4-5-1994

Eigenvalues and eigenvectors.

Definition. Let $A \in M_n(F)$. The **characteristic polynomial** of A is

$$p(\lambda) = \det(A - \lambda I).$$

(I is the $n \times n$ identity matrix.)

A root of the characteristic polynomial is called an **eigenvalue** (or a **characteristic value**) of A.

While the entries of A come from the field F, it makes sense to ask for the roots of $p(\lambda)$ in an extension field E of F. For example, if A is a matrix with real entries, you can ask for the eigenvalues of A in $\mathbb{R}$ or in $\mathbb{C}$.

Example. The matrix

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

has no eigenvalues in $\mathbb{R}$. Its eigenvalues in $\mathbb{C}$ are $\lambda = \pm i$. $\square$

Example. Let

$$A = \begin{bmatrix} 2 & -3 & 1 \\ 1 & -2 & 1 \\ 1 & -3 & 2 \end{bmatrix}.$$

You can use row and column operations to simplify the computation of $\det(A - \lambda I)$:

$$\begin{vmatrix} 2-\lambda & -3 & 1 \\ 1 & -2-\lambda & 1 \\ 1 & -3 & 2-\lambda \end{vmatrix} \underset{r_2 \to r_2 - r_3}{=} \begin{vmatrix} 2-\lambda & -3 & 1 \\ 0 & 1-\lambda & -1+\lambda \\ 1 & -3 & 2-\lambda \end{vmatrix} \underset{c_3 \to c_3 + c_2}{=} \begin{vmatrix} 2-\lambda & -3 & -2 \\ 0 & 1-\lambda & 0 \\ 1 & -3 & -1-\lambda \end{vmatrix}.$$

(Adding a multiple of a row or a column to a row or column, respectively, does not change the determinant.) Now expand by cofactors of the second row:

$$(1-\lambda)\left((2-\lambda)(-1-\lambda)+2\right) = (1-\lambda)(\lambda^2-\lambda) = -\lambda(\lambda-1)^2.$$

The eigenvalues are $\lambda = 0$, $\lambda = 1$ (double). $\square$

Example. A matrix $A \in M_n(F)$ is **upper triangular** if $A_{ij} = 0$ for $i > j$. Thus, the entries below the main diagonal are zero. (**Lower triangular** matrices are defined in an analogous way.)

The eigenvalues of a triangular matrix

$$A = \begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

are just the diagonal entries $\lambda_1, \ldots, \lambda_n$. (You can prove this by induction on n.) $\Box$

Remark. To find the eigenvalues of a matrix, you need to find the roots of the characteristic polynomial.

There are formulas for finding the roots of polynomials of degree ≤ 4. (For example, the quadratic formula gives the roots of a quadratic equation $ax^2 + bx + c = 0$.) However, Abel showed in the early part of the 19-th century that the general quintic is not solvable by radicals. (For example, $2x^5 - 5x^4 + 5$ is not solvable by radicals over $\mathbb{Q}$.) In the real world, the computation of eigenvalues often requires numerical approximation. $\Box$

Lemma. Let $A \in M_n(F)$. A is invertible if and only if the system $Ax = y$ has a unique solution for every $y \in F^n$.

Proof. If A is invertible, then $x = A^{-1}y$ is a solution to $Ax = y$, since $AA^{-1}y = y$. If $Ax' = y$, then $A^{-1}Ax = A^{-1}y$, so $x = A^{-1}y$. This proves that $x = A^{-1}y$ is the *unique* solution to the equation.

Conversely, suppose $Ax = y$ has a unique solution for every $y \in F^n$. For $i = 1, \ldots, n$, let $x_i \in F^n$ be the unique solution to $Ax = e_i$, where e_i is the i-th standard basis vector. Then

$$A \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ x_1 & x_2 & \cdots & x_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ e_1 & e_2 & \cdots & e_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}.$$

Let

$$X = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ x_1 & x_2 & \cdots & x_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}.$$

Then $AX = I$, so

$$|A||X| = |I| = 1.$$

Thus, $|A| \ne 0$. It follows that A is invertible. $\Box$

If λ is an eigenvalue of A, then $\det(A - \lambda I) = 0$. Hence, the $n \times n$ matrix $A - \lambda I$ is not invertible. It follows that $A - \lambda I$ must row reduce to a row reduced echelon matrix R with fewer than n leading coefficients. Thus, the system $Rx = 0$ has at least one free variable, and hence has infinitely many solutions. In particular, $Rx = 0$ — and therefore, $(A - \lambda I)x = 0$ — has at least *one nonzero solution.*

Definition. Let $A \in M_n(F)$, and let λ be an eigenvalue of A. An **eigenvector** (or a **characteristic vector**) of A for λ is a *nonzero* vector $v \in F^n$ such that

$$Av = \lambda v.$$

Equivalently,

$$(A - \lambda I)v = 0.$$

Example. Let

$$A = \begin{bmatrix} 2 & -3 & 1 \\ 1 & -2 & 1 \\ 1 & -3 & 2 \end{bmatrix}.$$

The eigenvalues are $\lambda = 0$, $\lambda = 1$ (double).

First, I'll find an eigenvector for $\lambda = 0$.

$$A - 0 \cdot I = \begin{bmatrix} 2 & -3 & 1 \\ 1 & -2 & 1 \\ 1 & -3 & 2 \end{bmatrix}.$$

I want $v = \langle a, b, c \rangle$ such that

$$\begin{bmatrix} 2 & -3 & 1 \\ 1 & -2 & 1 \\ 1 & -3 & 2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

You can solve the system by row reduction. Since the column of zeros on the right will never change, it's enough to row reduce the 3×3 matrix on the right.

$$\begin{bmatrix} 2 & -3 & 1 \\ 1 & -2 & 1 \\ 1 & -3 & 2 \end{bmatrix} \quad \to \quad \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

This says

$$\begin{array}{ccccc} a & & - & c & = & 0 \\ & b & - & c & = & 0 \end{array}$$

Therefore, $a = c$, $b = c$, and the eigenvector is

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} c \\ c \\ c \end{bmatrix} = c \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Notice that this is the usual algorithm for finding a basis for the solution space of a homogeneous system (or the null space of a matrix).

I can set c to any nonzero number. For example, $c = 1$ gives the eigenvector $\langle 1, 1, 1 \rangle$. Notice that there are infinitely many eigenvectors for this eigenvalue, but all of these eigenvectors are multiples of $\langle 1, 1, 1 \rangle$.

Likewise,

$$A - I = \begin{bmatrix} 1 & -3 & 1 \\ 1 & -3 & 1 \\ 1 & -3 & 1 \end{bmatrix} \quad \to \quad \begin{bmatrix} 1 & -3 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Hence, the eigenvectors are

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 3b - c \\ b \\ c \end{bmatrix} = b \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Taking $b = 1$, $c = 0$ gives $\langle 3, 1, 0 \rangle$; taking $b = 0$, $c = 1$ gives $\langle -1, 0, 1 \rangle$. This eigenvalue gives rise to two independent eigenvectors.

Note, however, that a double root of the characteristic polynomial *need not* give rise to two independent eigenvectors. $\square$

Lemma. Similar matrices have the same characteristic polynomial (and hence the same eigenvalues).

Proof.

$$PAP^{-1} - \lambda I = PAP^{-1} - \lambda PIP^{-1} = P(A - \lambda I)P^{-1}.$$

Therefore, the matrices $A - \lambda I$ and $PAP^{-1} - \lambda I$ are similar. Hence, they have the same determinant. The determinant of $A - \lambda I$ is the characteristic polynomial of A and the determinant of $PAP^{-1} - \lambda I$ is the characteristic polynomial of PAP^{-1}. $\square$

Definition. Let $T : V \to V$ be a linear transformation, where V is a finite-dimensional vector space. The **characteristic polynomial** of T is the characteristic polynomial of a matrix of T relative to a basis $\mathcal{B}$ of V.

The preceding lemma shows that this is independent of the choice of basis.

This shows that it makes sense to speak of the **eigenvalues** and **eigenvectors** of a *linear transformation* T.

Definition. A matrix $A \in M_n(F)$ is **diagonalizable** if A has n independent eigenvectors — that is, if there is a basis for F^n consisting of eigenvectors of A.

Proposition. $A \in M_n(F)$ is diagonalizable if and only if it is similar to a diagonal matrix.

Proof. Let $\{v_1, \ldots, v_n\}$ be n independent eigenvectors for A corresponding to eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$. Let T be the linear transformation corresponding to A:

$$T(v) = Av.$$

Since $Av_i = \lambda_i v_i$ for all i, the matrix of T relative to the basis $\mathcal{B} = \{v_1, \ldots, v_n\}$ is

$$[T]_{\mathcal{B},\mathcal{B}} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Now A is the matrix of T relative to the standard basis, so

$$[T]_{\mathcal{B},\mathcal{B}} = [\text{std} \to \mathcal{B}] \cdot A \cdot [\mathcal{B} \to \text{std}].$$

The matrix $[\mathcal{B} \to \text{std}]$ is obtained by building a matrix using the $v_1, \ldots, v_n$ as the columns. Then $[\text{std} \to \mathcal{B}] = [\mathcal{B} \to \text{std}]^{-1}$.

Hence,

$$\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = [\mathcal{B} \to \text{std}]^{-1} \cdot A \cdot [\mathcal{B} \to \text{std}].$$

Conversely, if D is diagonal, P is invertible, and $D = P^{-1}AP$, the columns $c_1, \ldots, c_n$ of P are independent eigenvectors for A. In fact, if

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix},$$

then $PD = AP$ says

$$\begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ c_1 & c_2 & \cdots & c_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = A \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ c_1 & c_2 & \cdots & c_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}.$$

Hence, $\lambda_i c_i = Ac_i$. □

Example. The matrix

$$A = \begin{bmatrix} 2 & -3 & 1 \\ 1 & -2 & 1 \\ 1 & -3 & 2 \end{bmatrix}$$

has independent eigenvectors $\langle 1, 1, 1 \rangle$, $\langle 3, 1, 0 \rangle$, $\langle -1, 0, 1 \rangle$.

Thus

$$\begin{bmatrix} 1 & 3 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 2 & -3 & 1 \\ 1 & -2 & 1 \\ 1 & -3 & 2 \end{bmatrix} \begin{bmatrix} 1 & 3 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} =$$

$$\begin{bmatrix} -1 & 3 & -1 \\ 1 & -2 & 1 \\ 1 & -3 & 2 \end{bmatrix} \begin{bmatrix} 2 & -3 & 1 \\ 1 & -2 & 1 \\ 1 & -3 & 2 \end{bmatrix} \begin{bmatrix} 1 & 3 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \square$$

Example. The matrix

$$B = \begin{bmatrix} -3 & 3 & -5 \\ 12 & -7 & 14 \\ 10 & -7 & 13 \end{bmatrix}$$

has eigenvalue $\lambda = 1$ (a triple root). The only eigenvectors are the nonzero multiples of $\langle -1, 2, 2 \rangle$. Hence, B is not diagonalizable. $\square$

Proposition. Let $T : V \to V$ be a linear operator on an n dimensional vector space. If $v_1, \ldots, v_n$ are eigenvectors corresponding to the *distinct* eigenvalues $\lambda_1, \ldots, \lambda_n$, then $\{v_1, \ldots, v_n\}$ is independent.

Proof. Suppose to the contrary that $\{v_1, \ldots, v_n\}$ is dependent. Let p be the smallest number such that the subset $\{v_1, \ldots, v_p\}$ is dependent. Then there is a nontrivial linear relation

$$a_1 v_1 + \cdots a_p v_p = 0.$$

Note that $a_p \neq 0$, else

$$a_1 v_1 + \cdots a_{p-1} v_{p-1} = 0.$$

This would contradict minimality of p.
Hence, I can rewrite the equation above in the form

$$v_p = b_1 v_1 + \cdots + b_{p-1} v_{p-1}.$$

Apply T to both sides, and use $T v_i = \lambda_i v_i$:

$$\lambda_p v_p = b_1 \lambda_1 v_1 + \cdots + b_{p-1} \lambda_{p-1} v_{p-1}.$$

On the other hand,

$$\lambda_p v_p = \lambda_p b_1 v_1 + \cdots + \lambda_p b_{p-1} v_{p-1}.$$

Subtract the last equation from the one before it to obtain

$$0 = b_1(\lambda_1 - \lambda_p) v_1 + \cdots + b_{p-1}(\lambda_{p-1} - \lambda_p) v_{p-1}.$$

Since the eigenvalues are distinct, the terms $\lambda_p - \lambda_i$ are nonzero. Hence, this is a linear relation in $v_1, \ldots, v_{p-1}$ which contradicts minimality of p — unless $b_1 = \cdots = b_{p-1} = 0$.

In this case, $v_p = 0$, which contradicts the fact that v_p is an eigenvector. Therefore, the original set must in fact be independent. $\square$

