Orthogonality

Definition:

Let $u = \begin{bmatrix} u_1 \\ . \\ . \\ u_n \end{bmatrix}$ and $v = \begin{bmatrix} v_1 \\ . \\ . \\ v_n \end{bmatrix}$ be two vectors in $\Re^n$.

Then the ***dot product*** of u and v is

$$u.v = u^T v = \begin{bmatrix} u_1 & . & . & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ . \\ . \\ v_n \end{bmatrix} = u_1 v_1 + ... + u_n v_n$$

Example: Let $u = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$ and $v = \begin{bmatrix} 4 \\ 6 \\ 2 \end{bmatrix}$.

$$u.v = u^T v = \begin{bmatrix} 3 & 1 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 6 \\ 2 \end{bmatrix} = 3.4 + 1.6 + 2.2 = 22$$

$$v.u = v^T u = \begin{bmatrix} 4 & 6 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} = 4.3 + 6.1 + 2.2 = 22$$

So, $u.v = v.u$

Properties:

Let u, v and w be vectors in $\Re^n$ and c be a scalar. Then,

1. $u.v = v.u$

2. $(u + v).w = u.w + v.w$

3. $(cu).v = c(u.v) = u.(cv)$

4. $u.u \geq 0$ and $u.u = 0 \Leftrightarrow u = 0$

Definition:

Let $u = \begin{bmatrix} u_1 \\ . \\ . \\ u_n \end{bmatrix}$ be a vector in $\Re^n$.

Then the ***length (norm)*** of u is defined as $\|u\| = \sqrt{u.u} = \sqrt{u_1^2 + ... + u_n^2}$.

A vector whose length is 1 is called a ***unit vector***.

For a given vector u, a unit vector in the direction of u is given by $\frac{u}{\|u\|}$.

(termed as normalizing a vector)

Example: Let $v = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$.

Find a the length of v and a unit vector v_1 in the direction of v.

Definition:

Let u and v be two vectors in $\Re^n$.

Then u and v are ***orthogonal*** to each other ($u \perp v$)

$$\text{if} \quad u.v = 0$$

Note: Zero vector is orthogonal to every vector.

Example:

1. Let u and v be two vectors in $\Re^n$. Show that

a) $u \perp v \Leftrightarrow \|u+v\|^2 = \|u\|^2 + \|v\|^2$

b) $u \perp v \Leftrightarrow \|u+v\|^2 = \|u-v\|^2$

Definition:

A set of non-zero vectors S is called an ***orthogonal set*** if each vector in S is orthogonal to every other vector in S.

i.e. $S = \{v_1, v_2, ...v_p\}$ is an orthogonal set $\Leftrightarrow$ $v_i.v_j = 0$ if $i \neq j$.

Example:

1. Does $\|u+v+z\|^2 = \|u\|^2 + \|v\|^2 + \|z\|^2$ imply that $\{u, v, z\}$ is an orthogonal set?

NOTE:

1. An orthogonal set in which each vector has length 1 is called an ***orthonormal set***.

2. A basis consisting of orthogonal vectors is called an ***orthogonal basis***.

3. A basis consisting of orthonormal vectors is called an ***orthonormal basis***.

Examples:

$\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\1\end{bmatrix}\right\}$ is an orthonormal basis for $\Re^3$.

$\left\{\begin{bmatrix}1\\-2\\3\end{bmatrix},\begin{bmatrix}-2\\2\\2\end{bmatrix},\begin{bmatrix}5\\4\\1\end{bmatrix}\right\}$ is an orthogonal basis for $\Re^3$.

$\left\{\begin{bmatrix}1/\sqrt{14}\\-2/\sqrt{14}\\3/\sqrt{14}\end{bmatrix},\begin{bmatrix}-2/\sqrt{12}\\2/\sqrt{12}\\2/\sqrt{12}\end{bmatrix},\begin{bmatrix}5/\sqrt{42}\\4/\sqrt{42}\\1/\sqrt{42}\end{bmatrix}\right\}$ is an orthonormal basis for $\Re^3$.

Theorem: Let $S=\{v_1,v_2,....v_p\}$ is an orthogonal set of nonzero vectors in $\Re^n$, then S is linearly independent.

Proof:

$S=\{v_1,v_2,....v_p\}$ is an orthogonal set of nonzero vectors.

Therefore, $v_i.v_j=0$ if $i\neq j$ and $\|v_1\|^2\neq 0$

Let $0=c_1v_1+c_2v_2+.......+c_pv_p$ (*) for some scalars $c_1,c_2,.....c_p$.

Then

$0.v_1=c_1v_1.v_1+c_2v_2.v_1+.......+c_pv_p.v_1$

$\Leftrightarrow 0=c_1v_1.v_1+0+.......+0$ $(v_i.v_j=0$ if $i\neq j)$

$\Leftrightarrow 0=c_1\|v_1\|^2 \quad \Leftrightarrow 0=c_1$ $(\|v_1\|^2\neq 0)$

Similarly, by taking the dot product to both the sides of (*) with vectors $v_2,....v_p$, we can show that $c_2=0,...,c_p=0$.

Thus S is linearly independent.

Theorem: Let $\{v_1, v_2, \ldots. v_n\}$ be an orthogonal basis for $\Re^n$.

Then for each $x \in \Re^n$,

$$x = c_1 v_1 + c_2 v_2 + \ldots\ldots + c_n v_n \text{ where}$$

$$c_i = \frac{x.v_i}{v_i.v_i}, \quad i = 1, 2, \ldots. n$$

Proof:

For a fixed $i, 1 \le i \le n$,

$$x.v_i = c_1 v_1.v_i + c_2 v_2.v_i + \ldots\ldots + c_n v_n.v_i = c_i v_i.v_i$$

$$\Leftrightarrow c_i = \frac{x.v_i}{v_i.v_i} \text{ (since } v_i.v_j = 0, \quad i \ne j)$$

Example:

Let $u_1 = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}$, $u_2 = \begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix}$, $u_3 = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}$ and $S = \{u_1, u_2, u_3\}$.

Express $x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ as a linear combination of the vectors in S.

Definition:

An $n \times n$ matrix Q whose columns form an orthonormal set is called an **orthogonal matrix**.

NOTE:

1. The rows of an orthogonal matrix is also orthonormal.

2. A square matrix Q is orthogonal if and only if $Q^{-1} = Q^T$.

3. If Q is an orthogonal matrix, then Q^{-1} is also orthogonal.

4. If Q is an orthogonal matrix, then $\det Q = \pm 1$.

5. If λ is an eigenvalue of an orthogonal matrix Q, then $|\lambda| = 1$.

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

Theorem: Let Q be an $n \times n$ matrix and let $x, y \in \Re^n$.

Then, the following statements are equivalent.

a) Q is orthogonal .

b) $\|Qx\| = \|x\|$ for every $x \in \Re^n$.

c) $Qx \cdot Qy = x \cdot y$ for every $x, y \in \Re^n$.

Example:

1. Let $U = \begin{bmatrix} -1/3 & 2/3 & 2/3 \\ 2/3 & -1/3 & 2/3 \\ 2/3 & 2/3 & -1/3 \end{bmatrix}$, $x = \begin{bmatrix} a_1 \\ b_1 \\ c_1 \end{bmatrix}$ and $y = \begin{bmatrix} a_2 \\ b_2 \\ c_2 \end{bmatrix}$.

Calculate $U^T U$, $\|Ux\|$, $Ux \cdot Uy$.

2. Show that $V = \begin{bmatrix} 1/\sqrt{5} & -2/\sqrt{5} & 0 \\ 0 & 0 & 1 \\ 2/\sqrt{5} & 1/\sqrt{5} & 0 \end{bmatrix}$ is an orthogonal matrix.

NOTE: For a matrix to be orthogonal it is not enough that the columns are *orthogonal*. For example, $V = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix}$

<u>Orthogonal Complement</u>

Definition:

Let U be a subspace of $\Re^n$.

1. If $z \in \Re^n$ is orthogonal to every vector in U, then z is said to be ***orthogonal*** to U.

2. The set of all z that are orthogonal to U is called the ***orthogonal complement*** of U. It is denoted by U^T.

Example:

1. Let $U = span\left\{\begin{bmatrix}1\\0\end{bmatrix}\right\}$. Then, $z = \begin{bmatrix}0\\1\end{bmatrix}$ is orthogonal to every vector in U and

$U^{\perp} = span\left\{\begin{bmatrix}0\\1\end{bmatrix}\right\}$.

2. Let $U = span\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}\right\}$. Then, $z = \begin{bmatrix}0\\0\\1\end{bmatrix}$ is orthogonal to every vector in U

and $U^{\perp} = span\left\{\begin{bmatrix}0\\0\\1\end{bmatrix}\right\}$.

3. Let $S = \left\{\begin{bmatrix}1\\2\\1\end{bmatrix}, \begin{bmatrix}1\\-1\\-1\end{bmatrix}\right\}$. Find $S^{\perp}$.

4. Show that if $y \perp u$ and $y \perp v$, then $y \perp (u+v)$.

Theorem: A vector x is in $W^{\perp}$ if and only if x is orthogonal to every vector in a set that spans W.

Theorem: Let A be an $m \times n$ matrix. Then,

$(RowA)^{\perp} = NulA$

$(ColA)^{\perp} = NulA^{T}$

Theorem: If W is a subspace of $\Re^{n}$, then $\dim W + \dim W^{\perp} = n$.

Example:

1. If x is in both W and $W^{\perp}$, then $x = O$.

2. Let $A = \begin{bmatrix} 3 & 2 \\ 6 & 4 \end{bmatrix}$. Find a vector which is in $(ColA)^{\perp}$.

3. Find the orthogonal complement $W^{\perp}$ of W and give a basis for $W^{\perp}$.

a) $W = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : 2x - 3y = 0 \right\}$.

b) $W = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : -x + 3y - 5z = 0 \right\}$

4. Find two linearly independent vectors which are orthogonal to $\begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix}$.

5. Find a basis for $W^{\perp}$, given that $W = span\left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix} \right\}$.

Orthogonal Projection of one vector onto another:

Definition: Let u be a non-zero vecto in $\Re^n$, and y be a vector in $\Re^n$.

$\hat{y} = \dfrac{y.u}{u.u}u$ is called the **orthogonal projection of y onto u.**

And $z = y - \hat{y}$ which is called the **component of y orthogonal to u.**

Example:

1. Let $y = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$ and $u = \begin{bmatrix} 7 \\ 1 \end{bmatrix}$. Write y as sum a vector in $span\{u\}$ and a vector orthogonal to u.

2. Let $y = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $u = \begin{bmatrix} 0 \\ 3 \\ 4 \end{bmatrix}$. Find the orthogonal projection of y onto u and the component of y orthogonal to u.

Orthogonal Projection of one vector onto a subspace of $\Re^n$:

Definition: Let W be a subspace of $\Re^n$ with orthogonal basis $\{u_1, u_2, ...u_p\}$ and let x be a vector in $\Re^n$.

Then, the orthogonal projection of x onto W is

$$proj_W x = \hat{x} = \frac{x.u_1}{u_1.u_1}u_1 + \frac{x.u_2}{u_2.u_2}u_2 + ... + \frac{x.u_p}{u_p.u_p}u_p$$

Theorem: Let W be a subspace of $\Re^n$. Then each vector $x \in \Re^n$ can be written uniquely in the form $x = \hat{x} + z$, where $\hat{x}$ is in W and z is in $W^{\perp}$.

In fact, $\hat{x}$ is the orthogonal projection of x onto W and $z = x - \hat{x}$.

$x = \hat{x} + z$ is called the orthogonal decomposition of x onto W.

Example:

1. Let $u_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, u_2 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$ and $W = span\{u_1, u_2\}$. Find the orthogonal decomposition of $x = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$ onto W.

2. Let $y = \begin{bmatrix} 6 \\ 3 \\ -2 \end{bmatrix}$, $u_1 = \begin{bmatrix} 3 \\ 4 \\ 0 \end{bmatrix}, u_2 = \begin{bmatrix} -4 \\ 3 \\ 0 \end{bmatrix}$.

Find the orthogonal projection of y onto the subspace spanned by the orthogonal vectors u_1 and u_2.

3. Let $y = \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix}$, $u_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, u_2 = \begin{bmatrix} -1 \\ 3 \\ -2 \end{bmatrix}$ and $W = span\{u_1, u_2\}$.

a) Find the orthogonal projection of y onto W.

b) Write y as the sum of a vector in W and a vector orthogonal to W.

Gram-Schimdt Orthogonalization process:

(The process of finding an orthogonal basis from any basis of a subspace of $\Re^n$)

Let $\{x_1, x_2, ...x_m\}$ is any basis of the subspace U of $\Re^n$. Define

$E_1 = X_1$,

$$E_2 = X_2 - \frac{X_2 \cdot E_1}{\|E_1\|^2} E_1$$

$$E_3 = X_3 - \frac{X_3 \cdot E_1}{\|E_1\|^2} E_1 - \frac{X_3 \cdot E_2}{\|E_2\|^2} E_2$$

.

.

.

$$E_k = X_k - \frac{X_k \cdot E_1}{\|E_1\|^2} E_1 - \frac{X_k \cdot E_2}{\|E_2\|^2} E_2 - ..... - \frac{X_k \cdot E_{k-1}}{\|E_{k-1}\|^2} E_{k-1}$$

for each $k = 2,3,.....m$.

Then,

a) $\{E_1, E_2, ...E_m\}$ is an orthogonal basis of U.

b) $span\{E_1, E_2, ...E_k\} = span\{X_1, X_2, ...X_k\}$ for each $k = 1,2,3,.....m$.

Example:

1. Let $X_1 = \begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix}, X_2 = \begin{bmatrix} 5 \\ 6 \\ 7 \end{bmatrix}$ be a basis for the subspace $W = span\{X_1, X_2\}$.

Find the orthogonal basis of W .

2. Let $X_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, X_2 = \begin{bmatrix} 3 \\ 2 \\ 0 \\ 1 \end{bmatrix}$ and $X_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$ be a basis for the subspace

$W = span\{X_1, X_2, X_3\}$. Find the orthogonal basis of W .

QR factorization:

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 with orthonormal columns and R is an invertible upper triangular matrix.

Alogorithm:

1. Suppose $A = \begin{bmatrix} C_1 & C_2 & . & . & . & C_n \end{bmatrix}$.

2. Apply Gram-Schimdt process to $C_1, C_2, ..., C_n$ to derive orthogonal columns $F_1, F_2, ..., F_n$.

3. Normalize the columns $F_1, F_2, ..., F_n$ to calculate $Q_1, Q_2, ..., Q_n$.

$$Q = \begin{bmatrix} Q_1 & Q_2 & . & . & . & Q_n \end{bmatrix}$$

4. (i) $R = Q^T A$

(ii) $R = \begin{bmatrix} \|F_1\| & C_2 \cdot Q_1 & C_3 \cdot Q_1 & . & . & . & C_n \cdot Q_1 \\ 0 & \|F_2\| & C_3 \cdot Q_2 & . & . & . & C_n \cdot Q_2 \\ 0 & 0 & \|F_3\| & . & . & . & C_n \cdot Q_3 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & . & \|F_n\| \end{bmatrix}$.

Example:

1. Let $A=\begin{bmatrix} 1 & -1 \\ 1 & 4 \\ 1 & -1 \\ 1 & 4 \end{bmatrix}$. Find the QR-factorization of the matrix A.

2. Let $A=\begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$. Find the QR-factorization of the matrix A.

Best approximation Theorem: Let W be a subspace of $\Re^n$ and y be a vector in $\Re^n$. Let $\hat{y}$ be the orthogonal projection of y onto W. Then $\hat{y}$ is the closest point in W to y.

i.e. $\|y-\hat{y}\| < \|y-v\|$ for all v in W distinct from $\hat{y}$.

Definition: The distance from a point y in $\Re^n$ to a subspace W is defined as the distance from y to the nearest point in W.

Example:

1. Let $y=\begin{bmatrix}2\\1\\2\\3\end{bmatrix}$, $u_1=\begin{bmatrix}1\\1\\-1\\-1\end{bmatrix}$, $u_2=\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix}$ and $W=span\{u_1,u_2\}$.

a) Find the nearest point in W to y.

b) Find the distance of y from W.

2. Find the point in the plane $2x+y-z=0$ that is closest to the point $P(2,-1,-3)$.

<u>Orthogonal Diagonalization of Symmetric matrix</u>

Definition: A square matrix A is ***orthogonally diagonalizable*** if there exists an orthogonal matrix Q and a diagonal matrix D s.t. $A = QDQ^T$

Definition: A symmetric matrix is a matrix s.t. $A^T = A$

Note:

1. A symmetric matrix is necessarily square.

2. Its main diagonal entries are arbitrary; but the entries on the opposite sides of the main diagonal occur in pairs.

$$\begin{bmatrix} * & \wedge & \vee \\ \wedge & \nabla & \circ \\ \vee & \circ & \Delta \end{bmatrix}$$

Theorem: (Spectral Theorem for Symmetric Matrix)

An $n \times n$ real symmetric matrix A has the following properties:

a. A has n real eigenvalues, counting the algebraic multiplicity.

(multiplicity of the eigenvalue as the root of the characteristic equation)

b. The geometric multiplicity (dimension of the eigenspace) for each eigenvalue equals to its algebraic multiplicity.

c. Eigenvectros corresponding to distinct eigenvalues are orthogonal.

d. A is orthogonally diagonalizable.

Definition: Let A be a $n \times n$ real symmetric matrix s.t. $A = QDQ^T$ where Q is an orthogonal matrix and D is a diagonal matrix. If $\lambda_1, \lambda_2, ..., \lambda_n$ are the diagonal entries of D and $q_1, q_2, ..., q_n$ are the columns of Q, then the **spectral decomposition** of A is given by

$$A = \lambda_1 q_1 q_1^T + \lambda_2 q_2 q_2^T + ... + \lambda_n q_n q_n^T$$

Example:

1. Orthogonally diagonalize the symmetric matrix $A = \begin{bmatrix} 7 & -4 & 4 \\ -4 & 5 & 0 \\ 4 & 0 & 9 \end{bmatrix}$.

Find the spectral decomposition of the matrix A.

2. Orthogonally diagonalize the symmetric matrix $A = \begin{bmatrix} 5 & -4 & -2 \\ -4 & 5 & 2 \\ -2 & 2 & 2 \end{bmatrix}$.

Find the spectral decomposition of the matrix A.

Quadratic Forms

Definition:

A ***quadratic form*** on $\Re^n$ is a function Q such that $Q(x) = x^T Ax$ for any $x \in \Re^n$, where A is an $n \times n$ symmetric matrix.

A is called the ***matrix of the quadratic form***.

Example:

1. Compute $Q(x) = x^T Ax$ for any $x \in \Re^n$, given that

i) $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ ii) $A = \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix}$ iii) $A = \begin{bmatrix} 1 & 3 & 4 \\ 3 & 2 & 5 \\ 4 & 5 & 3 \end{bmatrix}$.

2. For the given $Q(x) = x^T Ax$; $x \in \Re^n$, find the matrix of the quadratic form.

i) $Q(x) = 2x_1^2 - 6x_1x_2 + 5x_2^2$ ii) $Q(x) = x_1^2 + 2x_2^2 - 3x_3^2 + 2x_1x_2 - 4x_1x_3$

Change of variables:

Definition: If x represents a variable vector in $\Re^n$, then a change of variable is an equation of the form $x = Py$ or $y = P^{-1}x$ where P an invertible matrix is and y is a new variable vector in $\Re^n$.

The principal Axes theorem: Let A is a $n \times n$ symmetric matrix associated with the quadratic form $x^T Ax$ and $A = PDP^T$ s.t. P is an orthogonal matrix and D is a diagonal matrix. Then there is an orthogonal change of variable $x = Py$ that transforms the quadratic form $x^T Ax$ into a quadratic form $y^T Ay$ with no cross-product terms.

Definition/*Theorem*:

A quadratic form $f(x) = x^T Ax$, where A is an $n \times n$ symmetric matrix is classified as one of the following:

1. **positive definite** if $f(x) > 0$ for all $x \neq O$.

(if and only if all the eigenvalues of A are positive)

2. **positive semidefinite** if $f(x) \geq 0$ for all x.

(if and only if all the eigenvalues of A are nonnegative)

3. **negative definite** if $f(x) < 0$ for all $x \neq O$.

(if and only if all the eigenvalues of A are negative)

4. **negative semidefinite** if $f(x) \leq 0$ for all x.

(if and only if all the eigenvalues of A are nonositive)

5. **indefinite** if $f(x)$ taken on both positive and negative values.

(if and only if the eigenvalues of A are both positive and negative)

Example:

1. Make a change of variable that transforms the quadratic form $-3x_1^2 + 8x_1x_2 + 3x_2^2$ into a quadratic form with no cross product terms. Classify the quadratic form as positive definite, negative definite, indefinite or none of these.

2. Make a change of variable that transforms the quadratic form $7x_1^2 + 5x_2^2 + 9x_3^2 - 8x_1x_2 + 8x_1x_3$ into a quadratic form with no cross product terms. Classify the quadratic form as positive definite, negative definite, indefinite or none of these.

3. Make a change of variable that transforms the quadratic form $5x_1^2 + 5x_2^2 + 2x_3^2 - 8x_1x_2 - 4x_1x_3 + 4x_2x_3$ into a quadratic form with no cross product terms. Classify the quadratic form as positive definite, negative definite, indefinite or none of these.