(IV.A) Alternating Multilinear Functionals

The determinant of a $2 \times 2$ matrix with entries in $\mathbb{R}$, such as

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

is a familiar object. It has a couple of pretty properties, the first of which is “linearity in each column [vector] separately”:

$$\det \begin{pmatrix} \alpha_1 a_1 + \alpha_2 a_2 & b \\ \alpha_1 c_1 + \alpha_2 c_2 & d \end{pmatrix} = \alpha_1 \det \begin{pmatrix} a_1 & b \\ c_1 & d \end{pmatrix} + \alpha_2 \det \begin{pmatrix} a_2 & b \\ c_2 & d \end{pmatrix},$$

$$\det \begin{pmatrix} a & \beta_1 b_1 + \beta_2 b_2 \\ c & \beta_1 d_1 + \beta_2 d_2 \end{pmatrix} = \beta_1 \det \begin{pmatrix} a & b_1 \\ c & d_1 \end{pmatrix} + \beta_2 \det \begin{pmatrix} a & b_2 \\ c & d_2 \end{pmatrix},$$

which we shall call “bilinearity”. The second is the antisymmetry relation

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

(Both relations also hold with respect to the rows.) In fact what we will now show, is that the determinant is somehow “uniquely determined” by these two properties and the fact that

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

Uniquely determined amongst what set of objects? Basically, all functions on two copies of $\mathbb{R}^2$ — that is, functions sending any two vectors to a number. (Before we proceed, I note that you may replace $\mathbb{R}$ in what follows by any field of characteristic different from 2.)

So define a bilinear functional (or form) on a vector space $V$ to be a function

$$B : V \times V \to \mathbb{R}$$
satisfying (for any \( \vec{v}, \vec{w} \in V \) )
\[
B(a_1\vec{v}_1 + a_2\vec{v}_2, \vec{w}) = a_1B(\vec{v}_1, \vec{w}) + a_2B(\vec{v}_2, \vec{w}), \\
B(\vec{v}, c_1\vec{w}_1 + c_2\vec{w}_2) = c_1B(\vec{v}, \vec{w}_1) + c_2B(\vec{v}, \vec{w}_2).
\]
Call it alternating (or antisymmetric) if it also satisfies
\[
B(\vec{v}, \vec{w}) = -B(\vec{w}, \vec{v}).
\]
Notice that this implies
\[
B(\vec{v}, \vec{v}) = -B(\vec{v}, \vec{v}) \implies B(\vec{v}, \vec{v}) = 0.
\]
Now assume \( V = \mathbb{R}^2 \), \( B \) alternating bilinear, and set \( b := B(\hat{e}_1, \hat{e}_2) \).
Note that \( B(\hat{e}_2, \hat{e}_1) = -b \) and \( B(\hat{e}_1, \hat{e}_1) = B(\hat{e}_2, \hat{e}_2) = 0 \). Given
\[
A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \text{set } \vec{v}_1 = a_{11}\hat{e}_1 + a_{21}\hat{e}_2, \vec{v}_2 = a_{12}\hat{e}_1 + a_{22}\hat{e}_2
\]
so that \( \vec{v}_1 \) and \( \vec{v}_2 \) are just the columns of \( A \). Then
\[
B(\vec{v}_1, \vec{v}_2) = B(a_{11}\hat{e}_1 + a_{21}\hat{e}_2, \vec{v}_2) = a_{11}B(\hat{e}_1, \vec{v}_2) + a_{21}B(\hat{e}_2, \vec{v}_2) \\
= a_{11}(a_{12}B(\hat{e}_1, \hat{e}_1) + a_{22}B(\hat{e}_1, \hat{e}_2)) + a_{21}(a_{12}B(\hat{e}_2, \hat{e}_1) + a_{22}B(\hat{e}_2, \hat{e}_2)) \\
= a_{11}(a_{12} \cdot 0 + a_{22} \cdot b) + a_{21}(a_{12} \cdot (-b) + a_{22} \cdot 0) = b \cdot (a_{11}a_{22} - a_{21}a_{12});
\]
that is,
\[
B(\vec{v}_1, \vec{v}_2) = b \cdot \det(A).
\]

**Conclusion:** Any alternating multilinear functional \( B : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R} \) is determined uniquely by the number \( b = B(\hat{e}_1, \hat{e}_2) \). If \( b = 1 \), allowing \( B \) to act on the column vectors of a \( 2 \times 2 \) matrix \( A \) yields its determinant, \( \det A \).

\[
A : \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{n \text{ copies}} \to \mathbb{R}
\]
and showing that with the additional “normalization” requirement that $A(\hat{e}_1, \ldots, \hat{e}_n) = 1$, there is only one. The only difference is that this unique $A$, applied to the column vectors of a matrix $A$, will then be our definition of its determinant.

To establish precisely the kind of antisymmetry we need here it is convenient to introduce the “symmetric group” $S_n$ and sign function $sgn$. First, let a multi-index be any function from the set $\{1, \ldots, n\}$ into itself; they will be written $i = (i_1, \ldots, i_n)$. A 1-to-1 multi-index is called a permutation; they get their own notation $\sigma: i = \sigma \implies (i_1, \ldots, i_n) = (\sigma(1), \ldots, \sigma(n))$. The permutation sends $1 \mapsto \sigma(1), 2 \mapsto \sigma(2)$, etc. and we may compose them as we would any functions. For instance, let $\sigma = (32)$ (resp. $'\sigma = (13)$) be the permutation swapping 3 and 2 (resp. 1 and 3). Then the composition $(32)(13) = \sigma \circ '\sigma$ looks like $1, 2, 3 \rightarrow 3, 2, 1 \rightarrow 2, 3, 1$.

The set of all permutations $\sigma$ is called $S_n$; it contains $n!$ elements. For instance, the elements of $S_3$ are the permutations sending

$$1, 2, 3 \rightarrow 1, 2, 3 ; 1, 3, 2 ; 2, 1, 3 ; 2, 3, 1 ; 3, 1, 2 ; 3, 2, 1$$

Composition of permutations gives a product $S_n \times S_n \rightarrow S_n$, with identity simply the permutation that doesn’t do anything. Since every element $\sigma \in S_n$ (being a 1-to-1 function) has an inverse, $S_n$ is a group. It is a fact that every permutation is a composition of swaps of two elements (like $(13)$ and $(32)$). While there are different sequences of swaps one can follow to produce a permutation $\sigma$, one thing is the same no matter what route you follow: the parity of the number of swaps (that is, whether the number is even or odd). Therefore the function

$$sgn: S_n \rightarrow \{\pm 1\}$$

given by

$$\sigma \mapsto (-1)^{\# \text{ of swaps in } \sigma} =: sgn(\sigma)$$

is well-defined. The $sgn$ functions of the six permutations above ($\in S_3$) are (respectively) $+, -, -, +, +, -$. 
Now we are ready to state our three requirements on $A$. All $\vec{v}_i$ are arbitrary vectors $\in \mathbb{R}^n$.

- **Multilinearity:**
  \[
  A(a\vec{v}_1 + a'\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n) = aA(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n) + a'\vec{A}(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n)
  \]
  and the same in the remaining $n-1$ entries.

- **Antisymmetry:**
  \[
  A(\vec{v}_{\sigma(1)}, \vec{v}_{\sigma(2)}, \ldots, \vec{v}_{\sigma(n)}) = \text{sgn}(\sigma) \cdot A(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n);
  \]
  this has the implication that $A(\vec{v}_1, \ldots, \vec{v}_n) = 0$ if $\vec{v}_i = \vec{v}_j$ for $i \neq j$ (two distinct entries the same). To see this, simply let $\sigma$ be the swap $(i \ j)$. (We emphasize that the antisymmetry law must hold for any $n$ vectors $\{\vec{v}_i\}$ and all $\sigma \in S_n$, for $A$ to be considered antisymmetric or “alternating”).

- **Normalization:**
  \[
  A(\hat{e}_1, \ldots, \hat{e}_n) = 1.
  \]
  Combined with antisymmetry this gives that
  \[
  A(\hat{e}_{i_1}, \ldots, \hat{e}_{i_n}) = 0 \quad \text{unless $i$ is a permutation $\sigma$}
  \]
in which case
  \[
  A(\hat{e}_{\sigma(1)}, \ldots, \hat{e}_{\sigma(n)}) = \text{sgn}(\sigma) \cdot 1 = \text{sgn}(\sigma).
  \]

We are now ready to calculate $A$ on the columns $\vec{v}_j = \sum_i a_{ij} \hat{e}_i$ of

\[
A = \begin{pmatrix}
a_{11} & \cdots & a_{1n} \\
\vdots & \ddots & \vdots \\
a_{n1} & \cdots & a_{nn}
\end{pmatrix}.
\]

The result will show $A$ is uniquely determined, and will also give our definition of $\det(A)$. We start by using multilinearity to obtain

\[
A(\vec{v}_1, \ldots, \vec{v}_n) = A \left( \sum_{i_1} a_{i_11} \hat{e}_{i_1}, \ldots, \sum_{i_n} a_{i_n n} \hat{e}_{i_n} \right)
\]

\[
= \sum_{i=(i_1, \ldots, i_n)} a_{i_1 1} \cdots a_{i_n n} \cdot A(\hat{e}_{i_1}, \ldots, \hat{e}_{i_n}).
\]
If this is not completely clear, you may wish to check it step by step (using linearity in each entry separately) for \( n = 3 \). The next thing is to notice that we can throw out the terms for which \( i \) is not a permutation \( \sigma \), since then \( A(\hat{e}_1, \ldots, \hat{e}_n) = 0 \). So we end up with

\[
\sum_{\sigma \in S_n} a_{\sigma(1)} \cdots a_{\sigma(n)} \cdot A(\hat{e}_{\sigma(1)}, \ldots, \hat{e}_{\sigma(n)}) =
\sum_{\sigma \in S_n} \text{sgn}(\sigma) \cdot a_{\sigma(1)} \cdots a_{\sigma(n)} =: \det(A)
\]

and it’s no more complicated than that!

**Interpretation:** You should think of the series of entries

\[
\{ (\sigma(1), 1), \ldots, (\sigma(2), 2) \}
\]

in the matrix \( A \) as a graph of a discrete 1-to-1 function from columns to rows (for each column \( j \), choose the \( \sigma(j)^{th} \) entry). The determinant is then the sum over all such graphs, of the products of the matrix entries “lying on the graph” (namely \( a_{\sigma(1)} \) thru \( a_{\sigma(n)} \) ) times an appropriate sign. For \( 2 \times 2 \) matrices, there are only two such “graphs”: the diagonal and the anti-diagonal

\[
\begin{pmatrix} \circ & \circ \\ \circ & \circ \end{pmatrix}
\quad \text{and} \quad
\begin{pmatrix} \circ & \circ \\ \circ & \circ \end{pmatrix},
\]

while for \( 3 \times 3 \)’s you already have 6 (with the same series of signs +, −, −, +, +, − as given above):

\[
\begin{pmatrix} \circ & \circ & \circ \\ \circ & \circ & \circ \\ \circ & \circ & \circ \end{pmatrix},
\begin{pmatrix} \circ & \circ & \circ \\ \circ & \circ & \circ \\ \circ & \circ & \circ \end{pmatrix},
\begin{pmatrix} \circ & \circ & \circ \\ \circ & \circ & \circ \\ \circ & \circ & \circ \end{pmatrix},
\begin{pmatrix} \circ & \circ & \circ \\ \circ & \circ & \circ \\ \circ & \circ & \circ \end{pmatrix},
\begin{pmatrix} \circ & \circ & \circ \\ \circ & \circ & \circ \\ \circ & \circ & \circ \end{pmatrix},
\begin{pmatrix} \circ & \circ & \circ \\ \circ & \circ & \circ \\ \circ & \circ & \circ \end{pmatrix}
\]

and for \( n \times n \)’s there are \( n! \) (since of course these are just permutations). The way I wrote the \( 3 \times 3 \)’s should already suggest to you the Laplace expansion in the first column.
Exercises

(1) For an \( n \times n \) matrix \( A \), find the \( k \) for which \( \det(rA) = r^k \det(A) \) holds.

(2) Prove that the determinant of the Vandermonde matrix

\[
\begin{pmatrix}
1 & a & a^2 \\
1 & b & b^2 \\
1 & c & c^2
\end{pmatrix}
\]

is \( (b-a)(c-a)(c-b) \).

(3) An \( n \times n \) matrix \( A \) is called triangular if \( A_{ij} = 0 \) whenever \( i > j \) or if \( A_{ij} = 0 \) whenever \( i < j \). Prove that the determinant of a triangular matrix is the product \( A_{11}A_{22}\cdots A_{nn} \) of its diagonal entries.

(4) Let \( A \) be a \( 3 \times 3 \) matrix over the field of complex numbers. We form the matrix \( xI_{3} - A \) with polynomial entries, the \((i,j)\) entry of this matrix being the polynomial \( x\delta_{ij} - A_{ij} \). If \( f = \det(xI - A) \), show that \( f \) is a monic polynomial of degree 3. If we write \( f = (x - c_1)(x - c_2)(x - c_3) \) with \( c_i \in \mathbb{C} \), prove that \( \sum c_i = \text{tr}(A) \) and \( c_1c_2c_3 = \det(A) \).