Solution to problem 6 Due Tuesday November 8

Rui Zhang and T. Witten 11/12/05

properties of determinants A generalized determinant of a set of n arbitrary vectors $\mathbf{X}_1, ..., \mathbf{X}_n$ in an n-dimensional space is defined as a linear real-valued function $f(\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n)$ in each of the n vectors that changes sign when any adjacent pair $\mathbf{X}_i, \mathbf{X}_{i+1}$ are interchanged. The ordinary determinant is a specific case of a generalized determinant. If we have a basis $\mathbf{e}_1, ..., \mathbf{e}_n$, we can express $f(\mathbf{X}...)$ in terms of the $f(\mathbf{e}...)$ and the set of coefficients x_{ij} in $\mathbf{X}_i = \sum_{j=1}^{n} x_{ij}\mathbf{e}_j$. Thus the f of the vectors becomes a function on the matrix of x_{ij}. For a given set of $\mathbf{X}'s$ we define the corresponding matrix as $\underline{\mathrm{X}}$. This problem is a review of standard properties of determinants. Feel free to consult standard sources in writing out the answers to these questions.

a) Is it possible to make a sequences of adjacent interchanges that return to the original ordering yet that result in a net change of sign? Why not?

Solution: Of course it's not possible, otherwise the function f becomes multivalued, and is not defined properly. Now we need to show that, if there are N adjacent interchanges before the original ordering is returned, N must be an even number. For n different real numbers $x_1, x_2, ...x_n$, we consider the polynomial

$$P(x_1, x_2, ...x_n) = \prod_{i<j} (x_i - x_j). \tag{1}$$

Clearly P is nonzero for any set of distinct x's. We show below that it changes sign under all adjacent interchanges. We consider a sequence of N interchanges that restores the original order. If N were odd, we would have $P(x_1, ..., x_n) = -P(x_1, ..., x_n)$, so that $P(x_1, ..., x_n) = 0$. Since P is manifestly not 0, we conclude that N cannot be odd. If x_i and x_{i+1} are exchanged, the affected factors in P are $(x_k - x_i)(x_k - x_{i+1})$ with $k < i$, $(x_i - x_k)(x_{i+1} - x_k)$ with $k > i + 1$ and $(x_i - x_{i+1})$. Only the $(x_i - x_{i+1})$ factor changes sign; the others are unchanged. Thus P changes sign.

An important observation is that, for an interchange of a pair (x_i, x_j), where x_i, x_j are not necessarily adjacent, the polynomial P also changes its sign. (This is a little bit tricky: say we interchange a pair (x_i, x_j) and $i < j$. For $k < i < j$, the product $(x_k - x_i)(x_k - x_j)$ keeps the original sign after the interchange; for $i < k < j$, the product $(x_i - x_k)(x_k - x_j)$ does not change sign; for $i < j < k$, $(x_i - x_k)(x_j - x_k)$, the sign remains the same after the interchange; only the factor $(x_i - x_j)$ changes its sign after the interchange) Therefore the interchange of a pair (x_i, x_j) must be equivalent to an odd number of interchanges of adjacent pair. And thus the generalized determinant f changes its sign when any pair $\mathbf{X}_i, \mathbf{X}_j$ are interchanged. We may need this important result for the following proof.

b) What happens to f under a *cyclic* permutation in which each $\mathbf{X}_i$ is shifted one place to the right and $\mathbf{X}_n \to \mathbf{X}_1$?

Solution: A *cyclic* permutation consists of $n-1$ adjacent interchange $(n-1,n),(n-2,n-1),...,(1,2)$. Therefore under a *cyclic* permutation $f \to (-1)^{n-1} f$.

c) Explain how to relate the f of a product of two matrices $\underline{\mathbf{X}}$ and $\underline{\mathbf{Y}}$ to the f's of $\underline{\mathbf{X}}$ and $\underline{\mathbf{Y}}$.

Solution: For $f \equiv 0$, this is trivial. Now we consider the case $f \neq 0$, and both $\mathbf{X}$'s and $\mathbf{Y}$'s are linearly independent. Otherwise, the production matrix $\underline{\mathbf{X}}\underline{\mathbf{Y}}$ is also degenerate and the corresponding set of matrices are linearly dependent, and therefore $f(\underline{\mathbf{X}}) f(\underline{\mathbf{Y}}) = f(\underline{\mathbf{X}}\underline{\mathbf{Y}}) = 0$, as inferred from parts (d) and (e).

With respect the basis $\mathbf{e}_1,...,\mathbf{e}_n$, $\underline{\mathbf{X}}$=(x_{ij}), and we write $f(\underline{\mathbf{X}})$ and $f(\underline{\mathbf{X}}\underline{\mathbf{Y}})$ in terms of vectors, using the definition of matrix product. That is: $f(\underline{\mathbf{Y}}) = f(\mathbf{Y}_1, \mathbf{Y}_2, ..., \mathbf{Y}_n)$, and

$$\begin{aligned} f(\underline{\mathbf{X}}\underline{\mathbf{Y}}) &= f\left(\sum_{i_1=1}^{n} x_{1i_1}\mathbf{Y}_{i_1},\ \sum_{i_2=1}^{n} x_{2i_2}\mathbf{Y}_{i_2}, ..., \sum_{i_n=1}^{n} x_{ni_n}\mathbf{Y}_{i_n}\right) \\ &= \sum_{\sigma} x_{\sigma(1)1}x_{\sigma(2)2}, ..., x_{\sigma(n)n} f\left(\mathbf{Y}_{\sigma(1)}, \mathbf{Y}_{\sigma(2)}, ..., \mathbf{Y}_{\sigma(n)}\right). \end{aligned} \tag{2}$$

Here we define the set map $\sigma : \{1,2,...,n\} \to \{1,2,...,n\}$ to be a permutation, such that $\{1,2,...,n\}$ and $\{\sigma(1),\sigma(2),...\sigma(n)\}$ are the same set. Please note that $f(\mathbf{Y}_{i_1}, \mathbf{Y}_{i_2}, ..., \mathbf{Y}_{i_n}) = 0$ if two of $\mathbf{Y}_i$'s are the same according to the result of part (d).

Likewise, we can also express $f(\underline{\mathbf{X}})$ in terms of the basis $\mathbf{e}_1,...,\mathbf{e}_n$,

$$\begin{aligned} f(\underline{\mathbf{X}}) &= f\left(\sum_{i_1=1}^{n} x_{1i_1}\mathbf{e}_{i_1},\ \sum_{i_2=1}^{n} x_{2i_2}\mathbf{e}_{i_2}, ..., \sum_{i_n=1}^{n} x_{ni_n}\mathbf{e}_{i_n}\right) \\ &= \sum_{\sigma} x_{\sigma(1)1}x_{\sigma(2)2}, ..., x_{\sigma(n)n} f\left(\mathbf{e}_{\sigma(1)}, \mathbf{e}_{\sigma(2)}, ..., \mathbf{e}_{\sigma(n)}\right). \end{aligned} \tag{3}$$

According to the result in part (a), we know that

$$\frac{f\left(\mathbf{Y}_{\sigma(1)}, \mathbf{Y}_{\sigma(2)}, ..., \mathbf{Y}_{\sigma(n)}\right)}{f\left(\mathbf{e}_{\sigma(1)}, \mathbf{e}_{\sigma(2)}, ..., \mathbf{e}_{\sigma(n)}\right)} = \frac{f(\mathbf{Y}_1, \mathbf{Y}_2, ..., \mathbf{Y}_n)}{f(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)}. \tag{4}$$

Therefore, by comparing Eqs. (2) and (3), we have

$$\begin{aligned} f(\underline{\mathbf{X}}\underline{\mathbf{Y}}) &= \frac{f(\mathbf{Y}_1, \mathbf{Y}_2, ..., \mathbf{Y}_n)}{f(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)} \sum_{\sigma} x_{\sigma(1)1}x_{\sigma(2)2}, ..., x_{\sigma(n)n} f\left(\mathbf{e}_{\sigma(1)}, \mathbf{e}_{\sigma(2)}, ..., \mathbf{e}_{\sigma(n)}\right) \\ &= \frac{f(\underline{\mathbf{X}}) f(\underline{\mathbf{Y}})}{f(\underline{\mathbf{I}})}, \end{aligned} \tag{5}$$

where $\underline{\mathbf{I}}$ is the unit matrix with respect to the basis $\mathbf{e}_1,...,\mathbf{e}_n$.

d) What is the value of f if two of the $\mathbf{X}$'s are the same? proportional? linear dependent? Explain.

Solution: If two of the $\mathbf{X}$'s are the same, we can interchange these two $\mathbf{X}$'s, and according to the observation made in part (a), $f = -f$, and thus $f = 0$. If two of the $\mathbf{X}$'s are proportional, we can easily reduce this case to the case when two of the $\mathbf{X}$'s are the same using the linearity of f, and therefore $f = 0$. If one of the $\mathbf{X}$'s can be expressed as a linear combination of other $\mathbf{X}$'s, without loss of generality, suppose $\mathbf{X}_1 = \sum_{i=2}^{n} a_i \mathbf{X}_i$, then $f(\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n) = \sum_{i=2}^{n} a_i f(\mathbf{X}_i, \mathbf{X}_2, ..., \mathbf{X}_n) = 0$, according to previous result in the case when two of the $\mathbf{X}$'s are the same.

e) What is the value of f if the $\mathbf{X}$'s are not linearly independent?

Solution: I feel this question is not asked in a proper way, since it has already been answered in part (d). Instead, let's consider the case when $\mathbf{X}$'s are not linearly dependent. The function $f \equiv 0$ is certainly a generalized determinant, and this is the trivial case. If f is a nontrivial function, then we argue that $f(\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n) \neq 0$ if the $\mathbf{X}$'s are not linearly independent. Otherwise, there exists at least one set of linearly independent $\tilde{\mathbf{X}}$'s, such that $f\left(\tilde{\mathbf{X}}_1, \tilde{\mathbf{X}}_2, ..., \tilde{\mathbf{X}}_n\right)$. Since $\tilde{\mathbf{X}}$'s are linearly independent, they form a complete basis for the entire linear space. For arbitrary set of n $\mathbf{X}$'s, we have for $i = 1, 2, ...n$, $\mathbf{X}_i = \sum_{j=1}^{n} a_{ij} \tilde{\mathbf{X}}_j$, and using the linearity of f

$$f(\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n) = \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} ... \sum_{j_n=1}^{n} a_{1j_1} a_{2j_2} .. a_{nj_n} f\left(\tilde{\mathbf{X}}_{j_1}, \tilde{\mathbf{X}}_{j_2}, ..., \tilde{\mathbf{X}}_{j_n}\right) = 0, \tag{6}$$

therefore $f \equiv 0$, which is a contradiction.

f) $\underline{\mathbf{X}}^T$ means the transpose of $\underline{\mathbf{X}}$. How is $f\left(\underline{\mathbf{X}}^T\right)$ related to $f(\underline{\mathbf{X}})$? Why?

Solution: We argue that $f\left(\underline{\mathbf{X}}^T\right) = f(\underline{\mathbf{X}})$ .We need some preliminaries here.

For each permutation σ, we define the permutation σ^{-1}, such that $\sigma^{-1}\sigma = I$, where I is the identity permutation, $I(i) = i$, for $i = 1, 2, .., n$. It is easy to show that σ^{-1} is well defined, and is unique. According to the property of the function f, we can define the function $sign : \{\sigma\} \to \{0, 1\}$, such that

$$f\left(\mathbf{X}_{\sigma(1)}, \mathbf{X}_{\sigma(2)}, ..., \mathbf{X}_{\sigma(n)}\right) = (-1)^{sign(\sigma)} f(\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n). \tag{7}$$

We have

$$\begin{aligned} f(\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n) &= f\left(\mathbf{X}_{\sigma^{-1}\sigma(1)}, \mathbf{X}_{\sigma^{-1}\sigma(2)}, ..., \mathbf{X}_{\sigma^{-1}\sigma(n)}\right) \\ &= (-1)^{sign\left(\sigma^{-1}\right)} f\left(\mathbf{X}_{\sigma(1)}, \mathbf{X}_{\sigma(2)}, ..., \mathbf{X}_{\sigma(n)}\right) \\ &= (-1)^{sign\left(\sigma^{-1}\right) + sign(\sigma)} f(\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n), \end{aligned} \tag{8}$$

and thus $sign\left(\sigma^{-1}\right) = sign(\sigma)$,

$$f\left(\mathbf{X}_{\sigma(1)}, \mathbf{X}_{\sigma(2)}, ..., \mathbf{X}_{\sigma(n)}\right) = f\left(\mathbf{X}_{\sigma^{-1}(1)}, \mathbf{X}_{\sigma^{-1}(2)}, ..., \mathbf{X}_{\sigma^{-1}(n)}\right). \tag{9}$$

Now we are ready to show that $f\left(\underline{\mathbf{X}}^T\right) = f(\underline{\mathbf{X}})$. With respect to the basis $\mathbf{e}_1, ..., \mathbf{e}_n$, $f(\underline{\mathbf{X}})$ can

be expressed as (here we write down Eq. (3) again)

$$f\left(\underline{\mathbf{X}}\right)=\sum_{\sigma}x_{\sigma(1)1}x_{\sigma(2)2},...,x_{\sigma(n)n}f\left(\mathbf{e}_{\sigma(1)},\mathbf{e}_{\sigma(2)},...,\mathbf{e}_{\sigma(n)}\right). \tag{10}$$

Likewise

$$\begin{aligned}f\left(\underline{\mathbf{X}}^{T}\right)&=\sum_{\sigma}x_{1\sigma(1)}x_{2\sigma(2)},...,x_{n\sigma(n)}f\left(\mathbf{e}_{\sigma(1)},\mathbf{e}_{\sigma(2)},...,\mathbf{e}_{\sigma(n)}\right)\\&=\sum_{\sigma}x_{\sigma^{-1}\sigma(1)\sigma(1)}x_{\sigma^{-1}\sigma(2)\sigma(2)},...,x_{\sigma^{-1}\sigma(n)\sigma(n)}f\left(\mathbf{e}_{\sigma(1)},\mathbf{e}_{\sigma(2)},...,\mathbf{e}_{\sigma(n)}\right),\end{aligned} \tag{11}$$

In Eq. (11), we re-order $\sigma(1),\sigma(2),...,\sigma(n)$ in each product from 1 to n, and this re-ordering does not change the product, and therefore

$$\begin{aligned}f\left(\underline{\mathbf{X}}^{T}\right)&=\sum_{\sigma}x_{\sigma^{-1}(\sigma(1))\sigma(1)}x_{\sigma^{-1}(\sigma(2))\sigma(2)},...,x_{\sigma^{-1}(\sigma(n))\sigma(n)}f\left(\mathbf{e}_{\sigma(1)},\mathbf{e}_{\sigma(2)},...,\mathbf{e}_{\sigma(n)}\right)\\&=\sum_{\sigma}x_{\sigma^{-1}(1)1}x_{\sigma^{-1}(2)2},...,x_{\sigma^{-1}(n)n}f\left(\mathbf{e}_{\sigma(1)},\mathbf{e}_{\sigma(2)},...,\mathbf{e}_{\sigma(n)}\right)\\&=\sum_{\sigma}x_{\sigma^{-1}(1)1}x_{\sigma^{-1}(2)2},...,x_{\sigma^{-1}(n)n}f\left(\mathbf{e}_{\sigma^{-1}(1)},\mathbf{e}_{\sigma^{-1}(2)},...,\mathbf{e}_{\sigma^{-1}(n)}\right)\quad[\text{Eq. (9)}]\\&=\sum_{\sigma}x_{\sigma(1)1}x_{\sigma(2)2},...,x_{\sigma(n)n}f\left(\mathbf{e}_{\sigma(1)},\mathbf{e}_{\sigma(2)},...,\mathbf{e}_{\sigma(n)}\right).\ (\sigma^{-1}\rightarrow\sigma)\\&=f\left(\underline{\mathbf{X}}\right).\end{aligned} \tag{12}$$