EIGENVALUES & EIGENVECTORS

We say $\lambda$ is an eigenvalue of a square matrix $A$ if

$$Ax = \lambda x$$

for some $x \neq 0$. The vector $x$ is called an eigenvector of $A$, associated with the eigenvalue $\lambda$. Note that if $x$ is an eigenvector, then any multiple $\alpha x$ is also an eigenvector.

**EXAMPLES:**

$$\begin{bmatrix}
-7 & 13 & -16 \\
13 & -10 & 13 \\
-16 & 13 & -7
\end{bmatrix} \begin{bmatrix} 1 \\
-1 \\
1
\end{bmatrix} = -36 \begin{bmatrix} 1 \\
-1 \\
1
\end{bmatrix}$$

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

Knowing the eigenvalues and eigenvectors of a matrix $A$ will often give insight as to what is happening when solving systems or problems involving $A$. 
THE CHARACTERISTIC POLYNOMIAL

For an $n \times n$ matrix $A$, solving $Ax = \lambda x$ for a vector $x \neq 0$ is equivalent to solving the homogeneous linear system

$$(A - \lambda I)x = 0$$

This has a nonzero solution if and only if

$$\det(A - \lambda I) = 0$$

We can expand this determinant by minors, obtaining

$$f_A(\lambda) \equiv \det \begin{bmatrix} a_{1,1} - \lambda & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} - \lambda \end{bmatrix} = 0$$

$$f_A(\lambda) = (a_{1,1} - \lambda) \cdots (a_{n,n} - \lambda) + \text{terms of degree } \leq n - 2$$

$$= (-1)^n \lambda^n + (-1)^{n-1} \left(a_{1,1} + \cdots + a_{n,n}\right) \lambda^{n-1} + \text{terms of degree } \leq n - 2$$
We call $f_A(\lambda)$ the characteristic polynomial of $A$; and

$$f_A(\lambda) = 0$$

is called the characteristic equation for $A$. Since $f_A(\lambda)$ is a polynomial of degree $n$:

1. The matrix $A$ has at least one eigenvalue.
2. $A$ has at most $n$ distinct eigenvalues.

The multiplicity of $\lambda$ as a root of $f_A(\lambda) = 0$ is called the \textit{algebraic multiplicity} of $\lambda$. The number of independent eigenvectors associated with $\lambda$ is called the \textit{geometric multiplicity} of $\lambda$.

\textit{EXAMPLES}:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

has the eigenvalue $\lambda = 1$ with both the algebraic and geometric multiplicity equal to 2.

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

has the eigenvalue $\lambda = 1$ with algebraic multiplicity 2 and geometric multiplicity 1.
Let $A$ and $B$ be similar, meaning that for some non-singular matrix $P$,

$$B = P^{-1}AP$$

Being similar means that $A$ and $B$ represent the same linear transformation of the vector space $\mathbb{R}^n$ or $\mathbb{C}^n$, but with respect to different bases. Then

$$f_B(\lambda) = \det (B - \lambda I)$$

$$= \det (P^{-1}AP - \lambda I)$$

$$= \det (P^{-1} [A - \lambda I] P)$$

$$= \left[ \det P^{-1} \right] \det (A - \lambda I) [\det P]$$

$$= \det (A - \lambda I) = f_A(\lambda)$$

since $\det P^{-1} \det P = \det (P^{-1}P) = 1$. This says that the similar matrices have the same eigenvalues.
For the eigenvectors, write

\[ Ax = \lambda x \]

\[ (P^{-1}AP) P^{-1}x = \lambda P^{-1}x \]

\[ Bz = \lambda z \quad \text{with} \quad z = P^{-1}x \]

Thus there is a simple connection with the eigenvectors of similar matrices.

Since \( f_B(\lambda) = f_A(\lambda) \) for similar matrices \( A \) and \( B \), we have that their coefficients are invariant. In particular, note that

\[ f_A(0) = \det A \]

which is the constant term in \( f_A(\lambda) \). Thus similar matrices have the same determinant. In particular, introduce

\[ \text{trace } A = a_{1,1} + \cdots + a_{n,n} \]

It is the coefficient of \( \lambda^{n-1} \), except for sign. Similar matrices have the same trace and determinant.
Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of $A$, repeated according to their multiplicity. Then

$$f_A(\lambda) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda)$$

$$= (-1)^n \lambda^n + (-1)^{n-1} (\lambda_1 + \cdots + \lambda_n) \lambda^{n-1}$$

$$+ \cdots + (\lambda_1 \cdots \lambda_n)$$

Thus

$$\text{trace } A = \lambda_1 + \cdots + \lambda_n, \quad \det A = \lambda_1 \cdots \lambda_n$$

**EXAMPLE:** The eigenvalues of

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

satisfy

$$\lambda_1 + \lambda_2 = 5, \quad \lambda_1 \lambda_2 = -2$$
CANONICAL FORMS

By use of similarity transformations, we change a matrix $A$ into a similar matrix which is “simpler” in its form. These are often called *canonical forms*; and there is a large literature on such forms. The ones most used in numerical analysis are:

The Schur normal form
The principal axes form
The Jordan form
The singular value decomposition

The Schur normal form is less well-known, but is a powerful tool in deriving results in matrix algebra.
SCHUR’S NORMAL FORM. Let $A$ be a square matrix of order $n$ with coefficients from $\mathbb{C}$. Then there is a nonsingular unitary matrix $U$ for which

$$U^* AU = T, \quad U^* U = I$$

with $T$ an upper triangular matrix. Since $U$ is unitary, $U^* = U^{-1}$, and thus $T$ and $A$ are similar.

A proof by induction on $n$ is given in the text, on page 474.

Note that for a triangular matrix $T$, the characteristic polynomial is

$$f_T(\lambda) = \det \begin{bmatrix} t_{1,1} - \lambda & t_{1,2} & \cdots & t_{1,n} \\ 0 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & t_{n,n} - \lambda \end{bmatrix}$$

$$= (t_{1,1} - \lambda) \cdots (t_{n,n} - \lambda)$$

Thus the diagonal elements $\{t_{i,i}\}$ are the eigenvalues of $T$. 
**PRINCIPAL AXES THEOREM.** Let \( A \) be Hermitian (or self-adjoint), meaning
\[
A^* = A
\]
Then there is a nonsingular unitary matrix \( U \) for which
\[
U^* A U = D
\]
a diagonal matrix with only real entries. If the matrix \( A \) is real (and thus symmetric), the matrix \( U \) can be chosen as a real orthogonal matrix.

**PROOF:** Using the Schur normal form,
\[
U^* A U = T
\]
Now form the conjugate transpose of both sides
\[
T^* = (U^* A U)^*
= U^* A^* (U^*)^*
= U^* A^* U \quad \text{since} \quad (B^*)^* = B \quad \text{in general}
= U^* A U \quad \text{since} \quad A^* = A
= T
\]
Since $T^* = (\overline{T})^T$, the diagonal elements of $T$ are real. Moreover, $T = T^*$ implies $T$ must be a diagonal matrix, as asserted, and we write

$$D = \text{diag} [\lambda_1, \ldots, \lambda_n]$$

I will not show here that $A$ real implies $U$ real.

Consequences: Again, assume $A$ (and $U$) are $n \times n$; and let $V = \mathbb{R}^n$ or $\mathbb{C}^n$, depending on whether $A$ is real or complex. Write

$$U = [u_1, \ldots, u_n]$$

with $u_1, \ldots, u_n$ the orthogonal columns of $U$. Since these are elements of the $n$-dimensional space $V$, \{u_1, \ldots, u_n\} is an orthogonal basis of $V$. In addition, re-write $U^*AU = D$ as

$$AU = UD$$

$$A[u_1, \ldots, u_n] = [u_1, \ldots, u_n] \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$$
Matching corresponding columns on the two sides of this equation, we have

\[ A u_j = \lambda_j u_j, \quad j = 1, \ldots, n \]

Thus the columns \( u_1, \ldots, u_n \) are orthogonal eigenvectors of \( A \); and they form a basis for \( V \).

For a Hermitian matrix \( A \), the eigenvalues are all real; and there is an orthogonal basis for the associated vector space \( V \) consisting of eigenvectors of \( A \).

In dealing with such a matrix \( A \) in a problem, the basis \( \{ u_1, \ldots, u_n \} \) is often used in place of the original basis (or coordinate system) used in setting up the problem. With this basis, the use of \( A \) is clear:

\[
A \left( c_1 u_1 + \cdots + c_n u_n \right) = c_1 A u_1 + \cdots + c_n A u_n
\]

\[
= c_1 \lambda_1 u_1 + \cdots + c_n \lambda_n u_n
\]
SINGULAR VALUE DECOMPOSITION. This is a canonical form for general rectangular matrices. Let $A$ have order $n \times m$. Then there are unitary matrices $U$ and $V$, of respective orders $m \times m$ and $n \times n$, for which

$$V^*AU = F$$

with $F$ of order $n \times m$ and of the form

$$F = \begin{bmatrix}
\mu_1 & 0 & \cdots & \cdots \\
0 & \mu_2 & 0 & \cdots \\
\vdots & \ddots & \ddots & \ddots \\
 & & & \mu_r \\
& & & & 0 \\
& & & & \cdots 
\end{bmatrix}$$

The numbers $\mu_i$ are all real, with

$$\mu_1 \geq \mu_2 \geq \cdots \geq \mu_r > 0$$

This is proven in the text (p. 478).
JORDAN CANONICAL FORM

This is probably the best known canonical form of linear algebra textbooks. Begin by introducing matrices

\[ J_m(\lambda) = \begin{bmatrix}
\lambda & 1 & 0 & \cdots \\
0 & \lambda & \ddots & \ddots \\
\vdots & \ddots & \ddots & 1 \\
0 & \cdots & 0 & \lambda
\end{bmatrix} \]

The order is \( m \times m \), and all elements are zero except for \( \lambda \) on the diagonal and 1 on the superdiagonal. We refer to this matrix as a \textit{Jordan block}.

Let \( A \) have order \( n \), with elements chosen from \( \mathbb{C} \). Then there is a nonsingular matrix \( P \) for which

\[ P^{-1}AP = \begin{bmatrix}
J_{n_1}(\lambda_1) & 0 & \cdots & 0 \\
0 & J_{n_2}(\lambda_2) & \cdots & \vdots \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & J_{n_r}(\lambda_r)
\end{bmatrix} \]

The numbers \( \lambda_1, \ldots, \lambda_r \) are the eigenvalues of \( A \), and they need not be distinct.
Note that the Jordan block $J_m(\lambda)$ satisfies

$$J_m(0)^m = 0$$

Such a matrix is called nilpotent. We sometimes write the Jordan canonical form as

$$P^{-1}AP = D + N$$

with $D$ containing the diagonal elements of $P^{-1}AP$ and $N$ containing the remaining nonzero elements. In this case, $N$ is nilpotent with

$$N^q = 0$$

with $q = \max \{n_1, ..., n_r\} \leq n$.

The proof of the Jordan canonical form is very complicated, and it can be found in more advanced level books on linear algebra.