The Inverse Matrix and An Introduction to the Determinant

Throughout this assignment, we consider linear operators $\mathcal{L} : \mathbb{R}^m \to \mathbb{R}^m$, i.e. $\mathcal{L}(\mathbf{x}) \equiv A\mathbf{x}$ where A is a <u>square</u> $m \times m$ real matrix. Please note however that while not discussed directly here, all definitions and facts given below routinely extend to problems in which \mathbb{R}^m is replaced by \mathbb{C}^m and the square matrices associated to \mathcal{L} are complex valued.

The $m \times m$ identity matrix, denoted by I, is given by

$$I_{i,j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases} \implies I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Often, the entries of the identity matrix are denoted by $I_{i,j} = \delta_{i,j}$ where $\delta_{i,j}$ is called the *Kronecker delta*. You should verify the following is true. For any $m \times m$ matrix A, we have

$$AI = IA = A$$

This's why I is called the identity matrix. Also notice the columns of I are composed of the standard basis vectors for $\mathbb{R}^m = \operatorname{span}\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m\}$.

Next, for any $m \times m$ matrix A, I'll derive the following result.

If the *m* columns of *A* are linearly independent, there is an $m \times m$ matrix A^{-1} which satisfies

$$A A^{-1} = A^{-1} A = I.$$

 A^{-1} is called the *inverse matrix* of A. When A has an inverse we say A is *invertible*.

First, let me show you the range of A is all of \mathbb{R}^m . Since

$$\operatorname{Rang}(A) = \operatorname{span}\{\mathbf{a}_1, \dots, \mathbf{a}_m\},\$$

and by assumption the set $\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$ is independent, conclude that $\operatorname{Rang}(A)$ is an mdimensional subspace of the m-dimensional vector space \mathbb{R}^m . Therefore $\operatorname{Rang}(A) = \mathbb{R}^m$. Next, since the range of A is all of \mathbb{R}^m , conclude for any standard basis vector of \mathbb{R}^m , say \mathbf{e}_j , there is a (unique) vector, say $\mathbf{b}_j \in \mathbb{R}^m$, such that

$$A \mathbf{b}_j = \mathbf{e}_j$$
 for each $j = 1, \dots, m$.

Let *B* be the $m \times m$ matrix constructed with columns $\mathbf{b}_1, \ldots, \mathbf{b}_m$. This matrix *B* is the sought for *right inverse* of *A*, i.e. AB = I. (Remember the $m \times m$ identity *I* has columns $\mathbf{e}_1, \ldots, \mathbf{e}_m$.) I still need to show *B* is also the *left inverse* of *A*, i.e. BA = I. I claim the

columns of B must be independent. To see this, observe that for any scalars β_1, \ldots, β_m such that

$$\beta_1 \mathbf{b}_1 + \dots + \beta_m \mathbf{b}_m = \mathbf{0} \quad \Rightarrow \quad \mathbf{0} = A \left(\beta_1 \mathbf{b}_1 + \dots + \beta_m \mathbf{b}_m \right) = \beta_1 \mathbf{e}_1 + \dots + \beta_m \mathbf{e}_m.$$

But the standard basis vectors are clearly independent which says $\beta_1 = \cdots = \beta_n = 0$. Therefore, the columns of B are independent. Now that we know this fact, exactly the same argument as given above allows us to conclude there is a matrix $C \in \mathbb{R}^{m \times m}$ such that BC = I. Putting these together gives

$$AB = I$$

$$BC = I$$

$$\Rightarrow A(BC) = A(I) \Rightarrow (AB)C = A \Rightarrow C = A.$$

Finally, define $A^{-1} \equiv B$, and observe that we have shown $AA^{-1} = A^{-1}A = I$. You'll prove in an exercise below that when the columns of $A \in \mathbb{R}^{m \times m}$ are dependent then A can not have an inverse matrix. Specifically, you'll be asked to show:

A's columns are not independent. \Rightarrow There is no B satisfying AB = I.

This statement is logically equivalent to:

There is a B satisfying AB = I. \Rightarrow A's columns are independent.

Now, how do we compute A^{-1} ? By elimination of course. Let me show you by example. Consider the 4×4 matrix

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 1 & 1 & 2 \\ 1 & 1 & 2 & 3 \end{pmatrix}$$

Let's find A^{-1} if it exists. Write the augmented matrix attaching the 4×4 identity matrix. Eliminate the left side to upper triangular form if possible.

Swap row 3 and row 4 to get

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

This completes forward elimination. The left side of the augmented matrix is upper triangular with all pivots (the diagonal entries) nonzero. This tells us that A's columns are linearly independent. Therefore, the given matrix \underline{is} invertible. To finish computing the actual inverse, use backward elimination starting with column 4

Next, eliminate up column 3 and then column 2 to get

Γ1	L	0	0	0	2	-1	0	ך 0	
()	1	0	0	-1	1	1	-1	
()	0	1	0	1	0	-2	1	
[()	0	0	1	$\begin{vmatrix} 2\\ -1\\ 1\\ -1 \end{vmatrix}$	0	1	0	

The left side of the augmented matrix is now the <u>identity</u>, and so we're done. The right side is A^{-1} ,

$$A^{-1} = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 1 & 1 & -1 \\ 1 & 0 & -2 & 1 \\ -1 & 0 & 1 & 0 \end{pmatrix}$$

Here's a second example, but this time the matrix is not invertible. Let

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 1 & 1 & 2 \\ 1 & 1 & 1 & 3 \end{pmatrix},$$

write the augmented matrix attaching the identity and perform forward elimination

At this stage however, see that x_3 (the third column in the right augmented matrix) is a free variable. This tells us that A's columns are <u>not</u> linearly independent. Therefore, this matrix is <u>not</u> invertible.

Here are two other important facts you should know.

First, suppose $A \in \mathbb{R}^{m \times m}$, $B \in \mathbb{R}^{m \times m}$ are both invertible. Then the product AB is also invertible, and in particular

$$(AB)^{-1} = B^{-1}A^{-1}.$$

To see this, observe

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$
, and
 $(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I$.

Second, let $A \in \mathbb{R}^{m \times n}$. Note that this matrix need not be square. The *transpose* of A, denoted by A^T , is an $n \times m$ matrix obtained by switching A's rows with its columns. That is,

$$(A^T)_{i,j} = A_{j,i}$$
 for each $1 \le i \le n, \ 1 \le j \le m$.

For example,

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

Every student who's completed linear algebra knows the following. If the matrix product AB is defined, then

$$(AB)^T = B^T A^T.$$

To see this is true, take $A \in \mathbb{R}^{m_A \times n_A}$ and $B \in \mathbb{R}^{m_B \times n_B}$ with $n_A = m_B \equiv l$, and observe

$$(AB)_{i,j} = \sum_{k=1}^{l} A_{i,k} B_{k,j} \quad \Rightarrow \quad ((AB)^T)_{i,j} = (AB)_{j,i} = \sum_{k=1}^{l} A_{j,k} B_{k,i} = \sum_{k=1}^{l} B_{i,k}^T A_{k,j}^T.$$

The sum on the right represents the i, j th element of the product $B^T A^T$.

Here's an interesting fact concerning square matrices which can be deduced from the transpose product formula just given. A square matrix A has linearly independent columns if and only if its rows are linearly independent. You should be able to see this by using the transpose formula to verify $(A^{-1})^T = (A^T)^{-1}$. Of course, this fact is also implied by the fact that the row rank of a matrix is equal to its column rank.

1. Suppose $A \in \mathbb{R}^{m \times m}$ and there is another matrix $B \in \mathbb{R}^{m \times m}$ such that AB = I. In this exercise you'll prove this implies the column vectors making up A must be linearly independent.

- (a) Show the columns of B must be linearly independent.
- (b) Conclude B is invertible.
- (c) Show that given $A\mathbf{x} = \mathbf{0}$ we must have $\mathbf{x} = \mathbf{0}$.
- (d) Observe part (c) says the columns of A are independent.

Hints: (a) Do as I did on page 2: $\mathbf{0} = \beta_1 \mathbf{b}_1 + \cdots + \beta_m \mathbf{b}_m$ implies $\mathbf{0} = \beta_1 \mathbf{e}_1 + \cdots + \beta_m \mathbf{e}_m$. (b) I proved independent columns implies invertibility. (c) For any \mathbf{x} there is a \mathbf{y} such that $\mathbf{x} = B\mathbf{y}$. Use this and the given to conclude $\mathbf{y} = \mathbf{0}$. 2. Determine the inverse if it exists.

(a)
$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$
 (b) $\begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ (d) $\begin{pmatrix} 4 & 3 \\ 2 & 1 \end{pmatrix}$

3. Determine the inverse if it exists.

	(1)	1	1		(1	2	3
(a)	3	4	1	(b)	-	-2	-2	-5
(a)	$\backslash 2$	4	1/	. ,		2	2	$\begin{pmatrix} 3\\ -5\\ 8 \end{pmatrix}$
			-					-

4. Determine the inverse if it exists.

(a)	/1	1	1	1		/1	1	1	1
	1	2	1	1	(b)	1	2	1	1
	1	2	1	1		1	2	1	2
	\backslash_1	2	2	$_{3}/$		$\backslash 1$	2	2	$_{3}/$

The determinant of an $m \times m$ matrix A is given by the formula

$$\det(A) \equiv \sum_{\mathbf{j} \in \mathbf{P}_m} \operatorname{sgn}(\mathbf{j}) a_{1,j_1} a_{2,j_2} \cdots a_{m,j_m}.$$

 P_m above denotes the set containing all m! permutations of the sequence (1, 2, ..., m). For example, $P_3 = \{(1, 2, 3), (2, 3, 1), (3, 1, 2), (1, 3, 2), (2, 1, 3), (3, 2, 1)\}$. The symbol $\mathbf{j} = (j_1, j_2, ..., j_m)$ denotes one of these permutations. The sign of permutation \mathbf{j} , denoted by $\operatorname{sgn}(\mathbf{j})$, has value +1 if \mathbf{j} is obtained by an even number of interchanges of (1, 2, ..., m) and has value -1 if by an odd number. For example, when m = 3,

j	interchanges	$\mathrm{sgn}(\mathbf{j})$
(1, 2, 3)	0	+1
(2, 3, 1)	2	+1
(3, 1, 2)	2	+1
(1, 3, 2)	1	-1
(2, 1, 3)	1	$^{-1}$
(3, 2, 1)	1	-1

It's interesting to note that no matter how you perform interchanges to cast a given permutation **j** to (1, 2, ..., m), the value of sgn(**j**) remains invariant.

The determinant was first introduced and studied by Gottfried Leibniz, 1646–1716. The formula given above is often called the Leibniz determinant formula.

When m = 2, $P_2 = \{(1, 2), (2, 1)\}$, sgn(1, 2) = 1, sgn(2, 1) = -1, and so

$$\det(A) = a_{1,1} a_{2,2} - a_{1,2} a_{2,1}.$$

When m = 3, use the table above to arrive at

$$det(A) = (a_{1,1} a_{2,2} a_{3,3} + a_{1,2} a_{2,3} a_{3,1} + a_{1,3} a_{2,1} a_{3,2}) - (a_{1,1} a_{2,3} a_{3,2} + a_{1,2} a_{2,1} a_{3,3} + a_{1,3} a_{2,2} a_{3,1}).$$

You should memorize these two explicit formulae. I'll show you a trick in class to help.

Here are several important properties about the determinant you should know. These are all easily deduced from the Leibniz formula.

• If A has a zero row, i.e. $a_{i_*,j} = 0$ for every j, then det(A) = 0. This is true because for every $\mathbf{j} \in \mathbf{P}_m$ the product $a_{1,j_1}a_{2,j_2}\cdots a_{m,j_m} = a_{1,j_1}\cdots a_{i_*,j_{i_*}}\cdots a_{m,j_m} = 0$.

• If U is upper triangular, i.e. $u_{i,j} = 0$ for every i > j, then

$$\det(U) = u_{1,1}u_{2,2}\cdots u_{m,m}.$$

This is true because for every permutation $\mathbf{j} \, \underline{\text{except}} \, \mathbf{j} = (1, 2, \dots, m)$, there's at least one index i_* such that $i_* > j_{i_*}$. Since for all such permutations $u_{i_*,j_{i_*}} = 0$, and this says $u_{1,j_1}u_{2,j_2}\cdots u_{m,j_m} = 0$, the determinant sum collapses to a single term.

• If matrix A' is obtained from A by switching rows $i_1 \neq i_2$, then $\det(A') = -\det(A)$. To see this, let \mathbf{j}' be given by switching component i_1 with i_2 in \mathbf{j} . Clearly, $\operatorname{sgn}(\mathbf{j}') = -\operatorname{sgn}(\mathbf{j})$. Also, observe that the product $a'_{1,j_1}a'_{2,j_2}a'_{m,j_m} = a_{1,j'_1}a_{2,j'_2}\cdots a_{m,j'_m}$. Therefore,

$$\det(A') = \sum_{\mathbf{j} \in \mathbf{P}_m} \operatorname{sgn}(\mathbf{j}) a'_{1,j_1} a'_{2,j_2} \cdots a'_{m,j_m}$$
$$= \sum_{\mathbf{j}' \in \mathbf{P}_m} -\operatorname{sgn}(\mathbf{j}') a_{1,j'_1} a_{2,j'_2} \cdots a_{m,j'_m} = -\det(A)$$

• If A has two identical rows, then det(A) = 0. To see this, interchange the two identical rows and use the previous bulleted item, i.e. $A' = A \Rightarrow det(A) = det(A') = -det(A)$.

The next item is very important. In fact, it states the property about the determinant that likely led Leibniz to consider his particular formula in the first place.

• Suppose $m \times m$ matrices A, B and C are <u>identical</u> except in the i_* th row. Suppose the i_* th row of A is given by $a_{i_*,j} = \beta b_{i_*,j} + \gamma c_{i_*,j}$ for each $j = 1, \ldots, m$. Then

$$\det(A) = \beta \det(B) + \gamma \det(C).$$

The derivation is very simple.

$$det(A) = \sum_{\mathbf{j}\in\mathbf{P}_m} \operatorname{sgn}(\mathbf{j}) a_{1,j_1} \cdots \left(\beta \ b_{i_*,j_{i_*}} + \gamma \ c_{i_*,j_{i_*}}\right) \cdots a_{m,j_m}$$

$$= \beta \sum_{\mathbf{j}\in\mathbf{P}_m} \operatorname{sgn}(\mathbf{j}) a_{1,j_1} \cdots b_{i_*,j_{i_*}} \cdots a_{m,j_m} + \gamma \sum_{\mathbf{j}\in\mathbf{P}_m} \operatorname{sgn}(\mathbf{j}) a_{1,j_1} \cdots c_{i_*,j_{i_*}} \cdots a_{m,j_m}$$

$$= \beta \sum_{\mathbf{j}\in\mathbf{P}_m} \operatorname{sgn}(\mathbf{j}) \ b_{1,j_1} \cdots b_{i_*,j_{i_*}} \cdots b_{m,j_m} + \gamma \sum_{\mathbf{j}\in\mathbf{P}_m} \operatorname{sgn}(\mathbf{j}) \ c_{1,j_1} \cdots c_{i_*,j_{i_*}} \cdots c_{m,j_m}$$

$$= \beta \det(B) + \gamma \det(C).$$

This fact says the determinant is what's called a *multilinear* function of its rows.

Two immediate consequences of this are the following.

• Suppose A' is obtained from A by multiplying one of its rows by a scalar α . Then

$$\det(A') = \alpha \det(A).$$

• Suppose A' is obtained from A by adding any multiple of row i_2 to a different row i_1 . Then

$$\det(A') = \det(A).$$

To see this, let B be a matrix identical to A except in its i_1 th row take $b_{i_1,j} = a_{i_2,j}$. Use multilinearity, i.e.

$$a'_{i_1,j} = a_{i_1,j} + \beta \, b_{i_1,j} \quad (= a_{i_1,j} + \beta \, a_{i_2,j}), \quad \Rightarrow \quad \det(A') = \det(A) + \beta \det(B).$$

However, det(B) = 0 because its i_1 and i_2 rows are identical.

Here's why the determinant is so important. Recall by means of a succession of the three elementary row operations, see E1, E2 and E3 from page 5 on your homework 5, a matrix A can always be reduced to row echelon form. When A is square, its echelon form is an upper triangular square matrix, say U. From what we've shown about the determinant

$$A \sim U \Rightarrow \det(A) = \pm \alpha \det(U) \quad (\alpha \neq 0)$$

= $\pm \alpha (u_{1,1} \cdots u_{m,m}).$

The \pm corresponds to possible row interchanges during the elimination process, see row operation E1, and the scale factor α corresponds to possible nonzero row scalings, see row operation E2. Also recall from earlier that A has independent columns if and only if each diagonal entry in its row echelon form, U, is nonzero. Therefore,

A has independent columns $\iff \det(A) \neq 0.$

Here are two more important properties about the determinant you should all know. I'm not going to derive them here but will do so in a subsequent set of notes.

• Suppose A^T is the transpose of a square matrix A. Then $det(A^T) = det(A)$.

This tells us what was said in the bulleted items above concerning rows also applies to columns and vice versa. Finally:

• Suppose A and B are square $m \times m$ matrices. Then det(AB) = det(A) det(B).

Here are some, hopefully, illuminating examples.

Use the 2×2 determinant formula, i.e.

$$\det \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} = a_{1,1} a_{2,2} - a_{1,2} a_{2,1},$$

to verify multilinearity of the determinant by explicitly calculating both sides of

$$\det \begin{pmatrix} 2 \cdot 1 + 3 \cdot 2 & 2 \cdot 4 + 3 \cdot 5 \\ 6 & 7 \end{pmatrix} = 2 \det \begin{pmatrix} 1 & 4 \\ 6 & 7 \end{pmatrix} + 3 \det \begin{pmatrix} 2 & 5 \\ 6 & 7 \end{pmatrix}.$$

On the left side

$$\det \begin{pmatrix} 2 \cdot 1 + 3 \cdot 2 & 2 \cdot 4 + 3 \cdot 5 \\ 6 & 7 \end{pmatrix} = \det \begin{pmatrix} 8 & 23 \\ 6 & 7 \end{pmatrix} = 8 \cdot 7 - 23 \cdot 6 = -82.$$

On the right side

$$2 \det \begin{pmatrix} 1 & 4 \\ 6 & 7 \end{pmatrix} + 3 \det \begin{pmatrix} 2 & 5 \\ 6 & 7 \end{pmatrix} = 2(7 - 24) + 3(14 - 30) = -82.$$

Given

$$A \equiv \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \Rightarrow \quad \det(A) = (4-6) = -2$$

What is det(5A)? Answer: $5 \cdot 5(-2) = -50$. Notice the scalar-matrix product 5A is obtained by multiplying both row 1 and row 2 of A by 5.

Consider the 3×3 matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

Construct the matrix A' by subtracting row 2 from row 3 in A, and then again subtracting row 1 from row 2. That is,

$$A' = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 3 & 3 \\ 3 & 3 & 3 \end{pmatrix}.$$

Ask yourself why this implies det(A) = 0. Verify this fact directly via the 3×3 determinant formula, see the middle of page 6, applied to the original matrix A. (Not for me!) Here's one last example. Consider

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{pmatrix}.$$

Subtract 4 times row 1 from row 2 and 7 times row 1 from row 3 to get

$$A' = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -11 \end{pmatrix},$$

and then in A' subtract 2 times row 2 from row 3

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

Conclude det(A) = -3. Verify by applying the 3×3 determinant formula directly to A. (Again, not for me!)

5. If a matrix A is invertible, conclude $det(A^{-1}) = 1/det(A)$. Hint: $AA^{-1} = I$ and det(I) = 1. (This would make an easy exam question.)

6. You can conclude that A is invertible if and only if $\det(A) \neq 0$ by referring back to my remark "A has independent columns if and only if $\det(A) \neq 0$ " given in the middle of page 7. Suppose A and B are $m \times m$ matrices. Conclude the product C = AB is invertible if and only if both $\det(A) \neq 0$ and $\det(B) \neq 0$.

7. Compute the determinants of each of the following.

(a)
$$\begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix}$$
 (b) $\begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}$ (c) $\begin{pmatrix} 8 & 3 \\ 4 & 4 \end{pmatrix}$ (d) $\begin{pmatrix} -2 & -1 \\ 3 & 4 \end{pmatrix}$

8. Compute the determinants of each of the following. State the properties of the determinant you used.

$$\begin{array}{c} \text{(a)} & \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ 1 & 1 & 2 \end{pmatrix} \\ \text{(b)} & \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \\ \text{(c)} & \begin{pmatrix} 2 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 1 & 2 \end{pmatrix} \\ \text{(d)} & \begin{pmatrix} 3 \cdot 1 & 3 \cdot 2 & 3 \cdot 1 \\ 3 \cdot 1 & 3 \cdot 1 & 3 \cdot 2 \\ 3 \cdot 1 & 3 \cdot 1 & 3 \cdot 1 \end{pmatrix} \\ \text{(e)} & \begin{pmatrix} 1 & 2 & 1 \\ 5 \cdot 1 & 5 \cdot 1 & 5 \cdot 2 \\ 1 & 1 & 1 \end{pmatrix} \\ \text{(f)} & \begin{pmatrix} 1 & 2 & 1 \\ 2 + 1 & 2 + 1 & 3 + 2 \\ 1 & 1 & 1 \end{pmatrix} \\ \text{Answers I got: (a) 0, (b) 2, (c) 4, (d) 3^3 \cdot 1 = 27, (e) 5 \cdot 1 = 5, (f) 1 + 1 = 2. \end{array}$$