**Problem Set 4** due October 7, at 10 AM, on Gradescope (via Stellar)

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**: We will to reconstruct a matrix $A$ such that the general solution to the equation:

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

(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 $n = 2$, else it would not make sense to multiply by the $2 \times 1$ matrix $\begin{bmatrix}a\\b\end{bmatrix}$. Then, by the rules of matrix multiplication (see eq. 15, 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

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

so $m = 3$, i.e. $A$ is a $3 \times 2$ matrix.

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

(2) The solution is written in the form

$$v_{\text{gen}} = v_{\text{particular}} + w_{\text{general}}$$

(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

$$\begin{aligned} v_{\text{particular}} &= \begin{bmatrix}0\\1\end{bmatrix} \\ w_{\text{general}} &= \lambda\begin{bmatrix}1\\2\end{bmatrix}. \end{aligned}$$

Therefore, we must have

$$\begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} = A \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} A_{12} \\ A_{22} \\ A_{32} \end{bmatrix}.$$

That is, the second column of $A$ is $\begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$.

**Grading Rubric:** 3 points for identify the particular solution, 2 points for the correct computation.

(3) By the decomposition above, we know $\lambda \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is in $N(A)$. Therefore (for any $\lambda \neq 0$),

$$0 = A \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} A_{11} & 2 \\ A_{21} & -1 \\ A_{31} & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \lambda \begin{bmatrix} A_{11} + 2 \cdot 2 \\ A_{21} - 1 \cdot 2 \\ A_{31} + 3 \cdot 2 \end{bmatrix}.$$

From which we see the first column of $A$ must be $\begin{bmatrix} -4 \\ -2 \\ 6 \end{bmatrix}$ and

$$A = \begin{bmatrix} -4 & 2 \\ 2 & -1 \\ -6 & 3 \end{bmatrix}.$$

**Grading Rubric:** 5 points for the correct answer, 5 for the correct justification.

**Problem 2**: An $m \times n$ rank 1 matrix has the property that all its columns are multiples of each other, so they are of the form:

$$A = \left[ \; a_1\boldsymbol{b} \; \middle| \; a_2\boldsymbol{b} \; \middle| \; ... \; \middle| \; a_n\boldsymbol{b} \; \right]$$

for some non-zero vector $\boldsymbol{b} = \begin{bmatrix} b_1 \\ ... \\ b_m \end{bmatrix}$ and some scalars $a_1, ..., a_n$, 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* $3 \times 2$ *case for some intuition.* *(10 points)*

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

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

$$A = \begin{bmatrix} a_1 b_1 & a_2 b_1 & \dots & a_n b_1 \\ a_1 b_2 & a_2 b_2 & \dots & a_n b_2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1 b_m & a_2 b_m & \dots & a_n b_m \end{bmatrix}.$$

Then, by the rules of matrix multiplication, we see that for a $m \times 1$ matrix $\begin{bmatrix} b'_1 \\ ... \\ b'_m \end{bmatrix}$ and a $1 \times n$ matrix $\begin{bmatrix} a'_1 & ... & a'_n \end{bmatrix}$, their product is

$$\begin{bmatrix} b'_1 \\ ... \\ b'_m \end{bmatrix} \begin{bmatrix} a'_1 & ... & a'_n \end{bmatrix} = \begin{bmatrix} a'_1 b'_1 & a'_2 b'_1 & \dots & a'_n b'_1 \\ a'_1 b'_2 & a'_2 b'_2 & \dots & a'_n b'_2 \\ \vdots & \vdots & \ddots & \vdots \\ a'_1 b'_m & a'_2 b'_m & \dots & a'_n b'_m \end{bmatrix}.$$

We then see that we can simply write

$$A = \begin{bmatrix} b_1 \\ ... \\ b_m \end{bmatrix} \begin{bmatrix} a_1 & ... & a_n \end{bmatrix}$$

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

(2) By the above we have $A = \boldsymbol{b}\boldsymbol{a}^T$ where $\boldsymbol{a}$ and $\boldsymbol{b}$ are the column vectors given above. We can then simplify (using the rule that $(AB)^T = B^T A^T$))

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

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. $\|\boldsymbol{b}\|^2 = b_1^2 + \ldots + b_m^2$. Explicitly, the second term is

$$\boldsymbol{a}\boldsymbol{a}^T = \begin{bmatrix} a_1^2 & a_2 a_1 & \dots & a_n a_1 \\ a_1 a_2 & a_2^2 & \dots & a_n a_2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1 a_n & a_2 a_n & \dots & a_n^2 \end{bmatrix}.$$

**Grading Rubric:** 2 points for correctly taking the transpose, 3 points for identifying the dot product, 5 points for correct answer.

(3) From either the explicit form of $\boldsymbol{a}\boldsymbol{a}^T$ 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 $\boldsymbol{a}$. We therefore conclude the rank is 1.

**Grading Rubric:** 3 points for the correct answer, 2 for the justification.

**Problem 3**: Find a basis for the vector space spanned by the vectors:

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

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

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

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

$$\begin{aligned} \begin{bmatrix} 1 & 0 & 2 & 3 & -1 & 0 \\ 0 & 2 & -2 & 4 & 2 & 4 \\ -1 & -3 & 1 & 3 & -2 & -6 \\ 2 & 0 & 4 & 6 & 1 & 3 \end{bmatrix} &\xrightarrow{r_3+r_1, r_4-2r_1} \begin{bmatrix} 1 & 0 & 2 & 3 & -1 & 0 \\ 0 & 2 & -2 & -4 & 2 & 4 \\ 0 & -3 & 3 & 6 & -3 & -6 \\ 0 & 0 & 0 & 0 & 3 & 3 \end{bmatrix} \\ &\xrightarrow{r_3+\frac{3}{2}r_2} \begin{bmatrix} 1 & 0 & 2 & 3 & -1 & 0 \\ 0 & 2 & -2 & -4 & 2 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 3 \end{bmatrix} \\ &\xrightarrow{\frac{1}{2}r_2, \frac{1}{3}r_4, r_3\leftrightarrow r_4,} \begin{bmatrix} 1 & 0 & 2 & 3 & -1 & 0 \\ 0 & 1 & -1 & -2 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

Thus we see that after the row reduction columns 3 and 4 are in the span of columns 1 and 2, and column 6 is in the span of columns 1,2, and 5. This must therefore have been the case for the original matrix as well. That is, we find that columns 1,2, and 5 are a basis for the column space:

$$\begin{bmatrix} 1 \\ 0 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ -3 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ -2 \\ 1 \end{bmatrix}.$$

An alternative method is to use column operations. In this case, column reduction (without any column exchanges) results in

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ -1 & -\frac{3}{2} & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

from which we obtain a basis

$$\begin{bmatrix} 1 \\ 0 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ -3 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

which is the above but with the second colum, $c_2$, divided by 2, and the fifth, $c_5$, replaced with $c_5 - c_2 + c_1$.

**Grading Rubric:** 10 points for identifying a correct method. 10 points for correct computation.

**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. $\dim V + W < \dim V + \dim W$). *(5 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 $\boldsymbol{v}_1, \ldots \boldsymbol{v}_k, \boldsymbol{w}_1, \boldsymbol{w}_l$ spans if any vector in $V + W$ can be written as a linear combination of it. By definition, any vector $\boldsymbol{a}$ in $V + W$ can be written

$$\boldsymbol{a} = \boldsymbol{v_a} + \boldsymbol{w_a}$$

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

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

and

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

for the same reason. Therefore we have written

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

as a linear combination of $\boldsymbol{v}_1, \ldots \boldsymbol{v}_k, \boldsymbol{w}_1, \boldsymbol{w}_l$.

**Grading Rubric:** 3 points for writing $\boldsymbol{v_a}, \boldsymbol{w_a}$ in the given bases, 2 points for correct explanation.

(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 the definition of dimension, Lecture 9 Definition 10) that $\dim V = k$ and $\dim W = l$. By part (1) we have shown a set of $k + l$ 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:** 2.5 points each for the explanation and the example.

(3) The columns of $A+B$ are the sum of the columns of $A$ and the columns of $B$ respectively. In particular, they are linear combinations of the columns of $A$ and $B$. From Problem Set 3, Problem 4 (3), we know that the span of the columns of $A$ and the columns of $B$ together is the column space of the matrix

$$\begin{bmatrix} A & B \end{bmatrix}.$$

We then have

$$\text{span}\{\text{columns of } (A+B)\} \subseteq \text{span}\{\text{columns of } \begin{bmatrix} A & B \end{bmatrix}\}$$

which is to say

$$C(A+B) \text{ is contained in } C(\begin{bmatrix} A & B \end{bmatrix}) = C(A) + C(B).$$

(for the last equality, we have used the result of Problem Set 3 Problem 4(3)).

**Grading Rubric:** 3 points for a relation to $C(A)+C(B)$, 2 points for the correct relation (subspace of).

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

$$\text{rank}(A+B) = \dim C(A+B) \le \dim(C(A)+C(B)) \overset{by(2)}{\le} \dim C(A) + \dim C(B) = \text{rank}A + \text{rank}B.$$

**Grading Rubric:** 2.5 points each for the inequality of dimensions from being a subspace, and for correctly applying (2).

**Problem 5**: do problem 31 in Section 3.5 of the textbook. Make sure you explicitly describe the four subspaces for the matrix $A$ therein. *(20 points)*

**Solution:**

Suppose, as stated in the problem that two (block) matrices of size $m \times n$

$$A = \begin{bmatrix} I & F \\ 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} I & G \\ 0 & 0 \end{bmatrix}$$

have the same four fundamental subspaces, and both are in reduced row echelon form.

First we describe the four fundamental subspaces explicitly:

1. (The Column Space) Since the matrices are in reduced row echelon form, we can read off the column spaces

$$\begin{aligned} C(A) &= \text{span}\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k\} \\ C(B) &= \text{span}\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_l\} \end{aligned}$$

where $\boldsymbol{e}_i = (0, \ldots, 1, \ldots 0)$ where the 1 appears in the $i^{th}$ spot. Here, $k$ and $l$ are the size of the identity matrices in $A, B$ respectively (so in $A$ the $I$ is the $k \times k$ identity matrix).

2. (The Nullspace) Label the columns of $F$ (of which there are $n-k$) by $\mathbf{f}_1, \ldots, \mathbf{f}_{n-k}$. Likewise label the columns of $G$ (of which there are $n-l$) by $\mathbf{g}_1, \ldots, \mathbf{g}_{n-l}$. Since the matrices are in reduced row echelon form, we can immediate read off the nullspaces as

$$N(A) = \text{span}\left\{ \begin{bmatrix} -\mathbf{f}_1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} -\mathbf{f}_2 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots \begin{bmatrix} -\mathbf{f}_{n-k} \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \right\}$$

where $(f_i)$ is shorthand for the $i^{th}$ column vector of $F$ (which has $k$ entries). Likewise

$$N(B) = \text{span}\left\{ \begin{bmatrix} -\mathbf{g}_1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} -\mathbf{g}_2 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots \begin{bmatrix} -\mathbf{g}_{n-l} \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \right\}$$

3. (The Row Space) Since the matrices are in reduced row echelon form we can immediately see

$$R(A) = C(A^T) = \text{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ f_{11} \\ \vdots \\ f_{1,n-k} \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ f_{21} \\ \vdots \\ f_{2,n-k} \end{bmatrix}, \ldots \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ f_{k1} \\ \vdots \\ f_{k,n-k} \end{bmatrix} \right\}$$

and similarly for $R(B) = C(A^T)$ with $k$ replaced by $l$ and $f_{ij}$ replaced by $g_{ij}$.

4. (The Nullspace of $A^T$) We have

$$A^T = \begin{bmatrix} I & 0 \\ F^T & 0 \end{bmatrix}$$

from which we see the Nullspace contains vectors $\boldsymbol{e}_{k+1}, \boldsymbol{e}_{k+2}, \ldots, \boldsymbol{e}_m$. Since there are exactly $m-k$ of these, and we know the Nullspace of $A^T$ has dimension $m - \text{rank} = m-k$ (see top of page 181 in the textbook), these must be all of them. Therefore

$$N(A^T) = \text{span}\{\boldsymbol{e}_{k+1}, \ldots, \boldsymbol{e}_m\}$$

and likewise

$$N(B^T) = \text{span}\{\boldsymbol{e}_{l+1}, \ldots, \boldsymbol{e}_m\}.$$

Now we conclude that if the above four subspaces are equal, then $F = G$. From 1. we see that $k = l$, ie the identity matrices, and hence $F, G$ are the same size. Next, consider the row spaces. If the row spaces are equal, the first row of $A$ must be a linear combination of the rows of $B$. But since the first row of $A$ has an entry in the first column and zeros in the next $k-1$ columns, we see only the first row of $B$ can contribute to this linear combination. Since the top left entry is 1 for both, we conclude the first rows of $A$ and $B$ are equal. Likewise, the second row of $A$ must be a linear combination of the rows of $B$, but since the first and the third through $k^{th}$ entries are 0 of the second row are 0, we see only the second row of $B$ can contribute to the linear combination, and the scalar must again be 1. Therefore the second rows are equal as well. Proceeding in this fashion shows $F = G$.

**Grading Rubric:** 3 points for correct description of each subspace, 8 points for correct explanation of why $F = G$.