UNIT - 4

Orthogonalization , Eigen Values and Eigen Vectors

Chapter 3 – 3.4 (Gilbert Strang)
Chapter 5 – 5.1 , 5.2 (Gilbert Strang)
Chapter 2 - 2.13 to 2.15 (B S Grewal)
Chapter 28 – 28.9 (B S Grewal)

Book : Gilbert Strang

3.4 – Orthogonal Bases and Gram-Schmidt

5.1 –Introduction to Eigenvalues and Eigenvectors

5.2- Diagonalization of a matrix

Book – B S Grewal

2.13- Eigenvalues

14. -Properties of Eigenvalues

15.-Cayley- Hamilton Theorem 28.9-Determination of Eigenvalues by Iteration

Orthogonal Bases and Gram-Schmidt

The three topics basic to this section are

1.The definition and properties of an orthogonal matrix Q

2.The solution of Qx = b (both m=n and m > n)

3.The Gram- Schmidt process and A = QR factorization

Definition :

In an **orthogonal basis**, every vector is perpendicular to every other vector.

The coordinate axes are mutually orthogonal.

Mutually perpendicular unit vectors are called **Orthonormal** vectors.

For the vector space R^2,

1. The set (2, 0) , (0, 2) is an orthogonal basis.
2. The set (1, -2) , (2 , 1) is an orthogonal basis.
3. The set (1, 0), (0, 1) is an orthonormal basis.

- A matrix with Orthonormal columns will be called Q.
- A square matrix with Orthonormal columns is called an ***Orthogonal matrix*** denoted by Q.

Ex: Rotation matrix , any permutation matrix .

Note : The size of Q has to be square or tall.

Properties of Q

- If Q (square or rectangular) has orthonormal columns, then $Q^T Q = I$.
- An orthogonal matrix is a square matrix with orthonormal columns. Then Q^T is Q^{-1}.

- If Q is rectangular then Q^T is left inverse of Q.

- Multiplication by any Q preserves length. The norms of x and Qx are equal.

- Also, Q preserves inner products and angles, since $(Qx)^T(Qy) = x^TQ^TQy = x^Ty$.

If $q_1, q_2 \ldots . q_n$ are orthonormal basis of R^n then any vector b from R^n can be expressed as

$b = x_1 q_1 + x_2 q_2 + \ldots\ldots + x_n q_n$ …...Eqn(1)

Multiply both sides by q_1^T. Then $q_1^T b = x_1$.

Similarly, $x_2 = q_2^T b, \ldots\ldots, x_n = q_n^T b$.

Hence, $b = (q_1^T b) q_1 + (q_2^T b) q_2 + \ldots\ldots + (q_n^T b) q_n$

= sum of one dimensional projections on to q_i's.

The matrix form of equation (1) is $Qx = b$ and the solution of this system of equations is

$$x = Q^{-1} b = Q^T b$$

The rows of a square matrix are orthonormal whenever the columns are.

Orthonormal columns
Orthonormal rows

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

Rectangular Matrices with Orthonormal Columns

- If Q has orthonormal columns, the least- squares problem becomes easy.
- $Q^T Qx = Q^T b$ are the normal equations for the best solution -in which $Q^T Q = I$.
- $x = Q^T b$
- $p = Qx$ the projection of b is $(q_1^T b)q_1 + \dots + (q_n^T b)q_n$
- $p = QQ^T b$, the projection matrix is $P = Q Q^T$.

The Gram-Schmidt Process

- It is a process of converting linearly independent vectors into orthonormal vectors.

- Consider any 3 independent vectors a, b, c.
- Then the first orthonormal q_1= a/norm(a).

- If ‘b’ is perpendicular to the vector ‘a’ then q_2=b/norm(b) otherwiseB=b-(q_1^T b)q_1 and q_2=B/norm(B).
- If ‘c’ is perpendicular to the plane spanned by the vectors a and b then
- q_3=c/norm(c)

otherwise C=c-(q_1^Tc)q_1-(q_2^Tc) q_2 and

q_3=C/norm(C).

- This is the one idea of the whole Gram-Schmidt process, to subtract from every new vector its components in the directions that are already settled.
- That idea is used over and over again. When there is a fourth vector, we subtract away its components in the directions of q_1, q_2, q_3.

The Factorization A=QR

- We started with a matrix A, whose columns were a, b, c.
- We ended with a matrix Q, whose columns are q_1, q_2, q_3.
- A and Q are of order m by n when the n vectors are in m-dimensional space.

The whole factorization is

$$A = \begin{bmatrix} \\ a & b & c \\ \\ \end{bmatrix} = \begin{bmatrix} \\ q_1 & q_2 & q_3 \\ \\ \end{bmatrix} \begin{bmatrix} q_1^T a & q_1^T b & q_1^T c \\ & q_2^T b & q_2^T c \\ & & q_3^T c \end{bmatrix} .$$

A = Q R

Eigen values and Eigen vectors

Definition :

Let A be a square matrix of order n. If there exists a real or complex number λ and a non zero vector x such that $Ax = \lambda x$ then x is called the ***Eigen vector of A*** and λ is its corresponding ***Eigen value***.

Note :

- The vector x is in the null space of A- λ I.
- The number λ is chosen so that A- λ I has a null space.
- A- λ I must be singular.
- Det(A- λ I)=0 is called the *characteristic equation of A* and roots of this equation are called *characteristic roots or Eigen values or Latent roots*.

Corresponding to ‘n’ distinct Eigen values we get ‘n’ independent Eigen vectors. But when 2 or more eigen values are equal, it may or may not be possible to get linearly independent Eigen vectors corresponding to repeated roots.

Properties of Eigen Values and Eigen vectors

- If λ is an Eigen value of A with x as the corresponding Eigen vector then λ^2 is an Eigen value of A^2 with the same Eigen vector x.
- For a given Eigen vector x, there corresponds only one Eigen value λ.
- For a given Eigen value there corresponds infinitely many Eigen vectors.

- $\lambda = 0$ is an Eigen value of A , if and only if A is singular i.e $\det(A)=0$.
- If λ is an Eigen value of A with x as the Eigen vector then $1/\lambda$ is an Eigen value of A^{-1} provided A^{-1} exists.
- A and its transpose A^T have the same Eigen values.

- The Eigen values of a diagonal matrix are just the diagonal elements of the matrix.
- The Eigen values of an idempotent matrix are either zero or unity.
- The sum of the Eigen values of a matrix is the sum of the elements of the principal diagonal.
- The product of the Eigen values of a matrix A is equal to its determinant.

Procedure to find eigenvalues and eigenvectors

- Compute the determinant of A - λ I. With a λ subtracted along the diagonal, this determinant is a polynomial of degree n. It starts with $(-\lambda)^n$.
- Find the roots of this polynomial. The n roots are the eigenvalues of A.

For each eigenvalue λ, solve the equation $(A - \lambda I)x = 0$.
Since the determinant of $A - \lambda I$ is zero, there are solutions other than $x = 0$.
Those are the eigenvectors.

The Cayley-Hamilton Theorem

Statement:

Every square matrix satisfies its own characteristic equation .

Example:

Let

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

The characteristic polynomial is

$p(t) = \det(A - tI) = t^2 - 4t + 2 = 0$ and hence it can be verified

that

$$A^2 - 4A + 2I = 0.$$

Note

: If a matrix is invertible then we can find its inverse using Cayley-Hamilton Theorem.

Example: $A = \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix}$

Using Cayley-Hamilton theorem,

$A^2-4A+2I=(0)$. Therefore $A^{-1} =(4I-A)/2$.

Rayleigh's Power Method

- It is an iterative method of computing the numerically largest Eigen value of a matrix.

Procedure

- Let A be a square matrix of order n.
- Choose a initial vactor x_0.
- Compute Ax_0 and express $Ax_0 = \lambda_1 x_1$ where λ_1 is the numerically largest value in Ax_0.

- Compute Ax_1 and express $Ax_1 = \lambda_2 x_2$ where λ_2 is the numerically largest value in Ax_1.
- Repeat the procedure until the 2 consecutive values of λ are almost the same.

Diagonalization of a Matrix

Suppose the n by n matrix A has n linearly independent eigenvectors.
If these eigenvectors are the columns of a matrix S, then $S^{-1}AS$ is a diagonal matrix Λ.
The eigenvalues of A are on the diagonal of Λ

- If the matrix A has no repeated eigenvalues then its n eigenvectors are automatically independent .
- Therefore any matrix with distinct Eigen values can be diagonalized.
- The diagonalizing matrix S is not unique. An eigenvector x can be multiplied by a constant, and remains an eigenvector.

- Diagonalizability of A depends on enough eigenvectors.
- Invertibility of A depends on non zero eigen values.

Powers and Products

- If A is diagonalizable then $A=S\Lambda S^{-1}$.
- So $A^K = S\,\Lambda^K\,S^{-1}$.
- Diagonalizable matrices share the same eigenvector matrix S if and only if AB = BA.