4

Orthogonality and Least Squares

4.1 Inner Product

Let $u = \begin{bmatrix} u_1 \\ u_2 \\ . \\ . \\ . \\ u_n \end{bmatrix}$ and $v = \begin{bmatrix} v_1 \\ v_2 \\ . \\ . \\ . \\ v_n \end{bmatrix}$ be any two vectors in $\mathbb{R}^n$. Then the number $u^T v$ is called the **inner product** of u and v. This inner product is also commonly known as **dot product** and denoted by **u.v**.

4.1.1 Properties of Inner Product

1. $u.v = v.u$
2. $(u + v).w = u.w + v.w$
3. $(\alpha u).v = \alpha(u.v) = u.(cv)$
4. $u.u \geq 0$ and $u.u = 0 \iff u = 0$

4.1.2 The Length of a Vector

The length of a vector v is called the **norm** of v.
It is denoted by $\|v\|$ and defined by $\|v\| = \sqrt{v_1^2 + v_2^2 + ... + v_n^2}$ so that, $\|v\|^2 = v.v$ There are several kinds of norms actually, this particular norm is called **Euclidean norm**. For any scalar α, $\|\alpha v\| = |\alpha|\|v\|$. A vector whose length is unity is called a **unit vector**. If we divide a nonzero vector v by its length, we obtain a unit vector u. This process is called **normalizing** of the vector v.

Distance between vectors

For u and v in a vector space V, the distance between them is written as dist(u,v) and is defined as $dist(u, v) = \|u - v\|$.

4.2 Orthogonal Vectors

The two vectors u and v are orthogonal vectors if their dot product is zero,i.e $u.v = 0$. Observe that the zero vector is orthogonal to every vector as $0^T v = 0$ for all v.

Theorem 1 (The Pythagorean Theorem)
Two vectors u and v are orthogonal if and only if $\|u + v\|^2 = \|u\|^2 + \|v\|^2$.

4.2.1 Orthogonal Complement

- If a vector z is orthogonal to every vectors in a subspace W then, z is said to be orthogonal to W.
- The set of all vectors that are orthogonal to W is called the **orthogonal complement** of W. It is denoted by $W^\perp$. $W^\perp = \{z : \forall v \in W \ z.v = 0\}$

Theorem 2
1. A vector x is in $W^\perp$ if and only if x is orthogonal to every vector in a set that is spans W.
2. $W^\perp$ is also a subspace.
3. Row space is orthogonal complement of the Null space for a matrix.

4.2.2 Orthogonal Sets

A set of vectors $\{u_1, ...u_p\}$ in a vector space V is said to be **orthogonal set** if each pair of distinct vectors from the set is orthogonal,i.e for all $u_i, u_j \in V$ we have $u_i.u_j = 0$ whenever $i \neq j$.

Definition 1. An orthogonal basis for a vector space V is a basis for V that is an orthogonal set.

Theorem 1. Let $\{u_1, ..., u_p\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$. For each y in W, the coordinates of y with respect to the orthogonal basis : $y = c_1 u_1 + \ ... \ c_n u_n$ are given by $c_j = \dfrac{y.u_j}{u_j.u_j}$.

4.2.3 Orthogonal Projection

The coordinate c_j of y in Theorem 1 is actually orthogonal projection of y into the vector u_j. This can be generalized. For any given vector u. The orthogonal projection of a vector y on u is given by the formula $\hat{y} = \dfrac{y.u}{u.u}$

Or, it can be derived as follows using the inner-product. $(y - \alpha u)$ and u are orthogonal so, $(y - \alpha u).u = 0$. This gives us $\alpha = \dfrac{y.u}{u.u}$. For two dimensional vectors, another orthogonal component z can be easily obtained as by subtracting the projection from the vector y. $z = y - \hat{y}$.

4.3 Orthonormal Sets

A set $\{u_1, ..., u_p\}$ is an orthonormal set if it is an orthogonal set of unit vectors. And a basis of orthonormal set is a orthonormal basis. The simplest orthonormal basis is $\{e_1,\ ...\ e_n$ for $\mathbb{R}^n$.

Matrices whose columns form an orthonormal set are important in applications and in computer algorithms for matrix computations.

Theorem 2. An $m \times n$ matrix U has orthonormal columns if and only if $U^t U = I$