DETERMINANTS

1. Introduction

In these notes we discuss a simple tool for testing the non singularity of an $n \times n$ matrix that will be useful in our discussion of eigenvalues. Tis tool is the *determinant*. At the end of these notes, we will also discuss how the determinant can be used to solve equations (Cramer's Rule), and how it can be used to give a theoretically useful representation the inverse of a matrix (via the classical adjoint).

The Leibniz formula for the determinant of an $n \times n$ matrix A is

$$\det(A) = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} A_{i\sigma(i)}, \tag{1}$$

where $\mathcal{S}_n$ is the set of all permutations of the integers $\{1, 2, ..., n\}$. These permutations are functions that reorder this set of integers. The element in position i after the reordering σ is denoted $\sigma(i)$. For example, for n = 3, the original sequence $[1, 2, 3]$ might be reordered to $[2, 3, 1]$, with $\sigma(1) = 2,\ \sigma(2) = 3,\ \sigma(3) = 1$. The set of all such permutations (also known as the symmetric group on n elements) is denoted $\mathcal{S}_n$. For each permutation σ, $\operatorname{sgn}(\sigma)$ denotes the signature of σ; it is $+1$ for even σ and -1 for odd σ. Evenness or oddness can be defined as follows: the permutation is even (odd) if the new sequence can be obtained by an even number (odd, respectively) of switches of numbers. For example, starting from $[1, 2, 3]$ and switching the positions of 2 and 3 yields $[1, 3, 2]$, switching once more yields $[3, 1, 2]$, and finally, after a total of three (an odd number) switches, $[3, 2, 1]$ results. Therefore $[3, 2, 1]$ is an odd permutation. Similarly, the permutation $[2, 3, 1]$ is even since

$$[1, 2, 3] \to [2, 1, 3] \to [2, 3, 1],$$

an even number of switches.

The identity permutation is the unique element $\iota \in \mathcal{S}_n$ for which $\iota(i) = i$ for all $i = 1, 2, \dots, n$. Note that there are zero switches. We say that ι is even so that its signature is 1, $\operatorname{sgn}(\iota) = 1$. Observe that if I is the identity matrix, then

$$\det(I) = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} I_{i\,\sigma i} = \operatorname{sgn}(\iota) = 1.$$

Given $\sigma_1,\ \sigma_2 \in \mathcal{S}_n$, we can use composition $\sigma = \sigma_1 \circ \sigma_1$ to obtain another element of $\mathcal{S}_n$. The discussion above on the signature of a permutation tells us that $\operatorname{sgn}(\sigma_1 \circ \sigma) = \operatorname{sgn}(\sigma_1) \operatorname{sgn}(\sigma_2)$. Therefore, if $\tilde{A}$ is the matrix obtained from A by permuting its columns using the permutation

$\pi \in \mathcal{S}_n$, then Leibniz's formula (1) tells us that

$$\begin{aligned}
\det(\tilde{A}) &= \sum_{\sigma\in\mathcal{S}_n} \mathrm{sgn}\,(\sigma) \prod_{i=1}^{n} A_{i\sigma(\pi(i))} \\
&= \sum_{\sigma\in\mathcal{S}_n} \mathrm{sgn}\,(\sigma)\,[\mathrm{sgn}\,(\pi)]^2 \prod_{i=1}^{n} A_{i\sigma(\pi(i))} \\
&= \mathrm{sgn}\,(\pi) \sum_{\sigma\in\mathcal{S}_n} \mathrm{sgn}\,(\sigma)\,\mathrm{sgn}\,(\pi) \prod_{i=1}^{n} A_{i\sigma(\pi(i))} \\
&= \mathrm{sgn}\,(\pi) \sum_{\sigma\in\mathcal{S}_n} \mathrm{sgn}\,(\sigma\circ\pi) \prod_{i=1}^{n} A_{i(\sigma\circ\pi)(i)} \\
&= \mathrm{sgn}\,(\pi) \sum_{\sigma\in\mathcal{S}_n} \mathrm{sgn}\,(\sigma) \prod_{i=1}^{n} A_{i\sigma(i)} \\
&= \mathrm{sgn}\,(\pi)\det(A).
\end{aligned}$$

Since we can also write $\tilde{A} = AP_\pi$, where $P_\pi \in \mathbb{R}^{n\times n}$ is the permutation matrix corresponding to π, this gives us the formula

$$\det(AP_\pi) = \mathrm{sgn}\,(\pi)\det(A).$$

Taking $A = I$ and using the fact that $\det(I) = 1$, we obtain

$$\det(P_\pi) = \mathrm{sgn}\,(\pi)\det(I) = \mathrm{sgn}\,(\pi) \quad \forall\,\pi \in \mathcal{S}_n. \tag{2}$$

Therefore, for every permutation matrix $P \in \mathbb{R}^{n\times n}$ and matrix $A \in \mathbb{R}^{n\times n}$, we have

$$\det(AP) = \det(P)\det(A). \tag{3}$$

If $n = 2$, then $\mathcal{S}_2$ only has 2 elements:

$$[1,2] \xrightarrow{\iota} [1,2] \text{ and } [1,2] \xrightarrow{\sigma} [2,1],$$

where $\mathrm{sgn}\,(\iota) = 1$ and $\mathrm{sgn}\,(\sigma) = -1$. Therefore, given a 2×2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ its determinent is

$$\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = ad - bc\ .$$

2. Laplace's Formula for the Determinant

Lieibniz's formula fro the determinant, although extremely powerful, is very difficult to use in practical computations. For this reason, we discuss and alternative formula for its computation.

Theorem 1 (Laplace's Formula for the Determinant)**.** *Suppose $A \in \mathrm{R}^{n\times n}$. For every pair $i, j \in \{1, 2, \ldots, n\}$, define the matrix $A[i, j]$ to be the $(n-1)\times(n-1)$ submatrix of A obtain by deleting the ith row and the jth column. The for each $i_0, j_0 \in \{1, 2, \ldots, n\}$,*

$$\det(A) = \sum_{i=1}^{n} a_{ij_0}(-1)^{(i+j_0)}\det(A[i, j_0]) = \sum_{j=1}^{n} a_{i_0j}(-1)^{(i_0+j)}\det(A[i_0, j])\ .$$

The terms $C_{ij} = (-1)^{(i+j)} \det(A[i,j])$ are called the cofactors of the matrix A and the transpose of the matrix whose ijth component is C_{ij} is called the *classical adjoint* of A denoted $\mathrm{adj}\,(A) = [C_{ij}]^T$. The determinant satisfies the following properties.

Theorem 2 (Properties of the Determinant)**.** *Let $A, B \in \mathbb{R}^{n\times n}$.*

(1) $\det(A) = \det(A^T)$.
(2) *The determinant is a multi-linear function of its columns (rows). That is, if $A = [A_{\cdot 1},\ A_{\cdot 2}, \ldots,\ A_{\cdot n}]$, where $A_{\cdot j}$ is the jth column of A $(j = 1, \ldots, n)$, then for any vector $b \in \mathrm{R}^n$ and scalar $\lambda \in \mathbb{R}$*

$$\det\left([A_{\cdot 1}, \ldots,\ A_{\cdot j} + \lambda b, \ldots,\ A_{\cdot n}]\right) = \det\left([A_{\cdot 1}, \ldots,\ A_{\cdot j}, \ldots,\ A_{\cdot n}]\right) + \lambda \det\left([A_{\cdot 1}, \ldots,\ b, \ldots,\ A_{\cdot n}]\right) \ .$$

(3) *If any two columns (rows) of A coincide, then* $\det(A) = 0$.
(4) *For every $j_1, j_2 \in \{1, \ldots, n\}$ with $j_1 \neq j_2$ and $\lambda \in \mathbb{R}$,*

$$\det(A) = \det\left([A_{\cdot 1}, \ldots,\ A_{\cdot j_1} + \lambda A_{\cdot j_2}, \ldots,\ A_{\cdot n}]\right).$$

(5) *If A is singular, then* $\det(A) = 0$.

Proof. (1) This follows immediately from Laplace's formula for the determinant in Theorem 1.
(2) This follows immediately from Laplace's formula:

$$\begin{aligned}
\det\left([A_{\cdot 1}, \ldots,\ A_{\cdot j} + \lambda b, \ldots,\ A_{\cdot n}]\right) &= \sum_{i=1}^{n} (a_{ij} + \lambda b_i)(-1)^{(i+j)} \det(A[i,j]) \\
&= \sum_{i=1}^{n} a_{ij}(-1)^{(i+j_0)} \det(A[i,j]) + \lambda \sum_{i=1}^{n} b_i(-1)^{(i+j_0)} \det(A[i,j]) \\
&= \det\left([A_{\cdot 1}, \ldots,\ A_{\cdot j}, \ldots,\ A_{\cdot n}]\right) + \lambda \det\left([A_{\cdot 1}, \ldots,\ b, \ldots,\ A_{\cdot n}]\right) \ .
\end{aligned}$$

(3) The permutation $\pi_{ij} \in \mathcal{S}_n$ that interchanges i and j $(\pi_{ij}(i) = j,\ \pi_{ij}(j) = i,\ \pi_{ij}(k) = k\, \forall\, k \neq i, j)$ is odd since $\mathrm{sgn}\,(\pi_{ij}) = 2\,|i-j| - 1$ as long as $i \neq j$. Therefore, by (2), the permutation matrix P_{ij} which interchanges the columns $i \neq j$ has $\det(P_{ij}) = -1$. Now suppose that column i equals column $j \neq i$ in the matrix $A \in \mathbb{R}^{n\times n}$, then, by (3), $\det(A) = \det(AP_{ij}) = \det(A)\det(P_{ij}) = -\det(A)$. Hence $\det(A) = 0$.
(4) This follows immediately from Parts (2) and (3).
(5) If A is singular, then its columns are linearly dependent. That is, there is a non-trivial linear combination of its columns that give zero, or equivalently, there is some column j_0 that is a linear combination of the remaining columns, $A_{\cdot j_0} = \sum_{j\neq j_0} \lambda_j A_{\cdot j}$. Therefore, by Parts (2) and (3),

$$\begin{aligned}
\det(A) &= \det\left([A_{\cdot 1}, \ldots, A_{\cdot (j_0-1)}, \sum_{j\neq j_0} \lambda_j A_{\cdot j}, A_{\cdot (j_0+1)}, \ldots, A_{\cdot n}]\right) \\
&= \sum_{j\neq j_0} \lambda_j \det\left([A_{\cdot 1}, \ldots, A_{\cdot (j_0-1)}, A_{\cdot j}, A_{\cdot (j_0+1)}, , \ldots, A_{\cdot n}]\right) \\
&= 0\ .
\end{aligned}$$

□

We will also need two further properties for the determinant. These appear in the next theorem whose proof is omitted.

Theorem 3. *Let $A, B \in \mathbb{R}^{n\times n}$, $C \in \mathbb{R}^{m\times n}$, and $D \in \mathbb{R}^{m\times m}$. Then the following two formulas hold:*

$$\det(AB) = \det(A)\det(B) \tag{4}$$

$$\det\left(\begin{bmatrix} A & 0 \\ C & D \end{bmatrix}\right) = \det(A)\det(D)\ . \tag{5}$$

Note that (3) is a special case of (4). As an application of (4) we compute the determinant of A^{-1} when it exists:

$$1 = \det(I) = \det(AA^{-1}) = \det(A)\det(A^{-1}),$$

whenever A is nonsingular. That is, $\det(A^{-1}) = \det(A)^{-1}$.

3. Cramer's Rule

We now consider the system $Ax = b$ and Cramer's Rule. Cramer's Rule states that if A is nonsingular, then the unique solution to the system $Ax = b$ is given componentwise by

$$x_j = \frac{\det(A_j(b))}{\det(A)},\ j = 1, 2, \ldots, n,$$

where the matrix $A_j(b)$ is obtained from A by replacing the jth column of A by the vector b. The proof of Cramer's Rule follows easily from the properties of the determinant. Indeed, if $\bar{x}$ is the unique solution to the system $Ax = b$, then $b = A\bar{x} = \sum_{j=1}^{n} \bar{x}_j A_{\cdot j}$. Therefore, by Parts (2) and (3) of Theorem 2,

$$\begin{aligned}
\det(A_j(b)) &= \det\left([A_{\cdot 1}, \ldots, A_{\cdot(j-1)}, b, A_{\cdot(j+1)}, \ldots, A_{\cdot n}]\right) \\
&= \det\left([A_{\cdot 1}, \ldots, A_{\cdot(j-1)}, \sum_{i=1}^{n} \bar{x}_i A_{\cdot i}, A_{\cdot(j+1)}, \ldots, A_{\cdot n}]\right) \\
&= \sum_{i=1}^{n} \bar{x}_i \det\left([A_{\cdot 1}, \ldots, A_{\cdot(j-1)}, A_{\cdot i}, A_{\cdot(j+1)}, \ldots, A_{\cdot n}]\right) \\
&= \bar{x}_i \det\left([A_{\cdot 1}, \ldots, A_{\cdot(i-1)}, A_{\cdot i}, A_{\cdot(i+1)}, \ldots, A_{\cdot n}]\right) \\
&= \bar{x}_i \det(A)
\end{aligned}$$

giving Cramer's Rule.

Let us examine the expressions $\det(A_j(b))$ using Laplace's formula for the determinant:

$$\det(A_j(b)) = \sum_{i=1}^{n} b_i C_{ij} = C_{\cdot j}^T b,$$

where $C_{\cdot j}$ is the jth row of the classical adjoint $\mathrm{adj}\,(A)$. That is,

$$\bar{x} = \frac{1}{\det(A)} \mathrm{adj}\,(A)\, b.$$

Since this expression is valid for all choices of $b \in \mathrm{R}^n$, we must have

$$A^{-1} = \frac{1}{\det(A)} \mathrm{adj}\,(A)\,.$$

We call this the adjoint representaton of the inverse.