Numerical Methods

Course # PHY 203

Chapter #: 6: Least-Square Approximation

Lecture # 9.1: Orthonormality

Prepared by

Abu Mohammad Khan

Ch.6: Least-Square Approximation

- In this lecture, we will discuss about the orthogonality of the vectors.
- The main idea in this chapter is the following: We know a well-defined linear system has equal number of linear variables and numbers of conditions/equations.

$$\begin{array}{ccc} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}b_2 + \cdots + a_{2n}x_n & = & b_2 \\ \vdots & \vdots & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n & = & b_n \end{array}$$

- In matrix form , the above equation can be expressed as

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

- Here the $n \times 1$ column matrix x^T is the vector $\vec{x}$ which is unknown, and the right –hand side is also a constant vector $\vec{b}$ which in $n \times 1$ column matrix form.

- In other words, the transformation matrix, A, is a square matrix:

$$\boxed{A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}} \quad n \times n$$

- If $\det A \neq 0$, then there exits only one unique solution:

$$\boxed{x = A^{-1}b.}$$

- However, if we have a linear system where there are more equations to be satisfied than the number of variables, then how do we solve them. We call such a system an over-determined system.
- In such an over-determined system $x = A^{-1}b$ can not give any solution because the matrix A is not invertible (because it is no longer a square matrix).
- For an overdetermined system, the matrix A is now a $m \times n$ matrix. $\leftarrow \boxed{m > n}$
- Here m is the number of rows which is also same as the number of equations to be satisfied.
- And n is the number of columns, which is also the number of unknown variables in the system.

- Least-square approximation method is a way to find an approximate solution of an over-determined system where the transformation matrix A is no longer invertible.
- In an overdetermined system, we have the following matrix equations:

$$\boxed{m>n} \quad \underbrace{\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}}_{m\times n} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}_{n\times 1} = \underbrace{\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}}_{m\times 1} \iff \boxed{\underbrace{A}_{m\times n}\ \underbrace{x}_{n\times 1} = \underbrace{b}_{m\times 1}}$$

$q_1 \quad q_2 \quad q_n$

- Let's denote the $m \times n$ matrix by Q instead of A. $\leftarrow \boxed{Qx=b}$
- The first column of the matrix Q is q_1 and it is given by

$$q_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \qquad \text{which is a } m \times 1 \text{ column matrix.}$$

- Similarly, we write the other column matrices:

$$q_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \cdots, q_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}.$$

- So, we can express the matrix Q as

$$\boxed{Q = (\, q_1 \quad q_2 \quad \cdots \quad q_n \,)} \Leftrightarrow \boxed{Q = (q_1 | q_2 | \cdots | q_n)}$$

- Now we have a very important Theorem:
- **Theorem:** Let Q be an $m \times n$ matrix. The columns of Q form an orthonormal set if and only if

$$\boxed{Q^T Q = I_n.} \qquad I_n \equiv I_{n \times n} \text{ matrix}$$

- To understand the Theorem, we need to understand two concepts: one is the Orthogonality and the other is Unit vector (or norm one vector).

$Q^T \rightarrow n \times m \Rightarrow Q^T Q = (n \times m)(m \times n) = n \times n$ square matrix

- **Orthogonality**: Let $\vec{x}$ and $\vec{y}$ are two vectors in n-dimensional Euclidean vector space $\mathbb{R}^n$. That is: $\vec{x}, \vec{y} \in \mathbb{R}^n$.
- In matrix form, we identify an n-dimensional vector an $n \times 1$ column matrix.
- So, the inner product between any two vectors is expresses as

$$(1\times n)(n\times 1) = \boxed{(1\times 1)} = \vec{x}\cdot\vec{y} \equiv x^T y = \sum_{i=1}^{n} x_i\, y_i = \text{number}$$

$x = (\text{column})$

$x^T = (\text{row})$

which is known as the scalar or dot product in three dimensions.

- The inner product with itself is called the $\boxed{l_2\text{-norm},}$ or simply norm of the vector is given by:

$$|\vec{x}| = \sqrt{x^T x} = \sqrt{\vec{x}\cdot\vec{x}}$$

which is known as the absolute value or magnitude in three dimensions.

- If $\theta$is the angle between any two vectors, then we write:

$$x^T y = |\vec{x}|\,|\vec{y}|\cos\theta \equiv xy\cos\theta$$

$(1\times n)(n\times 1)$

- Two vectors, $\vec{x}$ and $\vec{y}$, are orthogonal if
$$x^T y = 0 \Longrightarrow \boxed{\theta = 90^0.}$$
- That is, they lie at right angles in $\mathbb{R}^n$. In three dimensions, we say that these two vector are perpendicular to each other.
- Let consider a set of n vectors in $\mathbb{R}^n$: $S = \{x_1, x_2, \cdots, x_n\}$.
- The set S is called the orthogonal set if
$$x_i^T x_j = 0 \ \forall\, i, j = 1, 2, \cdots, n \ \text{ with } \ i \neq j.$$
$\boxed{\vec{x}_i \cdot \vec{x}_j = 0 \text{ if } i \neq j}$
- Example: The set, $S = \{\hat{\imath}, \hat{\jmath}, \hat{k}\}$, is an orthogonal set of vectors in three dimensions because
$$\hat{\imath} \cdot \hat{\jmath} = 0. = \hat{j} \cdot \hat{k} = 0 = \hat{k} \cdot \hat{i}$$
- In matrix form, we can write the following:
$$\hat{\imath} = (1\ 0\ 0)^T, \text{ and } \hat{\jmath} = (0\ 1\ 0)^T.$$
$$\& \quad \hat{k} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = (0\ 0\ 1)^T$$

- On the other hand, in addition to orthogonality, we also have

$$\boxed{x_i^T x_i = 1,} \qquad \text{for each } i = 1,2,\cdots,n$$

We say that the vectors x_i has norm unity.

$\boxed{\vec{x}_i \cdot \vec{x}_i = 1} = |\vec{x}|$

- In other words, these are all unit vectors.
- The set of vectors in which every pair of vectors are orthogonal, and each vector has norm one, we say that the set is an orthonormal set.
- Mathematically, we write the orthonormal set as

$$\boxed{S = \{\, x_i \mid x_i \in \mathbb{R}^n, \boxed{x_i^T x_j = \delta_{ij}}, i,j = 1,2,\cdots,n\}.}$$

l_2- norm is one

vectors are norm 1 vectors

- Here the Kronecker delta δ_{ij} is defined as

$$\delta_{ij} = \begin{cases} 1, & i = j \Rightarrow \text{norm} = 1 \\ 0, & i \neq j \Rightarrow \text{orthogonal} \end{cases}$$

- The above concept will be necessary and very useful when in the next lecture, we discuss the QR-decomposition of an $m \times n$ matrix to solve a linear system.

- Example: The following set is orthonormal

$$S = \left\{ \frac{1}{\sqrt{5}}\,(2\,,1)^T, \frac{1}{\sqrt{5}}(1\,,-2)^T \right\}. \Rightarrow \text{is orthonormal.}$$

- Let $u = \frac{1}{\sqrt{5}}\,(2\,,1)^T$ and $v = \frac{1}{\sqrt{5}}\,(1\,,-2)^T$.
- We easily see that:

$$\vec{u}\cdot\vec{u} = u^T u = \left(\frac{1}{5}\right)(2^2 + 1^2) = 1.$$

$$(2\ \ 1)\begin{pmatrix}2\\1\end{pmatrix} = 2^2+1^2 = 5$$

$$\vec{v}\cdot\vec{v} = v^T v = \left(\frac{1}{5}\right)(1^2 + (-2)^2) = 1. \Rightarrow (1\ \ -2)\begin{pmatrix}1\\-2\end{pmatrix} = 1^2+(-2)^2 = 5$$

$$\vec{u}\cdot\vec{v} = \text{and } u^T v = \frac{1}{5}\,(2 \times 1\ - 1 \times 2) = 0. \quad (2\ \ 1)\begin{pmatrix}1\\-2\end{pmatrix} = 1\times 2 - 1\times 2 = 0$$

- Therefore, S is an orthonormal set in $\mathbb{R}^2$, and hence form a basis in two-dimensional vector space.