Topics in
Probability Theory and Stochastic Processes
Steven R. Dunbar

Eigenvalues, Eigenvectors and Normal Forms of Matrices

Rating
Mathematically Mature: may contain mathematics beyond calculus with proofs.
Question of the Day

What is the definition of an eigenvalue and an eigenvector of a matrix? What is the geometric meaning and interpretation of an eigenvalue and eigenvector for a matrix?

Key Concepts

1. The eigenvalues of a real symmetric matrix are real numbers.

2. Eigenvectors of a symmetric matrix corresponding to distinct eigenvalues are orthogonal.

3. Let $A$ be a matrix of order $n$ with elements from $\mathbb{C}$. Then there exists a unitary matrix $U$ such that

   $$T = U^*AU$$

   is upper triangular. The diagonal elements of $T$ are the eigenvalues of $A$.

4. Let $A$ be a Hermitian matrix of order $n$, so $A^* = A$. There is a unitary matrix $U$ for which

   $$U^*AU = D = \text{diag}[\lambda_1, \ldots, \lambda_n]$$

   is a diagonal matrix with diagonal elements which are the eigenvalues $\lambda_1, \ldots, \lambda_n$. If $A$ is real, then $U$ can be taken as orthogonal.
Vocabulary

1. A matrix is **stochastic** if the column sums $\sum_{i=1}^{n} P_{ij} = 1$ for $j = 1 \ldots n$. This is identical to saying that

$$1P = 1$$

2. Let $A$ be a matrix of order $n$ with elements from $\mathbb{C}$. Then there exists a unitary matrix $U$ such that

$$T = U^*AU$$

is upper triangular. The diagonal elements of $T$ are the eigenvalues of $A$. This similar matrix is called the **Schur Normal Form** of the matrix $A$.

---

Mathematical Ideas

This section provides proofs of some standard facts from linear algebra about eigenvalues, eigenvectors and normal forms of matrices. These facts are needed in the section on the fastest mixing Markov chain.

Some basic facts about eigenvalues of symmetric matrices.

**Lemma 1.** The eigenvalues of a real symmetric matrix are real numbers.

**Proof.** Let $P$ be a symmetric matrix, so that $P^T = P$. Let $\lambda$ be an eigenvalue of $P$ with corresponding eigenvector $x$. Denote the (complex) inner product on $\mathbb{C}^n$ as $(x, y) = \sum_{i=1}^{n} \bar{x}_i y_i$ where $\bar{x}_j$ is the complex conjugate. Note that
we use the complex inner product because $\lambda$ and $x$ are potentially complex. Then

\[
\lambda(x, x) = (x, \lambda x) \\
= (x, Px) \\
= (P^T x, x) \\
= (Px, x) \\
= (\lambda x, x) \\
= \bar{\lambda} (x, x).
\]

Hence, $\bar{\lambda} = \lambda$ and $\lambda$ must be real. \qed

**Lemma 2.** Let $x_k$ and $x_l$ be eigenvectors corresponding to distinct eigenvalues $\lambda_k \neq \lambda_l$ of the symmetric matrix $P$. Then

\[
(x_k, x_l) = 0
\]

That is, eigenvectors of a symmetric matrix corresponding to distinct eigenvalues are orthogonal.

**Proof.**

\[
\lambda_l (x_k, x_l) = (x_k, \lambda_l x_l) \\
= (x_k, Px_l) \\
= (P^T x_k, x_l) \\
= (Px_k, x_l) \\
= (\lambda_k x_k, x_l) \\
= \bar{\lambda}_k (x_k, x_l) \\
= \bar{\lambda}_k (x_k, x_l).
\]

Then

\[
(x_k, x_l) = 0.
\]

**Remark.** 1. These are standard proofs in numerical analysis, linear algebra, operator theory and many other places in mathematics. They are included here for completeness.
2. The same proofs show that for a Hermitian matrix \((P = P^*)\) the eigenvalues are real.

Remark. Actually, more is true and the next two theorems prove that a symmetric matrix has a complete set of orthogonal eigenvectors even if some eigenvalues are repeated. We will need this crucial fact in order to prove an essential inequality about the rate of mixing of Markov chains.

Theorem 3 (Schur Normal Form). Let \(A\) be a matrix of order \(n\) with elements from \(\mathbb{C}\). Then there exists a unitary matrix \(U\) such that

\[ T = U^*AU \]

is upper triangular. The diagonal elements of \(T\) are the eigenvalues of \(A\).

Proof. The proof is by induction on the order \(n\) of \(A\). The result is trivially true for \(k = 1\) using \(U = (1)\). We assume the result is true for all matrices of order \(n \leq k - 1\), and we will then prove it to be true of all matrices of order \(n = k\).

Let \(\lambda_1\) be an eigenvalue of \(A\) and let \(\mathbf{u}^{(1)}\) be an associated eigenvector with \(||\mathbf{u}^{(1)}||_2 = 1\). Beginning with \(\mathbf{u}^{(1)}\), pick an orthonormal basis for \(\mathbb{C}^k\). (Note that this can be done by filling out a basis with the standard unit vectors and then using the Gram-Schmidt orthogonalization procedure.) Call the basis so obtained \(\{\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \ldots, \mathbf{u}^{(k)}\}\). Define the \(k \times k\) column matrix

\[ P_1 = \begin{bmatrix} \mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \ldots, \mathbf{u}^{(k)} \end{bmatrix} \]

Note that \(P_1^*P_1 = I\), so that \(P_1^{-1} = P_1^*\). Define

\[ B_1 = P_1^*AP_1 \]

The claim is that

\[ B_1 = \begin{pmatrix} \lambda_1 & \alpha_2 & \ldots & \alpha_k \\ 0 & \ddots & \ddots & \vdots \\ \vdots & 0 & \ddots & \alpha_2 \\ 0 & \ldots & 0 & A_2 \end{pmatrix} \]

with \(A_2\) of order \(k-1\) and \(\alpha_2, \alpha_3, \ldots, \alpha_k\) some numbers. To prove this, simply
multiply:
\[
AP_1 = A \begin{bmatrix} u^{(1)}, u^{(2)}, \ldots, u^{(k)} \end{bmatrix} = \begin{bmatrix} Au^{(1)}, Au^{(2)}, \ldots, Au^{(k)} \end{bmatrix} = \begin{bmatrix} \lambda_1 u^{(1)}, v^{(2)}, \ldots, v^{(k)} \end{bmatrix}
\]

\[B_1 = P_1^* AP_1 = \begin{bmatrix} \lambda_1 P_1^* u^{(1)}, P_1^* v^{(2)}, \ldots, P_1^* v^{(k)} \end{bmatrix} = \begin{bmatrix} \lambda_1 e^{(1)}, w^{(2)}, \ldots, w^{(k)} \end{bmatrix}\]

The last equality follows from the constructed orthonormality of \(u^{(1)}\) and the associated construction of \(P_1\) and \(P_1^*\), with \(e = (1, 0, \ldots, 0)\) and \(w^{(j)} = P_1^* (j)\). Note that \(B_1\) has the desired form, and the claim is established.

By the induction hypothesis, there exists a unitary matrix \(\hat{P}_2\) of order \(k - 1\) for which
\[\hat{T} = \hat{P}_2^* A_2 \hat{P}_2\]
is upper triangular of order \(k - 1\). Define
\[
P_2 = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 \\ \vdots \\ 0 \\ \hat{P}_2 \end{pmatrix}
\]

Then \(P_2\) is unitary and
\[
P_2^* B_1 P_2 = \begin{pmatrix} \lambda_1 & \gamma_2 & \cdots & \gamma_k \\ 0 \\ \vdots \\ \hat{P}_2^* A_2 \hat{P}_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 & \gamma_2 & \cdots & \gamma_k \\ 0 \\ \vdots \\ \hat{T} \end{pmatrix}
\]
is an upper triangular matrix. Thus
\[T = P_2^* B_1 P_2 = P_2^* P_1^* AP_1 P_2 = (P_1 P_2)^* A (P_1 P_2)\]

Define \(U = P_1 P_2\), which is easily seen to be unitary. This completes the induction and the proof. \(\square\)
Theorem 4 (Principal Axes Theorem). Let \( A \) be a Hermitian matrix of order \( n \), so \( A^* = A \). Then \( A \) has \( n \) real eigenvalues \( \lambda_1, \ldots, \lambda_n \), not necessarily distinct, and \( n \) corresponding eigenvectors \( \{ \mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \ldots, \mathbf{u}^{(k)} \} \) that form an orthogonal basis for \( \mathbb{C}^n \). If \( A \) is real, the eigenvectors can be taken as real and they form an orthonormal basis for \( \mathbb{R}^n \). Finally, there is a unitary matrix \( U \) for which

\[
U^*AU = D = \text{diag}[\lambda_1, \ldots, \lambda_n]
\]

is a diagonal matrix with diagonal elements \( \lambda_1, \ldots, \lambda_n \). If \( A \) is real, then \( U \) can be taken as orthogonal.

Proof. From the Schur Normal Form theorem, there is a unitary matrix \( U \) with

\[
U^*AU = T
\]

with \( T \) upper triangular. Take the conjugate transpose of both sides to obtain

\[
T^* = (U^*AU)^* = U^*A^*(U^*)^* = U^*AU = T.
\]

Since \( T \) is upper triangular, then \( T^* \) is lower triangular, and since the two are equal, \( T \) must be a diagonal matrix

\[
T = \text{diag}[\lambda_1, \ldots, \lambda_n].
\]

Also, \( T^* = T \) involves complex conjugation of all elements of \( T \) and thus all diagonal elements of \( T \) must be real.

Write \( U \) as

\[
U = [\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \ldots, \mathbf{u}^{(n)}].
\]

Then \( T = U^*AU \) implies \( AU = UT \),

\[
A [\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \ldots, \mathbf{u}^{(n)}] = [\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \ldots, \mathbf{u}^{(n)}] \begin{pmatrix} \lambda_1 & 0 \\ \cdot & \ddots \\ 0 & \cdot & \lambda_n \end{pmatrix}
\]

\[
[A\mathbf{u}^{(1)}, A\mathbf{u}^{(2)}, \ldots, A\mathbf{u}^{(n)}] = [\lambda_1\mathbf{u}^{(1)}, \lambda_2\mathbf{u}^{(2)}, \ldots, \lambda_n\mathbf{u}^{(n)}]
\]

Hence,

\[
A\mathbf{u}^{(j)} = \lambda_j\mathbf{u}^{(j)} \quad j = 1, \ldots, n
\]

Since the columns of \( U \) are orthonormal, and since the dimension of \( \mathbb{C}^n \) is \( n \) these must form an orthonormal basis for \( \mathbb{C}^n \). \( \square \)
Lemma 5. If $P$ is a stochastic matrix, then $1$ is an eigenvector corresponding to eigenvalue $1$.

Proof. To say the matrix is stochastic is to say that the column sums $\sum_{i=1}^{n} P_{ij} = 1$ for each row $j = 1 \ldots n$. But this is identical to saying that

$$1^T P = 1^T$$

which is to say that $1$ is an left eigenvector corresponding to eigenvalue $1$ of $P$. \hfill \Box

Remark. Again, this is a completely standard and well-known fact from Markov chain analysis.

Lemma 6. The eigenvalues of a symmetric, stochastic matrix may be arranged in nonincreasing order:

$$1 = \lambda_1(P) \geq \lambda_2(P) \geq \cdots \geq \lambda_n(P) \geq -1$$

Proof. Since the eigenvalues are real, and one eigenvalue has value $1$, the proof must show that the remaining eigenvalues are between $-1$ and $1$. This follows from the Gershgorin disk theorem. The Gershgorin disk around the $j$-th diagonal element of $P$ is

$$|z - p_{jj}| \leq \sum_{i=1, i \neq j}^{n} |p_{ij}| = 1 - p_{jj}$$

The eigenvalues must lie in the union of these disks in the complex plane. However, we know that $0 \leq p_{jj} \leq 1$ and further that the eigenvalues must be real, so $-1 \leq \lambda_k \leq 1$. \hfill \Box

Remark. This is a specialized case of well-known facts from Frobenius theory about the eigenvalues and eigenspaces of non-negative and strictly positive matrices.

Sources

This section is completely standard material that can be found in most books on linear algebra or allied subjects such as numerical analysis, advanced calculus and engineering mathematics, or functional analysis.
Problems to Work for Understanding

Reading Suggestion:

References


Outside Readings and Links:

I check all the information on each page for correctness and typographical errors. Nevertheless, some errors may occur and I would be grateful if you would alert me to such errors. I make every reasonable effort to present current and accurate information for public use, however I do not guarantee the accuracy or timeliness of information on this website. Your use of the information from this website is strictly voluntary and at your risk.