MTH6140 Linear Algebra II

Coursework 5 Solutions

1. Let
$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$
be an upper triangular matrix. We have to show that
$$\det(A) = a_{11}a_{22}\cdots a_{nn}.$$

(a) Suppose that $a_{nn} = 0$. Then A has a row of zeros and $\det(A) = 0$. Clearly the right-hand side is also zero.

Now suppose that $a_{nn} \neq 0$. Multiply the row n by $a_{n,n}^{-1}$ (this row operation multiplies the determinant by $a_{n,n}^{-1}$) to get a matrix with $a_{n,n} = 1$. Now applying Type 1 row operations (which don't change the determinant) we can ensure that the rest of the entries in the final column are are zero.

Continuing this process $i = n-1, n-2, \ldots, 1$, either:

(i) On some iteration we find that $a_{i,i} = 0$. If this occurs, halt. Row i of the current matrix is composed entirely of zeros.

(ii) Otherwise, we succeed in reducing A to the identity matrix I_n. The sequence of row operations employed multiplied the determinant by $(a_{1,1}a_{2,2}\cdots a_{n,n})^{-1}$.

In case (i), $\det(A) = 0$ and $a_{1,1}a_{2,2}\cdots a_{n,n} = 0$ also. In case (ii), we have that
$$(a_{11}a_{22}\cdots a_{n,n})^{-1}\det(A) = \det(I) = 1,$$
so $\det(A) = a_{1,1}a_{2,2}\cdots a_{nn}$, as required.

(b) Let us use the cofactor expansion, along column 1. Also, we will use induction on n, assuming the result true for upper triangular matrices of size $(n-1)\times(n-1)$.

All entries in the first column of A are zero except possibly the first entry $a_{1,1}$; so there is only one term in the cofactor expansion, namely $a_{1,1}\det(A_{1,1})$, where A_{11} is the minor obtained from A by deleting the first row and column. But this is a lower triangular matrix with diagonal entries $a_{2,2}, \ldots, a_{n,n}$. By induction, $\det(A_{1,1}) = a_{22}\cdots a_{n,n}$, so that $\det(A) = a_{1,1}\cdots a_{n,n}$, as required.

(c) Finally, we try using the the sum-over-permutations formula. The only permutation that makes a non-zero contribution is the identity permutation. (First note that $\pi(n)$ must equal n, as $a_{n,n}$ is the only non-zero entry in the final row of the matrix; then $\pi(n-1)$ cannot be n so must be $n-1$, and so on. In general $\pi(i) = i$.) Since the sign of the identity permutation is $+1$, we deduce that the determinant of A is $a_{1,1}a_{2,2}\cdots a_{n,n}$.

2. (a) Taking the permutations π in the order $(1)(2)(3)$, $(1)(2,3)$, $(1,2)(3)$, $(1,2,3)$, $(1,3)(2)$, $(1,3,2)$ in the sum-over-permutations (Leibniz) formula we obtain

$$\det(A) = aei - afh - bdi + bfg + cdh - ceg.$$

We have used the fact that the sign of the identity is $+1$, of a transposition is -1 and of a 3-cycle is $+1$.

Using the Laplace expansion along the first row we have

$$\det(a) = a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix}.$$

Then, repeating with the smaller 2×2 matrices,

$$\begin{aligned}\det(A) &= a(ei - fh) - b(di - fg) + c(dh - eg)\\ &= aei - afh - bdi + bfg + cdh - ceg.\end{aligned}$$

(b) *Remark.* As seen in part (a) in the context of a 3×3 matrix, the terms of the Leibniz (sum-over-permutations) formula, and the terms obtained from the Laplace expansion along a row appear to be in 1-1 correspondence. In fact this is true for in general for $n \times n$ matrices. The correspondence between the two is done in detail in the Wikipedia entry on the Laplace expansion. The point of this part of the question is to hint at what goes on in the general case. Of course a similar identity applies to the other terms in the Laplace expansion.

First, let k be the number of cycles in σ. Then σ' has $k+1$ cycles, the extra one being the cycle (1). Thus $\text{sign}(\sigma') = (-1)^{n-(k+1)} = (-1)^{(n-1)-k} = \text{sign}(\sigma)$.

To verify the second identity, we just follow our noses. Let Σ denote the set of permutations of $\{2,3,\ldots,n\}$. Then

$$\begin{aligned}a_{1,1}K_{1,1}(A) &= a_{1,1}\sum_{\sigma\in\Sigma}\text{sign}(\sigma)a_{2,\sigma(2)}a_{3,\sigma(3)}\cdots a_{n,\sigma(n)}\\ &= \sum_{\sigma\in\Sigma}\text{sign}(\sigma)a_{1,\sigma'(1)}, a_{2,\sigma'(2)}a_{3,\sigma'(3)}\cdots a_{n,\sigma'(n)}\\ &= \sum_{\sigma\in\Sigma}\text{sign}(\sigma')a_{1,\sigma'(1)}, a_{2,\sigma'(2)}a_{3,\sigma'(3)}\cdots a_{n,\sigma'(n)}\\ &= \sum_{\pi\in S_n,\pi(1)=1}\text{sign}(\pi)a_{1,\pi(1)}\cdots a_{n,\pi(n)}.\end{aligned}$$

The second equality uses the definition of σ', and the third uses the fact we just proved, namely that $\operatorname{sign}(\sigma') = \operatorname{sign}(\sigma)$.

Remark. Another idea for verifying correctness of the Laplace expansion is to show that the expression $\sum_{j=1}^{n} a_{i,j} K_{i,j}(A)$ satisfies (D1)–(D3), and hence must be the determinant. The tricky case is (D2), and then only when the Laplace expansion is done along one of the two equal rows. One way out is to show that the result of doing a Laplace expansion along row i and then row j is same as that arising from the reverse order.

3. A Type 1 operation of the form $\mathrm{R}_i \leftarrow \mathrm{R}_i + c\mathrm{R}_j$ can be simulated by a sequence of Type 1′ and Type 2 operations: $\mathrm{R}_j \leftarrow c\mathrm{R}_j$, $\mathrm{R}_i \leftarrow \mathrm{R}_i + \mathrm{R}_j$ and $\mathrm{R}_j \leftarrow c^{-1}\mathrm{R}_j$. To verify this, suppose A is a matrix and that v_i and v_j are the ith and jth rows of A. (So v_i and v_j are row vectors rather than column vectors.) Applying the above sequence of operations to A causes rows i and j to evolve as follows:

$$(v_i, v_j) \to (v_i, cv_j) \to (v_i + cv_j, cv_j) \to (v_i + cv_j, v_j);$$

The overall effect is the same as a Type 1 operation.

A Type 3 operation of the form $\mathrm{R}_j \leftrightarrow \mathrm{R}_j$ can be simulated by a sequence of Type 1 and Type 2 operations: $\mathrm{R}_j \leftarrow \mathrm{R}_j + \mathrm{R}_i$, $\mathrm{R}_i \leftarrow (-1)\mathrm{R}_i$, $\mathrm{R}_i \leftarrow \mathrm{R}_i + \mathrm{R}_j$ and $\mathrm{R}_j \leftarrow \mathrm{R}_j - \mathrm{R}_i$. To verify this, suppose A, v_i and v_j are as above. Applying the above sequence of operations to A causes rows i and j to evolve as follows:

$$(v_i, v_j) \;\to\; (v_i, v_i + v_j) \;\to\; (-v_i, v_i + v_j) \;\to\; (v_j, v_i + v_j) \;\to\; (v_j, v_i).$$

The overall effect is the same as a Type 2 operation. Finally the Type 1 operations used in this sequence can be simulated by Type 1′ and Type 2 operations, as we saw earlier.

4. The adjugate of A is

$$\operatorname{Adj}(A) = \begin{bmatrix} 2 & -2 & 1 \\ -3 & 2 & -1 \\ -1 & 1 & -1 \end{bmatrix}.$$

Explanation of the first row of the matrix $A' = (a'_{i,j}) = \operatorname{Adj}(A)$:

$$a'_{1,1} = K_{1,1} = (-1)^{1+1} \begin{vmatrix} 1 & 1 \\ 0 & 2 \end{vmatrix} = 2, \quad a'_{1,2} = K_{2,1} = (-1)^{2+1} \begin{vmatrix} 1 & 0 \\ 0 & 2 \end{vmatrix} = -2 \quad \text{and}$$
$$a'_{1,3} = K_{3,1} = (-1)^{3+1} \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} = 1.$$

The other rows can be computed similarly.

By, e.g., Laplace expansion along the first row we have $\det(A) = 1 \times 2 - 1 \times 3 = -1$.

It is then easy to verify that

$$A \cdot \mathrm{Adj}(A) = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = -I_3 = \det(A) I_3.$$

5. (a) $p_A(x) = \det(xI - A) = x^3 - 1$ (e.g., by using the Leibniz formula). So

$$p_A(A) = A^3 - I = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}^3 - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I - I = O.$$

(b) Consider a general polynomial $m_A(x) = ax^2 + bx + c$ of degree at most 2. Note that

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \quad A^2 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Thus, substituting A for x in $m_A(x)$, we obtain

$$m_A(A) = a \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} + b \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} + c \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} c & b & a \\ a & c & b \\ b & a & c \end{bmatrix}.$$

So the only way $m_A(A)$ can equal O is for $a = b = c = 0$.

(c) It is convenient in this part to label rows and columns $0, 1, \ldots, n-1$ instead of $1, 2, \ldots, n$. Then the matrix A has a 1 at position (i, j) if and only if $j = i + 1 \pmod{n}$. By an easy direct calculation, the characteristic polynomial is

$$p_A(x) = \det(Ix - A) = \begin{vmatrix} x & -1 & 0 & \cdots & 0 \\ 0 & x & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & x & -1 \\ -1 & 0 & \cdots & 0 & x \end{vmatrix} = x^n - 1.$$

(There are two perumutations leading to non-zero terms in the Leibniz formula: one is the identity with sign 1, and the other is the cyclic permutation $(0, 1, \ldots, n-1)$, with sign $-(-1)^n$. The first of these yields the term x^n and the second the term $-(-1)^n(-1)^n = -1$.)

The effect of right multiplication by A is to cyclicly permute the columns right by one position. (Check this!) Thus, $A^n = IA^n = I$ and the matrix A satisfies its own characteristic polynomial.

Consider any polynomial $m_A(c) = c_{n-1}x^{n-1} + \cdots + c_1 x + c_0$ of degree less than n. The first row of the matrix

$$c_0 I + c_1 A + c_2 A^2 + \cdots + c_{n-1} A^{n-1}$$

is $[c_0, c_1, \ldots, c_{n-1}]$. Therefore it is clear that $m_A(A) = O$ only if $c_0 = c_1 = \cdots = c_{n-1} = 0$. In other words, the matrix A satisfies no non-trivial polynomial of degree less than n.

6. You just need to follow your nose and have faith:

$$\begin{aligned}
p_{A'}(x) &= \det(xI - A') \\
&= \det(xI - P^{-1}AP) \\
&= \det(P^{-1}(xI - A)P) \\
&= \det(P^{-1})\det(xI - A)\det(P) \\
&= \det(P^{-1})\det(P)\det(xI - A) \\
&= \det(I)\det(xI - A) \\
&= \det(xI - A) = p_A(x).
\end{aligned}$$