Appendix B : Determinants

B.1 Definition of the determinant

First, define $\mathbf{z_0}$ to be the n-component vector of increasing integers 1 to n,

$$\mathbf{z_0} \equiv \begin{pmatrix} 1 \\ 2 \\ \dots \\ n \end{pmatrix} . \tag{B.1.1}$$

Let **a** be some **permutation** (reordering) of these integers, so write

$$\mathbf{a} \equiv \begin{pmatrix} a_1 \\ a_2 \\ \dots \\ a_n \end{pmatrix} = A \begin{pmatrix} 1 \\ 2 \\ \dots \\ n \end{pmatrix} = A\,\mathbf{z_0} \tag{B.1.2}$$

where A is an nxn matrix which has 1's in the right places to create this permutation vector **a**. For example,

$$\mathbf{a} = A\mathbf{z_o} \qquad \leftrightarrow \qquad \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} .$$

The vector **a** can be obtained from the vector $\mathbf{z_0}$ by making $S_\mathbf{a}$ pairwise swaps of elements of $\mathbf{z_0}$. Although $S_\mathbf{a}$ is not unique, the number $(-1)^{S_\mathbf{a}}$ is unique. For example, to get from (1,2,3) to (1,3,2) one could swap the second pair so $S_\mathbf{a} = 1$, but one could then swap the first pair twice and then $S_\mathbf{a} = 3$. This number $(-1)^{S_\mathbf{a}}$ is of course ±1 and we shall call it the **parity** of the permutation **a**.

$$\text{Parity}(\mathbf{a}) \equiv (-1)^{S_\mathbf{a}} \tag{B.1.3}$$

Fact: $$\text{Parity}(C\mathbf{a}) = \text{Parity}(CA\mathbf{z_0}) = (-1)^{S_\mathbf{c}+S_\mathbf{a}} = \text{Parity}(C^{-1}\mathbf{a}) \tag{B.1.4}$$

Proof: Doing permutation CA involves first doing A with its $S_\mathbf{a}$ swaps, and then doing C with its $S_\mathbf{c}$ swaps, for a total of $S_\mathbf{a}+S_\mathbf{c}$ swaps. The inverse permutation C^{-1} obviously involves the same number of swaps as the permutation C.

Moving now toward the definition of the determinant of a matrix M, define

$$\Pi(\mathbf{a},\mathbf{b}; M) \equiv M_{a_1b_1}M_{a_2b_2} \dots\dots M_{a_nb_n} = \text{product of n factors} \tag{B.1.5}$$

Suppose $\mathbf{c} = C\mathbf{z_0}$ is some arbitrary permutation. Then:

Fact: $\Pi(C\mathbf{a},C\mathbf{b};\, M) = \Pi(\mathbf{a},\mathbf{b};\, M)$ (B.1.6)

Proof: Applying the same permutation to both the a_k and b_k indices of $M_{a_1b_1}M_{a_2b_2}$ ..... $M_{a_nb_n}$, just reorders the terms in the product but the product stays the same.

Fact: $\Pi(\mathbf{a},\mathbf{b};\, M^T) = \Pi(\mathbf{b},\mathbf{a};\, M)$ (B.1.7)

Proof: $\Pi(\mathbf{a},\mathbf{b};\, M^T) = M^T_{a_1b_1}M^T_{a_2b_2}$ .... $M^T_{a_nb_n} = M_{b_1a_1}M_{b_2a_2}$ .... $M_{b_na_n} = \Pi(\mathbf{b},\mathbf{a};\, M)$

We shall now **define the determinant** of an n x n matrix M in the following admittedly obscure manner (and later we will show that it reduces to more familiar forms),

$$\det(M) \equiv \frac{1}{n!}\,\Sigma_{\mathbf{a}}\,\Sigma_{\mathbf{b}}\,(-1)^{s_a+s_b}\,\Pi(\mathbf{a},\mathbf{b};\, M)$$

$$= \frac{1}{n!}\,\Sigma_{\mathbf{a}}\,\Sigma_{\mathbf{b}}\,(-1)^{s_a+s_b}\,M_{a_1b_1}M_{a_2b_2}\,.....\,M_{a_nb_n}\,. \qquad \text{(B.1.8)}$$

Here $\Sigma_{\mathbf{a}}$ means the sum over all permutations $\mathbf{a}$ of $\mathbf{z}_0$. In (B.1.8) the columns and rows of M are on a completely equal footing. Notice that $\mathbf{a}$ and $\mathbf{b}$ are in effect dummy summation indices. If we do $\mathbf{a}\leftrightarrow\mathbf{b}$, the expression for det(M) is unchanged. Thus we can rewrite (B.1.8) as

$$\det(M) = \frac{1}{n!}\,\Sigma_{\mathbf{b}}\,\Sigma_{\mathbf{a}}\,(-1)^{s_b+s_a}\;\;\Pi(\mathbf{b},\mathbf{a};\, M)\;\;. \qquad \text{(B.1.9)}$$

We are now ready for our first determinant theorem:

Theorem 1: $\det(M^T) = \det(M)$. Switching rows with columns does not change a determinant. (B.1.10)

Proof: $\det(M^T) = \frac{1}{n!}\,\Sigma_{\mathbf{b}}\,\Sigma_{\mathbf{a}}\,(-1)^{s_b+s_a}\;\;\Pi(\mathbf{b},\mathbf{a};\, M^T)$ // (B.1.9)

$= \frac{1}{n!}\,\Sigma_{\mathbf{b}}\,\Sigma_{\mathbf{a}}\,(-1)^{s_b+s_a}\;\;\Pi(\mathbf{a},\mathbf{b};\, M)$ // (B.1.7)

$= \det(M)$ // (B.1.8)

B.2 The permutation group and the permutation tensor

Definition. A set of elements $\{g_i\}$ form a **group** G if :

- g_ig_j is also in the group (closure)
- g_i^{-1} exists for each g_i where g_i^{-1} is an element of the group (inverse)
- $(g_ig_j)g_k = g_i(g_jg_k)$ (associative) (B.2.1)

Fact: $g_iG = G$ (the rearrangement theorem) (B.2.2)

This says that multiplication of all the elements of a group by an element g_i in the group creates a reordering of the group elements.

Proof. Consider $g_iG = g_i\,[g_1, g_2....g_n] = [g_ig_1, g_ig_2....g_ig_n]$ = set of n elements. Unless two elements are the same, this must exhaust the entire group. How do we know that g_ig_1 and g_ig_2 might not be the same? Apply g_i^{-1} from the left and that would say $g_1 = g_2$ which is not the case.

Fact: $\Sigma_g f(g) = \Sigma_g f(g_1g)$ if Σ_g runs over the entire group G (B.2.3)

Proof: In $\Sigma_g f(g_1g)$, as g runs over G, the argument $g' \equiv g_1g$ runs over G by the rearrangement theorem. Thus, the sum $\Sigma_g f(g_1g)$ is just a reordering of the terms in the sum $\Sigma_g f(g)$.

It is easy to show that the set of permutations of $\mathbf{z}_0$ is a group and therefore the above facts can be used. For example, the product of two permutations is a permutation, and every permutation clearly has an inverse, and (AB)C = A(BC) for any matrices. Fact (B.2.3) can be written

$$\Sigma_B f(B\mathbf{z}_o) = \Sigma_B f(CBz_0)$$

but this is the same as saying

$$\Sigma_b f(\mathbf{b}) = \Sigma_b f(C\mathbf{b}) \qquad \text{(B.2.4)}$$

Again, throwing in the permutation C merely causes a reordering of the sum.

Consider then:

$$\det(M) \equiv \frac{1}{n!}\,\Sigma_a \Sigma_b\,(-1)^{s_a+s_b}\,\Pi(\mathbf{a},\mathbf{b};M) \qquad \text{// (B.1.8)}$$

$$= \frac{1}{n!}\,\Sigma_a \Sigma_b\,(-1)^{s_a+s_b}\,\Pi(A^{-1}\mathbf{a},A^{-1}\mathbf{b};M) \qquad \text{// (B.1.6) with } C = A^{-1};\ A^{-1}\mathbf{a} = A^{-1}A\mathbf{z}_0 = \mathbf{z}_0$$

$$= \frac{1}{n!}\,\Sigma_a\,[\,\Sigma_b\ \text{Parity}(A^{-1}\mathbf{b})\ \Pi(\mathbf{z}_0,A^{-1}\mathbf{b};M)\,] \qquad \text{// (B.1.4) with } C = A^{-1} \text{ and } \mathbf{a} = \mathbf{b}$$

$$= \frac{1}{n!}\,\Sigma_a\,[\,\Sigma_b\ \text{Parity}(\mathbf{b})\ \Pi(\mathbf{z}_0,\mathbf{b};M)\,] \qquad \text{// (B.2.4) with } C = A^{-1} \text{ (key step)}$$

$$= \Sigma_b\ \text{Parity}(\mathbf{b})\ \Pi(\mathbf{z}_0,\mathbf{b};M) \qquad \text{// n! identical terms in } \Sigma_a$$

$$= \Sigma_a\ \text{Parity}(\mathbf{a})\ \Pi(\mathbf{z}_0,\mathbf{a};M) \qquad \text{// rename dummy sum variable}$$

$$= \Sigma_a\ \text{Parity}(\mathbf{a})\ M_{1a_1}M_{2a_2}\ .....\ M_{na_n} \qquad \text{// (B.1.5)} \qquad \text{(B.2.5)}$$

Since $\det(M^T) = \det(M)$ from Theorem 1 (B.1.10), we can also write this result as

$$\det(M) = \Sigma_{\mathbf{a}}\ \text{Parity}(\mathbf{a})\ M_{a_1 1} M_{a_2 2} \ldots\ldots M_{a_n n} \qquad \text{(B.2.6)}$$

Proof: $\det(M) = \det(M^T) = \Sigma_{\mathbf{a}}\ \text{Parity}(\mathbf{a})\ M^T_{1a_1} M^T_{2a_2} \ldots\ldots M^T_{na_n} = \Sigma_{\mathbf{a}}\ \text{Parity}(\mathbf{a})\ M_{a_1 1} M_{a_2 2} \ldots\ldots M_{a_n n}$

Here then are last two results:

$$\det(M) \equiv \Sigma_{\mathbf{a}}\ \text{Parity}(\mathbf{a})\ M_{1a_1} M_{2a_2} \ldots\ldots M_{na_n} \qquad \text{(B.2.7a)}$$
$$\det(M) \equiv \Sigma_{\mathbf{a}}\ \text{Parity}(\mathbf{a})\ M_{a_1 1} M_{a_2 2} \ldots\ldots M_{a_n n} \qquad \text{(B.2.7b)}$$

Definition: The **permutation tensor** $\varepsilon_{ijk..}$ (n subscripts) :

- $\varepsilon_{12\ldots n} = 1$
- for any index swap, ε changes sign: $\varepsilon_{..i..j..} = -\varepsilon_{..j..i..}$
- therefore, if any two indices are the same, ε = 0 $\varepsilon_{..i..i..} = -\varepsilon_{..i..i..} = 0$ (B.2.8)

Fact: $\Sigma_{\mathbf{a}}\ \text{Parity}(\mathbf{a})\ f(\mathbf{a}) = \Sigma_{a_1 a_2 \ldots a_n}\ \varepsilon_{a_1 a_2 \ldots a_n}\ f(a_1, a_2 \ldots a_n)$ (B.2.9)

Proof: Although the sum on the right includes all n^n terms, any terms with identical indices are removed since ε = 0 for such terms. Thus, the sum is really only over those n! terms which are permutations of 123,,n. For any such permutation, $\text{Parity}(\mathbf{a}) = \varepsilon_{a_1 a_2 \ldots a_n}$. Why? We know that $\text{Parity}(\mathbf{a}) = (-1)^{S_{\mathbf{a}}}$ where $S_{\mathbf{a}}$ is the number of swaps to get to **a** from $\mathbf{z}_0$. But it then takes $S_{\mathbf{a}}$ index swaps to get from $\varepsilon_{a_1 a_2 \ldots a_n}$ to $\varepsilon_{123..n}$ and that involves $S_{\mathbf{a}}$ negations, which again is $(-1)^{S_{\mathbf{a}}}$.

Using (B.2.9) in (B.2.7) then gives these two classic det(M) expressions:

Theorem 2: det(M) can be represented in these two ways:

$$\det(M) = \Sigma_{a_1 a_2 \ldots a_n}\ \varepsilon_{a_1 a_2 \ldots a_n}\ M_{1a_1} M_{2a_2} \ldots\ldots M_{na_n} \qquad \text{(B.2.10a)}$$
$$\det(M) = \Sigma_{a_1 a_2 \ldots a_n}\ \varepsilon_{a_1 a_2 \ldots a_n}\ M_{a_1 1} M_{a_2 2} \ldots\ldots M_{a_n n} \ . \qquad \text{(B.2.10b)}$$

In most applications, the following simpler notation suffices:

$$\det(M) = \Sigma_{abc..q}\ \varepsilon_{abc..q}\ M_{1a} M_{2b} M_{3c} \ldots M_{nq} \qquad \text{(B.2.11a)}$$
$$\det(M) = \Sigma_{abc..q}\ \varepsilon_{abc..q}\ M_{a1} M_{b2} M_{c3} \ldots M_{qn} \qquad \text{(B.2.11b)}$$

Theorem 3: For a square matrix M, adding a multiple of one row to another does not change det(M). The same is true for adding a multiple of one column to another. (B.2.12)

Proof: (rows) Suppose we replace $\mathbf{r}_3 \rightarrow \mathbf{r}_3 + \alpha\, \mathbf{r}_2$. This says $M_{3i} \rightarrow M_{3i} + \alpha\, M_{2i}$. Eq (B.2.11a) says :

$$\det(M') = \Sigma_{abc..q}\ \varepsilon_{abc..q}\ M_{1a} M_{2b} (M_{3c} + \alpha\, M_{2c}) \ldots M_{nq}$$

$$= \det(M) + \alpha\ \Sigma_{abc..q}\ \varepsilon_{abc..q}\ M_{1a} M_{2b} M_{2c} \ldots M_{nq} \ .$$

Since $M_{1a}M_{2b}M_{2c}$ ....M_{nq} is symmetric under b↔c while $\varepsilon_{abc..q}$ is anti-symmetric, the extra α term vanishes. That is to say, sum = $\Sigma_{bc}\ A_{bc}S_{bc}\ =\ \Sigma_{cb}\ A_{cb}S_{cb}\ =\ \Sigma_{cb}\ (-A_{bc})(S_{bc})$ = - sum = 0. Using (B.2.11b) this argument shows that $\mathbf{c}_3 \rightarrow \mathbf{c}_3 + \alpha\,\mathbf{c}_2$ similarly does not alter det(M).

Theorem 4: Swapping two rows (columns) of square matrix M causes det(M) → - det(M). (B.2.13)

Proof: Let's swap rows 1 and 3 in M to get M'. Then from (B.2.11a),

$$\det(M') = \Sigma_{abc..q}\,\varepsilon_{abc..q}\,M_{3a}M_{2b}M_{1c}....M_{nq}$$

$$= \Sigma_{abc..q}\,[-\varepsilon_{cba..q}]\,M_{3a}M_{2b}M_{1c}....M_{nq} \qquad \text{// swap a↔c on } \varepsilon$$

$$= -\,\Sigma_{abc..q}\,\varepsilon_{abc..q}\,M_{3c}M_{2b}M_{1a}....M_{nq} \qquad \text{// dummy rename a↔c}$$

$$= -\,\Sigma_{abc..q}\,\varepsilon_{abc..q}\,M_{1a}M_{2b}M_{3c}....M_{nq} \qquad \text{// reorder product}$$

$$= -\,\det(M)\ .$$

Using Theorem 1, we then know that $\det(M'^{T}) = -\det(M^{T})$, so swapping columns 1 and 3 negates det(M). An obvious corollary if the two swapped columns have identical data :

Corollary 4: If two rows (columns) of square matrix A are the same, then det(M) = 0. (B.2.14)

B.3 Minors, Cofactors, the Cofactor Expansions of det(M)

Definition: The **minor** of a matrix element M_{rs} is the determinant of the submatrix obtained by crossing out the rth row and the sth column. It is a challenge, however, to write this out in symbols. Let's start with a more detailed version of (B.2.11a), where the a_i are *column* summation indices,

$$\det(M) = \Sigma_{a_1a_2a_3a_4a_5\ldots a_n}\,\varepsilon_{a_1a_2a_3a_4a_5\ldots a_n}\,M_{1a_1}M_{2a_2}M_{3a_3}M_{4a_4}M_{5a_5}\,.....\,M_{na_n}\ . \qquad \text{(B.3.1)}$$

We shall make the following conjecture for the form of minor(M_{23})

$$\text{minor}(M_{32}) = (-1)^{3-2}\ \Sigma_{a_1a_2a_4a_5\ldots a_n}\ \ \varepsilon_{a_1a_22a_4a_5\ldots a_n}\ \ M_{1a_1}M_{2a_2}M_{4a_4}M_{5a_5}\,.....\,M_{na_n} \qquad \text{(B.3.2)}$$

Compared to det(M) shown in (B.3.1), we have made these changes :

- Removed the factor M_{3a_3} (since row 3 matrix elements cannot appear in minor(M_{32}))
- Removed the sum over a_3.
- Replaced a_3 by the number 2 on the ε tensor.
- added a sign factor $(-1)^{3-2}$.

Notice that since there is a 2 on the ε tensor, any time a summation index $a_i = 2$ there is no contribution since then the ε tensor has two indices the same. so in effect the value 2 has been removed from all the residual summations a_i. That is good, since column 2 is supposedly "crossed out" in minor(M_{32}). The factor $(-1)^{3-2}$ is added so that the "diagonal term" in the minor will be positive. This sign $(-1)^{3-2}$ gets used up if we slide the "2" to its natural position (position 2) on the ε tensor,

$$\text{minor}(M_{32}) = \Sigma_{a_1 a_2 a_4 a_5 \ldots a_n} \; \varepsilon_{a_1 2 a_2 a_4 a_5 \ldots a_n} \; M_{1a_1} M_{2a_2} M_{4a_4} M_{5a_5} \ldots\ldots M_{na_n} \; . \qquad \text{(B.3.3)}$$

The diagonal term in minor(M_{32}) is then positive, matching the mechanical crossing out method,

$$\varepsilon_{12345..n} \, M_{11} M_{22} M_{44} M_{55} \ldots\ldots M_{nn} \; = \; + M_{11} M_{22} M_{44} M_{55} \ldots\ldots M_{nn} \; . \qquad \text{(B.3.4)}$$

A more compact notation for (B.3.2) would be

$$\text{minor}(M_{32}) = (-1)^{3-2} \; \Sigma_{a_i, i \neq 3} \; \varepsilon_{a_i, a_3 = 2} \; \Pi_{i \neq 3} \, (M_{i, a_i}) \; . \qquad \text{(B.3.5)}$$

Starting over, we could show similarly that

$$\text{minor}(M_{42}) = (-1)^{4-2} \; \Sigma_{a_i, i \neq 4} \; \varepsilon_{a_i, a_4 = 2} \; \Pi_{i \neq 4} \, (M_{i, a_i}) \qquad \text{(B.3.6)}$$

We now have a sign factor $(-1)^{4-2}$ because the "2" on ε has to be slid 2 positions to get to its natural location (position 4 on ε). More generally we can say

$$\text{minor}(M_{s2}) = (-1)^{s-2} \; \Sigma_{a_i, i \neq s} \; \varepsilon_{a_i, a_s = 2} \; \Pi_{i \neq s} \, (M_{i, a_i}) \qquad \text{(B.3.7)}$$

where the slide is now s-2 places. Still more generally we find, replacing 2 by r,

$$\text{minor}(M_{sr}) = (-1)^{s \pm r} \; \Sigma_{a_i, i \neq s} \; \varepsilon_{a_i, a_s = r} \; \Pi_{i \neq s} \, (M_{i, a_i}) \; . \qquad \text{(B.3.8)}$$

This then is our "best form" expression for a minor of matrix element M_{sr}. Either sign will do, since

$$(-1)^{s-r} \; = \; (-1)^{s-r} (-1)^{2r} \; = (-1)^{s+r} \qquad \text{since } (-1)^{2r} = 1$$

Choosing the + sign in (B.3.8) and putting $(-1)^{s+r}$ on the left side we get,

$$(-1)^{s+r} \; \text{minor}(M_{sr}) = \; \Sigma_{a_i, i \neq s} \; \varepsilon_{a_i, a_s = r} \; \Pi_{i \neq s} \, (M_{i, a_i}) \; . \qquad \text{(B.3.9)}$$

The left side here is called the **cofactor** of M_{sr}, written cof(M_{sr}). Thus we have shown that

$$\text{cof}(M_{sr}) \; \equiv \; (-1)^{s+r} \; \text{minor}(M_{sr}) \; = \; \Sigma_{a_i, i \neq s} \; \varepsilon_{a_i, a_s = r} \; \Pi_{i \neq s} \, (M_{i, a_i}) \; . \qquad \text{(B.3.10)}$$

Fact: Neither minor(M_{sr}) nor cof(M_{sr}) are functions of the M_{sr} matrix elements of M! In (B.3.10) this is so because (1) row s is excluded in $\Pi_{i\neq s}(M_{i,a_i})$; (2) $a_i = r$ is excluded by the factor $\varepsilon_{a_i,a_s=r}$. More intuitively, this is so because to get minor(M_{sr}) we "cross out" row s and column r. Thus,

$$\frac{\partial \text{cof}(M_{sr})}{\partial M_{sr}} = \frac{\partial \text{minor}(M_{sr})}{\partial M_{sr}} = 0 \qquad \text{for any pair r,s in 1,2...n} \qquad \text{(B.3.11)}$$

Finally, suppose we replace r with an integer which we call a_s. Doing this causes

$$\varepsilon_{a_i,a_s=r} \rightarrow \varepsilon_{a_i,a_s=a_s} = \varepsilon_{a_i} = \varepsilon_{a_1a_2a_4a_5\ldots a_n} = \text{the normal } \varepsilon \text{ tensor form} \qquad \text{(B.3.12)}$$

Then (B.3.10) becomes the following,

$$\text{cof}(M_{sa_s}) \equiv (-1)^{s+a_s} \text{ minor}(M_{sa_s}) = \Sigma_{a_i,i\neq s}\ \varepsilon_{a_i}\ \Pi_{i\neq s}(M_{i,a_i}) \ . \qquad \text{(B.3.13)}$$

Now start again with det(M) of (B.3.1) and rewrite it in our compact notation:

$$\det(M) = \Sigma_{a_1a_2a_3a_4a_5\ldots a_n}\ \varepsilon_{a_1a_2a_3a_4a_5\ldots a_n}\ M_{1a_1}M_{2a_2}M_{3a_3}M_{4a_4}M_{5a_5} \ldots\ldots M_{na_n}$$

$$= \Sigma_{a_i}\ \varepsilon_{a_i}\ \Pi_i(M_{i,a_i}) \qquad \text{// next, extract the } a_s \text{ sum and its } M_{sa_s} \text{ factor :}$$

$$= \Sigma_{a_s}\ M_{sa_s}\ [\Sigma_{a_i,i\neq s}\ \ \varepsilon_{a_i}\ \Pi_{i\neq s}\ (M_{i,a_i})] \qquad \text{// next use (B.3.12) to get}$$

$$= \Sigma_{a_s}\ M_{sa_s}\ [\text{cof}(M_{sa_s})] \quad \text{// next, change summation index from } a_s \text{ to n}$$

$$= \Sigma_n\ M_{sn}\ \text{cof}(M_{sn}) \ . \qquad \text{// valid for any s in 1,2...n} \qquad \text{(B.3.14)}$$

This is the **cofactor expansion** of det(M) where one "works across row s" and n is a column index. Another form of this expansion is

$$\det(M) = \det(M^T) = \Sigma_n\ M^T_{sn}\ \text{cof}(M^T_{sn})$$

$$= \Sigma_n\ M_{ns}\ \text{cof}(M_{ns}) \qquad \text{(B.3.15)}$$

and in this form one "works down column s" where n is now a row index. We have just proven :

Theorem 5 (Cofactor Expansions):

$$\det(M) = \Sigma_n\ M_{sn}\ \text{cof}(M_{sn}) = \Sigma_n\ (\mathbf{r}_s)_n\ \text{cof}(M_{sn}) \qquad \text{work across row s} \qquad s = 1,2....n \qquad \text{(B.3.16a)}$$

$$\det(M) = \Sigma_n\ M_{ns}\ \text{cof}(M_{ns}) = \Sigma_n\ (\mathbf{c}_s)_n\ \text{cof}(M_{ns}) \qquad \text{work down column s} \qquad s = 1,2....n \qquad \text{(B.3.16b)}$$

B.4 Expressions for the inverse matrix M^{-1} and Cramer's Rule

Consider our expansion (B.3.15a) working across row s,

$$\det(M) = \Sigma_n M_{sn} \operatorname{cof}(M_{sn}) \tag{B.4.1}$$

Suppose we replace the elements of row s (M_{sn}) with the elements of some other row r in M (M_{rn}). In doing so we have created a new matrix, call it M'. Notice that $\operatorname{cof}(M'_{sn}) = \operatorname{cof}(M_{sn})$ *because*, although going from M to M' we have altered row s, we have not altered $\operatorname{cof}(M_{sn})$ since this depends only on the entries in all the rows *other than* row s (think "crossing out" row s for minor(M_{sn})). See Fact (B.3.11). Therefore we find,

$$\det(M') = \Sigma_n M'_{sn} \operatorname{cof}(M'_{sn}) = \Sigma_n M_{rn} \operatorname{cof}(M_{sn}) . \tag{B.4.2}$$

But by Corollary 4 (B.2.14) det(M') = 0 since two rows of M' are the same. Thus

$$0 = \Sigma_n M_{rn} \operatorname{cof}(M_{sn}) \qquad r \neq s \tag{B.4.3}$$

Combining this with (B.4.1) gives,

$$\Sigma_n M_{rn} \operatorname{cof}(M_{sn}) = \det(M)\delta_{r,s} . \tag{B.4.4}$$

Now for clarity define a cofactor matrix C in this manner

$$C_{sn} \equiv \operatorname{cof}(M_{sn}) \tag{B.4.5}$$

In matrix notation, we could define the matrix cof(M) to be matrix C and then

$$C = \operatorname{cof}(M) \qquad \Rightarrow \qquad C_{sn} = [\operatorname{cof}(M)]_{sn} = \operatorname{cof}(M_{sn}) . \tag{B.4.6}$$

Now (B.4.4) says

$$\Sigma_n M_{rn} C_{sn} = \det(M)\delta_{r,s}$$

or

$$\Sigma_n M_{rn} C^{T}_{ns} = \det(M)\delta_{r,s}$$

and finally in matrix notation,

$$MC^{T} = \det(M)\, 1$$

or

$$M \left[\frac{C^{T}}{\det(M)} \right] = 1 . \tag{B.4.7}$$

Therefore we find this classic square matrix inversion formula,

Theorem 6: $$M^{-1} = \frac{C^{T}}{\det(M)} = \frac{[\operatorname{cof}(M)]^{T}}{\det(M)} = \frac{\operatorname{cof}(M^{T})}{\det(M)} \tag{B.4.8}$$

To verify the last equality in (B.4.8) consider :

$= [\mathrm{cof}(M)]^{\mathbf{T}}{}_{\mathbf{ns}} = [\mathrm{cof}(M)]_{\mathbf{sn}}$ // meaning of transpose

$= \mathrm{cof}(M_{\mathbf{sn}})$ // (B.4.6)

$= \mathrm{cof}(M^{\mathbf{T}}{}_{\mathbf{ns}})$

$= [\mathrm{cof}(M^{\mathbf{T}})]_{\mathbf{ns}}$ // (B.4.6) applied to $M^{\mathbf{T}}$

and therefore we have this matrix identity,

$$[\mathrm{cof}(M)]^{\mathbf{T}} = [\mathrm{cof}(M^{\mathbf{T}})] \ . \tag{B.4.9}$$

Using (B.4.8) it is trivial to solve a non-singular ($\det M \neq 0$) system of linear equations $\mathbf{y} = M\mathbf{x}$:

$$\mathbf{y} = M\mathbf{x} \quad \Rightarrow \quad \mathbf{x} = M^{-1}\,\mathbf{y} = \frac{1}{\det(M)}\ \mathrm{cof}(M^{\mathbf{T}})\,\mathbf{y} = \frac{1}{\det(M)}\ [\mathrm{cof}(M)]^{\mathbf{T}}\,\mathbf{y} \ . \tag{B.4.10}$$

In components,

$$x_{\mathbf{s}} = \frac{1}{\det(M)}\ \Sigma_{\mathbf{n}}\ ([\mathrm{cof}(M)]^{\mathbf{T}})_{\mathbf{sn}}\, y_{\mathbf{n}} = \frac{1}{\det(M)}\,\Sigma_{\mathbf{n}}\ y_{\mathbf{n}}\,[\mathrm{cof}(M)]_{\mathbf{ns}} = \frac{1}{\det(M)}\,\Sigma_{\mathbf{n}}\ y_{\mathbf{n}}\ \mathrm{cof}(M_{\mathbf{ns}}) \ . \tag{B.4.11}$$

Now recall the cofactor expansion (B.3.16b),

$$\det(M) = \Sigma_{\mathbf{n}}\, M_{\mathbf{ns}}\, \mathrm{cof}(M_{\mathbf{ns}}) = \Sigma_{\mathbf{n}}\, (\mathbf{c}_{\mathbf{s}})_{\mathbf{n}}\, \mathrm{cof}(M_{\mathbf{ns}}) \ . \tag{B.4.12}$$

If one replaces column $\mathbf{c}_{\mathbf{s}}$ in M by $\mathbf{y}$ one gets

$\det(M[\mathbf{c}_{\mathbf{s}} \rightarrow \mathbf{y}]) = \Sigma_{\mathbf{n}}\, y_{\mathbf{n}}\, \mathrm{cof}(M_{\mathbf{ns}})$ // (B.4.12)

$= \det(M)\, x_{\mathbf{s}}$. // (B.4.11) (B.4.13)

Solving this for $x_{\mathbf{s}}$ we obtain another classic result,

Theorem 7 (Cramer's Rule):

$$\mathbf{y} = M\mathbf{x} \quad \Rightarrow \quad x_{\mathbf{s}} = \frac{\det(M[\mathbf{c}_{\mathbf{s}} \rightarrow \mathbf{y}])}{\det(M)} \tag{B.4.14}$$