 1 \end{pmatrix}$.

(ii) Compute $d = C^{-1}c$, i.e. solve $Cd = c$. We proceed by elimination, augmenting C with the right-hand side c:

$$\left(\begin{array}{c} \underbrace{\begin{matrix} 1 & 0 & 2 & 0 \\ 0 & 2 & 0 & 1 \\ 0 & 2 & -1 & 0 \\ 1 & 4 & 0 & 1 \end{matrix}}_{C} \underbrace{\begin{matrix} 3 \\ -2 \\ -1 \\ 1 \end{matrix}}_{c} \end{array}\right) \rightsquigarrow \begin{pmatrix} 1 & 0 & 2 & 0 & 3 \\ 0 & 2 & 0 & 1 & -2 \\ 0 & 2 & -1 & 0 & -1 \\ 0 & 4 & -2 & 1 & -2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 2 & 0 & 3 \\ 0 & 2 & 0 & 1 & -2 \\ 0 & 0 & -1 & -1 & 1 \\ 0 & 0 & -2 & -1 & 2 \end{pmatrix},$$

$$\rightsquigarrow \left(\begin{array}{c} \underbrace{\begin{matrix} 1 & 0 & 2 & 0 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & & 1 \end{matrix}}_{U} \begin{matrix} 3 \\ -2 \\ 1 \\ 0 \end{matrix} \end{array}\right)$$

which we then solve by backsubstitution (bottom to top), to get $d_4 = 0$, $d_3 = -1$, $d_2 = (-2 - d_4)/2 = -1$, $d_1 = 3 - 2d_3 = 5$, hence

$$d = \begin{pmatrix} 5 \\ -1 \\ -1 \\ 0 \end{pmatrix}.$$

(iii) Finally, compute $x = D^{-1}d$. Since D is a diagonal matrix, this is easy: we just divide d by the diagonal elements, yielding

$$x = \boxed{\begin{pmatrix} 5/2 \\ -1 \\ 1/2 \\ 0 \end{pmatrix}}.$$

(d) A is invertible, so $C(A)$ is all of $\mathbb{R}^4$. So any 4 linearly independent vectors work, e.g. the columns of B, C, or D, or simply the Cartesian basis (the columns of the 4×4 identity matrix I).