Problem Set 4 due March 18, at 10 PM, on Gradescope

Please list all of your sources: collaborators, written materials (other than our textbook and lecture notes) and online materials (other than Gilbert Strang's videos on OCW).

Give complete solutions, providing justifications for every step of the argument. Points will be deducted for insufficient explanation or answers that come out of the blue.

Problem 1:

Consider a matrix A such that the general solution to the equation:

$$A \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \end{bmatrix} \qquad \text{is} \qquad \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \lambda \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + \mu \begin{bmatrix} 0 \\ 3 \\ -1 \end{bmatrix}$$

for arbitrary numbers λ and μ.

(1) How many rows and columns does A have? *(5 points)*

(2) Based on the information in the equation above, what is the second column of A? *(5 points)*

(3) Find the entire matrix A. *(10 points)*

Solution: (1) Say A is an $m \times n$ matrix. We know that $n = 3$, else it would not make sense to multiply by the 3×1 matrix $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$. Then, by the rules of matrix multiplication (see (15) in Lecture 3):

$$\Big(m \times n \textbf{ matrix}\Big)\Big(n \times p \textbf{ matrix}\Big) = \Big(m \times p \textbf{ matrix}\Big)$$

we have:

$$\Big(m \times 3 \textbf{ matrix}\Big)\Big(3 \times 1 \textbf{ matrix}\Big) = \Big(2 \times 1 \textbf{ matrix}\Big)$$

so $m = 2$, i.e. A is a 2×3 matrix.

Grading Rubric: 2.5 points each for the correct number of rows and columns.

(2) The solution is written in the form:

$$\boldsymbol{v}_{\text{general}} = \boldsymbol{v}_{\text{particular}} + \boldsymbol{w}_{\text{general}}$$

(as in Fact 6, Lecture 8). Since the component of the solution in $N(A)$ is able to be scaled, but the particular solution is not, we conclude:

$$\boldsymbol{v}_{\text{particular}} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad \text{and} \qquad \boldsymbol{w}_{\text{general}} = \lambda \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + \mu \begin{bmatrix} 0 \\ 3 \\ -1 \end{bmatrix}$$

If we let:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

then the particular solution gives us:

$$\begin{bmatrix} 2 \\ -3 \end{bmatrix} = A \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a_{21} \\ a_{22} \end{bmatrix}$$

hence the second column of A is $\begin{bmatrix} 2 \\ -3 \end{bmatrix}$.

Grading Rubric:

- Correct answer with explanation of the fact that the second column is known to us from the particular solution *(5 points)*
- Correct answer with no useful explanation *(3 points)*
- Identified the particular solution, but did not use it to obtain the second column *(2 points)*
- Missing or incorrect argument *(0 points)*

(3) From part (2), we know that:

$$A = \begin{bmatrix} a_{11} & 2 & a_{13} \\ a_{21} & -3 & a_{23} \end{bmatrix}$$

From the particular solution, we know that the vectors $\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 3 \\ -1 \end{bmatrix}$ lie in the nullspace of A.

Hence:

$$A \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = 0 \quad \Rightarrow \begin{bmatrix} a_{11} + 4 \\ a_{21} - 6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$A \begin{bmatrix} 0 \\ 3 \\ -1 \end{bmatrix} = 0 \quad \Rightarrow \begin{bmatrix} 6 - a_{13} \\ -9 - a_{23} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

So we conclude that:

$$A = \begin{bmatrix} -4 & 2 & 6 \\ 6 & -3 & -9 \end{bmatrix}$$

Grading Rubric:

- Correct answer with explanation of the fact that the first and third columns are known to us from $\boldsymbol{w}_{\text{general}}$ *(10 points)*
- Correct answer and explanation, with minor algebra mistakes *(8 points)*
- Correct answer, but no explanation *(5 points)*

- Identified the role of $\boldsymbol{w}_{\text{general}}$, but did not use it to obtain the first and third columns *(3 points)*
- Missing or incorrect argument *(0 points)*

Problem 2:

An $m \times n$ matrix with rank 1 has the property that all its rows are multiples of each other, so they are of the form:

$$A = \begin{bmatrix} a_1 \boldsymbol{b}^T \\ \hline a_2 \boldsymbol{b}^T \\ \hline \dots \\ \hline a_m \boldsymbol{b}^T \end{bmatrix}$$

for some non-zero <u>row</u> vector $\boldsymbol{b}^T = \begin{bmatrix} b_1 & \dots & b_n \end{bmatrix}$ and some scalars $a_1, ..., a_m$, not all 0.

(1) Write A as the product of a $m \times 1$ matrix with an $1 \times n$ matrix. *(5 points)*

(2) Use part (1) to obtain a simple formula for the symmetric matrix $A^T A$ in terms of the a_i's and b_j's? *Hint: if you're stuck, try the case when A is 3×2 for some intuition.* *(10 points)*

(3) What is the rank of $A^T A$ from part (2)? Explain. *(5 points)*

Solution: (1) Written out fully, we have:

$$A = \begin{bmatrix} a_1 b_1 & a_1 b_2 & \dots & a_1 b_n \\ a_2 b_1 & a_2 b_2 & \dots & a_2 b_n \\ \vdots & \vdots & \ddots & \vdots \\ a_m b_1 & a_m b_2 & \dots & a_m b_n \end{bmatrix}$$

By the rules of matrix multiplication, we see that the explicit matrix above equals:

$$A = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} \begin{bmatrix} b_1 & \dots & b_n \end{bmatrix}$$

Grading Rubric: 2.5 points each for the correct row and column matrices.

(2) The formula above says that $A = \boldsymbol{a}\boldsymbol{b}^T$ where $\boldsymbol{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix}$ and $\boldsymbol{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$. We can then simplify, using the rule $(XY)^T = Y^T X^T$):

$$\begin{aligned} A^T A &= (\boldsymbol{a}\boldsymbol{b}^T)^T (\boldsymbol{a}\boldsymbol{b}^T) && (1) \\ &= \boldsymbol{b}\boldsymbol{a}^T \boldsymbol{a}\boldsymbol{b}^T && (2) \\ &= ||\boldsymbol{a}||^2 \boldsymbol{b}\boldsymbol{b}^T && (3) \end{aligned}$$

where in passing from the second to third lines we have used the formulation of the dot product in terms of matrix multiplication ($u \cdot v = u^T v$). In the last line, $\|\boldsymbol{b}\|^2$ denotes the length of $\boldsymbol{b}$ squared, i.e. $\|\boldsymbol{b}\|^2 = b_1^2 + \ldots + b_n^2$. Explicitly, the first term in (3) is a number and the second term in (3) is:

$$\boldsymbol{b}\boldsymbol{b}^T = \begin{bmatrix} b_1^2 & b_1 b_2 & \ldots & b_1 b_n \\ b_2 b_1 & b_2^2 & \ldots & b_2 b_n \\ \vdots & \vdots & \ddots & \vdots \\ b_n b_1 & b_n b_2 & \ldots & b_n^2 \end{bmatrix}$$

Grading Rubric:

- Correct computation and justification: either direct computation, or along the lines of equations (1)–(3) above *(10 points)*
- Correct computation, but final answer is not nicely factored (e.g. the a's and b's are intermingled) *(7 points)*
- Correct computation in a special case with $m, n \geq 2$ *(5 points)*
- Missing or incorrect computation *(0 points)*

(3) From either the explicit form of $\boldsymbol{b}\boldsymbol{b}^T$ or the fact that it is expressed as a product of an $n \times 1$ matrix with a $1 \times n$ matrix, we see all the columns are scalar multiples of $\boldsymbol{b}$. We therefore conclude its rank is 1. Since rank is unchanged when we multiply a matrix with the scalar $||\boldsymbol{a}||^2$, the answer is 1.

Grading Rubric:

- Correct answer with justification *(5 points)*
- Correct answer without justification *(3 points)*
- Incorrect answer *(0 points)*

Problem 3:

Find a basis for the vector space spanned by the vectors:

$$\begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} -3 \\ 6 \\ -3 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} -2 \\ 3 \\ -5 \\ 2 \end{bmatrix} \text{ and } \begin{bmatrix} 1 \\ -5 \\ -8 \\ 6 \end{bmatrix} \text{ and } \begin{bmatrix} 2 \\ -1 \\ 9 \\ -4 \end{bmatrix} \text{ and } \begin{bmatrix} 0 \\ 3 \\ 8 \\ -5 \end{bmatrix}$$

Explain your method. *(20 points)*

Solution: We may treat the vector space spanned by the set of vectors as the column space of the matrix:

$$A = \begin{bmatrix} 1 & -3 & -2 & 1 & 2 & 0 \\ -2 & 6 & 3 & -5 & -1 & 3 \\ 1 & -3 & -5 & -8 & 9 & 8 \\ 0 & 0 & 2 & 6 & -4 & -5 \end{bmatrix}$$

We can find the row echelon form of the matrix via row operations. Notice that although row operations do not preserve the column space, they **do** preserve linear relations between the columns. That is to say, if a column does not have a pivot after row reduction (so is linearly dependent on the columns to its left) then this must also have been the case before the row reduction. Row reduction gives:

$$\begin{bmatrix} 1 & -3 & -2 & 1 & 2 & 0 \\ -2 & 6 & 3 & -5 & -1 & 3 \\ 1 & -3 & -5 & -8 & 9 & 8 \\ 0 & 0 & 2 & 6 & -4 & -5 \end{bmatrix} \xrightarrow{r_2+2r_1} \begin{bmatrix} 1 & -3 & -2 & 1 & 2 & 0 \\ 0 & 0 & -1 & -3 & 3 & 3 \\ 1 & -3 & -5 & -8 & 9 & 8 \\ 0 & 0 & 2 & 6 & -4 & -5 \end{bmatrix} \xrightarrow{r_3-r_1}$$

$$\begin{bmatrix} 1 & -3 & -2 & 1 & 2 & 0 \\ 0 & 0 & -1 & -3 & 3 & 3 \\ 0 & 0 & -3 & -9 & 7 & 8 \\ 0 & 0 & 2 & 6 & -4 & -5 \end{bmatrix} \xrightarrow{r_3-3r_2} \begin{bmatrix} 1 & -3 & -2 & 1 & 2 & 0 \\ 0 & 0 & -1 & -3 & 3 & 3 \\ 0 & 0 & 0 & 0 & -2 & -1 \\ 0 & 0 & 2 & 6 & -4 & -5 \end{bmatrix} \xrightarrow{r_4+2r_2}$$

$$\begin{bmatrix} 1 & -3 & -2 & 1 & 2 & 0 \\ 0 & 0 & -1 & -3 & 3 & 3 \\ 0 & 0 & 0 & 0 & -2 & -1 \\ 0 & 0 & 0 & 0 & 2 & 1 \end{bmatrix} \xrightarrow{r_4+r_3} \begin{bmatrix} \boxed{1} & -3 & -2 & 1 & 2 & 0 \\ 0 & 0 & \boxed{-1} & -3 & 3 & 3 \\ 0 & 0 & 0 & 0 & \boxed{-2} & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The pivots (boxed) are on columns $1, 3, 5$. This means that all the other columns can be written as linear combinations of columns $1, 3, 5$. Since this is true after row reduction, it was also true before row reduction. So we conclude that the column space of A is spanned by its columns $1, 3, 5$, namely:

$$\begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -2 \\ 3 \\ -5 \\ 2 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2 \\ -1 \\ 9 \\ -4 \end{bmatrix}$$

Grading Rubric:

- Correct answer with justification *(20 points)*
- Correct answer with partial justification (e.g. showed a set of vectors which span the vector space in question, but didn't show that those vectors are linearly independent) *(10 points)*

We will accept other methods if they produce the correct answer, but only if the justification is complete and persuasive.

Problem 4:

For any two subspaces $V, W \subset \mathbb{R}^m$, we will write $V + W$ for the subspace consisting of all vectors of the form $\boldsymbol{v} + \boldsymbol{w}$, for arbitrary $\boldsymbol{v} \in V$ and $\boldsymbol{w} \in W$.

(1) If $\boldsymbol{v}_1, ..., \boldsymbol{v}_k$ is a basis of V and $\boldsymbol{w}_1, ..., \boldsymbol{w}_l$ is a basis of W, explain why the set $\boldsymbol{v}_1, ..., \boldsymbol{v}_k, \boldsymbol{w}_1, ..., \boldsymbol{w}_l$ spans the subspace $V + W$. *(5 points)*

(2) Show that the dimension of $V + W$ is $\leq$ the sum of dimensions of V and W. Give an example of when the inequality is strict (i.e. $\dim(V + W) < \dim V + \dim W$). *(10 points)*

(3) Given two matrices A and B of the same shape, what is the relation between $C(A + B)$ and the subspaces $C(A)$ and $C(B)$? *(5 points)*

(4) Use the previous parts to show that rank $(A + B) \leq$ rank $A +$ rank B. *(5 points)*

Solution: (1) The set $\boldsymbol{v}_1, \ldots \boldsymbol{v}_k, \boldsymbol{w}_1, \ldots \boldsymbol{w}_l$ spans the vector space $V + W$ means that any vector in $V + W$ can be written as a linear combination of $\boldsymbol{v}_1, \ldots \boldsymbol{v}_k, \boldsymbol{w}_1, \ldots \boldsymbol{w}_l$. By definition, any vector $\boldsymbol{a}$ in $V + W$ can be written:

$$\boldsymbol{a} = \boldsymbol{v} + \boldsymbol{w}$$

for some $\boldsymbol{v} \in V$, $\boldsymbol{w} \in W$. Then, since $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ are a basis of V (so in particular they span V) we can write:

$$\boldsymbol{v_a} = a_1\boldsymbol{v}_1 + \ldots + a_k\boldsymbol{v}_k$$

and since $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_l$ are a basis of W (so in particular they span) we can write:

$$\boldsymbol{w_a} = b_1\boldsymbol{w}_1 + \ldots + b_l\boldsymbol{w}_l$$

for some coefficients $a_1, \ldots, a_k, b_1, \ldots b_l$. Therefore we have written:

$$\boldsymbol{a} = \boldsymbol{v_a} + \boldsymbol{w_a} = a_1\boldsymbol{v}_1 + \ldots + a_k\boldsymbol{v}_k + b_1\boldsymbol{w}_1 + \ldots + b_l\boldsymbol{w}_l$$

which shows that $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_l$ span $V + W$.

Grading Rubric:

- Full explanation *(5 points)*
- Student explained what the problem is asking for, i.e. what it means to span $V + W$ *(3 points)*
- Unsatisfactory explanation *(0 points)*

(2) Since $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ and $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_l$ are bases of V and W, we see (from Definition 10 on dimension, Lecture 9) that $\dim V = k$ and $\dim W = l$. By part (1) we have shown a set of $k + l$ vectors spans $V + W$, and so we must have

$$\dim(V + W) \leq k + l = \dim V + \dim W.$$

This is because a basis can always be obtained from a spanning set by removing vectors until linear independence is achieved. The inequality above may be strict: consider the case of V being the $x - y$ plane and W being the $y - z$ plane in $\mathbb{R}^3$. That is, the span of $(1, 0, 0), (0, 1, 0)$ and the span of $(0, 1, 0), (0, 0, 1)$. In this case we have $\dim V = \dim W = 2$, but

$$\dim(V + W) = \dim \mathbb{R}^3 = 3 < 2 + 2 = 4.$$

Grading Rubric: 5 points each for the explanation and for the example.

(3) The columns of $A+B$ are the sum of the columns of A and the columns of B. Therefore, we conclude that:

$$C(A+B) \quad \text{is contained in} \quad C(A)+C(B)$$

Grading Rubric: 5 points for the inclusion above.

(4) By definition rank $(A+B) = \dim C(A+B)$. By part (3), $C(A+B)$ is a subspace of $C(A)+C(B)$. Therefore (since a subspace must have smaller dimension)

$$\text{rank } (A+B) = \dim C(A+B) \leq \dim(C(A)+C(B))$$

By part (2), the quantity above is less than or equal to:

$$\dim C(A) + \dim C(B) = \text{rank } A + \text{rank } B$$

Grading Rubric: 5 points for an explanation equivalent to the one above, or for any other complete justification.

Problem 5:

Does there exist a ..., as in each of (1), (2), (3) below? (*If the answer is yes, show such a matrix and explain why it has the required property. If the answer is no, explain why such a matrix doesn't exist*).

(1) 2×3 matrix with column space spanned by $\begin{bmatrix} 2 \\ 4 \end{bmatrix}$ and left nullspace spanned by $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$? *(5 points)*

(2) 3×2 matrix with row space spanned by $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and nullspace spanned by $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$? *(5 points)*

(3) 2×2 matrix with column space spanned by $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ and left nullspace spanned by $\begin{bmatrix} -2 \\ 3 \end{bmatrix}$? *(5 points)*

Solution: (1) No, because the column space is always orthogonal to the left nullspace, and:

$$\begin{bmatrix} 2 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 4 \neq 0$$

Grading Rubric: 5 points for a correct argument

(2) No, because the row space is always orthogonal to the nullspace, and:

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ -1 \end{bmatrix} = -1 \neq 0$$

Grading Rubric: 5 points for a correct argument

(3) Yes, it suffices to pick any matrix with column space spanned by $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$, e.g.:

$$A = \begin{bmatrix} 3 & 6 \\ 2 & 4 \end{bmatrix}$$

Because $\begin{bmatrix} -2 \\ 3 \end{bmatrix}$ is orthogonal to $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ and the latter spans the column space, then the former spans the left nullspace, as required (this is because $C(A)$ and $N(A^T)$ are orthogonal complements).

Grading Rubric:

- Correct matrix with explanation as to why the column space and left nullspace are as required *(5 points)*
- Correct matrix but without the aforementioned explanation *(3 points)*
- Incorrect matrix *(0 points)*