3: ROW OPERATIONS, THE DETERMINANT

STEVEN HEILMAN

Contents

1. Review

Theorem 1.1 (**Dimension Theorem/ Rank-Nullity Theorem**). *Let V, W be vector spaces over a field $\mathbf{F}$. Let $T: V \to W$ be linear. If V is finite-dimensional, then*

$$\mathrm{nullity}(T) + \mathrm{rank}(T) = \dim(V).$$

Theorem 1.2. *Let V, W be vector spaces over a field $\mathbf{F}$. Let $T: V \to W$ be a linear transformation. Assume that $\{v_1, \ldots, v_n\}$ spans V. Then $\{Tv_1, \ldots, Tv_n\}$ spans $R(T)$.*

Remark 1.3. Let A be an $m \times \ell$ matrix, let B be an $n \times m$ matrix. Then $L_B L_A = L_{BA}$.

Corollary 1.4. *An $m \times n$ matrix A is invertible if and only if the linear transformation $L_A: \mathbf{F}^n \to \mathbf{F}^m$ is invertible. Also, $(L_A)^{-1} = L_{A^{-1}}$.*

2. Row Operations

We begin our discussion of row operations on matrices with some examples.

Example 2.1 (**Type 1: Interchange two Rows**). For example, we can swap the first and third rows of the matrix

$$\begin{pmatrix} 1 & 2 \\ 3 & 5 \\ 0 & 8 \end{pmatrix}$$

to get

$$\begin{pmatrix} 0 & 8 \\ 3 & 5 \\ 1 & 2 \end{pmatrix}.$$

Define

$$E := \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Date: November 24, 2014.

Note that

$$E \begin{pmatrix} 1 & 2 \\ 3 & 5 \\ 0 & 8 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 5 \\ 0 & 8 \end{pmatrix} = \begin{pmatrix} 0 & 8 \\ 3 & 5 \\ 1 & 2 \end{pmatrix}.$$

Remark 2.2. E as defined above is invertible. In fact, $E = E^{-1}$. In general, if E is the $n \times n$ matrix that swaps two rows of an $n \times n$ matrix A, then EA is A with those two rows swapped. So $EEA = A$ for all $n \times n$ matrices A, so $EE = I_n$, i.e. E is invertible.

Example 2.3 (Type 2: Multiply a row by a nonzero scalar). For example, let's multiply the second row of the following matrix by 2.

$$\begin{pmatrix} 1 & 2 \\ 3 & 5 \\ 0 & 8 \end{pmatrix}.$$

We then get

$$\begin{pmatrix} 1 & 2 \\ 6 & 10 \\ 0 & 8 \end{pmatrix}.$$

Define

$$E := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Note that

$$E \begin{pmatrix} 1 & 2 \\ 3 & 5 \\ 0 & 8 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 5 \\ 0 & 8 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 6 & 10 \\ 0 & 8 \end{pmatrix}$$

Remark 2.4. E as defined above has inverse

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

In general, suppose E corresponds to multiplying the i^{th} row of a given matrix by $\alpha \in \mathbf{F}$, $\alpha \neq 0$. Then E is a matrix with ones on the diagonal, except for the i^{th} entry on the diagonal, which is α. And all other entries of E are zero. Then, we see that E^{-1} exists and is a matrix with ones on the diagonal, except for the i^{th} entry on the diagonal, which is α^{-1}. And all other entries of E^{-1} are zero. In particular, E is invertible.

Example 2.5 (Adding one row to another). Let's add two copies of the first row of the following matrix to the third row.

$$\begin{pmatrix} 1 & 2 \\ 3 & 5 \\ 0 & 8 \end{pmatrix}.$$

We then get

$$\begin{pmatrix} 1 & 2 \\ 3 & 5 \\ 2 & 12 \end{pmatrix}.$$

Define

$$E := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}.$$

Note that

$$E\begin{pmatrix} 1 & 2 \\ 3 & 5 \\ 0 & 8 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 5 \\ 0 & 8 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 5 \\ 2 & 12 \end{pmatrix}.$$

Remark 2.6. E as defined above has inverse

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}.$$

That is, adding 2 copies of row one to row three is inverted by adding -2 copies of row one to row three. In a similar way, a general row addition operator is seen to be invertible.

Remark 2.7 (Summary of Row Operations). The three row operations (Type 1, Type 2, and Type 3) are all invertible.

Remark 2.8 (Solving Systems of Linear Equations). Let A be an $m \times n$ matrix, let $x \in \mathbf{R}^n$ be a variable vector, and let $b \in \mathbf{R}^m$ be a known vector. Consider the system of linear equations

$$Ax = b.$$

Let E be any elementary row operation. Since E is invertible, finding a solution x to the system $Ax = b$ is equivalent to finding the solution x to the system $EAx = Eb$. By applying many elementary row operations, you have seen in a previous course how to solve the system $Ax = b$. That is, you continue to apply elementary row operations $E_1, \ldots, E_k$ such that $E_1 \cdots E_k A$ in in **row-echelon** form, and you then solve $E_1 \cdots E_k Ax = E_1 \cdots E_k b$. A matrix B is in row-echelon form if each row is either zero, or its left-most nonzero entry is 1, with zeros below the 1.

Remark 2.9 (Inverting a Matrix). Let A be an invertible $n \times n$ matrix. You learned in a previous course an algorithm for inverting A using elementary row operations. Below, we will prove that this algorithm works.

Remark 2.10 (Column Operations). In the above discussion, we could have also used column operations instead of row operations. Column operations would then correspond to multiplying the matrices E on the right side, rather than the left side. The invertibility of column operations would therefore still hold.

3. Rank of a Matrix

Let $T\colon V \to W$ be a linear transformation between two vector spaces. Recall that the rank of T, denoted by $\mathrm{rank}(T)$, is the dimension of $R(T)$, the range of T.

Lemma 3.1. *Let V, W be finite-dimensional vector spaces over a field $\mathbf{F}$. Assume that $\dim(V) = \dim(W) = n$. Let $T\colon V \to W$ be a linear transformation. Then T is invertible if and only if T has rank n.*

Proof. Suppose T is invertible. Then T is one-to-one. By the Dimension Theorem (Theorem 1.1), T has rank n.

Now, suppose T has rank n. Then, by the Dimension Theorem, $N(T) = \{0\}$, so T is one-to-one. Also, $R(T)$ is again a subspace of W of the same dimension as W, so we must have $R(T) = W$, so T is onto. Since T is both one-to-one and onto, T is invertible. □

Lemma 3.2. *Let V, W be finite-dimensional vector spaces over a field $\mathbf{F}$. Let $T: V \to W$ be an invertible linear transformation. Let $U \subseteq V$ be a subspace. Then $\dim(U) = \dim(T(U))$.*

Proof. Since $U \subseteq V$ is a subspace, U is a vector space. So, for each $u \in U$, define the map $T_U: U \to W$ by

$$T_U(u) := T(u).$$

Since T is linear, T_U is linear. Since T is one-to-one, T_U is one-to-one, so $N(T_U) = \{0\}$. So, the Dimension Theorem (Theorem 1.1) implies that $\dim R(T_U) = \dim(U)$. Since $R(T_U) = T_U(U) = T(U)$, we are done. □

Lemma 3.3 (**Isomorphisms Preserve Rank**). *Let U, V, W, X be vector spaces over a field $\mathbf{F}$. Let $T: V \to W$ be a linear transformation. Let $S: U \to V$ be an invertible linear transformation, and let $P: W \to X$ be an invertible linear transformation. Then*

$$\text{rank}(T) = \text{rank}(PT) = \text{rank}(TS) = \text{rank}(PTS).$$

Proof. We begin with the first equality. By the definition of range, $R(T) = T(V)$, and $R(PT) = PT(V)$. So,

$$R(PT) = PT(V) = P(T(V)) = P(R(T)).$$

So, $\text{rank}(PT) = \dim(P(R(T)))$. Since P is invertible, $\dim(P(R(T))) = \dim(R(T))$ by Lemma 3.2. So, $\text{rank}(PT) = \text{rank(T)}$.

We now prove that $\text{rank}(T) = \text{rank}(TS)$. Since $S: U \to V$ is invertible, S is onto. So, $S(U) = V$. By the definition of range,

$$R(TS) = TS(U) = T(S(U) = T(V).$$

So, $R(TS) = T(V) = R(T)$, so $\text{rank}(T) = \text{rank}(TS)$.

Finally, the equality $\text{rank}(PTS) = \text{rank}(TS)$ follows by applying the first equality to $T' := TS$. □

Definition 3.4 (**Rank of a Matrix**). Let A be a matrix. Then the **rank** of A is defined as $\text{rank}(L_A)$.

Lemma 3.5. *The rank of a matrix A is equal to the dimension of the space spanned by the columns of A.*

Proof. Suppose A is an $m \times n$ matrix. Let $(e_1, \ldots, e_n)$ be the standard basis of $\mathbf{F}^n$. Since this basis spans $\mathbf{F}^n$, the vectors $\{L_A(e_1), \ldots, L_A(e_n)\}$ span $R(L_A)$ by Theorem 1.2. But for each $i \in \{1, \ldots, n\}$, $L_A(e_i)$ is the i^{th} column of A. □

Remark 3.6. Suppose V and W are finite dimensional vector spaces. Let α, γ be ordered bases for V and let β, δ be ordered bases for W. Let $T: V \to W$ be a linear transformation. Recall that any two matrix representations $[T]_\alpha^\beta$ and $[T]_\gamma^\delta$ are related by the identity $[T]_\gamma^\delta = [I_W]_\beta^\delta [T]_\alpha^\beta [I_V]_\gamma^\alpha$. Also, two vector spaces of the same dimension are isomorphic. So, to compute the rank of T, it suffices to find any matrix representation A of T, and then to compute the rank of A. By Lemma 3.3, isomorphisms preserve rank, so any matrix representation A suffices. And we can compute the rank of A by row-reducing it into row-echelon form, and then applying the following lemma.

Lemma 3.7. *Let A be a matrix in row-echelon form. Then the rank of A is equal to the number of nonzero rows of A.*

Proof. Since A is in row-echelon form, there exists a positive integer k such that, each of the first k rows of A has one nonzero entry, while all subsequent rows of A are zero. So, the span of the columns of A are contained in the k-dimensional subspace

$$\{\begin{pmatrix} x_1 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{pmatrix} : x_1, \ldots, x_k \in \mathbf{F}\}. \qquad (*)$$

So, by Lemma 3.5, $\text{rank}(A) \leq k$. We now show that in fact $\text{rank}(A) = k$, as desired.

By Lemma 3.5, it suffices to show that the span of the columns of A contains the subspace $(*)$. To show this, let v be in the subspace $(*)$. Then, as a column vector, we have $v = (v_1, \ldots, v_k, 0, \ldots, 0)$, $v_i \in \mathbf{F}$ for all $i \in \{1, \ldots, k\}$. Consider the i^{th} row of A where $1 \leq i \leq k$. Since A is in row-echelon form, the i^{th} row first has several zeros, then a 1, then other entries afterwards. So, for each $i \in \{1, \ldots, k\}$, there exists $j(i)$ such that the $j(i)^{th}$ column of A has a 1 in the i^{th} row, and then zeros below that. So, beginning with $i = k$, we can subtract v_k copies of the $j(k)^{th}$ column of A from v, giving a vector with only $(k-1)$ nonzero entries. Then, setting $i = k-1$, we can subtract copies of the $j(k-1)^{st}$ column of A to get a vector with only $(k-2)$ nonzero entries. We continue in this way, and eventually we have eliminated all nonzero entries of v. That is, we have found an expression for v in terms of the columns $j(1), \ldots, j(k)$ of A. So, $\text{rank}(A) = k$, as desired.

□

Theorem 3.8. *Let A be an $m \times n$ matrix of rank r. Then, there exist a finite number of elementary row and column operations which, when applied to A, produce the matrix*

$$\begin{pmatrix} I_{r\times r} & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{pmatrix}.$$

Proof. We first use row reduction to put A into row-echelon form. So, after this row reduction, the first r rows of A have some zeros, and then a 1 with zeros below this 1. And the remaining $m-r$ rows are all zero. (In case $r = 0$, then we are done, so we may assume that $r > 0$.) Now, the first row of A has some zeros, then a 1 with zeros below this 1. So, by adding copies of the column that contains the entry 1 to each column to the right, the remaining entries of the first row can be made to be zero. And we still keep our matrix in row-echelon form. Now, the second row of A has some zeros, then a 1 with zeros above and below this 1. So, by adding copies of the column that contains this entry 1 to each column to the right, the remaining entries of the second row can be made to be zero. And once again, our matrix is still in row-echelon form. We then continue this procedure. The first r

rows then each have exactly one entry of 1, and all remaining entries in the matrix are zero. By swapping columns as needed, A is then put into the required form, as desired. $\square$

Corollary 3.9 (**A Factorization Theorem**). *Let A be an $m \times n$ matrix of rank r. Then, there exists an $m \times m$ matrix B and an $n \times n$ matrix C such that B is the product of a finite number of elementary row operations, C is the product of a finite number of elementary column operations, and such that*

$$A = B \begin{pmatrix} I_{r\times r} & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{pmatrix} C.$$

Proof. Let A be an $m \times n$ matrix of rank r. From Theorem 3.8, there exist a finite number of elementary row operations $E_1, \ldots, E_j$ and elementary column operations $F_1, \ldots, F_k$ such that

$$E_1 \cdots E_j A F_1 \cdots F_k = \begin{pmatrix} I_{r\times r} & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{pmatrix}. \qquad (*)$$

From Remarks 2.7 and 2.10, the matrices $E_1, \ldots, E_j$ and $F_1, \ldots, F_k$ are invertible, with inverses that are also elementary row and column operations, respectively. So, multiplying on the left of each side of $(*)$ by $B := E_j^{-1} \cdots E_1^{-1}$, and then multiplying on the right of each side of $(*)$ by $C := F_k^{-1} \cdots F_1^{-1}$, we deduce the theorem. $\square$

Lemma 3.10. *Let A be an $m \times n$ matrix. Let B be an $m \times m$ invertible matrix, and let C be an $n \times n$ invertible matrix. Then*

$$\mathrm{rank}(A) = \mathrm{rank}(BA) = \mathrm{rank}(AC) = \mathrm{rank}(BAC).$$

Proof. Since B is invertible, L_B is invertible with inverse $L_{B^{-1}}$, by Corollary 1.4. So, applying Remark 1.3 and Lemma 3.3,

$$\mathrm{rank}(L_A) = \mathrm{rank}(L_{BA}) = \mathrm{rank}(L_{AC}) = \mathrm{rank}(L_{BAC}).$$

Definition 3.4 then completes the proof. $\square$

Definition 3.11 (**Transpose**). Let A be an $m \times n$ matrix with entries A_{ij}, $1 \leq i \leq m$, $1 \leq j \leq n$. Then the **transpose** A^t of A is defined to be the $n \times m$ matrix with entries $(A^t)_{ij} := A_{ji}$, $1 \leq i \leq n$, $1 \leq j \leq m$.

Exercise 3.12. Let A be an $m \times n$ matrix. Let B be an $\ell \times m$ matrix. Show that $(BA)^t = A^t B^t$.

Remark 3.13. If A is an $n \times n$ invertible matrix, then $I_n^t = (AA^{-1})^t = (A^{-1})^t A^t$, so A^t is also invertible.

Lemma 3.14. *Let A be an $m \times n$ matrix with rank r. Then A^t also has rank r.*

Proof. From Theorem 3.9, there exists an invertible $m \times m$ matrix B and an invertible $n \times n$ matrix C such that

$$A = B \begin{pmatrix} I_{r\times r} & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{pmatrix} C.$$

Taking the transpose of both sides and applying Exercise 3.12,

$$A^t = C^t \begin{pmatrix} I_{r\times r} & 0_{r\times(m-r)} \\ 0_{(n-r)\times r} & 0_{(n-r)\times(m-r)} \end{pmatrix} B^t.$$

From Remark 3.13, C^t and B^t are invertible. So, Lemma 3.10 implies that A^t has rank r. $\square$

Corollary 3.15. *The rank of a matrix is equal to the dimension of the span of its rows.*

Proof. Apply Lemma 3.5 and Lemma 3.14. ☐

Lemma 3.16. *Let V be an n-dimensional vector space, and let W be an m-dimensional vector space. Let $T\colon V \to W$ be a linear transformation. Let α, β be finite bases for V, W respectively. Then* $\mathrm{rank}(T) = \mathrm{rank}([T]_\alpha^\beta)$.

Proof. Let $v \in V, w \in W$. The coordinate maps $\phi_\alpha\colon V \to \mathbf{F}^n$ and $\phi_\beta\colon W \to \mathbf{F}^m$ defined by $\phi_\alpha(v) := [v]^\alpha$, $\phi_\beta(w) := [w]^\beta$ are isomorphisms. Also, the map $L_{[T]_\alpha^\beta}\colon \mathbf{F}^n \to \mathbf{F}^m$ is a linear transformation. Beginning with the identity

$$[T(v)]^\beta = [T]_\alpha^\beta [v]^\alpha,$$

we then rewrite this as

$$\phi_\beta(T(v)) = L_{[T]_\alpha^\beta}\phi_\alpha(v).$$

Since this equality holds for all $v \in V$, we therefore have

$$\phi_\beta T = L_{[T]_\alpha^\beta}\phi_\alpha.$$

Since ϕ_β is invertible, we then get

$$T = \phi_\beta^{-1} L_{[T]_\alpha^\beta}\phi_\alpha.$$

So, applying Lemma 3.10 and Definition 3.4,

$$\mathrm{rank}(T) = \mathrm{rank}(\phi_\beta^{-1} L_{[T]_\alpha^\beta}\phi_\alpha) = \mathrm{rank}(L_{[T]_\alpha^\beta}) = \mathrm{rank}([T]_\alpha^\beta).$$

☐

Exercise 3.17. Show that an $m \times n$ matrix has rank at most $\min(m, n)$.

3.1. **Inverting a Matrix.**

Lemma 3.18. *Let A be an $n \times n$ matrix. Then A is invertible if and only if it is the product of elementary row and column operations.*

Proof. Suppose A is a product of elementary row and column operation matrices. From Remarks 2.7 and 2.10, A is a product of invertible matrices, so A is invertible.

Now, suppose A is invertible. Then $L_A\colon \mathbf{F}^n \to \mathbf{F}^n$ is invertible (with inverse $L_{A^{-1}}$). In particular, L_A is onto, so $\mathrm{rank}(L_A) = n$. By Definition 3.4, $\mathrm{rank}(A) = n$. Applying our Factorization Theorem (Theorem 3.9), there exists a finite number of elementary row operations $E_1, \ldots, E_j$ and elementary column operations $F_1, \ldots, F_k$ such that $A = E_1 \cdots E_j F_1 \cdots F_k$, as desired. ☐

Remark 3.19. Suppose A is an invertible matrix, and we have elementary row operations $E_1, \ldots, E_j$ such that

$$E_1 \cdots E_j A = I_n.$$

Multiplying both sides by A^{-1} on the right,

$$E_1 \cdots E_j I_n = A^{-1}.$$

So, to compute A^{-1} from A, it suffices to find row operations that turn A into the identity. And we then apply these operations to I_n to give A^{-1}. This is the algorithm for computing the inverse A^{-1} that you learned in a previous class.

4. The Determinant

There are a lot of nice things to say about the determinant, but we do not have sufficient time to discuss these things. We will therefore just state some facts about the determinant without proof, and then prove other things as consequences of these preliminary facts.

Let $A \in \mathbf{F}$. Then $\det(A) := A$.

Let A be a 2×2 matrix. That is, there exist $a, b, c, d \in \mathbf{F}$ such that

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

We then define $\det(A)$ so that

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

Let A be a 3×3 matrix. That is, there exist $a, b, c, d, e, f, g, h, i \in \mathbf{F}$ such that

$$A = \det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}.$$

We now define $\det(A)$ inductively so that

$$\det(A) := a \det \begin{pmatrix} e & f \\ h & i \end{pmatrix} - b \det \begin{pmatrix} d & f \\ g & i \end{pmatrix} + c \det \begin{pmatrix} d & e \\ g & h \end{pmatrix}.$$

Definition 4.1. More generally, if A is an $n \times n$ matrix, then for each $i, j \in \{1, \ldots, n\}$, let $\overline{A}_{ij}$ denote the $(n-1) \times (n-1)$ matrix formed by removing the i^{th} row and j^{th} column from A. Then, for any $i \in \{1, \ldots, n\}$, define

$$\det(A) := \sum_{j=1}^{n} (-1)^{i+j} A_{ij} \det(\overline{A}_{ij})$$

If A has columns $v_1, \ldots, v_n$, we write $\det(A) = \det(v_1, \ldots, v_n)$ to emphasize that the determinant is a function of the columns of A.

Remark 4.2 (Properties of the Determinant). Let $v_1, \ldots, v_n \in \mathbf{F}^n$.

(a) For all $\alpha \in \mathbf{F}$, for all $w \in \mathbf{F}^n$, for all $i \in \{1, \ldots, n\}$

$$\begin{aligned} &\det(v_1, \ldots, v_{i-1}, v_i + \alpha w, v_{i+1}, \ldots, v_n) \\ &\qquad = \det(v_1, \ldots, v_n) + \alpha \det(v_1, \ldots, v_{i-1}, w, v_{i+1}, \ldots, v_n). \qquad \text{(Multilinear)} \end{aligned}$$

(b) For all $i, j \in \{1, \ldots, n\}$ with $i \neq j$,

$$\det(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = -\det(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n). \qquad \text{(Alternating)}$$

(c) $\det(I_n) = 1$. (Normalized)

(d) For all $n \times n$ matrices A, B, we have $\det(AB) = \det(A)\det(B)$.

(e) For all $n \times n$ matrices A, we have $\det(A) = \det(A^t)$.

Theorem 4.3. *Suppose we have two functions F, G that map $v_1, \ldots, v_n \in \mathbf{F}^n$ to $\mathbf{F}$, both satisfying properties $(a), (b)$ and (c) above. Then $F = G$.*

Proof. Define $D(v_1, \ldots, v_n) := F(v_1, \ldots, v_n) - G(v_1, \ldots, v_n)$. We will show that $D = 0$. Since F, G both satisfy properties $(a), (b)$, D satisfies properties $(a), (b)$. Since F, G both satisfy property (c), we have $D(I_n) = 0$. Since D satisfies property (b) and $D(e_1, \ldots, e_n) = 0$, if $(e_{j(1)}, \ldots, e_{j(n)})$ is any permutation of the standard basis $(e_1, \ldots, e_n)$, we have

$$D(e_{j(1)}, \ldots, e_{j(n)}) = 0. \qquad (*)$$

Let $v_i \in \mathbf{F}^n$, and write $v_i = \sum_{j=1}^n \alpha_{ij} e_j$, $\alpha_{ik} \in \mathbf{F}$ for all $i, j \in \{1, \ldots, n\}$. Repeatedly applying property (a),

$$\begin{aligned} D(v_1, \ldots, v_n) &= D(\sum_{j=1}^n \alpha_{1j} e_j, v_2, \ldots, v_n) \\ &= \sum_{j=1}^n \alpha_{1j} D(e_j, v_2, \ldots, v_n) \\ &= \sum_{j_1=1}^n \alpha_{1j_1} D(e_{j_1}, \sum_{j=1}^n \alpha_{2j} e_j, \ldots, v_n) \\ &= \sum_{j_1=1}^n \sum_{j_2=1}^n \alpha_{1j_1} \alpha_{2j_2} D(e_{j_1}, e_{j_2}, \ldots, v_n) \\ &= \cdots = \sum_{j_1=1}^n \cdots \sum_{j_n=1}^n \alpha_{1j_1} \cdots \alpha_{nj_n} D(e_{j_1}, \ldots, e_{j_n}) \end{aligned}$$

And the final quantity is zero, by $(*)$, as desired. □

Theorem 4.3 can be used to show that various different definitions of the determinant all agree. Given some formula that should be equal to the determinant, it suffices to prove that this formula satisfies properties $(a), (b)$ and (c). For example, consider the following determinant formula you learned in Calc 3, for vectors $v_1, v_2, v_3 \in \mathbf{R}^3$:

$$\det(v_1, v_2, v_3) = v_1 \cdot (v_2 \times v_3).$$

Here $\cdot$ denotes the dot product, and $\times$ denotes the cross product. You could write the right side in coordinates to verify that it agrees with the left side. Or, you could verify that the right side satisfies properties $(a), (b)$ and (c), and then apply Theorem 4.3, instead giving a coordinate-free proof of the desired identity.

As another application of Theorem 4.3, we can show that property (d) of Remark 4.2 holds.

Theorem 4.4. *Assume that the determinant function satisfies properties* $(a), (b)$ *and* (c) *from Remark 4.2. Then the determinant function satisfies property* (d)*. For all* $n \times n$ *matrices* A, B*, we have* $\det(AB) = \det(A)\det(B)$.

Proof. Suppose $\det(A) \neq 0$. For $v_1, \ldots, v_n \in \mathbf{F}^n$, define

$$F(v_1, \ldots, v_n) := \det(Av_1, \ldots, Av_n)/\det(A).$$

Note that F then satisfies properties $(a), (b)$ and (c). So, we have by Theorem 4.3 that $F(B) = \det(AB)/\det(A) = \det(B)$. So, $\det(AB) = \det(A)\det(B)$, as desired.

In the case $\det(A) = 0$, define

$$F(v_1, \ldots, v_n) := \det(v_1, \ldots, v_n) + \det(Av_1, \ldots, Av_n).$$

Once again, F satisfies properties $(a), (b)$ and (c), so $F(B) = \det(B) = \det(B) + \det(AB)$. So, $\det(AB) = 0 = \det(A)\det(B)$. In any case, $\det(AB) = \det(A)\det(B)$, as desired. □

Theorem 4.5. *Let A be an $n \times n$ matrix. Then A is invertible if and only if* $\det(A) \neq 0$. *If A is invertible, then* $\det(A^{-1}) = (\det(A))^{-1}$.

Proof. Suppose A has rank r. From the Factorization Theorem (Theorem 3.9), A is the product of elementary row and column operations, and also a diagonal matrix D with r ones on the diagonal. If $r < n$, then $\det(D) = 0$, so $\det(A) = 0$ as well from property (d) of Remark 4.2. We have shown that, if A has rank less than n, then $\det(A) = 0$. Taking the contrapositive, if $\det(A) \neq 0$, then A has rank n. From Lemma 3.1, A is invertible if and only if A has rank n. So, if $\det(A) \neq 0$, then A is invertible.

We now prove the converse. Suppose A is invertible. From property (d) of Remark 4.2, $1 = \det(I_n) = \det(AA^{-1}) = \det(A)\det(A^{-1})$. So, $\det(A)$ must be nonzero. □

Corollary 4.6. *For any $n \times n$ matrix A,* $\det(A) = \det(A^t)$.

Proof. Suppose A has rank r. From the Factorization Theorem (Theorem 3.9), there exist elementary row operations $E_1, \ldots, E_j$ and elementary column operations $F_1, \ldots, F_k$, and there exists a diagonal matrix D with r ones on the diagonal such that

$$A = E_1 \cdots E_j D F_1 \cdots F_k \qquad (*).$$

Taking the transpose of $(*)$,

$$A^t = F_k^t \cdots F_1^t D E_j^t \cdots E_1^t. \qquad (**)$$

From Theorem 4.4 applied to $(*)$

$$\det(A) = \det(E_1) \cdots \det(E_j) \det(D) \det(F_1) \cdots \det(F_k).$$

By checking Type 1, 2 and 3 matrices from Examples 2.1, 2.3 and 2.5 directly, we see that $\det(E) = \det(E^t)$ for any elementary row or column operation E. So, applying Theorem 4.4 to $(**)$,

$$\begin{aligned}\det(A^t) &= \det(F_k^t) \cdots \det(F_1^t) \det(D) \det(E_j^t) \cdots \det(E_1^t)\\ &= \det(E_1) \cdots \det(E_j) \det(D) \det(F_1) \cdots \det(F_k) = \det(A).\end{aligned}$$

□

5. Appendix: Notation

Let A, B be sets in a space X. Let m, n be a nonnegative integers. Let $\mathbf{F}$ be a field.

$\mathbf{Z} := \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$, the integers

$\mathbf{N} := \{0, 1, 2, 3, 4, 5, \ldots\}$, the natural numbers

$\mathbf{Q} := \{m/n \colon m, n \in \mathbf{Z}, n \neq 0\}$, the rationals

$\mathbf{R}$ denotes the set of real numbers

$\mathbf{C} := \{x + y\sqrt{-1} : x, y \in \mathbf{R}\}$, the complex numbers

$\emptyset$ denotes the empty set, the set consisting of zero elements

$\in$ means "is an element of." For example, $2 \in \mathbf{Z}$ is read as "2 is an element of $\mathbf{Z}$."

$\forall$ means "for all"

$\exists$ means "there exists"

$\mathbf{F}^n := \{(x_1, \ldots, x_n) \colon x_i \in \mathbf{F}, \forall\, i \in \{1, \ldots, n\}\}$

$A \subseteq B$ means $\forall\, a \in A$, we have $a \in B$, so A is contained in B

$A \smallsetminus B := \{x \in A \colon x \notin B\}$

$A^c := X \smallsetminus A$, the complement of A

$A \cap B$ denotes the intersection of A and B

$A \cup B$ denotes the union of A and B

$C(\mathbf{R})$ denotes the set of all continuous functions from $\mathbf{R}$ to $\mathbf{R}$

$P_n(\mathbf{R})$ denotes the set of all real polynomials in one real variable of degree at most n

$P(\mathbf{R})$ denotes the set of all real polynomials in one real variable

$M_{m\times n}(\mathbf{F})$ denotes the vector space of $m \times n$ matrices over the field $\mathbf{F}$

I_n denotes the $n \times n$ identity matrix

5.1. **Set Theory.** Let V, W be sets, and let $f \colon V \to W$ be a function. Let $X \subseteq V$, $Y \subseteq W$.

$$f(X) := \{f(v) \colon v \in V\}.$$

$$f^{-1}(Y) := \{v \in V \colon f(v) \in Y\}.$$

The function $f \colon V \to W$ is said to be **injective** (or **one-to-one**) if: for every $v, v' \in V$, if $f(v) = f(v')$, then $v = v'$.

The function $f \colon V \to W$ is said to be **surjective** (or **onto**) if: for every $w \in W$, there exists $v \in V$ such that $f(v) = w$.

The function $f \colon V \to W$ is said to be **bijective** (or a **one-to-one correspondence**) if: for every $w \in W$, there exists exactly one $v \in V$ such that $f(v) = w$. A function $f \colon V \to W$ is bijective if and only if it is both injective and surjective.

Two sets X, Y are said to have the same **cardinality** if there exists a bijection from V onto W.

The **identity map** $I \colon X \to X$ is defined by $I(x) = x$ for all $x \in X$. To emphasize that the domain and range are both X, we sometimes write I_X for the identity map on X.

Let V, W be vector spaces over a field $\mathbf{F}$. Then $\mathcal{L}(V, W)$ denotes the set of linear transformations from V to W, and $\mathcal{L}(V)$ denotes the set of linear transformations from V to V.

UCLA Department of Mathematics, Los Angeles, CA 90095-1555
E-mail address: heilman@math.ucla.edu