Exam 1

Part I: True/False (5 points each; 25 points)

For each of the following mark as true or false.

a) _____ $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ for an $n \times n$ matrices A and B, where

$$\operatorname{tr}(C) = \sum_{i=1}^{n} C_{ii} = \text{the sum of the diagonal elements of } C.$$

This is true. This is just a computation. $(AB)_{ii} = \sum_{k=1}^{n} A_{ik}B_{ki}$, so

$$\operatorname{tr}(AB) = \sum_{i=1}^{n}\sum_{k=1}^{n} A_{ik}B_{ki}$$

and

$$\operatorname{tr}(BA) = \sum_{i=1}^{n}\sum_{k=1}^{n} B_{ik}A_{ki} = \sum_{k=1}^{n} A_{ki}B_{ik} = \operatorname{tr}(AB).$$

b) _____ $\operatorname{tr}(AB) = \operatorname{tr}(A)\operatorname{tr}(B)$ for an $n \times n$ matrices A, B, and C.

Interestingly, this is false, as an example can show. In fact, generating any three random 2×2 matrices with entries from $\{-1, 0, 1\}$ are likely to work. Try this using MATLAB: round(rand(2)). The first three matrices I got this way were:

$$A = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} \quad AB = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}$$

So, $\operatorname{tr}(AB) = -1$ while $\operatorname{tr}(A)\operatorname{tr}(B) = (-1)(0) = 0$.

c) _____ If W is a subspace of a vector space V and $\mathcal{B}$ is a basis for V, then B can be restricted to a basis for W.

This is false. Let $W = \operatorname{span}\{(1, 1)\} \subseteq \mathbb{R}^2 = V$. The standard basis for $\mathbb{R}^2$ can not be restricted to a basis for W.

d) _____ If W is a subspace of a vector space V, then there is a subspace U so that $V = W \oplus U$.

This notation is a little hard to find in your text: $V = U + W$ means that for all $\boldsymbol{v} \in V$, there is $\boldsymbol{u} \in U$ and $\boldsymbol{w} \in W$ so that $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w}$. $V = U \oplus V$ means $V = U + V$ **and** $U \cap W = \{\mathbf{0}\}$, equivalently, for every $\boldsymbol{v} \in V$, there is a **unique** $\boldsymbol{u} \in U$ and $\boldsymbol{w} \in W$ so that $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w}$.

This is true. Let $\mathcal{B}_W$ be a basis for W and extend $\mathcal{B}_W$ to $\mathcal{B}_V$ a basis for V. Then let $U = \mathrm{span}(\mathcal{B}_V - \cap B_W)$. It is clear that $V = W \oplus U$.

e) _____ For any $m \times n$ matrices A and B,

$$B = EA \text{ for some invertible } E \iff \mathrm{NS}(A) = \mathrm{NS}(B).$$

This is true. ($\Rightarrow$) is trivial, since if $B = EA$, then $B\boldsymbol{x} = \mathbf{0} \iff EA\boldsymbol{x} = \mathbf{0} \iff A\boldsymbol{x} = E^{-1}\mathbf{0} = 0$. ($\Leftarrow$) is discussed below in the "Proofs" section.

Part II: Definitions and Theorems (5 points each; 25 points)

a) Define what it means for a set of vectors $\mathcal{B} = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ from a real vector space V to span V.

$\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ **spans** V iff for all $\boldsymbol{v} \in V$, $\boldsymbol{v}$ is a linear combination of the vectors in $\mathcal{B}$, that is $\boldsymbol{v} = \sum_{i=1}^{n} \alpha_i \boldsymbol{v}_i$ for some coefficients $\alpha_i \in \mathbb{R}$.

b) Define what it means for a set of vectors $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ from a real vector space V to be linearly independent.

A set of vectors $\mathcal{B}$ is **linearly independent** iff $\sum_{i=1}^{n} \alpha_i \boldsymbol{v}_i = \boldsymbol{0}$, then $\alpha_i = 0$ for all i. Equivalently, any linear combination of the vectors that gives $\boldsymbol{0}$ must be trivial.

There are equivalent definitions, but essentially, this is what you need. You can say something like, for all i, $\boldsymbol{v}_i \notin \operatorname{span}\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_{i-1}, \boldsymbol{v}_{i+1}, \dots, \boldsymbol{v}_n\}$

c) Define what it means for a set of vectors $\mathcal{B} = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ to be a basis for a vector space V.

$\mathcal{B}$ has must be a linearly independent and span V.

d) State the Rank-Nullity Theorem.

If A is an $m \times n$ matrix, then $n = \dim(\mathrm{RS}(A)) + \dim(\mathrm{NS}(A)) = \operatorname{rank}(A) + \operatorname{nullity}(A)$.

e) What conditions must be checked to verify that $W \subseteq V$ is a subspace of a vector space. V

Closure under addition and scalar multiplication must be checked.

Part III: Computational (15 points each; 45 point)

a) Given that A is a 3×4 matrix and

$$\mathrm{NS}(A) = \mathrm{span}\left(\left\{\begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 0 \\ 1 \end{bmatrix}\right\}\right)$$

compute $\mathrm{rref}(A)$. Make sure to explain how you arrive at your result. You may use (a) from the "Proofs" part below.

Notation: $\mathrm{rref}(A)$ means the reduced row echelon form of A. This is unique, a general echelon form is not unique. From $\mathrm{rref}(A)$ there is a simple way to read off a basis for $\mathrm{NS}(A)$, this exercise asks you to reverse that process.

We know $A\boldsymbol{x} = \mathbf{0} \iff \mathrm{rref}(A)\boldsymbol{x} = \mathbf{0} \iff \boldsymbol{x} \in \mathrm{NS}(A)$. From what we are given we see $\boldsymbol{x} \in \mathrm{NS}(A)$ iff

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = r\begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix} + s\begin{bmatrix} 2 \\ 3 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} r + 2s \\ -2r + 3s \\ r \\ s \end{bmatrix}$$

Working backwards from what we usually do we see $x_4 = s, x_3 = r$, and so $x_2 = -2x_3 + 3x_4$ and $x_1 = x_3 + 2x_4$. This gives the system

$$\begin{aligned} x_1 - x_3 - 2x_4 &= 0 \\ x_2 + 2x_3 - 3x_4 &= 0 \end{aligned}$$

This corresponds to $B\boldsymbol{x} = \mathbf{0}$ for

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

B is rref and so by Proofs (a) $B = \mathrm{rref}(A)$.

b) For the same (unknown) A used in (a) for each of $\mathrm{RS}(A)$ and $\mathrm{CS}(A)$ find a basis if possible and explain how you know that you have found a basis; if it is not possible to find a basis, then explain why it is not.

$\mathrm{RS}(A)$: Here we know $\mathrm{RS}(A) = \mathrm{RS}(\mathrm{rref}(A))$ so the non-zero rows of $\mathrm{rref}(A)$ form a basis for $\mathrm{RS}(A)$.

$\mathrm{CS}(A)$: You know that the first two columns of $\mathrm{rref}(A)$ are where the pivots are and so the first two columns of A would be a basis for $\mathrm{CS}(A)$, but you have no way of finding these and you know nothing about $\mathrm{CS}(A)$ other than $\dim(\mathrm{CS}(A)) = 2$. For example, $\mathrm{rref}(A_1) = \mathrm{rref}(A_2) = B$ for

$$A_1 = \begin{bmatrix} 1 & 1 & 1 & -5 \\ 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & -3 \end{bmatrix} \text{ and } A_2 = \begin{bmatrix} 0 & 1 & -1 & -2 \\ 1 & 0 & -1 & -2 \\ 1 & 2 & 3 & -8 \end{bmatrix},$$

but

$$\mathrm{CS}(A_1) = \mathrm{span}\left(\left\{\left[\begin{smallmatrix}1\\1\\0\end{smallmatrix}\right], \left[\begin{smallmatrix}1\\0\\1\end{smallmatrix}\right]\right\}\right) \text{ and } \mathrm{CS}(A_2) = \mathrm{span}\left(\left\{\left[\begin{smallmatrix}0\\1\\1\end{smallmatrix}\right], \left[\begin{smallmatrix}1\\0\\2\end{smallmatrix}\right]\right\}\right),$$

and $\left[\begin{smallmatrix}1\\0\\2\end{smallmatrix}\right] \notin \mathrm{CS}(A_1)$ so $\mathrm{CS}(A_1) \neq \mathrm{CS}(A_2)$.

c) Show that the upper-triangular $n \times n$ matrices form a subspace of all $n \times n$ matrices and find a basis for this subspace.

Let U be the collection of upper-triangular $n \times n$ matrices, that is $A \in U \iff A(i,j) = 0$ for $n \geq i > j \geq 1$.

It is clear that $A + B \in U$ and $\alpha A \in U$ for any $A, B \in U$. Let $A, B \in U$:

$$(A+B)(i,j) = A(i,j) + B(i,j) = 0 + 0 = 0 \text{ for } n \geq i > j \geq 1$$

$$(\alpha A)(i,j) = \alpha A(i,j) = 0 \text{ for } n \geq i > j \geq 1$$

A basis for U is $\{E_{i,j} \mid 1 \leq i \leq j \leq n\}$ where $E_{i,j}(l,m) = \delta_{(i,j),(l,m)}$. It is clear that if $A \in U$, then $A = \sum_{1 \leq i \leq j \leq n} A(i,j) E_{i,j}$.

Note: $\delta_{a,b} = \begin{cases} 1 & \text{if } a = b \\ 0 & \text{otherwise} \end{cases}$.

Part IV: Proofs (15 points each; 60 points)

Provide complete arguments/proofs for the following.

a) Show that if A and B are 3×4 rref matrices, then

$$\boxed{A = B \iff \mathrm{NS}(A) = \mathrm{NS}(B).}$$

Note: the 3×4 is a red-herring, this holds for arbitrary $m \times n$ matrices. If this helps you , then just prove this more general result. Also notice that this actually gives

$$\mathrm{rref}(A) = \mathrm{rref}(B) \iff \mathrm{NS}(A) = \mathrm{NS}(B)$$

since $\mathrm{NS}(\mathrm{rref}(A)) = \mathrm{NS}(A)$, trivially.

($\Rightarrow$) is trivial. For ($\Leftarrow$) we prove the contrapositive, namely, suppose $A \neq B$, then $\mathrm{NS}(A) \neq \mathrm{NS}(B)$. To start recall $\mathrm{RS}(A) = \mathrm{NS}(A)^\perp = \mathrm{NS}(B)^\perp = \mathrm{RS}(B)$. We also know that if A is in rref form, then the non-zero rows of A are a basis for $\mathrm{RS}(A)$ and hence for $\mathrm{RS}(B)$.

Suppose $A \neq B$ and let i be the first row on which they differ. Let $r = \mathrm{rank}(A)$, this is $\dim(A)$ and as A is rref we know this is the number of non-zero rows of A.

Claim 1: The pivots (leading 1's) occur at the same places in A and B.

If this fails, let i be the first row where $R_i(A) \neq R_i(B)$, where $R_i(A) = i^{\text{th}}$ row of A. We may assume the first 1 of $R_i(A)$ occurs before the first 1 of $R_i(B)$. (Else just swap the roles of A and B.) Let k be the position of the pivot (leading 1) in $R_i(A)$. We know $R_i(A) = \sum_{i=1}^{r} c_j R_j(B)$, then clearly $c_j = 0$ for $j = 1, \ldots, i-1$. This is because all entries up to $k-1$ are 0's in $R_i(A)$ while all of $R_j(B)$ for $j < k$ have a leading 1 before k.

So $R_i(A) = \sum_{j=k}^{r} c_j R_j(B)$, but for $j \geq k$ we know the first k entries of $R_j(B)$ are 0's so it is impossible to get a 1 in the k^{th} position.

Claim 2: $R_i(A) = R_i(B)$ for all $i \leq r$.

This claim is actually trivial given the first. We know $R_i(A) = \sum_{j=1}^{r} c_j R_j(B)$, but $R_i(A)$ has 0 at all pivot places except for that of $R_i(B)$, so we must have $R_i(A) = c_i R_i(B)$. Since the leading non-zero element is 1 we have $C_i = 1$, so $R_i(A) = R_i(B)$.

b) **Prove:** For $m \times n$ matrices A and B define $A \underset{\text{row}}{\sim} B$ to mean that you can get from A to B by a series of elementary row operations. Use the $m \times n$ version of (a), namely: $\mathrm{rref}(A) = \mathrm{rref}(B) \iff \mathrm{NS}(A) = \mathrm{NS}(B)$ to show that

$$\boxed{A \underset{\text{row}}{\sim} B \iff \mathrm{rref}(A) = \mathrm{rref}(B)}$$

Remark: Using elementary matrices one can show

$$A \underset{\text{row}}{\sim} B \iff A = EB \text{ for some invertible matrix } E$$

This is done in the text and in my notes. So you get here

$$A \underset{\text{row}}{\sim} B \iff \text{rref}(A) = \text{rref}(B) \iff A = EB \text{ for some invertible matrix } E$$

($\Rightarrow$) Clearly if $A \underset{\text{row}}{\sim} B$, then $\text{NS}(A) = \text{NS}(B)$ since elementary row ops generate equivalent systems of equations. (This was the whole point of row operations in the first place!)

($\Leftarrow$), this is really trivial. Do a series of operations to get from A to $\text{rref}(A)$ and from B to $\text{rref}(B)$, then just reverse the series from B to $\text{rref}(B) = \text{rref}(A)$ to get back to B. Combining these you get a series of row ops that goes from A to $\text{rref}(A)$ and then from $\text{rref}(A) = \text{rref}(B)$ back to B.

c) **Prove:** Let A be an $m \times n$ matrix, $\mathbb{R}^n = \text{NS}(A) \oplus \text{RS}(A)$.

Recall: $V = U \oplus W$ means $V = U + W = \{\boldsymbol{u} + \boldsymbol{w} \mid \boldsymbol{u} \in U \text{ and } \boldsymbol{w} \in W\}$ and $U \cap W = \{\boldsymbol{0}\}$.

By the rank-nullity theorem if $\mathcal{B}$ a basis for $\text{RS}(A)$ and $\mathcal{C}$ a basis for $\text{NS}(A)$, then $\mathcal{B} \cup \mathcal{C}$ has size n.

If we can show that $\mathcal{B} \cup \mathcal{C}$ is linearly independent, then we have that $\mathcal{B} \cup \mathcal{C}$ is a basis for $\mathbb{R}^n$ and so $\text{RS}(A) \oplus \text{NS}(A) = \mathbb{R}^n$.

We just need to see that $\text{RS}(A) \cap \text{NS}(A) = \{\boldsymbol{0}\}$. The simplest thing here is to note that since $A\boldsymbol{x} = \boldsymbol{0}$ for an $\boldsymbol{x} \in \text{NS}(A)$, then $\boldsymbol{r}_i \perp \boldsymbol{x}$ where $\boldsymbol{r}_i^T$ is the i^{th} row of A. So $\text{RS}(A) \perp \text{NS}(A)$.

d) **Prove:** If A is an $n \times n$ matrix and $A^k = \boldsymbol{0}$ for any k, then $A^n = \boldsymbol{0}$.

Proof 1: To do this show

i) Show $\text{NS}(A^{m+1}) \supseteq \text{NS}(A^m)$ for all m.

ii) Show that if $\text{NS}(A^{m+1}) = \text{NS}(A^m)$, then $\text{NS}(A^n) = \text{NS}(A^m)$ for all $n \geq m$.

It is clear that $\text{NS}(A^{m+1}) \supseteq \text{NS}(A^m)$, since $A^m\boldsymbol{x} = \boldsymbol{0} \implies A(A^m\boldsymbol{x}) = \boldsymbol{0} \implies A^{m+1}\boldsymbol{x} = \boldsymbol{0}$. So (i) is shown,

For (ii) suppose $\text{NS}(A^m) = \text{NS}(A^{m+1})$, then $A^{m+2}\boldsymbol{x} = \boldsymbol{0} \implies A^{m+1}(A\boldsymbol{x}) = \boldsymbol{0} \implies A^m(A\boldsymbol{x}) = \boldsymbol{0} \implies A^{m+1}\boldsymbol{x} = \boldsymbol{0}$. So $\text{NS}(A^{m+2}) \subseteq \text{NS}(A^{m+1})$, but then $\text{NS}(A^{m+2}) = \text{NS}(A^{m+1}) = \text{NS}(A^m)$. Now just keep going to get $\text{NS}(A^k) = \text{NS}(A^m)$ for all $k \geq m$.

This means we have

$$\text{NS}(A^0) \subsetneq \text{NS}(A^1) \subsetneq \text{NS}(A^2) \subsetneq \cdots \subsetneq \text{NS}(A^{m-1}) \subsetneq \text{NS}(A^m) = \text{NS}(A^{m+1}) = \cdots$$

The m at which $\text{NS}(A^k) = \text{NS}(A^m)$ for all $m \geq k$ must itself be $\leq n$.

If $A^k = \boldsymbol{0}$ for any k, then $\text{NS}(A^k) = \mathbb{R}^n$ is maximal and thus $m \leq k$ and $\text{NS}(A^m) = \mathbb{R}^n$. Since $m \leq n$, $\text{NS}(A^n) = \mathbb{R}^n$ and so $A^n = \boldsymbol{0}$.

Proof 2: You can use induction. To do this we need to prove something that sounds slightly stronger:

P_n : For any $n \times n$ matrix A, if $A^m = \mathbf{0}$ for any $m > n$, then $A^n = 0$.

base case: $(\mathbf{n} = \mathbf{1})$ If $A^m = [a]^m = [a^m] = [0]$, for m>1, then $a = 0$, so $A^1 = [a] = [0]$ as needed.

inductive step: Suppose P_{n-1}: For any $m > n-1$, $A^m = \mathbf{0} \implies A^{n-1} = \mathbf{0}$ for all $(n-1) \times (n-1)$ matrices. We want to prove P_n.

Assume A is an $n \times n$ matrix and $A^m = 0$ for some $m > n$. Notice that $\ker(A) \neq \{\mathbf{0}\}$, since if $\ker(A) = \{0\}$, then $A : \mathbb{R}^n \to \mathbb{R}^n$ is injective and thus A^m is also injective, so $\ker(A^m) = \{\mathbf{0}\}$. This obviously contradicts $A^m = \mathbf{0}$.

Let $\boldsymbol{v}_1 \in \ker(A)$ and let $\mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ be a basis for $\mathbb{R}^n$. So letting $B = \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n \end{bmatrix}$ we have

$$A = BA'B^{-1}$$

where

$$A' = \begin{bmatrix} 0 & a'_{12} & a'_{13} & \cdots & a'_{1n} \\ 0 & & & & \\ \vdots & & \hat{A} & & \\ 0 & & & & \end{bmatrix}$$

where $\hat{A}$ is the indicated $(n-1) \times (n-1)$ submatrix of A'.

A' is the matrix of $L(\boldsymbol{x}) = A\boldsymbol{x}$ with respect to the basis $\mathcal{B}$. Notice that $A^m = \mathbf{0}$ means $L^m(\boldsymbol{x}) = \mathbf{0}$ for all $\boldsymbol{x}$ and hence $A'\boldsymbol{x} = \mathbf{0}$ for all $\boldsymbol{x}$, a finally this means $A'^m = \mathbf{0}$.

Notice that A' has the block form

$$\begin{bmatrix} 0 & \boldsymbol{b}^T \\ \mathbf{0} & \hat{A} \end{bmatrix}$$

and

$$(A')^2 = \begin{bmatrix} 0 & \boldsymbol{b}^T \\ \mathbf{0} & \hat{A} \end{bmatrix} \begin{bmatrix} 0 & \boldsymbol{b}^T \\ \mathbf{0} & \hat{A} \end{bmatrix} = \begin{bmatrix} 0 & \boldsymbol{b}^T\hat{A} \\ \mathbf{0} & \hat{A}^2 \end{bmatrix}$$

By a second, obvious induction,

$$(A')^k = \begin{bmatrix} 0 & \boldsymbol{b}^T\hat{A}^{k-1} \\ \mathbf{0} & \hat{A}^k \end{bmatrix}$$

We assume $\hat{A'}^m = \mathbf{0}$ so $\hat{A}^m = \mathbf{0}$ and by induction $\hat{A}^{n-1} = \mathbf{0}$ and thus

$$(A')^n = \begin{bmatrix} 0 & \boldsymbol{b}^T\hat{A}^{n-1} \\ \mathbf{0} & \hat{A}^n \end{bmatrix} = \mathbf{0}$$

Proof 3: This is not a proof, I would expect to see, it uses stuff we have not covered. However, it seems some students find this proof and then can't quite carry it out, so I will also indicate the main error.

It is true that is λ is an eigenvalue for A and $\boldsymbol{x}$ and eigenvector for λ, then $A^k\boldsymbol{x} = \lambda^n\boldsymbol{x} = 0\boldsymbol{x} = 0$, so $\lambda^n = 0$ and hence $\lambda = 0$. But this does not mean that $A = 0$. In fact, here is a matrix where $A^n = 0$,

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

You can think about what A "does"

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \underset{A}{\rightarrow} \begin{bmatrix} 0 \\ x_1 \\ \vdots \\ x_{n-1} \end{bmatrix} \underset{A}{\rightarrow} \cdots \underset{A}{\rightarrow} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ x_1 \end{bmatrix} \underset{A}{\rightarrow} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

So $A^n = 0$, but $A^m \neq 0$ for any $m < n$ and the only eigenvalue is $\lambda = 0$.

What you can argue is that if $p(t) = a_nx^n + \cdots a_1x + a_0$ is the characteristic function, then $p(A) = 0$, Cayley-Hamilton theorem. Since 0 is the only eigenvalue $p(x) = a_nx^n$, so we know $a_nA^n = 0$, thus $A^n = 0$.