Facts About Eigenvalues
By Dr David Butler

Definitions

Suppose A is an $n \times n$ matrix.

- An *eigenvalue* of A is a number λ such that $A\mathbf{v} = \lambda\mathbf{v}$ for some nonzero vector $\mathbf{v}$.
- An *eigenvector* of A is a nonzero vector $\mathbf{v}$ such that $A\mathbf{v} = \lambda\mathbf{v}$ for some number λ.

Terminology

Let A be an $n \times n$ matrix.

- The determinant $|\lambda I - A|$ (for unknown λ) is called the *characteristic polynomial* of A.
  (The zeros of this polynomial are the eigenvalues of A.)
- The equation $|\lambda I - A| = 0$ is called the *characteristic equation* of A.
  (The solutions of this equation are the eigenvalues of A.)
- If λ is an eigenvalue of A, then the subspace $E_\lambda = \{\mathbf{v} \mid A\mathbf{v} = \lambda\mathbf{v}\}$ is called the *eigenspace of A associated with λ.*
  (This subspace contains all the eigenvectors with eigenvalue λ, and also the zero vector.)
- An eigenvalue λ^* of A is said to have *multiplicity* m if, when the characteristic polynomial is factorised into linear factors, the factor $(\lambda - \lambda^*)$ appears m times.

Theorems

Let A be an $n \times n$ matrix.

- The matrix A has n eigenvalues (including each according to its multiplicity).
- The sum of the n eigenvalues of A is the same as the trace of A (that is, the sum of the diagonal elements of A).
- The product of the n eigenvalues of A is the same as the determinant of A.
- If λ is an eigenvalue of A, then the dimension of E_λ is at most the multiplicity of λ.
- A set of eigenvectors of A, each corresponding to a different eigenvalue of A, is a linearly independent set.
- If $\lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0$ is the characteristic polynomial of A,
  then $c_{n-1} = -\text{trace}(A)$ and $c_0 = (-1)^n|A|$.
- If $\lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0$ is the characteristic polynomial of A,
  then $A^n + c_{n-1}A^{n-1} + \cdots + c_1A + c_0I = O$. (The Cayley-Hamilton Theorem.)

Examples of Problems using Eigenvalues

Problem:
If λ is an eigenvalue of the matrix A, prove that λ^2 is an eigenvalue of A^2.

Solution:
Since λ is an eigenvalue of A, $A\mathbf{v} = \lambda\mathbf{v}$ for some $\mathbf{v} \neq \mathbf{0}$.
Multiplying both sides by A gives

$$\begin{aligned} A(A\mathbf{v}) &= A(\lambda\mathbf{v}) \\ A^2\mathbf{v} &= \lambda A\mathbf{v} \\ &= \lambda\lambda\mathbf{v} \\ &= \lambda^2\mathbf{v} \end{aligned}$$

Therefore λ^2 is an eigenvalue of A^2. ■

Problem:
Prove that the $n \times n$ matrix A and its transpose A^T have the same eigenvalues.

Solution:
Consider the characteristic polynomial of A^T: $|\lambda I - A^T| = |(\lambda I - A)^T| = |\lambda I - A|$ (since a matrix and its transpose have the same determinant). This result is the characteristic polynomial of A, so A^T and A have the same characteristic polynomial, and hence they have the same eigenvalues. ■

Problem:
The matrix A has $(1,2,1)^T$ and $(1,1,0)^T$ as eigenvectors, both with eigenvalue 7, and its trace is 2. Find the determinant of A.

Solution:
The matrix A is a 3×3 matrix, so it has 3 eigenvalues in total. The eigenspace E_7 contains the vectors $(1,2,1)^T$ and $(1,1,0)^T$, which are linearly independent. So E_7 must have dimension at least 2, which implies that the eigenvalue 7 has multiplicity at least 2.

Let the other eigenvalue be λ, then from the trace $\lambda+7+7 = 2$, so $\lambda = -12$. So the three eigenvalues are 7, 7 and -12. Hence, the determinant of A is $7 \times 7 \times -12 = -588$. ■

The sum and product of eigenvalues

Theorem: If A is an $n \times n$ matrix, then the sum of the n eigenvalues of A is the trace of A and the product of the n eigenvalues is the determinant of A.

Proof:

$$\text{Write } A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}.$$

Also let the n eigenvalues of A be $\lambda_1, \dots, \lambda_n$. Finally, denote the characteristic polynomial of A by $p(\lambda) = |\lambda I - A| = \lambda^n + c_{n-1}\lambda^{n-1} + \dots + c_1\lambda + c_0$. Note that since the eigenvalues of A are the zeros of $p(\lambda)$, this implies that $p(\lambda)$ can be factorised as $p(\lambda) = (\lambda - \lambda_1)\dots(\lambda - \lambda_n)$.

Consider the constant term of $p(\lambda)$, c_0. The constant term of $p(\lambda)$ is given by $p(0)$, which can be calculated in two ways:

Firstly, $p(0) = (0 - \lambda_1)\dots(0 - \lambda_n) = (-1)^n\lambda_1\dots\lambda_n$. Secondly, $p(0) = |0I - A| = |-A| = (-1)^n|A|$.

Therefore $c_0 = (-1)^n\lambda_1\dots\lambda_n = (-1)^n|A|$, and so $\lambda_1\dots\lambda_n = |A|$. That is, the product of the n eigenvalues of A is the determinant of A.

Consider the coefficient of λ^{n-1}, c_{n-1}. This is also calculated in two ways.

Firstly, it can be calculated by expanding $p(\lambda) = (\lambda - \lambda_1)\dots(\lambda - \lambda_n)$. In order to get the λ^{n-1} term, the λ must be chosen from $n-1$ of the factors, and the constant from the other. Hence, the λ^{n-1} term will be $-\lambda_1\lambda^{n-1} - \dots - \lambda\lambda^{n-1} = -(\lambda_1 + \dots + \lambda_n)\lambda^{n-1}$. Thus $c_{n-1} = -(\lambda_1 + \dots + \lambda_n)$.

Secondly, this coefficient can be calculated by expanding $|\lambda I - A|$:

$$|\lambda I - A| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \dots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \dots & \lambda - a_{nn} \end{vmatrix}$$

One way of calculating determinants is to multiply the elements in positions $1j_1$, $2j_2$, $\dots$, nj_n, for each possible permutation $j_1\dots j_n$ of $1\dots n$. If the permutation is odd, then the product is also multiplied by -1. Then all of these $n!$ products are added together to produce the determinant. One of these products is $(\lambda - a_{11})\dots(\lambda - a_{nn})$. Every other possible product can contain at most $n-2$ elements on the diagonal of the matrix, and so will contain at most $n-2$ λ's. Hence, when all of these other products are expanded, they will produce a polynomial in λ of degree at most $n-2$. Denote this polynomial by $q(\lambda)$.

Hence, $p(\lambda) = (\lambda - a_{11})\dots(\lambda - a_{nn}) + q(\lambda)$. Since $q(\lambda)$ has degree at most $n-2$, it has no λ^{n-1} term, and so the λ^{n-1} term of $p(\lambda)$ must be the λ^{n-1} term from $(\lambda - a_{11})\dots(\lambda - a_{nn})$. However, the argument above for $(\lambda - \lambda_1)\dots(\lambda - \lambda_n)$ shows that this term must be $-(a_{11} + \dots + a_{nn})\lambda^{n-1}$.

Therefore $c_{n-1} = -(\lambda_1 + \dots + \lambda_n) = -(a_{11} + \dots + a_{nn})$, and so $\lambda_1 + \dots + \lambda_n = a_{11} + \dots + a_{nn}$. That is, the sum of the n eigenvalues of A is the trace of A. $\blacksquare$

The Cayley-Hamilton Theorem

Theorem:
Let A be an $n \times n$ matrix. If $\lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0$ is the characteristic polynomial of A, then $A^n + c_{n-1}A^{n-1} + \cdots + c_1 A + c_0 I = O$.

Proof:
Consider the matrix $\lambda I - A$. If this matrix is multiplied by its adjoint matrix, the result will be its determinant multiplied by the identity matrix. That is,

$$(\lambda I - A)\mathrm{adj}(\lambda I - A) = |\lambda I - A| I \tag{1}$$

Consider the matrix $\mathrm{adj}(\lambda I - A)$. Each entry of this matrix is either the positive or the negative of the determinant of a smaller matrix produced by deleting one row and one column of $\lambda I - A$. The determinant of such a matrix is a polynomial in λ of degree at most $n-1$ (since removing one row and one column is guaranteed to remove at least one λ).

Let the polynomial in position ij of $\mathrm{adj}(\lambda I - A)$ be $b_{ij0} + b_{ij1}\lambda + \cdots + b_{ij(n-1)}\lambda^{n-1}$. Then

$$\begin{aligned}
\mathrm{adj}(\lambda I - A) &= \begin{bmatrix} b_{110} + b_{111}\lambda + \cdots + b_{11(n-1)}\lambda^{n-1} & \dots & b_{1n0} + b_{1n1}\lambda + \cdots + b_{1n(n-1)}\lambda^{n-1} \\ \vdots & \ddots & \vdots \\ b_{n10} + b_{n11}\lambda + \cdots + b_{n1(n-1)}\lambda^{n-1} & \dots & b_{nn0} + b_{nn1}\lambda + \cdots + b_{nn(n-1)}\lambda^{n-1} \end{bmatrix} \\
&= \begin{bmatrix} b_{110} & \dots & b_{1n0} \\ \vdots & \ddots & \vdots \\ b_{n10} & \dots & b_{nn0} \end{bmatrix} + \begin{bmatrix} b_{111}\lambda & \dots & b_{1n1}\lambda \\ \vdots & \ddots & \vdots \\ b_{n11}\lambda & \dots & b_{nn1}\lambda \end{bmatrix} + \cdots + \begin{bmatrix} b_{11(n-1)}\lambda^{n-1} & \dots & b_{1n(n-1)}\lambda^{n-1} \\ \vdots & \ddots & \vdots \\ b_{n1(n-1)}\lambda^{n-1} & \dots & b_{nn(n-1)}\lambda^{n-1} \end{bmatrix} \\
&= \begin{bmatrix} b_{110} & \dots & b_{1n0} \\ \vdots & \ddots & \vdots \\ b_{n10} & \dots & b_{nn0} \end{bmatrix} + \lambda \begin{bmatrix} b_{111} & \dots & b_{1n1} \\ \vdots & \ddots & \vdots \\ b_{n11} & \dots & b_{nn1} \end{bmatrix} + \cdots + \lambda^{n-1} \begin{bmatrix} b_{11(n-1)} & \dots & b_{1n(n-1)} \\ \vdots & \ddots & \vdots \\ b_{n1(n-1)} & \dots & b_{nn(n-1)} \end{bmatrix}
\end{aligned}$$

Denote the matrices appearing in the above expression by B_0, B_1, $\dots$, B_{n-1}, respectively so that

$$\begin{aligned}
\mathrm{adj}(\lambda I - A) &= B_0 + \lambda B_1 + \cdots + \lambda^{n-1} B_{n-1} \\
\text{Then } (\lambda I - A)\mathrm{adj}(\lambda I - A) &= (\lambda I - A)(B_0 + \lambda B_1 + \cdots + \lambda^{n-1} B_{n-1} \\
&= \lambda B_0 + \lambda^2 B_1 + \cdots + \lambda^n B_{n-1} \\
&\quad - AB_0 - \lambda A B_1 - \cdots - \lambda^{n-1} A B_{n-1} \\
&= -AB_0 + \lambda(B_0 - AB_1) + \cdots + \lambda^{n-1}(B_{n-2} - AB_{n-1}) + \lambda^n B_{n-1}
\end{aligned}$$

Next consider $|\lambda I - A| I$. This is the characteristic polynomial of A, multiplied by I. That is,

$$\begin{aligned}
|\lambda I - A| I &= (\lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0) I \\
&= \lambda^n I + c_{n-1}\lambda^{n-1} I + \cdots + c_1 \lambda I + c_0 I \\
&= c_0 I + \lambda(c_1 I) + \cdots + \lambda(c_{n-1} I) + \lambda^n I
\end{aligned}$$

Substituting these two expressions into the equation $(\lambda I - A)\mathrm{adj}(\lambda I - A) = |\lambda I - A| I$ gives

$$\begin{aligned}
& -AB_0 + \lambda(B_0 - AB_1) + \ldots + \lambda^{n-1}(B_{n-2} - AB_{n-1}) + \lambda^n B_{n-1} \\
= \quad & c_0 I \quad + \lambda(c_1 I) \qquad\quad + \ldots + \lambda^{n-1}(c_{n-1} I) \qquad\qquad + \lambda^n I
\end{aligned}$$

If the two sides of this equation were evaluated separately, each would be an $n \times n$ matrix with each entry a polynomial in λ. Since these two resulting matrices are equal, the entries in each position are also equal. That is, for each choice of i and j, the polynomials in position ij of the two matrices are equal. Since these two polynomials are equal, the coefficients of the matching powers of λ must also be equal. That is, for each choice of i, j and k, the coefficient of λ^k in position ij of one matrix is equal to the coefficient of λ^k in position ij of the other matrix. Hence, when each matrix is rewritten as sum of coefficient matrices multiplied by powers of λ (as was done above for $\operatorname{adj}(\lambda I - A)$), then for every k, the matrix multiplied by λ^k in one expression must be the same as the matrix multiplied by λ^k in the other.

In other words, we can equate the matrix coeffiicients of the powers of λ in each expression. This results in the following equations:

$$\begin{aligned}
c_0 I &= \qquad\quad - AB_0 \\
c_1 I &= B_0 \quad\; - AB_1 \\
&\vdots \\
c_{n-1} I &= B_{n-2} - AB_{n-1} \\
I &= B_{n-1}
\end{aligned}$$

Now right-multiply each equation by successive powers of A (that is, the first is multiplied by I, the second is multiplied by A, the third is multiplied by A^2, and so on until the last is multiplied by A^n). This produces the following equations:

$$\begin{aligned}
c_0 I &= \qquad\qquad\quad - AB_0 \\
c_1 A &= AB_0 \qquad\quad - A^2 B_1 \\
&\vdots \\
c_{n-1} A^{n-1} &= A^{n-1} B_{n-2} - A^n B_{n-1} \\
A^n I &= A^n B_{n-1}
\end{aligned}$$

Adding all of these equations together produces:

$$\begin{aligned}
&c_0 I + c_1 A + \cdots + c_{n-1} A^{n-1} + A^{n-1} = -AB_0 + AB_0 - A^2 B_1 + \cdots + A^{n-1} B_{n-2} - A^n B_{n-1} + A^n B_{n-1} \\
&c_0 I + c_1 A + \cdots + c_{n-1} A^{n-1} + A^{n-1} = O \qquad \blacksquare
\end{aligned}$$

Polynomials acted upon matrices

Theorem:
Let A be an $n \times n$ matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$ (including multiplicity). Let $g(x) = a_0 + a_1 x + \cdots + a_k x^k$ be a polynomial, and let $g(A) = a_0 I + a_1 A + \cdots + a_k A^k$. Then the eigenvalues of $g(A)$ are $g(\lambda_1), \ldots, g(\lambda_n)$ (including multiplicity).

Proof:
We will begin by showing that the determinant of $g(A)$ is $g(\lambda_1) \ldots g(\lambda_n)$.

By the fundamental theorem of algebra, the polynomial $g(x)$ can be factorised into k linear factors over the complex numbers. Hence we can write $g(x) = a_k(x - c_1) \ldots (x - c_k)$ for some complex numbers $c_1, \ldots, c_k$. Now a matrix commutes with all its powers, and with the identity, so it is also possible to write $g(A)$ as $g(A) = a_k(A - c_1 I) \ldots (A - c_k I)$.

Also, denote the characteristic polynomial of A by $p(\lambda) = |\lambda I - A|$. Since the eigenvalues of A are $\lambda_1, \ldots, \lambda_n$, the characteristic polynomial can be factorised as $p(\lambda) = (\lambda - \lambda_1) \ldots (\lambda - \lambda_n)$.

Consider the determinant of $g(A)$:

$$
\begin{aligned}
|g(A)| &= |a_k(A - c_1 I) \ldots (A - c_k I)| \\
&= (a_k)^n |A - c_1 I| \ldots |A - c_k I| \\
&= (a_k)^n | - (c_1 I - A)| \ldots | - (c_k I - A)| \\
&= (a_k)^n (-1)^n |c_1 I - A| \ldots (-1)^n |c_k I - A| \\
&= (a_k)^n (-1)^{nk} |c_1 I - A| \ldots |c_k I - A|
\end{aligned}
$$

Now $|c_i I - A|$ is $|\lambda I - A|$ with λ replaced by c_i, that is, it is the characteristic polynomial of A evaluated at $\lambda = c_i$. Thus $|c_i I - A| = p(c_i) = (c_i - \lambda_1) \ldots (c_i - \lambda_n)$.

$$
\begin{aligned}
\text{So, } |g(A)| &= (a_k)^n (-1)^{nk} p(c_1) \ldots p(c_k) \\
&= (a_k)^n (-1)^{nk} \times (c_1 - \lambda_1) \ldots (c_1 - \lambda_n) \\
&\qquad\qquad\quad \times \ldots \\
&\qquad\qquad\quad \times (c_k - \lambda_1) \ldots (c_k - \lambda_n) \\
&= (a_k)^n \times (\lambda_1 - c_1) \ldots (\lambda_n - c_1) \\
&\qquad\quad \times \ldots \\
&\qquad\quad \times (\lambda_1 - c_k) \ldots (\lambda_n - c_k) \\
&= (a_k)^n \times (\lambda_1 - c_1) \ldots (\lambda_1 - c_k) \\
&\qquad\quad \times \ldots \\
&\qquad\quad \times (\lambda_n - c_1) \ldots (\lambda_n - c_k) \\
&= \quad a_k(\lambda_1 - c_1) \ldots (\lambda_1 - c_k) \\
&\quad \times \ldots \\
&\quad \times a_k(\lambda_n - c_1) \ldots (\lambda_n - c_k) \\
&= g(\lambda_1) \times \cdots \times g(\lambda_n)
\end{aligned}
$$

The above argument shows that if $g(x)$ is any polynomial, then $|g(A)| = g(\lambda_1) \ldots g(\lambda_n)$.

Now we will show that the eigenvalues of $g(A)$ are $g(\lambda_1)$, …, $g(\lambda_n)$.

Let a be any number and consider the polynomial $h(x) = a - g(x)$. Then $h(A) = aI - g(A)$, and the argument above shows that $|h(A)| = h(\lambda_1) \dots h(\lambda_n)$. Substituting the formulas for $h(x)$ and $h(A)$ into this equation gives that $|aI - g(A)| = (a - g(\lambda_1)) \dots (a - g(\lambda_n))$.

Since this is true for all possible a, it can be concluded that as polynomials, $|\lambda I - g(A)| = (\lambda - g(\lambda_1)) \dots (\lambda - g(\lambda_n))$. But $|\lambda I - g(A)|$ is the characteristic polynomial of $g(A)$, which has been fully factorised here, so this implies that the eigenvalues of $g(A)$ are $g(\lambda_1)$, …, $g(\lambda_n)$. ■

Some corollaries:
Let A be an $n \times n$ matrix with eigenvalues λ_1, …, λ_n. Then:

- $2A$ has eigenvalues $2\lambda_1$, …, $2\lambda_n$.
- A^2 has eigenvalues λ_1^2, …, λ_n^2.
- $A + 2I$ has eigenvalues $\lambda_1 + 2$, …, $\lambda_n + 2$.
- If $p(\lambda)$ is the characteristic polynomial of A, then $p(A)$ has eigenvalues 0, …, 0.

Similar matrices

Definition:
Two matrices A and B are called *similar* if there exists an invertible matrix X such that $A = X^{-1}BX$.

Theorem:
Suppose A and B are similar matrices. Then A and B have the same characteristic polynomial and hence the same eigenvalues.

Proof:
Consider the characteristic polynomial of A:

$$\begin{aligned}
|\lambda I - A| &= |\lambda I - X^{-1}BX| \\
&= |\lambda X^{-1}IX - X^{-1}BX| \\
&= |X^{-1}(\lambda I - B)X| \\
&= |X^{-1}||\lambda I - B||X| \\
&= \frac{1}{|X|}|\lambda I - B||X| \\
&= |\lambda I - B|
\end{aligned}$$

This is the characteristic polynomial of B, so A and B have the same characteristic polynomial. Hence A and B have the same eigenvalues $\blacksquare$.

Multiplicity and the dimension of an eigenspace

Theorem:
If λ^* is an eigenvalue of A, then the multiplicity of λ^* is at least the dimension of the eigenspace E_{λ^*}.

Proof:
Suppose the dimension of E_{λ^*} is m and let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ form a basis for E_{λ^*}.

It is possible to find $n-m$ other vectors $\mathbf{u}_{m+1}, \ldots, \mathbf{u}_n$ so that $\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{u}_{m+1}, \ldots, \mathbf{u}_n$ form a basis for $\mathbb{R}^n$. Let X be the $n \times n$ matrix with these n basis vectors as its columns. This matrix X is invertible since its columns are linearly independent.

Consider the matrix $B = X^{-1}AX$. This matrix is similar to A and so it has the same characteristic polynomial as A. In order to describe the entries of B, we will first investigate AX.

$$\begin{aligned}
AX &= A[\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_m \mid \mathbf{u}_{m+1} \mid \cdots \mid \mathbf{u}_n] \\
&= [A\mathbf{v}_1 \mid \cdots \mid A\mathbf{v}_m \mid A\mathbf{u}_{m+1} \mid \cdots \mid A\mathbf{u}_n] \\
&= [\lambda^*\mathbf{v}_1 \mid \cdots \mid \lambda^*\mathbf{v}_m \mid A\mathbf{u}_{m+1} \mid \cdots \mid A\mathbf{u}_n] \quad \text{(since } \mathbf{v}_1, \ldots, \mathbf{v}_m \text{ are eigenvalues of } A\text{)} \\
B &= X^{-1}AX \\
&= X^{-1}[\lambda^*\mathbf{v}_1 \mid \cdots \mid \lambda^*\mathbf{v}_m \mid A\mathbf{u}_{m+1} \mid \cdots \mid A\mathbf{u}_n] \\
&= [X^{-1}(\lambda^*\mathbf{v}_1) \mid \cdots \mid X^{-1}(\lambda^*\mathbf{v}_m) \mid X^{-1}A\mathbf{u}_{m+1} \mid \cdots \mid X^{-1}A\mathbf{u}_n] \\
&= [\lambda^*X^{-1}\mathbf{v}_1 \mid \cdots \mid \lambda^*X^{-1}\mathbf{v}_m \mid X^{-1}A\mathbf{u}_{m+1} \mid \cdots \mid X^{-1}A\mathbf{u}_n]
\end{aligned}$$

Now consider $X^{-1}\mathbf{v}_i$. This is X^{-1} multiplied by the i'th column of X, and so it is the i'th column of $X^{-1}X$. However $X^{-1}X = I$, so its i'th column is the i'th standard basis vector $\mathbf{e}_i$. Thus:

$$\begin{aligned}
B &= [\lambda^*X^{-1}\mathbf{v}_1 \mid \cdots \mid \lambda^*X^{-1}\mathbf{v}_m \mid X^{-1}A\mathbf{u}_{m+1} \mid \cdots \mid X^{-1}A\mathbf{u}_n] \\
&= [\lambda^*\mathbf{e}_1 \mid \cdots \mid \lambda^*\mathbf{e}_m \mid X^{-1}A\mathbf{u}_{m+1} \mid \cdots \mid X^{-1}A\mathbf{u}_n] \\
&= \begin{bmatrix}
\lambda^* & 0 & \ldots & 0 & b_{1(m+1)} & \ldots & b_{1n} \\
0 & \lambda^* & \ldots & 0 & b_{2(m+1)} & \ldots & b_{2n} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & \lambda^* & b_{m(m+1)} & \ldots & b_{mn} \\
0 & 0 & \ldots & 0 & b_{(m+1)(m+1)} & \ldots & b_{(m+1)n} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 0 & b_{n(m+1)} & \ldots & b_{nn}
\end{bmatrix}
\end{aligned}$$

$$\text{So, } \lambda I - B = \begin{bmatrix}
\lambda-\lambda^* & 0 & \ldots & 0 & -b_{1(m+1)} & \ldots & -b_{1n} \\
0 & \lambda-\lambda^* & \ldots & 0 & -b_{2(m+1)} & \ldots & -b_{2n} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & \lambda-\lambda^* & -b_{m(m+1)} & \ldots & -b_{mn} \\
0 & 0 & \ldots & 0 & \lambda-b_{(m+1)(m+1)} & \ldots & -b_{(m+1)n} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 0 & -b_{n(m+1)} & \ldots & \lambda-b_{nn}
\end{bmatrix}$$

If the determinant of this matrix $\lambda I - B$ is expanded progressively along the first m columns, it results in the following:

$$
\begin{aligned}
|\lambda I - B| &= (\lambda - \lambda^*)\dots(\lambda - \lambda^*) \begin{vmatrix} \lambda - b_{(m+1)(m+1)} & \dots & -b_{(m+1)n} \\ \vdots & \ddots & \vdots \\ -b_{n(m+1)} & \dots & \lambda - b_{nn} \end{vmatrix} \\
&= (\lambda - \lambda^*)^m \begin{vmatrix} \lambda - b_{(m+1)(m+1)} & \dots & -b_{(m+1)n} \\ \vdots & \ddots & \vdots \\ -b_{n(m+1)} & \dots & \lambda - b_{nn} \end{vmatrix}
\end{aligned}
$$

Hence, the factor $(\lambda - \lambda^*)$ appears at least m times in the characteristic polynomial of B (it may appear more times because of the part of the determinant that is as yet uncalculated). Since A and B have the same characteristic polynomial, the factor $(\lambda - \lambda^*)$ appears at least m times in the characteristic polynomial of A. That is, the multiplicity of the eigenvalue λ^* is at least m. ■

The product of two matrices

Theorem:
Let A be an $m \times n$ matrix and let B be an $n \times m$ matrix, with $n \geq m$. Then the n eigenvalues of BA are the m eigenvalues of AB with the extra eigenvalues being 0.

Proof:
Consider the $(m+n) \times (m+n)$ matrices:

$$M = \begin{pmatrix} O_{n\times n} & O_{n\times m} \\ A & AB \end{pmatrix}, \quad N = \begin{pmatrix} BA & O_{n\times m} \\ A & O_{m\times m} \end{pmatrix}$$

$$\text{Also let} \quad X = \begin{pmatrix} I_{n\times n} & B \\ O_{m\times n} & I_{m\times m} \end{pmatrix}$$

Then

$$\begin{aligned} \text{Then} \quad XM &= \begin{pmatrix} I & B \\ O & I \end{pmatrix} \begin{pmatrix} O & O \\ A & AB \end{pmatrix} \\ &= \begin{pmatrix} BA & BAB \\ A & AB \end{pmatrix} \\ \text{And} \quad NX &= \begin{pmatrix} BA & O \\ A & O \end{pmatrix} \begin{pmatrix} I & B \\ O & I \end{pmatrix} \\ &= \begin{pmatrix} BA & BAB \\ A & AB \end{pmatrix} \end{aligned}$$

So $XM = NX$. Now X is an upper triangular matrix with every entry on the diagonal equal to 1. Therefore it is invertible. Hence we can multiply both sides of this equation by X^{-1} to get $M = X^{-1}NX$. Thus M and N are similar and so have the same characteristic polynomial.

Consider the characteristic polynomial of each:

$$\begin{aligned} |\lambda I - M| &= \left| \lambda I - \begin{pmatrix} O_{n\times n} & O_{n\times m} \\ A & AB \end{pmatrix} \right| \\ &= \left| \begin{pmatrix} \lambda I_{n\times n} & O_{n\times m} \\ -A & \lambda I_{m\times m} - AB \end{pmatrix} \right| \\ &= |\lambda I_{n\times n}||\lambda I_{m\times m} - AB| \\ &= \lambda^n |\lambda I_{m\times m} - AB| \end{aligned}$$

$$\begin{aligned} |\lambda I - N| &= \left| \lambda I - \begin{pmatrix} BA & O_{n\times m} \\ A & O_{m\times m} \end{pmatrix} \right| \\ &= \left| \begin{pmatrix} \lambda I_{n\times n} - AB & O_{n\times m} \\ -A & \lambda I_{m\times m} \end{pmatrix} \right| \\ &= |\lambda I_{n\times n} - BA||\lambda I_{m\times m}| \\ &= \lambda^m |\lambda I_{m\times m} - BA| \end{aligned}$$

Since M and N have the same characteristic polynomial,

$$\begin{aligned} |\lambda I - M| &= |\lambda I - N| \\ \lambda^n |\lambda I_{m\times m} - AB| &= \lambda^m |\lambda I_{m\times m} - BA| \\ \lambda^{n-m} |\lambda I_{m\times m} - AB| &= |\lambda I_{m\times m} - BA| \end{aligned}$$

So the characteristic polynomial of BA is the same as the characteristic polynomial of AB, but multiplied by λ^{n-m}. Hence BA has all of the eigenvalues of AB, but with $n - m$ extra zeros. ■

A proof in finite geometry with a surprising use of eigenvalues

Preliminaries:

- A **finite projective plane** is a collection of finitely many points and finitely many lines such that
  - Every two points are contained in precisely one line.
  - Every two lines share precisely one point.
  - There are at least three points on every line.
  - Not all the points are on the same line.
- For a finite projective plane, there is a number q called the **order** such that:
  - There are $q+1$ points on every line.
  - There are $q+1$ lines through every point.
  - There are q^2+q+1 points in total.
  - There are q^2+q+1 lines in total.
- A **polarity** of a finite projective plane is a one-to-one map σ which maps points to lines and lines to points, so that if the point P is on the line ℓ, then the point $\sigma(\ell)$ is on the line $\sigma(P)$, and also for any point or line X, $\sigma(\sigma(X)) = X$.
- An **absolute point** with respect to a polarity σ of a projective plane is a point P such that P is on the line $\sigma(P)$.

Theorem: A polarity of a finite projective plane plane must have an absolute point.

Proof: Let σ be a polarity of a finite projective plane of order q. Denote the points in the projective plane by P_1, P_2, $\ldots$, P_{q^2+q+1}, and denote the line $\sigma(P_i)$ by ℓ_i for each $i = 1, \ldots, q^2+q+1$. Note that since σ is a polarity, then $\sigma(\ell_i) = \sigma(\sigma(P_i)) = P_i$ for any i.

Create a $(q^2+q+1) \times (q^2+q+1)$ matrix A as follows: if the point P_i is on the line ℓ_j, then put a 1 in entry ij of the matrix A, otherwise, put a 0 in position ij. The matrix A is called an incidence matrix of the projective plane.

Since σ is a polarity, if P_i is on ℓ_j, then $\sigma(P_i)$ is on $\sigma(\ell_j) = \sigma(\sigma(P_j)) = P_j$. Hence if there is a 1 in position ij, then there is also a 1 in position ji. Thus the matrix A is symmetric and $A^T = A$.

Now an abolute point is a point P_i such that P_i is on $\sigma(P_i)$. That is, an absolute point is a point such that P_i is on ℓ_i. Hence, if σ has an absolute point, then there is a 1 on the diagonal of A. Therefore, the number of absolute points of σ is the sum of the diagonal elements of A. That is, it is the trace of A.

To find the trace of A, we may instead find the sum of the eigenvalues of A.

Firstly note that every row of A contains precisely $q+1$ entries that are 1 since each point lies on $q+1$ lines. Hence the rows of A all add to $q+1$. Therefore when A is multiplied by $(1,1,\ldots,1)^T$, the result is $(q+1,\ldots,q+1)^T$. This means that $(1,\ldots,1)^T$ is an eigenvector of A with eigenvalue $q+1$.

Consider the matrix A^2. If A has eigenvalues of λ_1, …, λ_{q^2+q+1}, then A^2 will have eigenvalues of λ_1^2, …, $\lambda_{q^2+q+1}^2$. Hence information about the eigenvalues of A can be obtained from the eigenvalues of A^2.

Since $A = A^T$, we have that $A^2 = AA = A^TA$. The ij element of A^TA is the dot product of the ith row of A^T with the jth column of A, which is the dot product of the ith column of A with the jth column of A. Now each column of A represents a line of the projective plane and has a 1 in the position of each of the points on this line. Two lines share precisely one point, and so two columns of A are both 1 in precisely one position. Hence the dot product of two distinct columns of A is 1. One the other hand, any line contains $q+1$ points, and therefore any column of A has $q+1$ entries equal to 1. Hence the dot product of a column of A with itself is $q+1$.

Therefore, the diagonal entries of A^2 are $q+1$ and the other entries of A^2 are 1. That is $A^2 = J+qI$, where J is a matrix with all of its entries equal to 1.

Now the matrix J is equal to the product of two vectors $(1,\ldots,1)^T(1,\ldots,1)$. Multiplying these vectors in the opposite order gives a 1×1 matrix: $(1,\ldots,1)(1,\ldots,1)^T = [q^2+q+1]$. This 1×1 matrix has eigenvalue q^2+q+1. Now for any two matrices such that AB and BA are both defined, the eigenvalues of the larger matrix are the same as the eigenvalues of the smaller matrix with the extra eigenvalues all being zero. Thus the q^2+q+1 eigenvalues of J must be q^2+q+1, 0, …, 0.

Consider the polynomial $g(x) = x+q$. The matrix $g(J) = J+qI = A^2$. Therefore the eigenvalues of A^2 are $g(q^2+q+1)$, $g(0)$ …, $g(0)$. That is, the eigenvalues of A^2 are q^2+2q+1, q, …, q. One of the eigenvalues of A is $q+1$, and so the remaining eigenvalues of A must all be $\sqrt{q}$ or $-\sqrt{q}$

Suppose there are k eigenvalues equal to $-\sqrt{q}$ and q^2+q-k equal to $\sqrt{q}$. Then the trace of A is equal to $-k\sqrt{q}+(q^2+q-k)\sqrt{q}+q+1 = \sqrt{q}(q^2+q-2k)+q+1$

If $q = p^2$ for some natural number p, then the trace is $p(q^2+q-2k)+p^2+1$, which is one more than a multiple of p, and so it cannot be 0.

If q is not a square number, then $\sqrt{q}$ is irrational and so if $q^2+q-2k \neq 0$ then the trace of A must be irrational also. However, the entries of A are all 0 or 1, so its trace is an integer. Hence $q^2+q-2k = 0$ and the trace is $q+1$.

In either case, the trace is not 0, so the polarity must have an absolute point. ■