1 Matrix Multiplication

Proposition 1. *If $T : \mathcal{U} \to \mathcal{V}$ and $S : \mathcal{V} \to W$ are two linear transformations, then the composite $S \circ T : \mathcal{U} \to \mathcal{W}$ is a linear transformation.*

Proof. We need to check that $S \circ T$ distributes over addition and scalar multiplication.

$$S \circ T(\vec{v} + \vec{w}) = S(T(\vec{v}) + T(\vec{w})) = S \circ T(\vec{v}) + S \circ T(\vec{w}).$$

$$S \circ T(c\vec{v}) = S(cT(\vec{v})) = cS \circ T(\vec{v}).$$

□

Definition 1. If $A \in M(m, n)$ and $B \in M(n, p)$ with columns $\vec{v}_1, \ldots, \vec{v}_p$, then we define the ***product*** matrix $AB \in M(m, p)$ by

$$AB = \begin{bmatrix} A\vec{v}_1 & A\vec{v}_2 & \cdots & A\vec{v}_p \end{bmatrix}.$$ [1]

Proposition 2. *If $S : \mathbb{R}^n \to \mathbb{R}^m$ is represented by an $m \times n$ matrix A, and $T : \mathbb{R}^p \to \mathbb{R}^n$ is represented by an $n \times p$ matrix B, then $S \circ T : \mathbb{R}^p \to \mathbb{R}^m$ can be represented by the $m \times p$ product matrix AB.*

Proof. Recall that the j^{th} column of the matrix representation of $S \circ T$ is given by $(S \circ T)(\hat{e}_j)$. Thus, we need to show that this equals the j^{th} columns of AB. Let $\vec{v}_1, \ldots, \vec{v}_p$ be the columns of B, then we have

$$(S \circ T)(\hat{e}_j) = S(B\hat{e}_j) = S(\vec{v}_j) = A\vec{v}_j,$$

Which is the j^{th} column of AB by definition. □

Here are is another way to compute the entries of AB.

Proposition 3. *Let $\vec{r}_1, \ldots, \vec{r}_m \in M(1, n)$ be the rows of $A = [a_{ik}] \in M(m, n)$ and let $\vec{v}_1, \ldots, \vec{v}_p \in M(n, 1)$ be the columns of $B = [b_{kj}] \in M(n, p)$. Then,*

$$ij^{th} \textit{ entry of } AB = \vec{r}_i\, \vec{v}_j = \vec{r}_i^{\,t} \cdot \vec{v}_j = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

The ij^{th} entry of AB is the dot product of the i^{th} row of A and the j^{th} column of B.

Proof. By definition, the j^{th} column of AB is $A\vec{v}_j$. Reviewing the definition of $A\vec{v}_j$, we see that the i^{th} entry of $A\vec{v}_j$ is indeed $\sum_{k=1}^{n} a_{ik} b_{kj}$, which equals $\vec{r}_i^{\,t} \cdot \vec{v}_j$. □

[1] Matrix multiplication is ***not*** defined if the number of columns of A doesn't equal the number fo rows of B

Example 1.

$$\begin{bmatrix}\cos(\alpha) & -\sin(\alpha)\\ \sin(\alpha) & \cos(\alpha)\end{bmatrix}\begin{bmatrix}\cos(\beta) & -\sin(\beta)\\ \sin(\beta) & \cos(\beta)\end{bmatrix} =$$

$$= \begin{bmatrix}\cos(\alpha)\cos(\beta)-\sin(\alpha)\sin(\beta) & -\cos(\alpha)\sin(\beta)-\sin(\alpha)\cos(\beta)\\ \sin(\alpha)\cos(\beta)+\cos(\alpha)\sin(\beta) & -\sin(\alpha)\sin(\beta)+\cos(\alpha)\cos(\beta)\end{bmatrix} =$$

$$= \begin{bmatrix}\cos(\alpha+\beta) & -\sin(\alpha+\beta)\\ \sin(\alpha+\beta) & \cos(\alpha+\beta)\end{bmatrix}$$

Notice that these matrices correspond to counter-clockwise rotation around the origin in $\mathbb{R}^2$ of angle α and β respectively. As you can see, their composition is then a rotation by angle $\alpha+\beta$.

Example 2 (Row reduction as matrix multiplication on the left)**.** Notice that the three row reduction operations on a matrix A can be expressed through multiplication on the left by what are called ***elementary matrices***. If $A \in M(4,n)$ with rows $\vec{r}_1, \ldots, \vec{r}_4$. Let

$$D_2(s) = \begin{bmatrix}1&0&0&0\\0&s&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix} \qquad T_{14} = \begin{bmatrix}0&0&0&1\\0&1&0&0\\0&0&1&0\\1&0&0&0\end{bmatrix} \qquad L_{31}(s) = \begin{bmatrix}1&0&0&0\\0&1&0&0\\s&0&1&0\\0&0&0&1\end{bmatrix}.$$

Notice that $D_2(s)A$ is the matrix there the second row of A is scaled by s. Also, $T_{14}A$ is the matrix where rows $\vec{r}_1$ and $\vec{r}_4$ are interchanged. Lastly, $L_{31}(s)$ is the matrix where $\vec{r}_3$ is replaced by $\vec{r}_3 + s\vec{r}_1$.

Since matrix multiplication is a representation of composition of linear transformations. We automatically have the following result.

Proposition 4. *Matrix multiplication is*

- *associative:* $A(BC) = (AB)C$
- *left distributative:* $A(B+C) = AB + AC$
- *right distributative:* $(A+B)C = AC + BC$
- *scalar multiplicative:* $A(cB) = c(AB) = (cA)B$

Note. Matrix multiplication need not commute, i.e. it can be $AB \neq BA$. First, one may not be defined. Next, AB and BA need not be the same size matrix. Even if they are the same size, need not be equal:

$$\begin{bmatrix}1&2\\3&4\end{bmatrix}\begin{bmatrix}1&0\\1&1\end{bmatrix} = \begin{bmatrix}3&2\\7&4\end{bmatrix}, \qquad \begin{bmatrix}1&0\\1&1\end{bmatrix}\begin{bmatrix}1&2\\3&4\end{bmatrix} = \begin{bmatrix}1&2\\4&6\end{bmatrix}.$$

1.1 Linear transformations from $\mathbb{R}^2$ to itself

Consider a linear transformations T on $\mathbb{R}^2$. Let T be represented by a 2×2 matrix A, where the columns of A are $T(\hat{e}_1)$ and $T(\hat{e}_2)$. Here are some key examples:

- ***Scaling:*** A is a diagonal matrix, $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$, $a, b \neq 0$.

- ***Rotation:*** Rotations preserve the angle between two vectors and also the length of each vector. A rotation of degree α has the form

$$A = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$

  Check that product of two rotation matrices is again a rotation matrix.

- ***Projection:*** A projection onto the line spanned by the ***unit*** vector $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ (can be generalized to $\mathbb{R}^n$) takes the form

$$T(\vec{w}) = (\vec{w} \cdot \vec{v})\vec{v}, \qquad A = \begin{bmatrix} v_1^2 & v_1 v_2 \\ v_1 v_2 & v_2^2 \end{bmatrix}$$

  . Note that $\ker(T)$ is the line perpendicular to $\vec{v}$ and $\operatorname{im}(T) = \operatorname{span}\{\vec{v}\}$.

  Note. It is obvious that rotation and projection matrices need not commute. For instance, let $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ be the rotation matrix counter clockwise by 90 degrees, and $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ be projection onto the x-axis $\vec{e}_1$. Then, the vector $\vec{e}_1$ would be mapped by AB to $\vec{e}_2$, but by BA to $\vec{0}$. Indeed

$$AB = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \qquad BA = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix}.$$

- ***Reflection:*** A reflection across a line spanned by the ***unit*** vector $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ (can be generalized to $\mathbb{R}^n$) takes the form

$$T(\vec{w}) = 2(\vec{w} \cdot \vec{v})\vec{v} - \vec{w}, \qquad A = \begin{bmatrix} 2v_1^2 - 1 & 2v_1 v_2 \\ 2v_1 v_2 & 2v_2^2 - 1 \end{bmatrix}.$$

  For example,

$$A = \begin{bmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{bmatrix}, \qquad \vec{v} = \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix}$$

- ***Shear:*** Transformations in plane with the property that there is a unit vector $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ such that $T(\vec{v}) = \vec{v}$ and $T(\vec{w}) - \vec{w}$ is a multiple of $\vec{v}$ for all $\vec{w}$. If $\vec{u} \cdot \vec{v} = 0$, then

  $$T(\vec{w}) = \vec{w} + (\vec{u} \cdot \vec{w})\vec{v}, \qquad \vec{u} = c \begin{bmatrix} v_2 \\ -v_1 \end{bmatrix}, \qquad A = \begin{bmatrix} cv_1v_2 + 1 & -cv_1^2 \\ cv_2^2 & -cv_1v_2 + 1 \end{bmatrix}.$$

  For example,

  $$A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \qquad \vec{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \vec{u} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

2 Inverses

Recall that for a linear transformation $T : \mathcal{V} \to \mathcal{W}$, where $\mathcal{V}$ and $\mathcal{W}$ are ***finite dimensional***, the following are true

- T is injective (a.k.a. one-to-one) if and only if $\ker(T)$ is trivial, if and only if $\text{rank}(T) = \dim(\mathcal{V})$.
- T is surjective (a.k.a. onto) if and only if $\text{rank}(T) = \dim(\mathcal{W})$.
- If $\dim(\mathcal{V}) = \dim(\mathcal{W})$, then T is injective if and only if T is surjective.

The goal of this section is to discuss the idea of "undoing" a linear transformation. We will first do this in the language on linear transformations.

Definition 2. A linear transformation $T : \mathcal{V} \to \mathcal{W}$ between two vectors spaces is said to be ***invertible*** if there exists a map $S : \mathcal{W} \to \mathcal{V}$, such that $T \circ S = I_{\mathcal{W}}$ and $S \circ T = I_{\mathcal{V}}$, where $I_{\mathcal{V}}$ and $I_{\mathcal{W}}$ are identity maps on $\mathcal{V}$ and $\mathcal{W}$ (i.e. $I_{\mathcal{V}}(X) = X$ for all $X \in \mathcal{V}$ and $I_{\mathcal{W}}(Y) = Y$ for all $Y \in \mathcal{W}$.

Proposition 5. *T is invertible if and only if T is bijective (i.e. one-to-one and onto).*

Proof. ($\Rightarrow$) Assume T is invertible and let S be an inverse. We see that

(i) T is surjective: Let $X_Y = S(Y)$ for a given $Y \in \mathcal{W}$, then $T(X_Y) = T(S(Y) = I_{\mathcal{W}}(Y) = Y$, so $Y \in \text{im}T$ and T is surjective.

(ii) T is injective: Assume $T(X_1) = T(X_2)$, then $X_1 = I_{\mathcal{V}}(X_1) = S(T(X_1)) = S(T(X_2)) = I_{\mathcal{V}}(X_2) = X_2$, so $X_1 = X_2$ and T is injective.

($\Leftarrow$) Conversely, assume T is bijective. Then, for any $Y \in \mathcal{W}$, by definition, there exists a ***unique*** X such that $T(X) = Y$. Define $S(Y) = X$. We see that $(T \circ S)(Y) = T(X) = Y$ and $(S \circ T)(X) = S(Y) = X$, so S is indeed an inverse. $\square$

If $T : \mathcal{V} \to \mathcal{W}$ is an invertible linear transformation, then we say $\mathcal{V}$ and $\mathcal{W}$ are ***isomorphic*** and T is an ***isomorphism*** between $\mathcal{V}$ and $\mathcal{W}$.

Proposition 6. *If $T : \mathcal{V} \to \mathcal{W}$ is a linear map, with $\mathcal{V}$ and $\mathcal{W}$* **finite dimensional***, the following are* **equivalent**

- *T is invertible.*
- *T is an isomorphim.*
- *T is bijective.*
- $\operatorname{rank}(T) = \dim \mathcal{W}$ *and* $\operatorname{null}(T) = 0$
- $\dim(\mathcal{V}) = \dim(\mathcal{W})$ *and T is injective.*
- $\dim(\mathcal{V}) = \dim(\mathcal{W})$ *and* $\operatorname{null}(T) = 0$.
- $\dim(\mathcal{V}) = \dim(\mathcal{W}) = \operatorname{rank}(T)$.

There are many more properties about invertible transformations that we need to know. One particular question that should stand out, is an inverse also a linear map? It is, as shown below. You might also wonder if inverses are unique.

Proposition 7. *(i) The inverse of an invertible linear transformation must be a linear transformation.*

(ii) The inverse of an invertible linear transformation is unique. We will let T^{-1} denote the unique inverse of T.

(iii) $(T^{-1})^{-1} = T$ and, in particular, T^{-1} is invertible.

(iv) For invertible linear transformations T_1 and T_2, $(T_1 \circ T_2)^{-1} = T_2^{-1} \circ T_1^{-1}$

Proof. (i) Let S be an inverse of T. For $Y_1, Y_2 \in \mathcal{W}$, we know there are $X_1, X_2 \in \mathcal{V}$ such that $T(X_i) = Y_i$ and $S(Y_i) = X_i$. We check that S is linear by

$$S(a_1Y_1 + a_2Y_2) = S(a_1T(X_1) + a_2T(X_2)) \xlongequal{T \text{ is lin.}} S(T(a_1X_1 + a_2X_2))$$

$$\xlongequal{S \circ T = I_{\mathcal{V}}} a_1X_1 + a_2X_2 = a_1S(Y_1) + a_2S(Y_2).$$

(ii) Suppose $T : \mathcal{V} \to \mathcal{W}$ has inverses S and R, then

$$R = I_{\mathcal{V}} \circ R = (S \circ T) \circ R = S \circ (T \circ R) = S \circ I_{\mathcal{W}} = S.$$

Hence, $R = S$. We will let T^{-1} denote the unique inverse of T.

(iii) Suppose $T : \mathcal{V} \to \mathcal{W}$ has inverse $S = T^{-1}$. So, $S \circ T = I_{\mathcal{V}}$ and $T \circ S = I_{\mathcal{W}}$. By definition, T must be S^{-1}, and therefore $(T^{-1})^{-1} = T$.

(iv) Suppose $T : \mathcal{V} \to \mathcal{W}$ and $S : \mathcal{W} \to \mathcal{U}$ are invertible linear transformations. Check that

$$T^{-1} \circ S^{-1}) \circ (S \circ T) = T^{-1} \circ (S^{-1} \circ S) \circ T = T^{-1} \circ I_{\mathcal{W}} \circ T = T^{-1} \circ T = I_{\mathcal{V}}$$
$$(S \circ T) \circ (T^{-1} \circ S^{-1}) = S \circ (T \circ T^{-1}) \circ S^{-1} = S \circ I_{\mathcal{W}} \circ S^{-1} = S \circ S^{-1} = I_{\mathcal{U}}.$$

Hence $S \circ T$ has $T^{-1} \circ S^{-1}$ as its unique inverse.

□

2.1 Inverses and matrices

Notice that in order for $T : \mathbb{R}^n \to \mathbb{R}^m$ to be invertible, we need, at the least, that $m = n$ (but, of course, this is ***not enough***). For $A, B \in M(n, n)$, if $T_B = T_A^{-1}$ then we must have $AB = I_n = BA$, were I_n is the $n \times n$ identity matrix.

Definition 3. Let A be an $n \times n$ square matrix A. A matrix A^{-1} is the ***inverse*** of A if and only $AA^{-1} = A^{-1}A = I_n$. The matrix A is called ***invertible*** if A^{-1} exists.

It follows from the proposition above that, A^{-1} is unique, and

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

Proposition 8. *If A is invertible, then there is a unique solution to $A\vec{x} = \vec{b}$ given by $\vec{x} = A^{-1}\vec{b}$. In particular, $A\vec{x} = \vec{0}$ can not have non-zero solution.*

Proof. Since $A(A^{-1}\vec{b}) = (AA^{-1})\vec{b} = \vec{b}$, $A^{-1}\vec{b}$ is a solution. $A\vec{x} = \vec{0}$ has solution $A^{-1}\vec{0} = \vec{0}$. □

Proposition 9. *If A is invertible, then*

$$(A^{-1})^t = (A^t)^{-1}.$$

Proof. Since $AA^{-1} = I$, apply transpose to both sides, we have $(A^{-1})^t A^t = I$, so $(A^t)^{-1} = (A^{-1})^t$. □

So how to we actually compute A^{-1}? If A is an $n \times n$ invertible matrix, we want to solve $A\vec{x} = \vec{y}$ for $\vec{x}$ in terms of $\vec{y}$. We can do this using Gauss-Jordan elimination because of the following fact.

Proposition 10. *If $A \in M(n, n)$ is an invertible matrix, then* $\operatorname{rref}(A) = I_n$. *In particular, A is non-singular.*

Since $AA^{-1} = I_n$, the columns of A^{-1} are simply solutions to $A\vec{x}_i = \hat{e}_i$, where $\hat{e}_i$ are the standard basis vectors. Hence, Gauss-Jordan elimination ***should*** find A^{-1} by reducing $[A \mid I_n]$ to $[I_n \mid A^{-1}]$.

Example 3. Solve for $\vec{x}$ in terms of $\vec{y}$ where

$$A\vec{x} = \begin{bmatrix} 3 & 1 & 3 \\ 3 & 3 & 3 \\ 2 & 3 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \implies \begin{matrix} 3x_1 & + & x_2 & + & 3x_3 \\ 3x_1 & + & 3x_2 & + & 3x_3 \\ 2x_1 & + & 3x_2 & + & 3x_3 \end{matrix} = \begin{matrix} y_1 & + & 0 & + & 0 \\ 0 & + & y_2 & + & 0 \\ 0 & + & 0 & + & y_3 \end{matrix}$$

Start with the following augmented matrix

$$[A \mid I_3] = \left[\begin{array}{ccc|ccc} 3 & 1 & 3 & 1 & 0 & 0 \\ 3 & 3 & 3 & 0 & 1 & 0 \\ 2 & 3 & 3 & 0 & 0 & 1 \end{array}\right] \xrightarrow{E_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}} \left[\begin{array}{ccc|ccc} 3 & 3 & 3 & 0 & 1 & 0 \\ 3 & 1 & 3 & 1 & 0 & 0 \\ 2 & 3 & 3 & 0 & 0 & 1 \end{array}\right] \xrightarrow{E_2 = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}$$

$$\left[\begin{array}{ccc|ccc} 3 & 3 & 3 & 0 & 1 & 0 \\ 0 & -2 & 0 & 1 & -1 & 0 \\ 2 & 3 & 3 & 0 & 0 & 1 \end{array}\right] \xrightarrow{E_3 = \begin{bmatrix} 1/3 & 0 & 0 \\ 0 & -1/2 & 0 \\ -2/3 & 0 & 1 \end{bmatrix}} \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 0 & 1/3 & 0 \\ 0 & 1 & 0 & -1/2 & 1/2 & 0 \\ 0 & 1 & 1 & 0 & -2/3 & 1 \end{array}\right] \xrightarrow{E_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}}$$

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 0 & 1/3 & 0 \\ 0 & 1 & 0 & -1/2 & 1/2 & 0 \\ 0 & 0 & 1 & 1/2 & -7/6 & 1 \end{array}\right] \xrightarrow{E_5 = \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & -1/2 & 1/2 & 0 \\ 0 & 0 & 1 & 1/2 & -7/6 & 1 \end{array}\right] = [I_3 \mid A^{-1}]$$

So

$$\begin{matrix} x_1 & + & 0 & + & 0 \\ 0 & + & x_2 & + & 0 \\ 0 & + & 0 & + & x_3 \end{matrix} = \begin{matrix} 0 & + & y_2 & - & y_3 \\ -y_1/2 & + & y_2/2 & + & 0 \\ y_1/2 & - & 7y_2/6 & + & y_3 \end{matrix}$$

In particular, suppose E_i's are the matrices corresponding to the elementary row operations used to obtain $\text{rref}(A)$, then

$$E_k E_{k-1} \cdots E_2 E_1 A = I_n.$$

This ***suggests*** that $B = E_k E_{k-1} \cdots E_2 E_1$ is the inverse of A. However, this only tells us that $BA = I_n$, but why would $AB = I_n$? ***Remember, in general,*** $AC \neq CA$ ***for arbitrary square matrices*** A, C. However, we have (without proof)

Proposition 11. *For $A, B \in M(n, n)$, if $BA = I_n$, then $AB = I_n$. Similarly, if $AB = I_n$, then $BA = I_n$. In particular, if B is a left inverse or B is right inverse then $B = A^{-1}$.*

Returning to finding A^{-1} by row reducing $[A \mid I_n]$, we showed that $E_k E_{k-1} \cdots E_2 E_1 A = I_n$ for some sequence of elementary matrices E_i. It follows from the corollary above that $A^{-1} = E_k E_{k-1} \cdots E_2 E_1$. In particular, we have

Corollary 1. *$A \in M(n,n)$ is an invertible matrix if and only if* $\operatorname{rref}(A) = I_n$ *(i.e. if and only if A is non-singular).*

Proof. We already saw the forward (i.e. $\Rightarrow$) direction in the earlier exercise. For the backward direction (i.e. $\Leftarrow$), we just showed that $A^{-1} = E_k E_{k-1} \cdots E_2 E_1$ above. $\square$

All this work shows the non-trivial fact that $A E_k E_{k-1} \cdots E_2 E_1 = I_n$.

3 Determinants

Definition 4. The ***determinant*** of an $n \times n$ matrix $A = [a_{ij}]$ is defined by recursive ***Laplace (cofactor) expansion***: fix a column j, and for each entry a_{ij}, take the $(n-1) \times (n-1)$ matrix A_{ij} which does not contain the j^{th} column and i^{th} row of A. A_{ij} is called a ***minor***, and $(-1)^{i+j} \det(A_{ij})$ is called a ***cofactor***. We have

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}).$$

Example 4.

$$\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = ad - bc, \qquad \det\left(\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}\right) = aei + bfg + cdh - ceg - ahf - bdi$$

Example 5. The determinant of a diagonal or triangular matrix is the product of diagonal entries because you can always find a column with only one entry and the recursively apply this to successive minors.

3.1 Properties of the determinant

- $\det(A^t) = \det(A)$.
- If a matrix contains a column (or row) of zeros, then

  $$\det(\left[\vec{v}_1, \ldots, \vec{v}_{k-1}, \vec{0}, \vec{v}_{k+1}, \ldots, \vec{v}_n\right]) = 0$$

  To see this, just cofactor expand along the 0 column.

- det is a multi-linear function in the sense that for each column, we have respectively ***column additivity*** and ***column scalar*** properties

  $$\det(\left[\vec{v}_1, \ldots, \vec{v}_{k-1}, \vec{v}, \vec{v}_{k+1}, \ldots, \vec{v}_n\right]) + \det(\left[\vec{v}_1, \ldots, \vec{v}_{k-1}, \vec{w}, \vec{v}_{k+1}, \ldots, \vec{v}_n\right]) =$$
  $$= \det(\left[\vec{v}_1, \ldots, \vec{v}_{k-1}, \vec{v} + \vec{w}, \vec{v}_{k+1}, \ldots, \vec{v}_n\right])$$

  and

  $$\det(\left[\vec{v}_1, \ldots, \vec{v}_{k-1}, k\vec{v}, \vec{v}_{k+1}, \ldots, \vec{v}_n\right]) = k \det(\left[\vec{v}_1, \ldots, \vec{v}_{k-1}, \vec{v}, \vec{v}_{k+1}, \ldots, \vec{v}_n\right]).$$

  The same holds for each row vector. This follows directly from the Leibniz definition or cofactor expansion of determinant.

  For columns, this property is know as (column) ***multilinearity***. ***Warning:*** the determinant is ***not*** linear, i.e. $\det(A+B) \neq \det(A)+\det(B)$ almost always.

- The determinant of a matrix with two identical columns/rows is 0.

- ***Column interchange property:*** If B is obtained from A by swapping two columns/rows, then $\det(A) = -\det(B)$. To see why, let $A = \left[\vec{v}_1, \ldots, \vec{v}_n\right]$ and assume that we are swapping the i^{th} and j^{th} columns to get B. Consider the matrix C where we replace both the i^{th} and j^{th} columns of A with $\vec{v}_i+\vec{v}_j$. Then,

  $$\begin{aligned} 0 = \det(C) &= \det(\left[\ldots, \vec{v}_i + \vec{v}_j, \ldots, \vec{v}_i + \vec{v}_j, \ldots\right]) \\ &= \det(\left[\ldots, \vec{v}_i, \ldots, \vec{v}_i + \vec{v}_j, \ldots\right]) + \det(\left[\ldots, \vec{v}_j, \ldots, \vec{v}_i + \vec{v}_j, \ldots\right]) \\ &= \det(\left[\ldots, \vec{v}_i, \ldots, \vec{v}_i, \ldots\right]) + \det(\left[\ldots, \vec{v}_i, \ldots, \vec{v}_j, \ldots\right]) \\ &+ \det(\left[\ldots, \vec{v}_j, \ldots, \vec{v}_i, \ldots\right]) + \det(\left[\ldots, \vec{v}_j, \ldots, \vec{v}_j, \ldots\right]) \\ &= 0 + \det(A) + \det(B) + 0 \end{aligned}$$

  Remark. Notice that we used the property that having identical rows/columns implies that the determinant is zero. We could have gone in the other direction: assume the column interchange property to show that the determinant of a matrix with two identical columns is zero. Indeed, suppose two identical columns are swapped in A to produce B, then $\det(A) = -\det(B)$, but $A = B$, so $\det(A) = -\det(B) = -\det(A) = 0$. We would, of course, need to find another proof of the column interchange property to avoid circular reasoning.

  Thus, these two properties are logically equivalent. They are also known as ***alternating property*** of the determinant.

- If B is obtained from A by adding a multiple of one column/row to another, then $\det(A) = \det(B)$. This follows by linearity and the alternating properties.

- If $c_1, \ldots, c_k$ are the factors used to scale different rows and s is the number of row swaps used in deriving $\mathrm{rref}(A)$, then $\det(A) = (-1)^s c_1 \cdots c_k \det(\mathrm{rref}(A))$. Here, "scale" means to multiply a row by $1/c_i$.

$$\det\left(\begin{bmatrix} 0 & 0 & 0 & 1 \\ 2 & 3 & 4 & 5 \\ 0 & 6 & 7 & 8 \\ 0 & 0 & 9 & 10 \end{bmatrix}\right) = (-1)^3 \det\left(\begin{bmatrix} 2 & 3 & 4 & 5 \\ 0 & 6 & 7 & 8 \\ 0 & 0 & 9 & 10 \\ 0 & 0 & 0 & 1 \end{bmatrix}\right) = -108.$$

- An $n \times n$ matrix A is invertible if and only if $\det(A) \neq 0$. We can see this because $\det(A)$ and $\det(\mathrm{rref}(A))$ are either simultaneously zero or nonzero.

- ***Multiplicativity:*** $\det(AB) = \det(A)\det(B)$. To see this, we first assume that A is invertible. Then, Gauss-Jordan reduces $[A \mid AB]$ to $[I \mid A^{-1}AB] = [I \mid B]$. By applying the previous bullet point, we see that the determinant has changed by a factor of $\det(A)$ on both sides of the augmented matrix in the process. So $\frac{\det(AB)}{\det(A)} = \det(B)$. Lastly, if A is not invertible, then both AB and A are not invertible, so $0 = \det(AB) = \det(A)$.

- If A is invertible, $\det(A^{-1}) = \frac{1}{\det(A)}$. We simply apply multiplicativity to AA^{-1}.

3.2 Area and volumes

Recall that $\det(I_n) = 1$ and that det is a multilinear and alternating function on the space of n-tuples of column vectors (or on the space of n-tuples of row vectors).

Theorem 12. *The determinant function* det *is the unique alternating multilinear function f from $M(n,n) \to \mathbb{R}$ satisfying*

- $f(I_n) = 1$.

- *f is multi-linear in the sense that it satisfies row additivity and scalar properties*

- *f is alternating in the sense of the row exchange property.*

The proof uses induction and the fact that elementary row operations do not alter the value of an alternating multilinear function Uniqueness is always a nice thing to have in mathematics, so the uniqueness property of the determinant is no exception.

For example, consider the map area : $\mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ where $\mathrm{area}(\vec{v}_1, \vec{v}_2)$ is the ***signed*** area of the parallelogram with sides $\vec{v}_1$ and $\vec{v}_2$. Here, signed means that the area is positive if $\vec{v}_1$ is the "right" side of the parallelogram and $\vec{v}_2$ is the "left." It is negative in the other case. It follows that $\mathrm{area}(\vec{v}_1, \vec{v}_2) = -\mathrm{area}(\vec{v}_2, \vec{v}_1)$, i.e. area is alternating.

Using some planar geometry, you should be able to show that $\mathrm{area}(\vec{v}_1, \vec{v}_2 + \vec{w}) = \mathrm{area}(\vec{v}_1, \vec{v}_2) + \mathrm{area}(\vec{v}_1, \vec{w})$. This also hold for the first coordinate (using either the same argument or using the alternating property). For $c \in \mathbb{R}$, we also see that $\text{area}(\vec{v}_1, c\vec{v}_2) = c\,\text{area}(\vec{v}_1, \vec{v}_2)$. Note, if $c < 0$, then you change the sign of the area. Lastly, $\text{area}\left(\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}\right) = 1$. Thus, it follows by the above Theorem that this signed area is the determinant. Notice that $\text{area}(\vec{v}_1, \vec{v}_2) = \text{sign}(\det(\begin{bmatrix}\vec{v}_1, \vec{v}_2\end{bmatrix}))|\text{area}(\vec{v}_1, \vec{v}_2)|$, where the absolute volume of the signed area function is just the area.

In general, this lets us define

$$\text{vol}(\vec{v}_1, \ldots, \vec{v}_n) = \text{sign}(\det(\begin{bmatrix}\vec{v}_1, \ldots, \vec{v}_n\end{bmatrix}))\,|\text{vol}(\vec{v}_1, \ldots, \vec{v}_n)|,$$

where $|\text{vol}(\vec{v}_1, \ldots, \vec{v}_n)|$ is the volume of the parallelepiped in $\mathbb{R}^n$ defined by the vectors $\vec{v}_1, \ldots, \vec{v}_n$. One can similarly argue that vol is multilinear and alternating. Since $\text{vol}(\hat{e}_1, \ldots, \hat{e}_n) = 1$, we see that vol = det. Thus, determinant gives signed volume.[2]

Proposition 13. *Let P be the parallelepiped in $\mathbb{R}^n$ defined by the vectors $\vec{v}_1, \ldots, \vec{v}_n$. Consider the image $T_A(P)$ for some invertible $A \in M(n,n)$. Then,*

$$|\text{vol}(T_A(P))| = |\det(A)||\text{vol}(P)|.$$

In particular, if $\det(A) = \pm 1$, then $T_A(P)$ and P have the same (unsigned) volume.

Proof. Notice that $|\text{vol}(T_A(P))| = \det(\begin{bmatrix}T_A(\vec{v}_1), \ldots, T_A(\vec{v}_n)\end{bmatrix})$. Let $B = \begin{bmatrix}\vec{v}_1, \ldots, \vec{v}_n\end{bmatrix}$, then $\begin{bmatrix}T_A(\vec{v}_1), \ldots, T_A(\vec{v}_n)\end{bmatrix} = AB$ by definition. Lastly, since $|\det(B)| = |\text{vol}(P)|$ and $\det(AB) = \det(A)\det(B)$, we see that $|\text{vol}(T_A(P))| = |\det(A)||\text{vol}(P)|$. □

3.3 Orientation

If you are familiar with the right-hand rule in physics, you may have wondered if something like it works in higher dimensions. For a list $\vec{v}_1, \ldots, \vec{v}_n \in \mathbb{R}^n$ of n linearly independent vectors we define the ***orientation of*** $\vec{v}_1, \ldots, \vec{v}_n$ for be the sign of $\det(\begin{bmatrix}\vec{v}_1, \ldots, \vec{v}_n\end{bmatrix})$. That is, $\vec{v}_1, \ldots, \vec{v}_n$ is ***positively oriented*** if $\det(\begin{bmatrix}\vec{v}_1, \ldots, \vec{v}_n\end{bmatrix}) > 0$ and negatively oriented otherwise. Notice, if you are given $\vec{v}_1, \ldots, \vec{v}_{n-1}$ and need to choose between $\pm\vec{v}_n$ (e.g. to define a "canonical" direction of some potential or field), then you can choose between $\pm\vec{v}_n$ by forcing the determinant to be positive.

3.4 Trace

Another important invariant of a matrix is its ***trace***.

Definition 5. The ***trace*** of an $n \times n$ matrix A denoted $\text{tr}(A)$ is defined to be

$$\text{tr}(A) = \sum_i a_{ii}.$$

[2] I am glossing over how one geometrically computes volume in $\mathbb{R}^n$, but it is similar to how you go from areas in $\mathbb{R}^2$ to volumes in $\mathbb{R}^3$.

Proposition 14. *Trace is a linear transformation from $M(n,n)$ to $\mathbb{R}$, i.e.*

$$\operatorname{tr}(kA) = k\operatorname{tr}(A), \qquad \operatorname{tr}(A+B) = \operatorname{tr}(A) + \operatorname{tr}(B).$$

Furthermore,

$$\operatorname{tr}(AB) = \operatorname{tr}(BA).$$

Proof. It is easy to check linearity. For the last property, we can see that the i^{th} diagonal entry of AB is $\sum_{j=1}^{n} a_{ij}b_{ji}$. Similarly, the i^{th} diagonal entry of BA is $\sum_{j=1}^{n} b_{ij}a_{ji}$.

$$\operatorname{tr}(AB) = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}b_{ji}\right) \xlongequal[\text{reorder mult.}]{\text{sum in diff. order}} \sum_{j=1}^{n}\left(\sum_{i=1}^{n} b_{ji}a_{ij}\right) \xlongequal{\text{relabel } i\leftrightarrow j} \sum_{i=1}^{n}\left(\sum_{j=1}^{n} b_{ij}a_{ji}\right) = \operatorname{tr}(BA).$$

□