PROOFS

ABSTRACT. The following are proofs you should be familiar with for the midterm and final exam. On both the midterm and final exam there will be a proof to write out which will be similar to one of the following proofs.

1. PROOFS YOU ARE RESPONSIBLE FOR ON THE MIDTERM AND FINAL

Theorem 1.1. *For A,B nxn matrices, and s a scalar*

(1) $Tr(A+B) = Tr(A) + Tr(B)$
(2) $Tr(sA) = sTr(A)$
(3) $Tr(A^T) = Tr(A)$

Proof — Note that the (i,i)-th entry of $A+B$ is $a_{ii}+b_{ii}$, the (i,i)-th entry of cA is ca_{ii}, and the (i,i)-th entry of A^T is a_{ii}. Then consider the following:

$$\begin{array}{lrcl}
(a) & Tr(A+B) & = & (a_{11}+b_{11}) + ... + (a_{nn}+b_{nn}) \\
 & & = & (a_{11}+...+a_{nn}) + (b_{11}+...+b_{nn}) \\
 & & = & Tr(A) + Tr(B) \\
(b) & Tr(cA) & = & (ca_{11}+...+ca_{nn}) \\
 & & = & c(a_{11}+...+a_{nn}) \\
 & & = & cTr(A) \\
(c) & Tr(A^T) & = & a_{11}+...+a_{nn} \\
 & & = & Tr(A)
\end{array}$$

□

Theorem 1.2. *If you have a set of n vectors $S = \{v_1, ..., v_n\}$ in $\mathbb{R}^m$ where $n > m$, then the set S of vectors is linearly dependent.*

Proof — Begin by constructing the following mxn matrix

$$A = \begin{pmatrix} v_1 & \dots & v_n \\ | & \dots & | \end{pmatrix}$$

Since this matrix has m rows and there can be at most one pivot per row, it follows that $Rk(A) \leq m < n =$ number of columns. By a theorem from the book (Thm 1.8) we know that the the columns of a matrix are linearly independent if and only if the rank of the matrix is equal to the number of columns. The negation of this statement is that the columns of a matrix are linearly dependent if and only if the rank of the matrix is strictly less than the number of columns (by definition it is never possible for the rank of a matrix to be strictly more than the number of columns as there is at most one pivot per column). However, above we have shown that for the matrix A, the rank is in fact strictly less than the number of columns thereby immediately implying that the columns of A are linearly dependent or in other words the set $S = \{v_1, ..., v_n\}$ in $\mathbb{R}^m$ is linearly dependent.

□

Theorem 1.3. *If A, B are invertible nxn matrices, then:*

(1) A^{-1} *is invertible with inverse A*
(2) *AB is invertible with inverse* $B^{-1}A^{-1}$
(3) A^T *is invertible with inverse* $(A^{-1})^T$

Proof —
Part (1): Since A is invertible, it follows that $\exists\ A^{-1}$ such that: $AA^{-1} = A^{-1}A = I_n$. However, by definition this immediately implies that A^{-1} is invertible with inverse A.

Part (2): Since A,B are invertible, it follows that $\exists\ A^{-1}, B^{-1}$ such that: $AA^{-1} = A^{-1}A = I_n = BB^{-1} = B^{-1}B$. Then consider the following computations:

$$\begin{aligned} AB(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n \\ (B^{-1}A^{-1})AB &= B^{-1}(A^{-1}A)B = B^{-1}I_nB = B^{-1}B = I_n \end{aligned}$$

Which implies that AB is invertible with inverse $B^{-1}A^{-1}$.

Part (3): Since A is invertible, it follows that $\exists\ A^{-1}$ such that: $AA^{-1} = A^{-1}A = I_n$. Then consider the following computations:

$$\begin{aligned} A^T(A^{-1})^T &= (A^{-1}A)^T = I_n^T = I_n \\ (A^T)^{-1}A^T &= (AA^{-1})^T = I_n^T = I_n \end{aligned}$$

Which implies that A^T is invertible with inverse $(A^{-1})^T$.

$\square$

Theorem 1.4. *Let A be an mxn matrix and B an nxp matrix, then* $(AB)^T = B^TA^T$

Proof — Let the matrices $A = (a_{ij})$, $B = (b_{ij})$ i.e. the notation is that the (i,j)th entry of a matrix M is given by m_{ij}. Then note that $A^T = (a_{ji})$, $B^T = (b_{ji})$ Then consider the following to complete the proof:

$$\begin{aligned} (AB_{ij})^T &= (\sum_{k=1}^{n} a_{ik}b_{kj})^T \\ &= (\sum_{k=1}^{n} a_{jk}b_{ki}) \\ &= (\sum_{k=1}^{n} b_{ki}a_{jk}) \\ &= ((B^TA^T)_{ij}) \end{aligned}$$

$\square$

Note: (AB_{ij}) stands for the (i,j)th entry in the AB matrix, and similarly, $((B^TA^T)_{ij})$ stands for the (i,j)th entry in the B^TA^T matrix.

Theorem 1.5. *Let A be an invertible nxn matrix. Then* A^{-1} *is unique.*

Proof — Assume that there are two inverses: $A^{-1}, \overline{A^{-1}}$. Since they are both inverses, we have the following:

$$\begin{aligned} AA^{-1} = I_n = A\overline{A^{-1}} &\implies A^{-1}(AA^{-1}) = A^{-1}(I_n) = A^{-1}(A\overline{A^{-1}}) \\ &\implies (A^{-1}A)A^{-1} = A^{-1} = (A^{-1}A)\overline{A^{-1}} \\ &\implies I_nA^{-1} = A^{-1} = I_n\overline{A^{-1}} \\ &\implies A^{-1} = A^{-1} = \overline{A^{-1}} \\ &\implies A^{-1} = \overline{A^{-1}} \end{aligned}$$

Thereby completing the proof.

□

Theorem 1.6. *Let A be an nxn matrix and let B be the nxn matrix gotten by interchanging the the ith and jth rows of A. Then $-det(A) = det(B)$.*

Proof — By induction. For the base case, consider the case where n=2, then direct computation shows that:

$$-det(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = -(ad - bc) = bc - ad = det(\begin{pmatrix} c & d \\ a & b \end{pmatrix})$$

Next let $n \geq 3$ and assume the result is true for (n-1)x(n-1) matrices, and we will complete the induction by proving the result for nxn matrices. Let

$$A = \begin{pmatrix} a_{11} & \dots & \dots & \dots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{j1} & \dots & \dots & \dots & a_{jn} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{k1} & \dots & \dots & \dots & a_{kn} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \dots & \dots & \dots & a_{nn} \end{pmatrix}, \qquad B = \begin{pmatrix} a_{11} & \dots & \dots & \dots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{k1} & \dots & \dots & \dots & a_{kn} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{j1} & \dots & \dots & \dots & a_{jn} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \dots & \dots & \dots & a_{nn} \end{pmatrix}$$

i.e. B is obtained from A by interchanging the jth and kth rows. Assume that that row $p \neq j, k$ (we can always do this since in this case we have $n \geq 3$, so there must exist some row p that is being left fixed). Then consider the following computation to complete the proof (where the determinant of both matrices is calculated by going across the pth row):

$$\begin{aligned} -det(A) &= -\sum_{m=1}^{n} a_{pm}(C_{pm}) \\ &= -\sum_{m=1}^{n} a_{pm}(-(C'_{pm})) \qquad \text{[by the induction assumption for (n-1)x(n-1) matrices]} \\ &= \sum_{m=1}^{n} a_{pm}(C'_{pm}) \\ &= det(B) \end{aligned}$$

Notation: In the computation above,

$C_{pm} = (-1)^{p+m} det((n-1)x(n-1))$ minor of the A with the pth row and mth columns removed)

$C'_{pm} = (-1)^{p+m} det((n-1)x(n-1))$ minor of the B with the pth row and mth columns removed)

□

Theorem 1.7. *Let A be an nxn invertible matrix, then* $det(A^{-1}) = \frac{1}{det(A)}$

Proof — First note that the identity matrix is a diagonal matrix so its determinant is just the product of the diagonal entries. Since all the entries are 1, it follows that $det(I_n) = 1$. Next consider the following computation to complete the proof:

$$\begin{aligned} 1 &= det(I_n) = det(AA^{-1}) \\ &= det(A)det(A^{-1}) && \text{[using the fact that } det(AB) = det(A)det(B)\text{]} \\ \Longrightarrow & \; det(A^{-1}) = \frac{1}{det(A)} && \text{[Note that } det(A) \neq 0 \text{ since A is invertible]} \end{aligned}$$

□

2. Proofs that you are responsible for on the Final only

Theorem 2.1. *Similar matrices have the same eigenvalues with the same multiplicities.*

Proof — Let A and B be similar nxn matrices. That is, there exists an invertible nxn matrix P such that $B = P^{-1}AP$. Since the eigenvalues of a matrix are precisely the roots of the characteristic equation of a matrix, in order to prove that A and B have the same eigenvalues, it suffices to show that they have the same characteristic polynomials (and hence the same characteristic equations). To see this consider the following:

$$\begin{aligned} \chi_B(t) &= det(B - tI_n) \\ &= det(P^{-1}AP - tP^{-1}(I_n)P) \\ &= det(P^{-1}AP - P^{-1}(tI_n)P) \\ &= det(P^{-1}(A - tI_n)P) \\ &= det(P^{-1})det(A - tI_n)det(P) \\ &= \frac{1}{det(P)}det(A - tI_n)det(P) \\ &= \frac{1}{det(P)}det(P)det(A - tI_n) \\ &= det(A - tI_n) \\ &= \chi_A(t) \end{aligned}$$

Corollary 2.2. *Let A and B be similar 2x2 matrices, then* $Tr(A) = Tr(B)$ *and* $det(A) = det(B)$ *[Note: This results holds true for nxn matrices as well]*

Proof — Since A and B are similar, in the proof of theorem 2.1 we saw that A and B have the same characteristic polynomial. However for an arbitrary 2x2 matrix $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, we know that the characteristic polynomial is given by

$$\chi_M(t) = t^2 - (a + d)t + (ad - bc) = t^2 - Tr(M)t + det(M)$$

Thus since the similar 2x2 matrices A and B have the same characteristic polynomials, this implies that

$$t^2 - Tr(A)t + det(A) = t^2 - Tr(B)t + det(B)$$

Then equating the terms of the equivalent polynomials, it follows that:

$$Tr(A) = Tr(B) \qquad \text{and} \qquad det(A) = det(B)$$

Theorem 2.3. *Let $\vec{v}, \vec{w}, \vec{z} \in \mathbb{R}^n$ be orthonormal vectors. Prove that $\{\vec{v}, \vec{w}, \vec{z}\}$ are linearly independent*

Proof — Assume that there are coefficients $a, b, c \in \mathbb{R}$ such that

$$a\vec{v} + b\vec{w} + c\vec{z} = \vec{\mathbf{0}} \tag{2.1}$$

Then since $\vec{v}, \vec{w}, \vec{z} \in \mathbb{R}^n$ are orthonormal, we have that

$$\vec{v} \cdot \vec{w} = \vec{v} \cdot \vec{z} = \vec{w} \cdot \vec{z} = 0 \tag{2.2}$$

$$\vec{v} \cdot \vec{v} = \vec{w} \cdot \vec{w} = \vec{z} \cdot \vec{z} = 1 \tag{2.3}$$

Using facts 2.2 and 2.3 in conjunction with equation 2.1 it follows that:

$$\begin{aligned}
a &= a + 0 + 0 = a(\vec{v} \cdot \vec{v}) + b(\vec{v} \cdot \vec{w}) + c(\vec{v} \cdot \vec{z}) = \vec{v} \cdot (a\vec{v}) + \vec{v} \cdot (b\vec{w}) + \vec{v} \cdot (c\vec{z}) \\
&= \vec{v} \cdot (a\vec{v} + b\vec{w} + c\vec{z}) = \vec{v} \cdot (\vec{\mathbf{0}}) = 0 \\
b &= 0 + b + 0 = a(\vec{w} \cdot \vec{v}) + b(\vec{w} \cdot \vec{w}) + c(\vec{w} \cdot \vec{z}) = \vec{w} \cdot (a\vec{v}) + \vec{w} \cdot (b\vec{w}) + \vec{w} \cdot (c\vec{z}) \\
&= \vec{w} \cdot (a\vec{v} + b\vec{w} + c\vec{z}) = \vec{w} \cdot (\vec{\mathbf{0}}) = 0 \\
c &= 0 + 0 + c = a(\vec{z} \cdot \vec{v}) + b(\vec{z} \cdot \vec{w}) + c(\vec{z} \cdot \vec{z}) = \vec{z} \cdot (a\vec{v}) + \vec{z} \cdot (b\vec{w}) + \vec{z} \cdot (c\vec{z}) \\
&= \vec{z} \cdot (a\vec{v} + b\vec{w} + c\vec{z}) = \vec{z} \cdot (\vec{\mathbf{0}}) = 0
\end{aligned}$$

Hence $a = b = c = 0 \implies$ by definition that $\{\vec{v}, \vec{w}, \vec{z}\}$ are linearly independent.

□

E-mail address: HaroldSultan@gmail.com