The Characteristic Polynomial

1. Coefficients of the characteristic polynomial

Consider the eigenvalue problem for an $n \times n$ matrix A,

$$A\vec{\boldsymbol{v}} = \lambda\vec{\boldsymbol{v}}\,, \qquad \vec{\boldsymbol{v}} \neq 0\,, \tag{1}$$

where the trivial solution, $\vec{\boldsymbol{v}} = 0$ is excluded. That is, the zero vector is *not* an eigenvector. The solution to this problem consists of identifying all possible values of λ (called the eigenvalues), and the corresponding non-zero vectors $\vec{\boldsymbol{v}}$ (called the eigenvectors) that satisfy eq. (1). Noting that $\mathbf{I}\vec{\boldsymbol{v}} = \vec{\boldsymbol{v}}$, where $\mathbf{I}$ is the $n \times n$ identity matrix, one can rewrite eq. (1) as

$$(A - \lambda\mathbf{I})\vec{\boldsymbol{v}} = 0\,. \tag{2}$$

This is a set of n homogeneous equations, where the n unknowns are the components of the vector $\vec{\boldsymbol{v}} = (v_1, v_2, \ldots v_n)$. If $A - \lambda\mathbf{I}$ is an invertible matrix, then one can simply multiply both sides of eq. (2) by $(A-\lambda\mathbf{I})^{-1}$ to conclude that $\vec{\boldsymbol{v}} = 0$ is the unique solution. Since the zero vector is not an eigenvector, one must demand that $A-\lambda\mathbf{I}$ is not invertible in order to find non-trivial solutions to eq. (2). That is, we demand that,

$$p(\lambda) \equiv \det(A - \lambda\mathbf{I}) = 0\,. \tag{3}$$

Eq. (3) is called the *characteristic equation*. Evaluating the determinant yields an nth order polynomial in λ, called the *characteristic polynomial*, which we have denoted above by $p(\lambda)$. The possible solutions to $p(\lambda) = 0$ yield the eigenvalues of the matrix A.

The determinant in eq. (3) can be evaluated by the usual methods. It takes the form,

$$\begin{aligned} p(\lambda) = \det(A - \lambda\mathbf{I}) &= \begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix} \\ &= (-1)^n \left[\lambda^n + c_1\lambda^{n-1} + c_2\lambda^{n-2} + \cdots + c_{n-1}\lambda + c_n\right]\,, \end{aligned} \tag{4}$$

where $A = [a_{ij}]$. The coefficients c_i are to be computed by evaluating the determinant. Note that we have identified the coefficient of λ^n to be $(-1)^n$. This arises from one term in the determinant that is given by the product of the diagonal elements. Evaluating the determinant via the expansion in cofactors (see, e.g., Section 3 of the class handout entitled *Determinant and the adjugate*), one can quickly verify that $(-1)^n\lambda^n$ is the only term in the characteristic polynomial that is proportional to λ^n. It is then convenient to factor out the $(-1)^n$ before defining the coefficients c_i.

Two of the coefficients are easy to obtain. Note that eq. (4) is valid for any value of λ. If we set $\lambda = 0$, then eq. (4) yields:

$$p(0) = \det\, A = (-1)^n c_n\,.$$

Noting that $(-1)^n(-1)^n = (-1)^{2n} = +1$ for any integer n, it follows that,

$$\boxed{c_n = (-1)^n \det\, A\,.} \tag{5}$$

One can also easily work out c_1 by evaluating the determinant in eq. (4) using the cofactor expansion. This yields a characteristic polynomial of the form,[1]

$$p(\lambda) = \det(A - \lambda\mathbf{I}) = (a_{11} - \lambda)(a_{22} - \lambda)\cdots(a_{nn} - \lambda) + c'_2\lambda^{n-2} + c'_3\lambda^{n-3} + \cdots + c'_n\,. \tag{6}$$

The term $(a_{11} - \lambda)(a_{22} - \lambda)\cdots(a_{nn} - \lambda)$ on the right hand side of eq. (6) is the product of the diagonal elements of $A - \lambda\mathbf{I}$. As explained in footnote 1, *none* of the remaining terms that arise in the cofactor expansion of $\det(A - \lambda\mathbf{I})$ [denoted by $c'_2\lambda^{n-2} + c'_3\lambda^{n-3} + \cdots + c'_n$ in eq. (6)] are proportional to λ^n or λ^{n-1}. Moreover,

$$\begin{aligned}
(a_{11} - \lambda)(a_{22} - \lambda)\cdots(a_{nn} - \lambda) &= (-\lambda)^n + (-\lambda)^{n+1}\left[a_{11} + a_{22} + \cdots + a_{nn}\right] + \cdots\,, \\
&= (-1)^n\left[\lambda^n - \lambda^{n-1}(\mathrm{Tr}\ A) + \cdots\right]\,,
\end{aligned} \tag{7}$$

where $\cdots$ contains terms that are proportional to λ^p, where $p \leq n - 2$. This means that the terms in the characteristic polynomial that are proportional to λ^n and λ^{n-1} arise *solely* from the term $(a_{11} - \lambda)(a_{22} - \lambda)\cdots(a_{nn} - \lambda)$. As shown in eq. (7), the term proportional to $-(-1)^n\lambda^{n-1}$ is the trace of A, which is defined to be equal to the sum of the diagonal elements of A. Comparing eqs. (4) and (6), it follows that,

$$\boxed{c_1 = -\mathrm{Tr}\ A\,.} \tag{8}$$

Expressions for $c_2, c_3, \ldots, c_{n-1}$ are more complicated. For example, eqs. (4) and (6) yield

$$c_2 = \sum_{\substack{i=1 \\ i<j}}^{n}\sum_{j=1}^{n} a_{ii}a_{jj} + c'_2\,. \tag{9}$$

An explicit expression for c'_2 will be given below eq. (38).

In conclusion, the general form for the characteristic polynomial, $p(\lambda) \equiv \det(A - \lambda\mathbf{I})$, is given by,

$$p(\lambda) = (-1)^n\left[\lambda^n - \lambda^{n-1}\,\mathrm{Tr}\ A + c_2\lambda^{n-2} + \cdots + (-1)^{n-1}c_{n-1}\lambda + (-1)^n\det\ A\right]. \tag{10}$$

Explicit formulae for $c_2, c_3, \ldots, c_{n-1}$ in terms of traces of powers of the matrix A are presented in Appendix A and in terms of the matrix elements of A in Appendix B.

[1] In computing the cofactor of the ij matrix element of $A - \lambda\mathbf{I}$, one removes row i and column j of $A - \lambda\mathbf{I}$ and evaluates the determinant of the remaining submatrix [multiplied by the sign factor $(-1)^{i+j}$]. Except for the product of diagonal elements, there is always one factor of λ in each of the rows and columns that has been removed. This implies that the maximal power one can achieve outside of the product of diagonal elements of $A - \lambda\mathbf{I}$ is λ^{n-2}.

By the fundamental theorem of algebra, an nth order polynomial equation of the form $p(\lambda) = 0$ possesses precisely n solutions (called roots), which we shall denoted by $\lambda_1, \lambda_2, \ldots, \lambda_n$. These are the eigenvalues of A, and they may be real or complex. If a root is non-degenerate (i.e., only one root has a particular numerical value), then we say that the root has multiplicity one—it is called a *simple root.* If a root is degenerate (i.e., more than one root has a particular numerical value), then we say that the root has multiplicity p, where p is the number of roots with that same value—such a root is called a *multiple root.* For example, a double root (as its name implies) arises when precisely two of the roots of $p(\lambda)$ are equal. In the counting of the n roots of $p(\lambda)$, multiple roots are counted according to their multiplicity.

One can always factor a polynomial in terms of its roots.[2] Thus, eq. (4) implies that:

$$p(\lambda) = (-1)^n (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)\,, \tag{11}$$

where multiple roots appear according to their multiplicity.[3] Multiplying out the n factors above yields,

$$\begin{aligned} p(\lambda) = (-1)^n \Bigg\{ \lambda^n - \lambda^{n-1} \sum_{i=1}^{n} \lambda_i + \lambda^{n-2} \sum_{\substack{i=1 \\ i<j}}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j + \ldots \\ + (-1)^k \lambda^{n-k} \sum_{\substack{i_1=1 \\ i_1<i_2<\cdots<i_k}}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_k=1}^{n} \underbrace{\lambda_{i_i} \lambda_{i_2} \cdots \lambda_{i_k}}_{k\ \text{factors}} + \cdots + (-1)^n \lambda_1 \lambda_2 \cdots \lambda_n \Bigg\}. \end{aligned} \tag{12}$$

Comparing with eq. (10), it immediately follows that:

$$\boxed{\begin{aligned} \text{Tr } A &= \sum_{i=1}^{n} \lambda_i = \lambda_1 + \lambda_2 + \cdots + \lambda_n\,, && \text{(13a)} \\ \det A &= \lambda_1 \cdot \lambda_2 \cdot \lambda_3 \cdots \lambda_n\,. && \text{(13b)} \end{aligned}}$$

The coefficients $c_2, c_3, \ldots, c_{n-1}$ are also determined by the eigenvalues. In general,

$$c_k = (-1)^k \sum_{\substack{i_1=1 \\ i_1<i_2<\cdots<i_k}}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_k=1}^{n} \underbrace{\lambda_{i_1} \lambda_{i_2} \cdots \lambda_{i_k}}_{k\ \text{factors}}\,, \qquad \text{for } \; k = 1, 2, \ldots, n\,. \tag{14}$$

For example, if $k = 2 \leq n$ then eq. (14) yields,

$$c_2 = \lambda_1 \lambda_2 + \lambda_1 \lambda_3 + \ldots + \lambda_1 \lambda_n + \lambda_2 \lambda_3 + \ldots + \lambda_2 \lambda_n + \ldots + \lambda_{n-1} \lambda_n\,.$$

[2]In practice, it may not be possible to explicitly determine the roots algebraically. Indeed, unlike quadratic, cubic, and quartic polynomials, the roots of a general polynomial of degree 5 or higher cannot be obtained algebraically in terms of a finite number of additions, subtractions, multiplications, divisions, and root extractions (due to a famous result known as Abel's impossibility theorem). Of course, one can always determine the roots numerically.

[3]This means that if multiple degenerate roots exist, then some of the λ_i among $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ will be equal, and hence will appear more than once on the right hand side of eq. (11).

2. The Cayley-Hamilton Theorem

Theorem: Given an $n \times n$ matrix A, the characteristic polynomial is defined by $p(\lambda) = \det(A - \lambda\mathbf{I}) = (-1)^n\,[\lambda^n + c_1\lambda^{n-1} + c_2\lambda^{n-2} + \cdots + c_{n-1}\lambda + c_n]$, it follows that[4]

$$p(A) = (-1)^n\left[A^n + c_1A^{n-1} + c_2A^{n-2} + \cdots + c_{n-1}A + c_n\mathbf{I}\right] = \mathbf{0}\,, \tag{15}$$

where $A^0 \equiv \mathbf{I}$ is the $n \times n$ identity matrix and $\mathbf{0}$ is the $n \times n$ zero matrix.

False proof: The characteristic polynomial is $p(\lambda) = \det(A - \lambda\mathbf{I})$. Setting $\lambda = A$, we get $p(A) = \det(A - A\mathbf{I}) = \det(A - A) = \det(\mathbf{0}) = 0$. This "proof" does not make any sense. In particular, $p(A)$ is an $n \times n$ matrix, but in this false proof we obtained $p(A) = 0$ where 0 is a number.

Correct proof: Recall that the adjugate of M, denoted by adj M, is the transpose of the matrix of cofactors. Moreover, using eq. (29) of the class handout entitled *Determinant and the Adjugate*, it follows that $M(\text{adj } M) = \mathbf{I}\det M$ for any matrix M. In particular, setting $M = A - \lambda\mathbf{I}$, it follows that

$$(A - \lambda\mathbf{I})\,\text{adj}(A - \lambda\mathbf{I}) = p(\lambda)\mathbf{I}\,, \tag{16}$$

where $p(\lambda) = \det(A - \lambda\mathbf{I})$. Since $p(\lambda)$ is an nth-order polynomial, it then follows from eq. (16) that $\text{adj}(A - \lambda\mathbf{I})$ is a matrix polynomial of order $n - 1$. Thus, we can write:

$$\text{adj}(A - \lambda\mathbf{I}) = B_0 + B_1\lambda + B_2\lambda^2 + \cdots + B_{n-1}\lambda^{n-1}\,,$$

where $B_0,\ B_1, \ldots, B_{n-1}$ are $n \times n$ matrices (whose explicit forms are not required in these notes). Inserting the above result into eq. (16) and using eq. (4), one obtains:

$$(A - \lambda\mathbf{I})(B_0 + B_1\lambda + B_2\lambda^2 + \cdots + B_{n-1}\lambda^{n-1}) = (-1)^n\left[\lambda^n + c_1\lambda^{n-1} + \cdots + c_{n-1}\lambda + c_n\right]\mathbf{I}\,. \tag{17}$$

Eq. (17) is true for any value of λ. Consequently, the coefficient of λ^k on the left-hand side of eq. (17) must equal the coefficient of λ^k on the right-hand side of eq. (17), for $k = 0, 1, 2, \ldots, n$. As a result, the following $n + 1$ equations must be satisfied,

$$AB_0 = (-1)^n c_n\mathbf{I}\,, \tag{18}$$

$$-B_{k-1} + AB_k = (-1)^n c_{n-k}\mathbf{I}\,, \qquad k = 1, 2, \ldots, n-1\,, \tag{19}$$

$$-B_{n-1} = (-1)^n\mathbf{I}\,. \tag{20}$$

Using eqs. (18)–(20), we can evaluate the matrix polynomial $p(A)$ as a telescoping sum,

$$\begin{aligned}
p(A) &= (-1)^n\left[A^n + c_1A^{n-1} + c_2A^{n-2} + \cdots + c_{n-1}A + c_n\mathbf{I}\right] \\
&= AB_0 + (-B_0 + B_1A)A + (-B_1 + B_2)A^2 + \cdots + (-B_{n-2} + B_{n-1}A)A^{n-1} - B_{n-1}A^n \\
&= \mathbf{0}\,,
\end{aligned}$$

which completes the proof of the Cayley-Hamilton theorem.

[4] In the expression for $p(\lambda)$, we interpret c_n to mean $c_n\lambda^0$. Thus, when evaluating $p(A)$, the coefficient c_n multiplies $A^0 \equiv \mathbf{I}$.

It is instructive to illustrate the Cayley-Hamilton theorem for 2×2 matrices. In this case,

$$p(\lambda) = \lambda^2 - \lambda \mathrm{Tr}\ A + \det\ A\,.$$

Hence, by the Cayley-Hamilton theorem,

$$p(A) = A^2 - A\,\mathrm{Tr}\ A + \mathbf{I}\det\ A = 0\,.$$

Let us take the trace of this equation. Since Tr $\mathbf{I} = 2$ for the 2×2 identity matrix,

$$\mathrm{Tr}(A^2) - (\mathrm{Tr}\ A)^2 + 2\det\ A = 0\,.$$

It follows that

$$\boxed{\det\ A = \tfrac{1}{2}\left[(\mathrm{Tr}\ A)^2 - \mathrm{Tr}(A^2)\right]\,, \qquad \text{for any}\ \ 2\times 2 \text{ matrix } A\,.}$$

One can easily verify the validity of this formula for any 2×2 matrix.

One final application of the Cayley-Hamilton theorem is noteworthy. This theorem provides a new way to evaluate the inverse of a matrix. Assuming that $\det A \neq 0$, then one can multiply both sides of eq. (15) by A^{-1} and solve for A^{-1}. Using $c_n = (-1)^n \det\ A$ [cf. eq. (5)], the end result is

$$A^{-1} = \frac{(-1)^{n+1}}{\det A}\left[A^{n-1} + c_1 A^{n-2} + \ldots + c_{n-2}A + c_{n-1}\mathbf{I}\right]. \tag{21}$$

As a check, one can evaluate eq. (21) in the case of $n = 2$. After employing $c_1 = -\mathrm{Tr}\ A$ [cf. eq. (8)], it follows that for $n = 2$ and $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$,

$$A^{-1} = -\frac{1}{\det A}\left[A - \mathbf{I}\,\mathrm{Tr}\ A\right] = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\,, \tag{22}$$

as expected. To evaluate the inverse of an $n \times n$ matrix for $n > 2$, one must evaluate the coefficients c_2, $c_3, \ldots c_{n-1}$. These coefficients can be evaluated in terms of traces of powers of the matrix A, as shown in Appendix A below.

Appendix A: Identifying the coefficients of the characteristic polynomial of A in terms of traces of powers of A

The characteristic polynomial of an $n \times n$ matrix A is given by:

$$p(\lambda) = \det(A - \lambda\mathbf{I}) = (-1)^n\left[\lambda^n + c_1\lambda^{n-1} + c_2\lambda^{n-2} + \cdots + c_{n-1}\lambda + c_n\right]\,.$$

In Section 1, we identified,

$$c_1 = -\mathrm{Tr}\ A\,, \qquad c_n = (-1)^n \det\ A\,. \tag{23}$$

One can also derive expressions for $c_2, c_3, \ldots, c_{n-1}$ in terms of traces of powers of A. Although we will not need these expressions in Physics 116A, they are useful for some applications. Hence, you may may wish to take note of these results for future reference.

In this Appendix, I will exhibit the relevant results without proofs (which can be found in the references at the end of these notes). Let us introduce the notation:

$$t_k \equiv \mathrm{Tr}(A^k)\,. \tag{24}$$

The t_k can be expressed in terms of the eigenvalues of A. In particular, note that if $A\vec{\boldsymbol{v}} = \lambda\vec{\boldsymbol{v}}$, then it follows that $A^k\,\vec{\boldsymbol{v}} = \lambda^k\,\vec{\boldsymbol{v}}$. As a result, eqs. (13a) and (24) yield,

$$t_k = \lambda_1^k + \lambda_2^k + \ldots + \lambda_n^k\,. \tag{25}$$

Likewise, the c_k can be expressed in terms of the eigenvalues of A as shown explicitly in eq. (14). Then, the following set of recursive equations can be derived,

$$t_1 + c_1 = 0 \quad \text{and} \quad t_k + c_1 t_{k-1} + \cdots + c_{k-1}t_1 + kc_k = 0\,, \qquad k = 2,3,\ldots,n\,. \tag{26}$$

These equations are called *Newton's identities*. Two different proofs of these identities can be found in Refs. [1,2].

The equations exhibited in eq. (26) are called recursive, since one can solve for the c_k in terms of the traces $t_1, t_2, \ldots, t_k$ iteratively by starting with $c_1 = -t_1$, and then proceeding step by step by solving the equations with $k = 2,3,\ldots,n$ in successive order. This recursive procedure yields expressions for the c_k in terms of traces of powers of A,

$$c_1 = -t_1\,, \tag{27}$$

$$c_2 = \tfrac{1}{2}(t_1^2 - t_2)\,, \tag{28}$$

$$c_3 = -\tfrac{1}{6}t_1^3 + \tfrac{1}{2}t_1t_2 - \tfrac{1}{3}t_3\,, \tag{29}$$

$$c_4 = \tfrac{1}{24}t_1^4 - \tfrac{1}{4}t_1^2t_2 + \tfrac{1}{3}t_1t_3 + \tfrac{1}{8}t_2^2 - \tfrac{1}{4}t_4\,, \tag{30}$$

and so on for $k > 4$. The results above can be summarized by the following equation (see, e.g., Ref. [3]),

$$c_k = -\frac{t_k}{k} + \frac{1}{2!}\sum_{\substack{i=1\\ i+j=k}}^{k-1}\sum_{j=1}^{k-1}\frac{t_it_j}{ij} - \frac{1}{3!}\sum_{\substack{i=1\\ i+j+\ell=k}}^{k-2}\sum_{j=1}^{k-2}\sum_{\ell=1}^{k-2}\frac{t_it_jt_\ell}{ij\ell} + \cdots + \frac{(-1)^k t_1^k}{k!}\,, \qquad k = 1,2,\ldots,n\,. \tag{31}$$

Note that by using $c_n = (-1)^n \det A$, one obtains a general expression for the determinant in terms of traces of powers of A [cf. eq. (24)],

$$\det A = (-1)^n\left[-\frac{t_n}{n} + \frac{1}{2!}\sum_{\substack{i=1\\ i+j=n}}^{n-1}\sum_{j=1}^{n-1}\frac{t_it_j}{ij} - \frac{1}{3!}\sum_{\substack{i=1\\ i+j+k=n}}^{n-2}\sum_{j=1}^{n-2}\sum_{k=1}^{n-2}\frac{t_it_jt_k}{ijk} + \cdots + \frac{(-1)^n t_1^n}{n!}\right]. \tag{32}$$

For example, using eq. (32), one obtains the following explicit expressions for $n = 2$ and $n = 3$, respectively,

$$\det A = \tfrac{1}{2}\left[(\mathrm{Tr}\,A)^2 - \mathrm{Tr}(A^2)\right]\,, \qquad \text{for any } 2\times 2 \text{ matrix}\,,$$

$$\det A = \tfrac{1}{6}\left[(\mathrm{Tr}\,A)^3 - 3(\mathrm{Tr}\,A)\,\mathrm{Tr}(A^2) + 2\,\mathrm{Tr}(A^3)\right]\,, \qquad \text{for any } 3\times 3 \text{ matrix}\,.$$

BONUS MATERIAL

One can derive another closed-form expression for the c_k. To see how to do this, let us write out the Newton identities explicitly. Eq. (26) for $k = 1, 2, \ldots, n$ yields the following n equations,

$$
\begin{aligned}
c_1 &= -t_1 \,, \\
t_1 c_1 + 2c_2 &= -t_2 \,, \\
t_2 c_1 + t_1 c_2 + 3c_3 &= -t_3 \,, \\
\vdots \qquad &\qquad \vdots \\
t_{k-1} c_1 + t_{k-2} c_2 + \cdots + t_1 c_{k-1} + k c_k &= -t_k \,, \\
\vdots \qquad &\qquad \vdots \\
t_{n-1} c_1 + t_{n-2} c_2 + \cdots + t_1 c_{n-1} + n c_n &= -t_n \,,
\end{aligned}
$$

where $t_k \equiv \mathrm{Tr}(A^k)$. Consider the first k equations above (for any value of $k = 1, 2, \ldots, n$). This is a system of linear equations for $c_1,\, c_2, \ldots, c_k$, which can be written in matrix form:

$$
\begin{pmatrix}
1 & 0 & 0 & \cdots & 0 & 0 \\
t_1 & 2 & 0 & \cdots & 0 & 0 \\
t_2 & t_1 & 3 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
t_{k-2} & t_{k-3} & t_{k-4} & \cdots & k-1 & 0 \\
t_{k-1} & t_{k-2} & t_{k-3} & \cdots & t_1 & k
\end{pmatrix}
\begin{pmatrix}
c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_{k-1} \\ c_k
\end{pmatrix}
=
\begin{pmatrix}
-t_1 \\ -t_2 \\ -t_3 \\ \vdots \\ -t_{k-1} \\ -t_k
\end{pmatrix} .
$$

Applying Cramer's rule, we can solve for c_k in terms of $t_1,\, t_2, \ldots, t_k$ [3]:

$$
c_k = \frac{
\begin{vmatrix}
1 & 0 & 0 & \cdots & 0 & -t_1 \\
t_1 & 2 & 0 & \cdots & 0 & -t_2 \\
t_2 & t_1 & 3 & \cdots & 0 & -t_3 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
t_{k-2} & t_{k-3} & t_{k-4} & \cdots & k-1 & -t_{k-1} \\
t_{k-1} & t_{k-2} & t_{k-3} & \cdots & t_1 & -t_k
\end{vmatrix}
}{
\begin{vmatrix}
1 & 0 & 0 & \cdots & 0 & 0 \\
t_1 & 2 & 0 & \cdots & 0 & 0 \\
t_2 & t_1 & 3 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
t_{k-2} & t_{k-3} & t_{k-4} & \cdots & k-1 & 0 \\
t_{k-1} & t_{k-2} & t_{k-3} & \cdots & t_1 & k
\end{vmatrix}
} . \tag{33}
$$

Note that the denominator is the determinant of a lower triangular matrix, which is equal to the product of its diagonal elements (i.e., it is equal to $k!$).

From eq. (33), it then follows that,

$$c_k = \frac{1}{k!}\begin{vmatrix} 1 & 0 & 0 & \cdots & 0 & -t_1 \\ t_1 & 2 & 0 & \cdots & 0 & -t_2 \\ t_2 & t_1 & 3 & \cdots & 0 & -t_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ t_{k-2} & t_{k-3} & t_{k-4} & \cdots & k-1 & -t_{k-1} \\ t_{k-1} & t_{k-2} & t_{k-3} & \cdots & t_1 & -t_k \end{vmatrix}.$$

It is convenient to multiply the kth column of the above matrix by -1, and then move the kth column over to the first column (which requires a series of $k-1$ interchanges of adjacent columns). These operations multiply the determinant by (-1) and $(-1)^{k-1}$ respectively, leading to an overall sign change of $(-1)^k$. Hence, our final result is:[5]

$$c_k = \frac{(-1)^k}{k!}\begin{vmatrix} t_1 & 1 & 0 & 0 & \cdots & 0 \\ t_2 & t_1 & 2 & 0 & \cdots & 0 \\ t_3 & t_2 & t_1 & 3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ t_{k-1} & t_{k-2} & t_{k-3} & t_{k-4} & \cdots & k-1 \\ t_k & t_{k-1} & t_{k-2} & t_{k-3} & \cdots & t_1 \end{vmatrix}, \qquad k = 1, 2, \ldots, n\,, \tag{34}$$

which is equivalent to eq. (31).

One can test this formula by evaluating the first three cases $k = 1, 2, 3$:

$$c_1 = -t_1\,, \qquad\qquad c_2 = \frac{1}{2!}\begin{vmatrix} t_1 & 1 \\ t_2 & t_1 \end{vmatrix} = \tfrac{1}{2}(t_1^2 - t_2)\,,$$

$$c_3 = -\frac{1}{3!}\begin{vmatrix} t_1 & 1 & 0 \\ t_2 & t_1 & 2 \\ t_3 & t_2 & t_1 \end{vmatrix} = \tfrac{1}{6}\left[-t_1^3 + 3t_1t_2 - 2t_3\right]\,,$$

which coincide with the previously stated results [cf. eqs. (27)–(29)]. Finally, setting $k = n$ in eq. (34) yields $c_n = (-1)^n \det A$, which provides a formula for the determinant of the $n \times n$ matrix A in terms of traces of powers of A [cf. eq. (24)],

$$\det\, A = \frac{1}{n!}\begin{vmatrix} t_1 & 1 & 0 & 0 & \cdots & 0 \\ t_2 & t_1 & 2 & 0 & \cdots & 0 \\ t_3 & t_2 & t_1 & 3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ t_{n-1} & t_{n-2} & t_{n-3} & t_{n-4} & \cdots & n-1 \\ t_n & t_{n-1} & t_{n-2} & t_{n-3} & \cdots & t_1 \end{vmatrix}, \tag{35}$$

which is equivalent to eq. (32). Indeed, one can check that our previous results for the determinants of a 2×2 matrix and a 3×3 matrix, exhibited below eq. (32), are recovered.

[5]Eq. (34) is derived in section 4.1 on p. 20 of Ref. [4]. However, the determinantal expression given in Ref. [4] for $\sigma_k \equiv (-1)^k c_k$ contains a typographical error—the diagonal series of integers, $1, 1, 1, \ldots, 1$, appearing just above the main diagonal of σ_k should be replaced by $1, 2, 3, \ldots, k-1$.

Appendix B: Identifying the coefficients of the characteristic polynomial of A in terms of its matrix elements

Given an $n \times n$ matrix $A = [a_{ij}]$, eqs. (5) and (8) provide expressions for c_1 and c_n in terms of the trace and determinant of A, respectively, which again are repeated here,

$$c_1 = -\mathrm{Tr}\, A\,, \qquad\qquad c_n = (-1)^n \det A\,. \tag{36}$$

Related expressions exist for the c_k ($k = 2, 3, \ldots, n-1$). In particular, Ref. [5] shows that $(-1)^k c_k$ is equal to the sum of the determinants of the $k \times k$ principal submatrices[6] of A in all possible ways [there are $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ terms in the sum]. That is,

$$c_k = (-1)^k \sum_{\substack{i_1=1 \\ i_1<i_2<\cdots<i_k}}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_k=1}^{n} \det \begin{pmatrix} a_{i_1 i_1} & a_{i_1 i_2} & \cdots & a_{i_1 i_k} \\ a_{i_2 i_1} & a_{i_2 i_2} & \cdots & a_{i_2 i_k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_k i_1} & a_{i_k i_2} & \cdots & a_{i_k i_k} \end{pmatrix}, \qquad \text{for} \quad k = 1, 2, \ldots, n\,. \tag{37}$$

It is easy to verify that the $k = 1$ and $k = n$ cases of eq. (37) reduce to the results given in eq. (36). As a more nontrivial example, note that for $k = 2$ we obtain,

$$c_2 = \sum_{\substack{i=1 \\ i<j}}^{n} \sum_{j=1}^{n} \det \begin{pmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{pmatrix} = \sum_{\substack{i=1 \\ i<j}}^{n} \sum_{j=1}^{n} (a_{ii}a_{jj} - a_{ij}a_{ji})\,. \tag{38}$$

Comparing with eq. (9), we can identify $c_2' = -\sum_{1\leq i<j\leq n} a_{ij}a_{ji}$. Indeed a careful analysis of $\det(A - \lambda\mathbf{I})$ will yield the quoted result, where c_2' is defined in eq. (6).

Finally, it is straightforward but tedious to check that eq. (38) is equivalent to

$$c_2 = \tfrac{1}{2}\left[(\mathrm{Tr}\, A)^2 - \mathrm{Tr}(A^2)\right],$$

which was previously obtained in eq. (28).

REFERENCES

1. Ján Mináč, *Newton's Identities Once Again!*, *The American Mathematical Monthly* **110**, 232-234 (2003).

2. Dan Kalman, *A Matrix Proof of Newton's Identities*, *Mathematics Magazine* **73**, 313–315 (2000).

3. H.K. Krishnapriyan, *On Evaluating the Characteristic Polynomial through Symmetric Functions*, *J. Chem. Inf. Comput. Sci.* **35**, 196–198 (1995).

4. V.V. Prasolov, *Problems and Theorems in Linear Algebra* (American Mathematical Society, Providence, RI, 1994).

5. See pp. 494–495 of Carl D. Meyer, *Matrix Analysis and Applied Linear Algebra* (SIAM, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2000).

[6] A $k \times k$ principal submatrix of the $n \times n$ matrix A is obtained by removing $n - k$ rows and columns of A such that the set of row indices that remain is the same as the set of column indices that remain.