1 Eigen$\cdots$

1.1 Eigenvalues and Eigenvectors

The **characteristic polynomial** of A is

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

Solutions for $\chi_A(\lambda) = 0$ are **eigenvalues**. For each eigenvalues, we have a vector space

$$E_\lambda = \text{Nul }(A - \lambda I)$$

, called the **eigenspace** associated with the eigenvalue λ, which is not a zero-dimensional space because $\det M = 0$ if and only if $M\mathbf{x} = \mathbf{0}$ has nontrivial solutions. Except the zero vector, all vectors in E_λ are called **eigenvectors** associated with the eigenvalue λ.

1.2 Diagonalization

Suppose that we have a basis of $\mathbb{R}^n$ consisting of eigenvectors ; $\mathcal{B} = \{v_1, \cdots, v_n\}$ associated with eigenvalues λ_1, $\cdots$, λ_n respectively. λ_i's do not need to be distinct. If we choose a new coordinate system using $\mathcal{B}$, it is really simple to explain the linear transformation A since A is applied to a $\mathcal{B}$-coordinate $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ resulting in $\begin{pmatrix} \lambda_1 x_1 \\ \lambda_2 x_2 \\ \vdots \\ \lambda_n x_n \end{pmatrix}$. In other words,

$$A(x_1 v_1 + \cdots + x_n v_n) = \lambda_1 x_1 v_1 + \cdots + \lambda_n x_n v_n$$

Equivalently,

$$A \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda_1 v_1 & \lambda_2 v_2 & \cdots & \lambda_n v_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Deleting $\begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}$ part and defining P as the matrix obtained by attaching v_1, $\cdots$, v_n side by side, we get

$$AP = PD$$

for $D = \text{diag}(\lambda_1, \cdots, \lambda_n)$. Since $\mathcal{B}$ is a basis, P is invertible so that we can multiply P^{-1} on the left in both sides.

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

In other words, this can be interpreted as

There is a basis $\mathcal{B}$ such that the $\mathcal{B}$-matrix for A is diagonal.

since $P^{-1}\square P$ is the $\mathcal{C}$-matrix of $\square$ (or **Change of Basis** from $\mathcal{E}$ (the standard basis) to $\mathcal{C}$) where $\mathcal{C} = \{Pe_1, \cdots, Pe_n\}$.

Remember that we have easily proven that

A is diagonalizable if and only if $\mathbb{R}^n$ has a basis consisting of eigenvectors.

1.3 Similarity and Change of Basis

A said to be **similar** to B if and only if there exists an invertible matrix P such that

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

Shortly, we use the notation $\sim$ for similarity. There are some obvious properties of $\sim$.

1) $\sim$ is reflective in the sense that $A \sim A$ for every matrix A.

$$I^{-1}AI = A$$

2) $\sim$ is symmetric in the sense that $A \sim B$ implies $B \sim A$.

$$P^{-1}AP = B \text{ implies } (P^{-1})^{-1}BP^{-1} = A$$

3) $\sim$ is transitive in the sense that $A \sim B$ and $B \sim C$ implies $A \sim C$.

$$P^{-1}AP = B \text{ and } Q^{-1}BQ = C \text{ implies } (PQ)^{-1}A(PQ) = C$$

And also, if $A \sim B$, then

$$\det A = \det B, \quad \text{tr } A = \text{tr } B, \quad \chi_A(\lambda) = \chi_B(\lambda)$$

, and A and B have the same eigenvalues, counting multiplicities.

Here is one tautology ;

A is diagonalizable if and only if A is similar to some diagonal matrix D.

In this sense, the invertible matrix P plays a role of a **change of basis**-matrix.

2 Inner Product

2.1 Inner Product Space

A vector space V EQUIPPED WITH one more binary operation, called the inner product $\langle , \rangle$, is called an **inner product space** $(V, \langle , \rangle)$. Here, $\langle , \rangle : V \times V \to \mathbb{R}$ should satisfy below 4 conditions :

1) $\langle u, v \rangle = \langle v, u \rangle$

2) $\langle u + u', v \rangle = \langle u, v \rangle + \langle u', v \rangle$

3) $\langle cu, v \rangle = c\langle u, v \rangle$

4) $\langle u, u \rangle \geq 0$ always. And $\langle u, u \rangle = 0$ if and only if $u = 0$.

The MOST STANDARD INNER PRODUCT, called the dot product, in $\mathbb{R}^n$ (obviously) satisfies all the conditions above.

In an inner product, we define **length**, **distance**, and **orthogonality** using $\langle , \rangle$ as followings :

a. Length of v : $||v|| = \sqrt{\langle v, v \rangle}$

b. Distance of u and v : $\text{dist}(u, v) = ||u - v||$

c. Orthogonal u and v : $\langle u, v \rangle = 0$.

We can show that $\langle u, v \rangle = 0$ if and only if $||u + v||^2 = ||u||^2 + ||v||^2$ and this is called the **Pythagorean Theorem**.

The last thing I would like to mention is

$$W^\perp = \{v \in V : \langle v, w \rangle = 0 \text{ for all } w \in W\}$$

for a subspace W of V. This $W^\perp$ is called the **orthogonal complement** of W. Immediately, we have

$$(\text{Row } A)^\perp = \text{Nul } A \quad \text{and} \quad (\text{Col } A)^\perp = \text{Nul } A^T$$

2.2 Orthogonal Projections

An **orthogonal set** is a set of vectors which are mutually orthogonal to each other. An **orthogonal basis** is an **orthogonal set** which is a basis. Given an orthogonal basis $\{u_1, \cdots, u_p\}$ of W, every vector w in W can be expressed as

$$w = \frac{\langle w, u_1 \rangle}{\langle u_1, u_1 \rangle} u_1 + \frac{\langle w, u_2 \rangle}{\langle u_2, u_2 \rangle} u_2 + \cdots + \frac{\langle w, u_p \rangle}{\langle u_p, u_p \rangle} u_p$$

We can use this formula to find the orthogonal projection of a vector v (maybe outside of W) onto W.

Suppose that the orthogonal projection of v is $\hat{v}$. Then, $v - \hat{v}$ should be orthogonal to W since $\hat{v}$ is the ORTHOGONAL PROJECTION. It means that $\langle v - \hat{v}, u_i \rangle = 0$ for all i so that $\langle v, u_i \rangle = \langle \hat{v}, u_i \rangle$. Hence,

$$\begin{aligned}\hat{v} &= \frac{\langle \hat{v}, u_1 \rangle}{\langle u_1, u_1 \rangle} u_1 + \frac{\langle \hat{v}, u_2 \rangle}{\langle u_2, u_2 \rangle} u_2 + \cdots + \frac{\langle \hat{v}, u_p \rangle}{\langle u_p, u_p \rangle} u_p \\ &= \frac{\langle v, u_1 \rangle}{\langle u_1, u_1 \rangle} u_1 + \frac{\langle v, u_2 \rangle}{\langle u_2, u_2 \rangle} u_2 + \cdots + \frac{\langle v, u_p \rangle}{\langle u_p, u_p \rangle} u_p\end{aligned}$$

So, we have induced the formula for the orthogonal projection $\hat{v}$ of v onto W given an orthogonal basis.

What if the given basis is not orthogonal? We need to change the given basis into an orthogonal basis that spans the same space. For this, we need **Gram-Schmidt Orthogonalization Process**.

2.3 Orthogonalizing a basis : Gram-Schmidt

Given a basis $\{v_1, \cdots, v_p\}$ of a vector space V, we can find an orthogonal basis $\{u_1, \cdots, u_p\}$ such that

$$\begin{aligned}\text{Span}\{v_1\} &= \text{Span}\{u_1\} \\ \text{Span}\{v_1, v_2\} &= \text{Span}\{u_1, u_2\} \\ &\vdots \\ \text{Span}\{v_1, v_2, \cdots, v_p\} &= \text{Span}\{u_1, u_2, \cdots, u_p\}\end{aligned}$$

I believe that you already know the formula.

2.4 Projection map as a linear transformation

One more thing about the orthogonal projection is that it is a linear transformation from V to W. Imagine the picture of projecting $\mathbb{R}^3$ vectors onto a plane that goes through the origin point. It is straightforward that the PROJECTION MAP is a linear transformation from $\mathbb{R}^3$ to P(the plane). Hence, there should exist a matrix that corresponds to this linear transformation.

A projection map corresponds to a matrix B (such that $v \mapsto Bv$) and B is called the **projection matrix**.

One of properties is $B^2 = B$ since the orthogonal projection of a vector in W is again itself. In fact, if you define

$$B = u_1 u_1^T / ||u_1||^2 + u_2 u_2^T / ||u_2||^2 + \cdots + u_p u_p^T / ||u_p||^2$$

then, $Bv = \hat{v}$.

2.5 Best Approximation Theorem

Basically, what we are interested in is ;

To find the closest point in a subspace W from a point (possibly inside but usually) outside of W.

Just remember how we have proven that

All least-squares solutions for $A\mathbf{x} = \mathbf{b}$ are the same as all solutions for $A^T A\mathbf{x} = A^T\mathbf{b}$.