Coefficients of the Characteristic Polynomial

LOUIS L. PENNISI
University of Illinois at Chicago
Box 4348, Chicago, Il 60680

This note derives the determinantal formulas for the coefficients in the characteristic polynomial of a matrix. In some areas of mathematics and other fields of science one has to determine the characteristic polynomial

$$P(\lambda)=\det(\underline{A}-\lambda\underline{I})=b_0\lambda^n+b_1\lambda^{n-1}+b_2\lambda^{n-2}+\cdots+b_{n-1}\lambda+b_n, \tag{1}$$

of a matrix $\underline{A}$, where $\underline{A}=(a_{ij})$ is an $n\times n$ matrix over the field of real numbers (or complex numbers) and $\underline{I}$ is the $n\times n$ identity matrix. Once $P(\lambda)$ is determined, the sought for eigenvalues λ_i, $i=1,2,\ldots,n$, of $\underline{A}$ may be obtained by solving the characteristic equation $P(\lambda)=0$. When $n\leqslant 3$, usually one obtains $P(\lambda)$ by expanding the $\det(\underline{A}-\lambda\underline{I})$ by minors. When $n\geqslant 3$, this task may become somewhat laborious. For a method of determining the b_i's in (1), see [**1**], where the following recursive formula is found:

$$\begin{aligned}&b_0=(-1)^n,\qquad b_1=-(-1)^nT_1,\qquad b_2=-\frac{1}{2}\left[b_1T_1+(-1)^nT_2\right],\\&b_3=-\frac{1}{3}\left[b_2T_1+b_1T_2+(-1)^nT_3\right],\ldots,\\&b_n=-\frac{1}{n}\left[b_{n-1}T_1+b_{n-2}T_2+\cdots+b_1T_{n-1}+(-1)^nT_n\right],\end{aligned} \tag{2}$$

where $T_1,T_2,\ldots,T_n$ denote the traces of the matrices $\underline{A},\underline{A}^2,\ldots,\underline{A}^n$, respectively. (The *trace* of an $n\times n$ matrix $\underline{B}$ is the sum of the elements on its main diagonal.)

We shall now present an alternative method for determining the coefficients of the characteristic polynomial. This method will involve the expansion of determinants of order 1 through n rather than computing the traces of the matrices $\underline{A},\underline{A}^2,\ldots,\underline{A}^n$.

For our purpose, we shall consider the polynomial

$$g(t)=\det(\underline{K}+t\underline{I})\quad(\text{the }+\text{ sign intended}), \tag{3}$$

where $\underline{K}=(k_{ij})$ is an $n\times n$ matrix over the field of real numbers (or complex numbers). We now ask, what is the coefficient C_r of t^r in (3)? To separate the occurrences of t, let

$$f(t_1,t_2,\ldots,t_n)=\det\left(\underline{K}+\operatorname{diag}(t_1,t_2,\ldots,t_n)\right).$$

Then $g(t)=f(t,t,\ldots,t)$, and C_r is the sum of the coefficients of the terms of total degree r in $f(t_1,t_2,\ldots,t_n)$. The fact that $f(t_1,t_2,\ldots,t_n)$ is of degree 1 in each t_i separately simplifies the accounting:

$$C_r=\sum_{i_1<i_2<\cdots<i_r}\frac{\partial^r}{\partial t_{i_1}\partial t_{i_2}\cdots\partial t_{i_r}}f(t_1,t_2,\ldots,t_n)\Bigg|_{\substack{t_1=0\\t_2=0\\\vdots\\t_n=0}}$$

where $1\leqslant i_1,\ i_r\leqslant n$, and $0\leqslant r\leqslant n$.

Thus, the C_r's are now expressed in terms of the elements of the matrix $\underline{K}$. Let D be the determinant as a function of the n^2 entries:

$$D(k_{11},k_{12},\ldots,k_{nn})=\begin{vmatrix}k_{11}&\cdots&k_{1n}\\\cdot&\cdots&\cdot\\\cdot&\cdots&\cdot\\\cdot&\cdots&\cdot\\k_{n1}&\cdots&k_{nn}\end{vmatrix}.$$

Then

$$C_r = \sum_{i_1 < i_2 < \cdots < i_r} \frac{\partial^r}{\partial k_{i_1 i_1} \partial k_{i_2 i_2} \cdots \partial k_{i_r i_r}} D. \tag{4}$$

Because of the cluster of n^2 variables, it may be wise to spell out the transition to formula (4). We can reduce this cluster by suppressing the dependence of D on its off-diagonal variables k_{ij} where $i \neq j$. This suppression is feasible because these off-diagonal variables are parameters that remain fixed throughout this whole discussion. So let us write

$$f(t_1, t_2, \ldots, t_n) = D^*(k_{11} + t_1, k_{22} + t_2, \ldots, k_{nn} + t_n)$$

where D^* is a function of only n variables $k_{11}, k_{22}, \ldots, k_{nn}$ defined by

$$D^*(k_{11}, k_{22}, \ldots, k_{nn}) = D(k_{11}, k_{12}, \ldots, k_{nn}).$$

In this notation, it is manifest that

$$\frac{\partial^r f}{\partial t_{i_1} \partial t_{i_2} \cdots \partial t_{i_r}} = D^*_{k_{i_1 i_1} k_{i_2 i_2} \cdots k_{i_r i_r}}(k_{11} + t_1, k_{22} + t_2, \cdots, k_{nn} + t_n),$$

where the subscripts denote partial derivatives of D^*.

The rule for expansion of a determinant by a row (or column) tells us that $\partial D / \partial k_{ij}$ is the (i, j) cofactor of $\underline{K}$; when $i = j$ (on the diagonal), this is the (i, i)-minor. Thus, the partial derivative in (4) is just the subdeterminant of $\underline{K}$ resulting from crossing out the rows and columns numbered $i_1, i_2, \ldots, i_r$.

EXAMPLE. Let us determine $g(t) = \det(\underline{K} + t\underline{I})$, where

$$\underline{K} = \begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix}.$$

Solution.

$$D(k_{11}, k_{12}, \ldots, k_{33}) = \begin{vmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{vmatrix}. \tag{5}$$

By definition:

$$C_0 = D(k_{11}, k_{12}, \ldots, k_{33}) \tag{6}$$

$$C_1 = \sum_{i_1} \frac{\partial}{\partial k_{i_1 i_1}} D = \frac{\partial D}{\partial k_{11}} + \frac{\partial D}{\partial k_{22}} + \frac{\partial D}{\partial k_{33}} \tag{7}$$

$$= \begin{vmatrix} 1 & k_{12} & k_{13} \\ 0 & k_{22} & k_{23} \\ 0 & k_{32} & k_{33} \end{vmatrix} + \begin{vmatrix} k_{11} & 0 & k_{13} \\ k_{21} & 1 & k_{23} \\ k_{31} & 0 & k_{33} \end{vmatrix} + \begin{vmatrix} k_{11} & k_{12} & 0 \\ k_{21} & k_{22} & 0 \\ k_{31} & k_{32} & 1 \end{vmatrix}$$

$$= \begin{vmatrix} k_{22} & k_{23} \\ k_{32} & k_{33} \end{vmatrix} + \begin{vmatrix} k_{11} & k_{13} \\ k_{31} & k_{33} \end{vmatrix} + \begin{vmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{vmatrix}$$

$$C_2 = \sum_{i_1 < i_2} \frac{\partial^2}{\partial k_{i_1 i_1} \partial k_{i_2 i_2}} D = \frac{\partial^2 D}{\partial k_{11} \partial k_{22}} + \frac{\partial^2 D}{\partial k_{11} \partial k_{33}} + \frac{\partial^2 D}{\partial k_{22} \partial k_{33}} \tag{8}$$

$$= \begin{vmatrix} 1 & 0 & k_{13} \\ 0 & 1 & k_{23} \\ 0 & 0 & k_{33} \end{vmatrix} + \begin{vmatrix} 1 & k_{12} & 0 \\ 0 & k_{22} & 0 \\ 0 & k_{32} & 1 \end{vmatrix} + \begin{vmatrix} k_{11} & 0 & 0 \\ k_{21} & 1 & 0 \\ k_{31} & 0 & 1 \end{vmatrix}$$

$$= k_{33} + k_{22} + k_{11}$$

$$C_3 = \sum_{i_1 < i_2 < i_3} \frac{\partial^3}{\partial k_{i_1 i_1} \partial k_{i_2 i_2} \partial k_{i_3 i_3}} D = \frac{\partial^3 D}{\partial k_{11} \partial k_{22} \partial k_{33}} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1. \tag{9}$$

Thus, our desired polynomial $g(t)$ is given by

$$g(t) = C_0 + C_1 t + C_2 t^2 + C_3 t^3, \tag{10}$$

where C_0, C_1, C_2, and C_3 are given by (6)–(9), respectively.

Taking

$$\underline{K} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 2 & 2 \\ 3 & 2 & 3 \end{pmatrix},$$

we have:

$$C_0 = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 2 & 2 \\ 3 & 2 & 3 \end{vmatrix} = 4, \qquad C_1 = \begin{vmatrix} 2 & 2 \\ 2 & 3 \end{vmatrix} + \begin{vmatrix} 1 & 3 \\ 3 & 3 \end{vmatrix} + \begin{vmatrix} 1 & 2 \\ 4 & 2 \end{vmatrix} = 2 - 6 - 6 = -10,$$

$$C_2 = 3 + 2 + 1 = 6, \qquad C_3 = 1.$$

Substituting these values in (10), we obtain

$$g(t) = -4 - 10t + 6t^2 + t^3.$$

REMARK. The relationship between $g(t)$ of (3) and $P(\lambda)$ of (1) is given by $g(-\lambda) = P(\lambda)$, where going from (1) to (3), we have put $\underline{K} = \underline{A}$ and $t = -\lambda$.

Reference

[1] L. A. Zadeh and C. A. Desoer, Linear System Theory: The State Approach, McGraw-Hill, New York, 1963, pp. 303–305.

Matrices as Sums of Invertible Matrices

N. J. Lord
Tonbridge School
Kent TN9 1JP England

While it is a trivial truism that not every matrix is invertible, it does not seem to be well known that every matrix can be expressed as the sum of two invertible matrices. The proof of this makes a good exercise in elementary linear algebra and, although a direct proof is short, we have expanded the discussion to indicate several contrasting lines of attack.

For convenience we shall adopt the following notation:

$\mathbb{F}$ will denote the field under consideration;

q will denote the number of elements of $\mathbb{F}$ if $\mathbb{F}$ is finite;

$M(n,\mathbb{F})$ will denote the ring of $n \times n$ matrices with entries in $\mathbb{F}$, where $n > 1$;

$G(n,\mathbb{F})$ will denote the group of invertible matrices in $M(n,\mathbb{F})$;

I (or I_n, for emphasis) will denote the $n \times n$ identity matrix.

The theorem that we are going to prove then is as follows: