THE UNIVERSITY OF SYDNEY

MATH2961 Linear Algebra and Vector Calculus (advanced)

(http://www.maths.usyd.edu.au/u/UG/SM/MATH2961/)

Semester 1, 2009 Lecturers: R. Howlett and J. Parkinson

Linear Algebra Assignment

1. Let n be positive integer, let k and ℓ be distinct numbers in the set $\{1, 2, \ldots, n\}$, and let $\tau \in S_n$ be the transposition $(k\ \ell)$. Define

$$P = \{ \sigma \in S_n \mid \sigma(k) > \sigma(\ell) \},$$
$$Q = \{ \sigma \in S_n \mid \sigma(k) < \sigma(\ell) \}.$$

(*i*) Prove that P and Q are complementary subsets of S_n and that $f(\sigma) = \sigma\tau$ defines a bijective function from P to Q.

(*ii*) Let $A = (a_{ij})$ be an $n \times n$ matrix over a field F and suppose that the k-th row of A equals the ℓ-th row of A. Show that if $\sigma \in S_n$ and $\rho = \sigma\tau$ then

$$\varepsilon(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)} = -\varepsilon(\rho)a_{1\rho(1)}a_{2\rho(2)}\cdots a_{n\rho(n)},$$

and using Part (*i*) deduce that $\det(A) = 0$.

Solution.

(*i*) Let $\sigma \in S_n$ be arbitrary. Suppose first that $\sigma \notin P$. Then $\sigma(k) \not> \sigma(l)$, and so $\sigma(k) \le \sigma(l)$. If $\sigma(k) = \sigma(l)$ then $k = l$, since σ is injective (since permutations, by definition, are bijective functions). But $k \ne l$, by hypothesis, and so we conclude that $\sigma(k) \ne \sigma(l)$. Hence $\sigma(k) < \sigma(l)$, whence $\sigma \in Q$. Conversely, if $\sigma \in Q$ then $\sigma(k) < \sigma(l)$, whence $\sigma(k) \not> \sigma(l)$, and $\sigma \notin P$. So an element of S_n is in Q if and only if it is not in P. That is, the subsets P and Q of S_n are complementary.

For the next bit we make use of the following facts (which are are standard and should be familiar): composition of functions is associative, a function is bijective if and only if it has an inverse, and the composite of two bijective functions is bijective.

Let $\sigma \in S_n$ be arbitrary and let $\rho = \sigma\tau$. Then $\rho \in S_n$ (since it is the composite of two bijective functions), and we find that

$$\rho(k) = (\sigma\tau)(k) = \sigma(\tau(k)) = \sigma(l)$$

and

$$\rho(l) = (\sigma\tau)(l) = \sigma(\tau(l)) = \sigma(k).$$

If $\sigma \in P$ then $\sigma(k) > \sigma(l)$, whence $\rho(k) < \rho(l)$ and $\rho \in Q$; hence $f(\sigma) = \sigma\tau$ defines a function from P to Q. On the other hand, if $\sigma \in Q$ then $\sigma(k) < \sigma(l)$, whence $\rho(k) > \rho(l)$ and $\rho \in P$; so $g(\sigma) = \sigma\tau$ defines a function from Q to P. Now if $\sigma \in P$ then

$$(gf)(\sigma) = g(f(\sigma)) = g(\sigma\tau) = (\sigma\tau)\tau = \sigma(\tau\tau) = \sigma\mathtt{i} = \sigma,$$

and if $\rho \in Q$ then

$$(fg)(\sigma) = f(g(\sigma)) = f(\sigma\tau) = (\sigma\tau)\tau = \sigma(\tau\tau) = \sigma\mathtt{i} = \sigma.$$

Thus fg is the identity function on Q and gf the identity function on Q. So f has an inverse (namely g), and so f is bijective.

(*ii*) Note first that the assumption that the k-th row of A equals the ℓ-th row of A means that $a_{kj} = a_{\ell j}$ for all $j \in \{1, 2, \ldots, n\}$. Now let $\sigma \in S_n$ and put $\rho = \sigma\tau$. If $j \in \{1, 2, \ldots, n\}$ and $j \notin \{k, \ell\}$ then $\rho(j) = (\sigma\tau)(j) = \sigma(\tau(j)) = \sigma(j)$, since by definition $\tau(j) = j$ for $j \notin \{k, \ell\}$. Thus $a_{j\rho(j)} = a_{j\sigma(j)}$ for all such j. Moreover, since $\tau(k) = \ell$ and $\tau(\ell) = k$ we find that

$$\rho(k) = (\sigma\tau)(k) = \sigma(\tau(k)) = \sigma(\ell)$$

and

$$\rho(\ell) = (\sigma\tau)(\ell) = \sigma(\tau(\ell)) = \sigma(k),$$

giving $a_{k\rho(k)} = a_{k\sigma(\ell)} = a_{\ell\sigma(\ell)}$ and $a_{\ell\rho(\ell)} = a_{\ell\sigma(k)} = a_{k\sigma(k)}$, since the k-th and ℓ-th rows of A are equal. It follows that

$$\prod_{j=1}^{n} a_{j\rho(j)} = \Big(\prod_{j\notin\{k,\ell\}} a_{j\rho(j)}\Big)a_{k\rho(k)}a_{\ell\rho(\ell)}$$
$$= \Big(\prod_{j\notin\{k,\ell\}} a_{j\sigma(j)}\Big)a_{\ell\sigma(\ell)}a_{k\sigma(k)} = \prod_{j=1}^{n} a_{j\sigma(j)}.$$

Now since $\varepsilon(\rho) = \varepsilon(\sigma\tau) = \varepsilon(\sigma)\varepsilon(\tau) = -\varepsilon(\sigma)$ (since transpositions are odd permutations) it follows that

$$\varepsilon(\rho)a_{1\rho(1)}a_{2\rho(2)}\cdots a_{n\rho(n)}, = -\varepsilon(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)}$$

as required. Before proceeding we note that this can be rewritten as

$$\varepsilon(f(\sigma))\prod_{j=1}^{n} a_{jf(\sigma)(j)} = -\varepsilon(\sigma)\prod_{j=1}^{n} a_{j\sigma(j)}$$

where f is the function defined in Part (*i*).

By definition $\det(A) = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{j=1}^{n} a_{j\sigma(j)}$, and since P and Q are complementary subsets of S_n it follows that

$$\det(A) = \sum_{\sigma \in P} \varepsilon(\sigma) \prod_{j=1}^{n} a_{j\sigma(j)} + \sum_{\rho \in Q} \varepsilon(\rho) \prod_{j=1}^{n} a_{j\rho(j)}.$$

Since the function f in Part (i) is bijective, each $\rho \in Q$ is uniquely expressible as $f(\sigma)$ for some σ in P, enabling us to write the second sum in the above expression as a sum over $\sigma \in P$ instead of a sum over $\rho \in Q$. We obtain

$$\det(A) = \sum_{\sigma \in P} \varepsilon(\sigma) \prod_{j=1}^{n} a_{j\sigma(j)} + \sum_{\sigma \in P} \varepsilon(f(\sigma)) \prod_{j=1}^{n} a_{jf(\sigma)(j)}.$$

This is zero, since by the preceding paragraph we see that each term in the first sum cancels the corresponding term in the second sum.

2. An $n \times n$ symmetric matrix S with entries in $\mathbb{R}$ (the field of all real numbers) is said to be *positive definite* if ${}^{t}xSx > 0$ for all nonzero $x \in \mathbb{R}^n$. (Recall that if $x \in \mathbb{R}^n$ then x is a column vector, which can be regarded as an $n \times 1$ matrix, and the transpose of x, denoted ${}^{t}x$, is therefore a $1 \times n$ matrix. So ${}^{t}xSx$ is a 1×1 matrix, which is essentially just a real number.)

(i) Show that the identity matrix is positive definite.

(ii) Show that if S is positive definite and A is an $n \times n$ matrix such that ${}^{t}A + A = S$ then A is nonsingular. (Hint: You may use the fact that a square matrix is nonsingular if its null space is zero. Thus you should assume that $Ax = 0$ and prove that $x = 0$.)

Solution.

(i) Let $x \in \mathbb{R}^n$ be nonzero. If S is the identity matrix then

$${}^{t}xSx = {}^{t}xx = \sum_{i=1}^{n} x_i^2,$$

where x_i is the i-th entry of x. Since $x_i^2 \geq 0$ for all i and since at least one of these terms is nonzero (since x is nonzero) it follows that ${}^{t}xSx > 0$. Since x was chosen as an arbitrary nonzero element of $\mathbb{R}^n$ this proves that the identity matrix is positive definite.

(ii) Assume that S is positive definite and that ${}^{t}A + A = S$. Let x be an arbitrary element of the null space of A. Then $Ax = \underset{\sim}{0}$, where $\underset{\sim}{0}$ denotes the zero column vector. Transposing this and using the fact that transposing reverses

products, it follows that ${}^{t}x{}^{t}A = {}^{t}(Ax) = {}^{t}\underset{\sim}{0}$, the zero row vector. Multiplying this equation on the right by x gives

$${}^{t}x{}^{t}Ax = {}^{t}\underset{\sim}{0} = 0. \qquad (1)$$

But we also have that $Ax = \underset{\sim}{0}$, and multiplying this on the left by ${}^{t}x$ gives

$${}^{t}xAx = {}^{t}x0 = 0. \qquad (2)$$

Adding (1) and (2) gives

$${}^{t}xSx = {}^{t}x({}^{t}A + A)x = {}^{t}x({}^{t}Ax + Ax) = {}^{t}x{}^{t}Ax + {}^{t}xAx = 0 + 0 = 0,$$

and since S is positive definite it follows that $x = \underset{\sim}{0}$. Now since x was chosen as an arbitrary element of the null space of A it follows that the null space of A is $\{\underset{\sim}{0}\}$. Hence A is nonsingular, as required.