Diagonalisation

Monday 21.08.2023 (Room M1)

1 Finding Eigenvectors

When a matrix has repeated eigenvalues, the process of finding the associated eigenvectors can sometimes be intricate. We discuss the concept of generalized eigenvectors and provide a systematic method to determine them.

For a matrix A of size $n \times n$ and its eigenvalue λ, we define:

- **Algebraic Multiplicity** k_λ: The number of times λ appears as a root of the characteristic polynomial $p(t) = \det(A - tI)$.
- **Geometric Multiplicity** p_λ: The number of linearly independent eigenvectors associated with λ.

For a $n \times n$ matrix A, if $\{\lambda_1, \lambda_2, \cdots \lambda_m\}$ are eigenvalues with coresponding algebraic multiplicities of $\{k_{\lambda_1}, k_{\lambda_2}, \cdots k_{\lambda_m}\}$, it means the eigenvalue λ_i is repeated k_{λ_i} times. There are primarily two cases to consider while finding the corresponding eigenvectors.

1.1 Case 1: Complete Set of Linearly Independent Eigenvectors

In the ideal scenario, $p_{\lambda_i} = k_{\lambda_i}$ linearly independent eigenvectors are associated with λ_i. The eigenvectors can be found by solving the homogeneous system:

$$(A - \lambda_i I)\mathbf{v} = \mathbf{0}. \tag{1}$$

In this case since $p_{\lambda_1} + p_{\lambda_2} + ... + p_{\lambda_m} = k_{\lambda_1} + k_{\lambda_2} + ... + k_{\lambda_m} = n$, we will have n linearly independent eigenvectors.

Lemma 1.1 *Let A be a matrix, and λ_1 and λ_2 be two distinct eigenvalues of A with corresponding eigenvectors v_1 and v_2 respectively. Then v_1 and v_2 are linearly independent.*

Proof 1.2 *Assume for the sake of contradiction that v_1 and v_2 are linearly dependent. This implies that one is a scalar multiple of the other. Without loss of generality, let*

$$v_1 = cv_2$$

for some nonzero scalar c.

Given $Av_1 = \lambda_1 v_1$ and $Av_2 = \lambda_2 v_2$, using our assumption we have:

$$Av_1 = \lambda_1 v_1 = \lambda_1(cv_2) = c\lambda_1 v_2 \quad \textit{(i)}$$

From the given condition on v_2, we get:

$$Av_1 = Acv_2 = cAv_2 = c\lambda_2 v_2 \quad \textit{(ii)}$$

Comparing (i) and (ii), since the left-hand sides are identical, we get:

$$c\lambda_1 v_2 = c\lambda_2 v_2$$

Dividing both sides by the nonzero scalar $k = cv_2$, we find:

$$\lambda_1 = \lambda_2$$

This is a contradiction as λ_1 and λ_2 are distinct eigenvalues. Therefore, our assumption is false and v_1 and v_2 are linearly independent.

Example 1.3 *Consider the matrix:*

$$A = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}$$

The eigenvalues are $\lambda_1 = 2, \lambda_2 = 3, \lambda_3 = 4$, and $\lambda_4 = 5$. Each eigenvalue has its respective eigenvectors being the standard basis vectors in $\mathbb{R}^4$.

The eigenvalues and their corresponding eigenvectors are:

$$\lambda_1 = 2 \Rightarrow N(A - 2I) = \left\{ v \in \mathbb{R}^4 \middle| \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} v = 0 \right\} = \left\{ a \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \middle| a \in \mathbb{R} \right\} \Rightarrow \mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\lambda_2 = 3 \Rightarrow N(A - 3I) = \left\{ v \in \mathbb{R}^4 \middle| \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} v = 0 \right\} = \left\{ a \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \middle| a \in \mathbb{R} \right\} \Rightarrow \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

$$\lambda_3 = 4 \Rightarrow N(A - 4I) = \left\{ v \in \mathbb{R}^4 \middle| \begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} v = 0 \right\} = \left\{ a \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \middle| a \in \mathbb{R} \right\} \Rightarrow \mathbf{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

$$\lambda_4 = 5 \Rightarrow N(A - 5I) = \left\{ v \in \mathbb{R}^4 \middle| \begin{pmatrix} -3 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} v = 0 \right\} = \left\{ a \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \middle| a \in \mathbb{R} \right\} \Rightarrow \mathbf{v}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Example 1.4 *Consider the matrix:*

$$B = \begin{pmatrix} 6 & 0 & 0 & 1 \\ 0 & 6 & 1 & 0 \\ 0 & 1 & 6 & 0 \\ 1 & 0 & 0 & 6 \end{pmatrix}$$

The matrix has $\lambda_1 = 5$ and $\lambda_2 = 7$ as the only eigenvalue with 4 linearly independent eigenvectors.

The eigenvalue is $\lambda_1 = 5$ with corresponding eigenvectors:

$$N(B-5I) = \left\{ v \in \mathbb{R}^4 \middle| \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} v = 0 \right\} = \left\{ a \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} \middle| a, b \in \mathbb{R} \right\}$$

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}.$$

The eigenvalue is $\lambda_2 = 7$ with corresponding eigenvectors:

$$N(B-7I) = \left\{ v \in \mathbb{R}^4 \middle| \begin{pmatrix} -1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix} v = 0 \right\} = \left\{ a \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \middle| a, b \in \mathbb{R} \right\}$$

$$\mathbf{v}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \qquad \mathbf{v}_4 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}.$$

1.2 Case 2: Generalized Eigenvectors

However, in some cases, there might be fewer than n linearly independent eigenvectors. This leads us to the concept of generalized eigenvectors. Let's denote the number of linearly independent eigenvectors found as p_i (where $p_i < k_i$) coresponding to Eigenvalues λ_i. It means if only p_i linearly independent eigenvectors corresponding to λ_i are found, then $k_i - p_i$ generalized eigenvectors are to be looked for:

1. Find the usual eigenvectors by solving:
$$(A - \lambda_i I)\mathbf{v} = \mathbf{0}.$$
Assume p eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \cdots \mathbf{v}_p$ are found.
2. For the first generalized eigenvector, solve:
$$(A - \lambda_i I)\mathbf{v}_{p+1} = \mathbf{v}_1,$$
where $\mathbf{v}_1$ is a previously found eigenvector. A non-trivial solution $\mathbf{v}_{p+1}$ represents the first generalized eigenvector.
3. If more generalized eigenvectors are needed, continue this process by solving:
$$(A - \lambda_i I)\mathbf{v} = \mathbf{v}_2,$$
to find $\mathbf{v}_3$, and so on, up to $\mathbf{v}_k$, Now we have $k-p$ generalised eigenvectors $\mathbf{v}_{p+1}, \mathbf{v}_{p+2}, \cdots \mathbf{v}_k$.

Lemma 1.5 *Let A be a matrix and λ be an eigenvalue of A with eigenvector v_1 such that $(A - \lambda I)v_1 = 0$. If $(A - \lambda I)v_2 = v_1$, then v_1 and the generalised eigenvector v_2 are linearly independent.*

Proof 1.6 *Assume for contradiction that v_1 and v_2 are linearly dependent. This means that one is a scalar multiple of the other. Without loss of generality, let:*

$$v_2 = cv_1$$

for some non-zero scalar c. Inserting this into our given equation, we obtain:

$$(A - \lambda I)v_2 = (A - \lambda I)(cv_1) = c(A - \lambda I)v_1 = 0$$

However, $(A - \lambda I)v_2 = v_1$, therefore:

$$0 = v_1$$

This is a contradiction, as eigenvectors are non-zero by definition.

Given the contradiction, our initial assumption that v_1 and v_2 are linearly dependent is incorrect. Hence, v_1 and v_2 must be linearly independent.

Example 1.7 *Consider the matrix:*

$$C = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Here, $\lambda = 1$ has only 2 linearly independent eigenvectors, with the remaining 2 being generalized eigenvectors. The eigenvectors and generalized eigenvectors are:

$$N(C - I) = \left\{ v \in \mathbb{R}^4 \middle| \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} v = 0 \right\} = \left\{ a \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \middle| a, b \in \mathbb{R} \right\}$$

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

Generalized eigenvectors:

$$(C - I)\mathbf{v}_3 = \mathbf{v}_1 \quad \Rightarrow \quad \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \mathbf{v}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \Rightarrow \quad \mathbf{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$

$$(C - I)\mathbf{v}_4 = \mathbf{v}_2 \quad \Rightarrow \quad \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \mathbf{v}_4 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad \Rightarrow \quad \mathbf{v}_4 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

Example 1.8

$$D = \begin{pmatrix} 5 & 1 & 0 & 0 \\ 0 & 5 & 1 & 0 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 5 \end{pmatrix}$$

For this matrix, the eigenvalue $\lambda = 5$ yields just 1 true eigenvector, and the other 3 are generalized eigenvectors.

$$N(D-5I) = \left\{ v \in \mathbb{R}^4 \middle| \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} v = 0 \right\} = \left\{ a \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \middle| a \in \mathbb{R} \right\}$$

The eigenvalue $\lambda = 5$ yields:

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Generalized eigenvectors:

$$(C-5I)\mathbf{v}_2 = \mathbf{v}_1 \quad \Rightarrow \quad \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \Rightarrow \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

$$(C-5I)\mathbf{v}_3 = \mathbf{v}_2 \quad \Rightarrow \quad \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \mathbf{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \Rightarrow \quad \mathbf{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$

$$(C-5I)\mathbf{v}_4 = \mathbf{v}_3 \quad \Rightarrow \quad \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \mathbf{v}_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \quad \Rightarrow \quad \mathbf{v}_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

2 Diagonalisable Matrices

The diagonalizability of a matrix depends on its eigenvectors and eigenvalues. When eigenvalues are repeated, the conditions for diagonalizability can be described in terms of the algebraic and geometric multiplicities of these eigenvalues.

2.1 Conditions for Diagonalizability

A matrix A is diagonalizable if and only if the following conditions hold:

1. For all eigenvalues λ of A, the geometric multiplicity equals the algebraic multiplicity:

$$k_\lambda = p_\lambda$$

2. The sum of the geometric multiplicities of all eigenvalues is equal to the size of the matrix:

$$\sum_\lambda p_\lambda = n$$

   where the sum runs over all distinct eigenvalues of A.

Then:

Case 1: A has k_λ linearly independent eigenvectors corresponding to λ. In this case, $p_\lambda = k_\lambda$ and the matrix is diagonalizable.

Case 2: A has fewer than k linearly independent eigenvectors corresponding to λ. This means $m_g(\lambda) < m_a(\lambda)$, and the matrix is not diagonalizable.

Therefore a matrix A is diagonalizable if it falls only in **Case 1**.

3 Diagonalization

If a matrix A can be diagonalized, it means that it can be written in the form:

$$A = PDP^{-1}$$

Where D is a diagonal matrix and P is the matrix whose columns are the eigenvectors of A. The diagonal elements of D are the eigenvalues of A.

Here's a step-by-step procedure to diagonalize a matrix:

1. **Find the Eigenvalues:**
   Start by finding the determinant of the matrix subtracted by a scalar multiple of the identity matrix.
   $$|A - \lambda I| = 0$$
   Solve this polynomial equation to find the eigenvalues λ.
2. **Find the Eigenvectors:**
   For each eigenvalue λ, substitute it into the equation:
   $$(A - \lambda I)\mathbf{v} = 0$$
   Solve for the vector $\mathbf{v}$. This vector is the eigenvector associated with the eigenvalue λ.
3. **Form Matrix P:**
   Construct a matrix P whose columns are the eigenvectors you found. Ensure the order of columns matches the order you'll place the eigenvalues in the diagonal matrix D.
4. **Form the Diagonal Matrix D:**
   Place the eigenvalues on the diagonal of the matrix D. The order should match with the order of eigenvectors in P.
5. **Verify Diagonalization:**
   Check that $A = PDP^{-1}$.

Note: Not all matrices can be diagonalized. A matrix is diagonalizable if and only if the number of independent eigenvectors is equal to the size of the matrix (i.e., the matrix has a full set of linearly independent eigenvectors).

The matrices in examples 1.3 and 1.4 can be diagonalized, but the matrices in examples 1.7 and 1.8 are not diagonalizable.

Example 3.1 *Consider the matrix B in example* 1.4*, we calculated its eigenvalues and eigenvectors, we seek to decompose it in terms of a matrix P containing the eigenvectors and a diagonal matrix D containing the eigenvalues.*

The matrix B is:

$$B = \begin{pmatrix} 6 & 0 & 0 & 1 \\ 0 & 6 & 1 & 0 \\ 0 & 1 & 6 & 0 \\ 1 & 0 & 0 & 6 \end{pmatrix}$$

Eigenvalues and their corresponding eigenvectors are:

For $\lambda_1 = 5$:

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}$$

For $\lambda_2 = 7$:

$$\mathbf{v}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{v}_4 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$

Construct the matrix P using the eigenvectors as columns:

$$P = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & -1 & 0 & 1 \\ -1 & 0 & 1 & 0 \end{pmatrix}$$

The diagonal matrix D is:

$$D = \begin{pmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & 7 \end{pmatrix}$$

Thus, the decomposition of B is given by $PDP^{-1} = B$.

$$PDP^{-1} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & -1 & 0 & 1 \\ -1 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 7 & 0 \\ 0 & 0 & 0 & 7 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & 0 & 0 & -\frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix} = \begin{pmatrix} 6 & 0 & 0 & 1 \\ 0 & 6 & 1 & 0 \\ 0 & 1 & 6 & 0 \\ 1 & 0 & 0 & 6 \end{pmatrix}$$

4 Diagonalization of Symmetric Matrices

Definition 4.1 *A matrix A is said to be symmetric if it equals its transpose, i.e., $A = A^T$.Symmetric matrices have a unique diagonalization property.*

Definition 4.2 (Orthogonal Matrix) *A matrix P is said to be orthogonal if its columns are mutually orthogonal. Formally, if P is of size $n \times n$, then for any columns p_i and p_j of P (where $i \neq j$), the inner product is zero:*

$$p_i^T . p_j = 0$$

Definition 4.3 (Orthonormal Matrix) *A matrix P is said to be orthonormal if its columns are unit vectors and P is orthogonal. This implies:*

$$p_i^T.p_i = 1$$

for every column p_i of P.

Theorem 4.4 *An orthonormal matrix P has the property that:*

$$P^{-1} = P^T$$

Gram-Schmidt Process

Given a set of linearly independent vectors $v_1, v_2, \ldots, v_k$, the Gram-Schmidt process produces an orthonormal set $u_1, u_2, \ldots, u_k$.

1. Begin with the first vector:
$$u_1 = \frac{v_1}{\|v_1\|}$$

2. For $j = 2, \ldots, k$:
$$\begin{aligned} w_j &= v_j - \sum_{i=1}^{j-1}(v_j.u_i)u_i \\ u_j &= \frac{w_j}{\|w_j\|} \end{aligned}$$

Example 4.5 *Given vectors:*

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$v_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

We apply the Gram-Schmidt process:

1. For the first vector:

$$u_1 = \frac{v_1}{\|v_1\|} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

2. For the second vector:

$$\begin{aligned} w_2 &= v_2 - (v_2.u_1)u_1 \\ &= \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - (1)\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \end{aligned}$$

Normalize w_2:

$$u_2 = \frac{w_2}{\|w_2\|} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Result:

$$u_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$u_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

4.1 Properties

- All eigenvalues of a symmetric matrix are real numbers.
- The eigenvectors corresponding to distinct eigenvalues of a symmetric matrix are orthogonal. It means that if $\lambda_i \neq \lambda_j$, then corresponding eigenvectors are orthogonal.
- For repeated λ_i, we can find eigenvectors that are orthogonal, using Gram Schmidt method.
- For all λ_i, the algebraic and geometric orders are equal, $p_{\lambda_i} = k_{\lambda_i}$.
- A symmetric matrix is orthogonally diagonalizable. This means that there exists an orthogonal matrix P such that $A = PDP^T$, where D is a diagonal matrix containing the eigenvalues of A.

4.2 Orthogonal Diagonalization

Given a symmetric matrix A, we can find a set of orthonormal eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. The matrix P whose columns are these eigenvectors will be an orthogonal matrix, and $P^T AP$ will be diagonal.

The steps for orthogonal diagonalization are:

1. Find the eigenvalues of A.
2. For each eigenvalue, find an eigenvector.
3. Use the Gram-Schmidt process to orthonormalize the set of eigenvectors.
4. Form matrix P using the orthonormal eigenvectors as columns.
5. $A = PDP^T$ will be diagonal.

Example

Recall the matrix

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

(i) Find the eigenvalues and eigenvectors of A.

(ii) Is A diagonalizable?

(iii) Find an orthonormal matrix P such that $A = PDP^T$, where D is a diagonal matrix.

Solution

(i), (ii) Observe that A is a real symmetric matrix. By the above theorem, we know that A is diagonalizable. i.e. we will be able to find a sufficient number of linearly independent eigenvectors. The eigenvalues of A were: $\lambda_1 = -1$, $\lambda_2 = 2$. We found two linearly independent eigenvectors corresponding to $\lambda_1 = -1$:

$$\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

And one eigenvector corresponding to $\lambda_2 = 2$:

$$\mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Since A is a real symmetric matrix, eigenvectors corresponding to distinct eigenvalues are orthogonal. However, the eigenvectors corresponding to eigenvalue $\lambda_1 = -1$, $\mathbf{v}_1$ and $\mathbf{v}_2$ are not orthogonal to each other. We have to use the Gram Schmidt process.

Using the Gram Schmidt process:

$$\mathbf{u}_1 = \frac{\mathbf{v_1}}{\|\mathbf{v_1}\|} = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

$$\mathbf{w}_2 = \mathbf{v}_2 - (\mathbf{v}_2.\mathbf{u}_1)\mathbf{u}_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1/2 \\ -1/2 \\ 1 \end{bmatrix}$$

$$\mathbf{u}_2 = \frac{\mathbf{w_2}}{\|\mathbf{w_2}\|} = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$$

Normalize the vectors: $\mathbf{u_1}$ and $\mathbf{u_2}$ are normalised but we need to normalise $\mathbf{v_3}$:

$$\mathbf{u}_3 = \frac{\mathbf{v_3}}{\|\mathbf{v_3}\|} = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Now we have a set of orthonormal eigenvectors corresponding to eigenvalues $\{-1, -1, 2\}$ as follows:

$$\left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}, \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

Thus we have an orthonormal matrix:

$$P = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{bmatrix}$$

and the corresponding diagonal matrix

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

Verification $A = PDP^T$:

$$PDP^T = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{6} & -1/\sqrt{6} & 2/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix} = A$$