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 $dist(u, v) = \|u - v\|$.

4.1.3 Exercise

1. Compute the following for the given vectors:
$u = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, v = \begin{bmatrix} 4 \\ 6 \end{bmatrix}, w = \begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix}, x = \begin{bmatrix} 6 \\ -2 \\ 3 \end{bmatrix}$

   a). $u.u, \;\; v.u, \;\; \frac{v.u}{u.u}, \;\; \|v\|$ b). $w.w, \;\; x.w, \;\; \frac{x.w}{w.w}, \;\; \|x\|$

2. Find the distance between u and v, and w and x.
3. Use matrix product and transpose definition to verify, property-2 and 3 of the inner product.
4. Explain why $u.u \geq 0$. When is $u.u = 0.?$

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 Exercise

1. Verify parallelogram law: $\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2$.
2. Suppose y is orthogonal to u and v. Show that y is orthogonal to every w in $Span\{u, v\}$.
3. Let W be a subspace of $\mathbb{R}^n$, then show that $W^\perp$ a subspace of $\mathbb{R}^n$.
4. Show that if x is in both W and $W^\perp$, then $x = 0$.

4.2.3 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 $i \neq j$.

Theorem 3. Any orthogonal set is a linearly independent set.

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

Theorem 4. 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 : $y = c_1u_1 + \ ... \ c_nu_n$ are given by $c_j = \dfrac{y.u_j}{u_j.u_j}$.

4.2.4 Orthogonal Projection

The coordinate c_j of y in Theorem 4 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 $\hat{y} = \dfrac{y.u}{u.u}u$

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, Important

Remark. A vector having length unity is called unit vector. Given a vector v if we divide it by its magnitude which its length we obtain a unit vector along v, i.e $\dfrac{v}{\|v\|}$.

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.

4.3.1 Unitary Matrix

A real matrix U is said to be unitary if $U^tU = I$. If U is complex then, $\overline{U}^{\,t}U = I$, where $\overline{U}^{\,t}$ is denoted by $U^\dagger$.

Theorem 5. An $m \times n$ matrix U has orthonormal columns if and only if $U^tU = I$

4.3.2 Exercise

1. Determine whether the following sets of vectors are orthogonal or not.

   (a) $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -5 \\ -2 \\ 1 \end{bmatrix}$

2. Show that $\{u_1, u_2\}$ is an orthogonal basis of $\mathbb{R}^2$ and find the coordinates of x in terms of this basis.

   (a) $u_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, u_2 = \begin{bmatrix} 6 \\ 4 \end{bmatrix}, x = \begin{bmatrix} 9 \\ -7 \end{bmatrix}$

   (b) $u_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, u_2 = \begin{bmatrix} 2 \\ -6 \end{bmatrix}, x = \begin{bmatrix} -6 \\ 3 \end{bmatrix}$

   (c) Compute the orthogonal projection of $\begin{bmatrix} 1 \\ 7 \end{bmatrix}$ onto to the line through $\begin{bmatrix} -4 \\ 2 \end{bmatrix}$ and the origin.

   (d) Compute the distance of $y \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ to the line through $u = \begin{bmatrix} 8 \\ 6 \end{bmatrix}$ and the origin.

3. Take a square unitary matrix of order 2 and a vector $x \in \mathbb{R}^2$ and verify that $\|Ux\| = \|x\|$. What can you infer from this property of a unitary matrix?

4. Show that the Hadamard quantum gate defined by $H = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$ is a unitary matrix.

4.4 The Gram-Schmidt Process

The Gram-Schmidt Process is a simple algorithm for producing an orthogonal or orthonormal basis for any nonzero subspace of $\mathbb{R}^n$. It uses the concept of **projection**. First we describe this process when a basis consisting of two vectors $\{x_1, x_2\}$ are given. Now we construct an orthogonal basis $\{v_1, v_2\}$ using the two vectors: x_1 and x_2 as follows:

1. Let $v_1 = x_1$.

2. Draw the orthogonal projection from x_2 to v_1, which is given by $\dfrac{x_2.v_1}{v_1.v_1} v_1$. Now, from the theory of orthogonal projection given above, the vector perpendicular to v_1 is given by $v_2 = x_2 - \dfrac{x_2.v_1}{v_1.v_1} v_1$.

For a given basis of p vectors $\{x_1, x_2, ..., x_p\}$, we can continue the process of drawing the orthogonal projections and subtraction to get orthogonal vectors as follows:

1. To obtain v_3 draw the orthogonal projection from x_3 onto the Spanv_1, v_2 and subtract this from x_3: $v_3 = x_3 - \dfrac{x_3.v_1}{v_1.v_1} v_1 - \dfrac{x_3.v_2}{v_2.v_2} v_2$.
2. $v_p = x_p - \dfrac{x_p.v_1}{v_1.v_1} v_1 - \dfrac{x_p.v_2}{v_2.v_2} v_2 - ... - \dfrac{x_p.v_{p-1}}{v_1.v_{p-1}} v_{p-1}$.

Remark. After constructing orthogonal basis orthonormal basis can be easily constructed from the orthogonal basis. *How?*

4.5 Exercise

1. Find an orthonormal basis from the given vectors:

a). $\begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 8 \\ 5 \\ 6 \end{bmatrix}$ b). $\begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} 5 \\ 6 \\ 7 \end{bmatrix}$

2. Find an orthogonal basis for the column space of the matrix:

a). $\begin{bmatrix} 3 & -5 & 1 \\ 1 & 1 & 1 \\ -1 & 5 & -2 \\ 3 & -7 & 8 \end{bmatrix}$ b). $\begin{bmatrix} -1 & 6 & 6 \\ 3 & -8 & 3 \\ 1 & -2 & 6 \\ 1 & -4 & -3 \end{bmatrix}$

3. Find the QR factorization of the matrices given in the previous question.

4.5.1 QR Factorization

Theorem 6. If A is an $m \times n$ matrix with linearly independent columns, then A can be factored as $A = QR$, where Q is an $m \times n$ matrix whose columns form an **orthonormal basis for** $Col(A)$ and R is an $n \times n$ **upper triangular matrix with positive entries on its diagonal**.

Proof. Given linearly independent columns of the matrix A, we can construct Q from the Gram-Schmidt process or any process. Then Q being real unitary matrix we have $Q^t\,Q = I$. Then $Q^t\,A = Q^t\,QR \implies R = Q^t\,A$. □

4.6 Extra Materials

The angle between the two vectors in plane and space can be generalized by the inner product. For two vectors $u, v \in \mathbb{R}^n$ we have, $u.v = \|u\|\,\|v\|\,cos\theta$, where θ is the angle between the two vectors: u, v.