Ortogonalidad

Credit: This notes are 100% from chapter 5 of the book entitled *Linear Algebra. A Modern Introduction* by David Poole. Thomson. Australia. 2006.

Orthogonality in $\mathbb{R}^n$

In this section, we will generalize the notion of orthogonality or vectors in $\mathbb{R}^n$ from two vectors to sets of vectors. In doing so, we will see that two properties make the standard basis $\{\bar{e}_1, \cdots, \bar{e}_n\}$ of $\mathbb{R}^n$ easy to work with: (i) any two distinct vectors in the set are orthogonal and (ii) each vector is a unit vector.

Orthogonal and Orthonormal Sets of Vectors

Orthogonal set: A set of vectors $\{\bar{v}_1, \cdots, \bar{v}_n\}$ in $\mathbb{R}^n$ is called orthogonal set if all pairs of distinct vectors in the set are orthogonal-that is, if

$$\bar{v}_i \cdot \bar{v}_j = 0 \tag{1}$$

whenever $i \neq j$, for $i, j = 1, \cdots, k$.

Theorem 5.1 If $\{\bar{v}_1, \cdots, \bar{v}_n\}$ is an orthogonal set of nonzero vectors in $\mathbb{R}^n$, then these vectors are LI.
Proof: See book, pag. 366.

Orthogonal basis: An orthogonal basis for a subspace W of $\mathbb{R}^n$ is a basis of W that is an orthogonal set.

Example 5.3: Find an orthogonal basis for the subspace W of $\mathbb{R}^3$ given by

$$W = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : x - y + 2z = 0 \right\} \tag{2}$$

A basis for W is given by

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} y - 2z \\ y \\ z \end{bmatrix} = y \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \tag{3}$$

but, they are not orthogonal.

Let us take one of them,

$$\bar{v} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \tag{4}$$

an orthogonal vector $\bar{w}$ to $\bar{v}$ must verifies

$$\bar{v} \cdot \bar{w} = v^T \bar{w} = [1\ 1\ 0] \begin{bmatrix} x \\ y \\ z \end{bmatrix} = x + y = 0 \tag{5}$$

Since w must also belong to the plane W, it must verifies

$$x - y + 2z = 0 \tag{6}$$

Solving the LSE

$$x + y = 0 \tag{7}$$

$$x - y + 2z = 0 \tag{8}$$

we get $y = -x$, $2z = -x + y = -2x \Rightarrow z = -x$, then

$$\bar{w} = t \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} \tag{9}$$

with t in $\mathbb{R}$. Finally,

$$\bar{v} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \text{and} \quad \bar{w} = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} \tag{10}$$

form and orthogonal basis for the subspace W with dimension two.

Theorem 5.2: Let $\{\bar{v}_1, \cdots, \bar{v}_k\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$ and let $\bar{w}$ be any vector in W. Then the unique scalars $c_1, \cdots, c_k$ such that

$$\bar{w} = c_1\bar{v}_1 + \cdots + c_k\bar{v}_k \tag{11}$$

are given by

$$c_i = \frac{\bar{w} \cdot \bar{v}_i}{\bar{v}_i \cdot \bar{v}_i} \tag{12}$$

for $i = 1, \cdots, k$.
Proof: Since $\{\bar{v}_1, \cdots, \bar{v}_k\}$ is a basis for W, we known from Theorem 3.29 that there are unique scalars $c_1, \cdots, c_k$ such that $\bar{w} = c_1\bar{v}_1 + \cdots + c_k\bar{v}_k$, then

$$\bar{w} \cdot \bar{v}_i = (c_1\bar{v}_1 + \cdots + c_k\bar{v}_k) \cdot \bar{v}_i \tag{13}$$

since $\bar{v}_j \cdot \bar{v}_i = \delta_{ji}\bar{v}_i \cdot \bar{v}_i$ results

$$\bar{w} \cdot \bar{v}_i = c_i(\bar{v}_i \cdot \bar{v}_i) \Rightarrow c_i = \frac{\bar{w} \cdot \bar{v}_i}{\bar{v}_i \cdot \bar{v}_i} \tag{14}$$

Orthonormal basis: A set of vectors in $\mathbb{R}^n$ is an orthonormal set if it is an orthogonal set of unit vectors. An orthonormal basis for a subspace W of $\mathbb{R}^n$ is a basis of W that is an orthonormal set.

Theorem 5.3: Let $\{\bar{q}_1, \cdots, \bar{q}_k\}$ be an orthonormal basis for a subspace W of $\mathbb{R}^n$ and let $\bar{w}$ be any vector in W. Then

$$\bar{w} = (\bar{w} \cdot \bar{q}_1)\bar{q}_1 + \cdots + (\bar{w} \cdot \bar{q}_k)\bar{q}_k \tag{15}$$

and this representation is unique.

Orthogonal Matrices

In this section we will examine the properties of matrices whose columns form an orthonormal set.

Theorem 5.4: The columns of an $m \times n$ matrix Q form an orthonormal set if and only $Q^TQ = I_n$.
Proof: We need to show that $(Q^TQ)_{ij} = \delta_{ij}$. Let $\bar{q}_i$, denote the ith column of Q. The (i, j) entry of Q^TQ is the dot product of the ith row of Q^T and the jth column of Q: $(Q^TQ)_{ij} = \bar{q}_i \cdot \bar{q}_j$. Now the columns Q form and orthonormal set if $\bar{q}_i \cdot \bar{q}_j = \delta_{ij} = (Q^TQ)_{ij}$.

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

Theorem 5.5: A square matrix Q is orthogonal if and only if $Q^{-1} = Q^T$.
Proof: By Theorem 5.4, Q is orthogonal if and only if $Q^TQ = I_n$. By Theorem 3.13 this is true if and only if Q is invertible and $Q^{-1} = Q^T$.

Exercise for the student in class: Show that the following two matrices are orthogonal and find their inverses:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix} \tag{16}$$

Theorem 5.6: Let Q be an $n \times n$ matrix. The following statements are equivalent:

a. Q is orthogonal

b. $\| Q\bar{x} \| = \| \bar{x} \|$ for every $\bar{x}$ in $\mathbb{R}$

c. $Q\bar{x} \cdot Q\bar{y} = \bar{x} \cdot \bar{y}$ for every $\bar{x}$ and $\bar{y}$ in $\mathbb{R}$

Proof: See book, pag. 372.

Theorem 5.7: If Q is an orthogonal matrix, then its rows form an orthonormal set.
Proof: From Theorem 5.5 $Q^{-1} = Q^T$. Therefore $(Q^T)^{-1} = (Q^{-1})^{-1} = Q = (Q^T)^T$ so Q^T is an orthogonal matrix. Thus, the columns of Q^T-which are the rows of Q-form an orthogonal set.

Theorem 5.8: Let Q be an orthogonal matrix.

a. Q^{-1} is orthogonal

b. det $Q = \pm 1$

c. If λ is an eigenvalue of Q, the $|\lambda| = 1$.

d. If Q_1 and Q_2 are orthogonal $n \times n$ matrices, then so is Q_1Q_2

Proof: See book, pag. 373.

Orthogonal Complements and Orthogonal Projections

In this section we will generalize the concepts of normal vector to a plane and will extend the concept of the projection of one vector onto another.

Orthogonal Complements

A normal vector $\bar{n}$ to a plane is orthogonal to every vector in that plane. If the plane passes through the origin, then it is a subspace W of $\mathbb{R}^3$, as is span($\bar{n}$). Hence, we have two subspaces of $\mathbb{R}^3$ with the property that every vector of one is orthogonal to every vector of the other.

Orthogonal Complement: Let W be a subspace of $\mathbb{R}^n$. We say that a vector $\bar{v}$ in $\mathbb{R}^n$ is orthogonal to W if $\bar{v}$ is orthogonal to every vector in W. The set of all vectors that are orthogonal to W is called the orthogonal complement of W, denoted $W^{\perp}$,

$$W^{\perp} = \{\bar{v} \text{ in } \mathbb{R}^n : \bar{v} \cdot \bar{w} = 0 \text{ in } W\} \tag{17}$$

Theorem 5.9: Let W be a subspace of $\mathbb{R}^n$.

a. $W^{\perp}$ is a subspace of $\mathbb{R}^n$.

b. $(W^{\perp})^{\perp} = W$

c. $W \cap W^{\perp} = \{\bar{0}\}$

d. If $W = span(\bar{w}_1, \cdots, \bar{w}_k)$, then $\bar{v}$ is in $W^{\perp}$ if and only if $\bar{v} \cdot \bar{w}_i = 0$ for all $i = 1, \cdots, k$

Proof: See book, pag. 376.

Theorem 5.10: Let A be an $m \times n$ matrix. Then the orthogonal complement of the row space of A is the null space of A

$$(row(A))^{\perp} = null(A) \tag{18}$$

and the orthogonal complement of the column space of A is the null space of A^T

$$(col(A))^{\perp} = null(A^T) \tag{19}$$

Proof: If $\bar{x}$ is a vector in $\mathbb{R}^n$, then $\bar{x}$ is in $(row(A))^{\perp}$ if and only if $\bar{x}$ is orthogonal to every row of A. But this is true if and only if $A\bar{x} = \bar{0}$, which is equivalent to $\bar{x}$ being in $null(A)$. To show the second identity we replace A by A^T and use the fact that $row(A^T) = col(A)$.

Exercise for the student in class(example 5.10): Consider the subspace W of $\mathbb{R}^5$ spanned by $\bar{w}_1, \bar{w}_2$ and $\bar{w}_3$

$$\bar{w}_1 = \begin{bmatrix} 1 \\ -3 \\ 5 \\ 0 \\ 5 \end{bmatrix}, \quad \bar{w}_2 = \begin{bmatrix} -1 \\ 1 \\ 2 \\ -2 \\ 3 \end{bmatrix}, \quad \bar{w}_3 = \begin{bmatrix} 0 \\ -1 \\ 4 \\ -1 \\ 5 \end{bmatrix} \tag{20}$$

The subspace W is the same as the column space of A, with

$$A = \begin{bmatrix} 1 & -1 & 0 \\ -3 & 1 & -1 \\ 5 & 2 & 4 \\ 0 & -2 & -1 \\ 5 & 3 & 5 \end{bmatrix} \tag{21}$$

Using $W^{\perp} = (col(A))^{\perp} = null(A^T)$, find a basis for $W^{\perp}$.

By reducing

$$[A^T|\bar{0}] = \left[\begin{array}{rrrrr|r} 1 & -3 & 5 & 0 & 5 & 0 \\ -1 & 1 & 2 & -2 & 3 & 0 \\ 0 & -1 & 4 & -1 & 5 & 0 \end{array}\right] \to \left[\begin{array}{rrrrr|r} 1 & 0 & 0 & 3 & 4 & 0 \\ 0 & 1 & 0 & 1 & 3 & 0 \\ 0 & 0 & 1 & 0 & 2 & 0 \end{array}\right] \tag{22}$$

then

$$x_1 + 3x_4 + 4x_5 = 0 \tag{23}$$
$$x_2 + x_4 + 3x_5 = 0 \tag{24}$$
$$x_3 + 2x_5 = 0 \tag{25}$$

then

$$\bar{x} = \begin{bmatrix} -3x_4 - 4x_5 \\ -x_4 - 3x_5 \\ -2x_5 \\ x_4 \\ x_5 \end{bmatrix} = x_4 \begin{bmatrix} -3 \\ -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -4 \\ -3 \\ -2 \\ 0 \\ 1 \end{bmatrix} \tag{26}$$

Then, a basis for $W^{\perp}$ is

$$\left\{ \begin{bmatrix} -3 \\ -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -4 \\ -3 \\ -2 \\ 0 \\ 1 \end{bmatrix} \right\} \tag{27}$$

Fundamental subspaces: an $m \times n$ matrix A has four subspaces, which are called fundamental subspaces:

1. $row(A)$
2. $null(A)$
3. $col(A)$
4. $null(A^T)$

The first two are orthogonal complements in $\mathbb{R}^n$, and the last two are orthogonal complements in $\mathbb{R}^m$

Transformation between subspaces of A: The $m \times n$ matrix A defines a LT from $\mathbb{R}^n$ into $\mathbb{R}^m$ whose range is $col(A)$. This transformation sends $null(A)$ to $\bar{0}$ in $\mathbb{R}^m$.

Exercise for the student in class(example 5.9): Find the fundamental subspaces of the matrix A, with

$$A = \begin{bmatrix} 1 & 1 & 3 & 1 & 6 \\ 2 & -1 & 0 & 1 & -1 \\ -3 & 2 & 1 & -2 & 1 \\ 4 & 1 & 6 & 1 & 3 \end{bmatrix} \tag{28}$$

Basis for row(A)

By reducing the matrix to its echelon form we get

$$R = \begin{bmatrix} 1 & 0 & 1 & 0 & -1 \\ 0 & 1 & 2 & 0 & 3 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{29}$$

By Theorem 3.20, row(A)=row(R). From the reduce matrix R one can see that the first three rows are LI. Then a basis for the row space of A is

$$\{[1\ 0\ 1\ 0\ -1], [0\ 1\ 2\ 0\ 3][0\ 0\ 0\ 1\ 4]\} \tag{30}$$

Basis for the col(A)

We obtain the column basis by selecting the vectors from the matrix A which correspond to the column of the reduced matrix R which correspond to the heads (pivots). The justification of this procedure is as follows: considering the system $A\bar{x} = 0$, the reduction from A to R represents a dependence relation among the columns of A. Since the elementary row operations do not affect the solution set, if A is row equivalent to R, the columns of A have the same dependence relationships as the columns of R. Then the columns $\bar{a}_1, \bar{a}_2, \bar{a}_4$ from a basis for the $col(A)$,

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ -3 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -2 \\ 1 \end{bmatrix} \right\} \tag{31}$$

Basis for $null(A)$

We have to find the solution of the homogeneous system $A\bar{x} = 0$ from the augmented matrix of A, $[A|\bar{0}] = \left[\begin{array}{ccccc|c} 1 & 1 & 3 & 1 & 6 & 0 \\ 2 & -1 & 0 & 1 & -1 & 0 \\ -3 & 2 & 1 & -2 & 1 & 0 \\ 4 & 1 & 6 & 1 & 3 & 0 \end{array}\right]$

From the previous calculation we have

$$[R|\bar{0}] = \left[\begin{array}{ccccc|c} 1 & 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \tag{32}$$

then,

$$x_1 + x_3 - x_5 = 0 \tag{33}$$
$$x_2 + 2x_3 + 3x_5 = 0 \tag{34}$$
$$x_4 + 4x_5 = 0 \tag{35}$$

Since the leading 1s are in columns 1, 2 and 4, we solve for x_1, x_2 and x_4. Let us renamed $x_3 = s$ and $x_5 = t$, then,

$$\bar{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = s \begin{bmatrix} -1 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ -3 \\ 0 \\ -4 \\ 1 \end{bmatrix} \tag{36}$$

Then, the following vectors form a basis for $null(A)$

$$\left\{ \begin{bmatrix} -1 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -3 \\ 0 \\ -4 \\ 1 \end{bmatrix} \right\} \tag{37}$$

Basis for $null(A^T)$

We have to find the solution of the homogeneous system $A^T\bar{x} = 0$ from the augmented matrix of A,

$$[A^T|\bar{0}] = \left[\begin{array}{cccc|c} 1 & 2 & -3 & 4 & 0 \\ 1 & -1 & 2 & 1 & 0 \\ 3 & 0 & 1 & 6 & 0 \\ 1 & 1 & -2 & 1 & 0 \\ 6 & -1 & 1 & 3 & 0 \end{array}\right] \rightarrow \left[\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 6 & 0 \\ 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \tag{38}$$

then,

$$y_1 + y_4 = 0 \tag{39}$$

$$y_2 + 6y_4 = 0 \tag{40}$$

$$y_3 + 3y_4 = 0 \tag{41}$$

then

$$\bar{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} -y_4 \\ -6y_4 \\ -3y_4 \\ y_4 \end{bmatrix} \tag{42}$$

Then, the following vectors form a basis for $null(A^T)$

$$\left\{ \begin{bmatrix} -1 \\ -6 \\ -3 \\ 1 \end{bmatrix} \right\} \tag{43}$$

Orthogonal Projections

Let W be subspace of $\mathbb{R}^n$ and let $\{\bar{u}_1, \cdots, \bar{u}_k\}$ be an orthogonal basis for W. For any vector $\bar{v}$ in $\mathbb{R}^n$, the orthogonal projection of $\bar{v}$ onto W is defined as

$$proj_W(\bar{v}) = \left(\frac{\bar{u}_1 \cdot \bar{v}}{\bar{u}_1 \cdot \bar{u}_1}\right)\bar{u}_1 + \cdots + \left(\frac{\bar{u}_k \cdot \bar{v}}{\bar{u}_k \cdot \bar{u}_k}\right)\bar{u}_k \tag{44}$$

The component of $\bar{v}$ orthogonal to W is the vector

$$perp_W(\bar{v}) = \bar{v} - proj_W(\bar{v}) \tag{45}$$

The $proj_W(\bar{v})$ can be written in terms of projections onto single vectors, i.e. one-dimensional subspace, then

$$proj_W(\bar{v}) = proj_{\bar{u}_1}(\bar{v}) + \cdots + proj_{\bar{u}_k}(\bar{v}) \tag{46}$$

Geometric interpretation of Theorem 5.2: As a special case of the definition of $proj_W(\bar{v})$ we can give a geometric interpretation to the Theorem 5.2 which states that if $\bar{w}$ is in the subspace W of $\mathbb{R}^n$, which has orthogonal basis $\{\bar{v}_1, \cdots, \bar{v}_k\}$ then

$$\begin{aligned} \bar{w} &= \left(\frac{\bar{v}_1 \cdot \bar{w}}{\bar{v}_1 \cdot \bar{v}_1}\right) \bar{v}_1 + \cdots + \left(\frac{\bar{v}_k \cdot \bar{w}}{\bar{v}_k \cdot \bar{v}_k}\right) \bar{v}_k & (47) \\ &= proj_{\bar{v}_1}(\bar{w}) + \cdots + proj_{\bar{v}_k}(\bar{w}) & (48) \end{aligned}$$

Thus, $\bar{w}$ is decomposed into a sum of orthogonal projections onto mutually orthogonal one-dimensional subspaces of W.

Comment: The definition above seems to depend on the choice of orthogonal basis; that is, a different basis $\{\bar{v}'_1, \cdots, \bar{v}'_k\}$ for W we would have a different $proj_W(\bar{v})$ and $perp_W(\bar{v})$. This is not the case.

Exercise for the student in class: Let W be the plane in $\mathbb{R}^3$ with equation $x - y + 2z = 0$ and let $\bar{v} = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}$. Find the orthogonal projection of $\bar{v}$ onto W and the component of $\bar{v}$ orthogonal to W. (this is the same subspace as the Example 5.3)

Orthogonal projection of $\bar{v}$ onto W

- First we find two vectors which expand the subspace

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} y - 2z \\ y \\ z \end{bmatrix} = y \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \tag{49}$$

  Then, the two vectors are

$$\bar{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \bar{v}_2 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \tag{50}$$

  Since, $\bar{v}_1$ and $\bar{v}_2$ are not orthogonal we take one of them and search one orthogonal include in the plane.

- Next we search for a vector which is orthogonal to $\bar{v}_1$

$$0 = \bar{v}_1 \cdot \bar{v}_3 = (\bar{v}_1)^T \bar{v}_3 = [1\ 1\ 0] \begin{bmatrix} x \\ y \\ z \end{bmatrix} \Rightarrow x + y = 0 \tag{51}$$

- Next we make that the found vector belong to the plane $x - y + 2z = 0$ by solving the system

$$\begin{aligned} x + y &= 0 \quad (52) \\ x - y + 2z &= 0 \quad (53) \end{aligned}$$

We get $y = -x$, $2z = -x + y = -2x \Rightarrow z = -x$, then

$$\bar{v}_3 = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} \quad (54)$$

In order to compare with the book, let us take

$$\bar{v}_3 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \quad (55)$$

- Now that we have the orthonormal vectors $\bar{v}_1$ and $\bar{v}_3$ we project the vector $\bar{v}$ in this orthogonal basis

$$proy_W(\bar{v}) = \left(\frac{\bar{v}_1 \cdot \bar{v}}{\bar{v}_1 \cdot \bar{v}_1}\right)\bar{v}_1 + \left(\frac{\bar{v}_3 \cdot \bar{v}}{\bar{v}_3 \cdot \bar{v}_3}\right)\bar{v}_3 \quad (56)$$

with

$$\begin{aligned} \bar{v}_1 \cdot \bar{v} &= 2 \quad (57) \\ \bar{v}_3 \cdot \bar{v} &= -2 \quad (58) \\ \bar{v}_1 \cdot \bar{v}_1 &= 2 \quad (59) \\ \bar{v}_3 \cdot \bar{v}_3 &= 3 \quad (60) \end{aligned}$$

then

$$proy_W(\bar{v}) = \frac{2}{2}\bar{v}_1 + \frac{-2}{3}\bar{v}_3 \quad (61)$$

$$proy_W(\bar{v}) = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \frac{2}{3}\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \quad (62)$$

$$proy_W(\bar{v}) = \begin{bmatrix} 5/3 \\ 1/3 \\ -2/3 \end{bmatrix} \quad (63)$$

Component of $\bar{v}$ orthogonal to W

$$perp_W(\bar{v}) = \bar{v} - proy_W(\bar{v}) = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix} - \begin{bmatrix} 5/3 \\ 1/3 \\ -2/3 \end{bmatrix} = \begin{bmatrix} 4/3 \\ -4/3 \\ 8/3 \end{bmatrix} \quad (64)$$

Theorem 5.11: The Orthogonal Decomposition Theorem. Let W be a subspace of $\mathbb{R}^n$ and let $\bar{v}$ be a vector in $\mathbb{R}^n$. Then there are unique vectors $\bar{w}$ and $\bar{w}^\perp$ in $W^\perp$ such that

$$\bar{v} = \bar{w} + \bar{w}^\perp \tag{65}$$

Proof: See book, pag. 381.

Example: from the previous example we have

$$\bar{v} = \bar{w} + \bar{w}^\perp \tag{66}$$

with

$$\bar{w} = proj_W(\bar{v}) \tag{67}$$

and

$$\bar{w}^\perp = perp_W(\bar{v}) \tag{68}$$

By making the sum we get

$$\bar{w} + \bar{w}^\perp = \begin{bmatrix} 5/3 \\ 1/3 \\ -2/3 \end{bmatrix} + \begin{bmatrix} 4/3 \\ -4/3 \\ 8/3 \end{bmatrix} = \begin{bmatrix} 9/3 \\ -3/3 \\ 6/3 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix} = \bar{v} \tag{69}$$

Theorem 5.13: If W is a subspace of $\mathbb{R}^n$, then

$$dim\, W + dim\, W^\perp = n \tag{70}$$

Proof: See book, pag. 383.

The Gram-Schmidt Process and the QR Factorization

In this section, we present a simple method for constructing an orthogonal/orthonormal basis for any subspace of $\mathbb{R}^n$.

The Gram-Schmidt Process

We want to find an orthonormal basis for a subspace W of $\mathbb{R}^n$ from an arbitrary basis $\{\bar{x}_1, \cdots, \bar{x}_k\}$ of W.

Motivation: Example 5.12. Let $W = span(\bar{x}_1, \bar{x}_2\}$, where $\bar{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\bar{x}_2 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$.
Construct an orthogonal basis for W.

Let us start from $\bar{x}_1$ (starting from $\bar{x}_2$ we would get a different pair of orthogonal vectors, **Check it!!**), then

- Let us define $\bar{v}_1 = \bar{x}_1$
- The projection of $\bar{x}_2$ over $\bar{x}_1$ is parallel to $\bar{x}_1$: $proj_{\bar{x}_1}(\bar{x}_2)$
- Then, the orthogonal to $proj_{\bar{x}_1}(\bar{x}_2)$ is also orthogonal to $\bar{x}_1$: $perp_{\bar{x}_1}(\bar{x}_2) = \bar{x}_2 - proj_{\bar{x}_1}(\bar{x}_2)$

- Then, $\bar{v}_2 = perp_{\bar{x}_1}(\bar{x}_2) = \bar{x}_2 - proj_{\bar{x}_1}(\bar{x}_2)$

$$\bar{v}_2 = \bar{x}_2 - \left(\frac{\bar{x}_1 \cdot \bar{x}_2}{\bar{x}_1 \cdot \bar{x}_1}\right)\bar{x}_1 \quad (71)$$

$$= \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} - \left(\frac{-2}{2}\right)\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \quad (72)$$

Then $\{\bar{v}_1, \bar{v}_2\}$ is an orthogonal set for W. The normalized basis reads $\{\bar{q}_1, \bar{q}_2\}$ with

$$\bar{q}_1 = \frac{\bar{v}_1}{||\bar{v}_1||} \quad (73)$$

$$\bar{q}_2 = \frac{\bar{v}_2}{||\bar{v}_2||} \quad (74)$$

where $||\bar{v}_1|| = \sqrt{2}$ and $||\bar{v}_2|| = \sqrt{3}$. Then

$$\bar{q}_i \cdot \bar{q}_j = \delta_{ij} \quad (75)$$

Check it!!

Theorem 5.15: The Gram-Schmidt Process. Let $\{\bar{v}_1, \cdots, \bar{v}_k\}$ be a basis for a subspace W of $\mathbb{R}^n$ and define the following:

$$\begin{array}{ll} \bar{v}_1 = \bar{x}_1 & W_1 = span(\bar{x}_1) \\ \bar{v}_2 = \bar{x}_2 - \left(\frac{\bar{v}_1 \cdot \bar{x}_2}{\bar{v}_1 \cdot \bar{v}_1}\right)\bar{v}_1 & W_2 = span(\bar{x}_1, \bar{x}_2) \\ \vdots & \\ \bar{v}_k = \bar{x}_k - \left(\frac{\bar{v}_1 \cdot \bar{x}_k}{\bar{v}_1 \cdot \bar{v}_1}\right)\bar{v}_1 - \cdots - \left(\frac{\bar{v}_{k-1} \cdot \bar{x}_k}{\bar{v}_{k-1} \cdot \bar{v}_{k-1}}\right)\bar{v}_{k-1} & W_k = span(\bar{x}_1, \cdots, \bar{x}_k) \end{array} \quad (76)$$

Then for each $i = 1, \cdots, k$, $\{\bar{v}_1, \cdots, \bar{v}_i\}$ is an orthogonal basis for W_i. In particular, $\{\bar{v}_1, \cdots, \bar{v}_k\}$ is an orthogonal basis for W.
Proff: See book, pag. 387.

Comments related to Theorem 5.15

- The theorem states that every subspace of $\mathbb{R}^n$ has an orthogonal basis and it gives an algorithm for constructing such a basis.
- If we require an orthonormal basis we normalize the orthogonal vectors produced by the Gram-Schmidt Process. That is, for each i, we replace $\bar{v}_i$ by the unit vector $\bar{q}_i = \bar{v}_i/||\bar{v}_i||$.

Exercise for the student in class (Example 5.13): Apply the Gram-Schmidt Process (starting from $\bar{x}_1$) to construct an orthonormal basis for the subspace $W = span(\bar{x}_1, \bar{x}_2, \bar{x}_3)$ of $\mathbb{R}^4$, where

$$\bar{x}_1 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \quad \bar{x}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \quad \bar{x}_3 = \begin{bmatrix} 2 \\ 2 \\ 1 \\ 2 \end{bmatrix} \quad (77)$$

Solution:

- First we must check that $\bar{x}_1, \bar{x}_2, \bar{x}_3$ are LI. In this case they are LI.
- If They are not, we through away any of LD vectors.
- Next we apply the GS procedure. We should get

$$\bar{v}_1 = \bar{x}_1 \tag{78}$$

$$\bar{v}_2 = \bar{x}_2 - \left(\frac{\bar{v}_1 \cdot \bar{x}_2}{\bar{v}_1 \cdot \bar{v}_1}\right)\bar{v}_1 = \begin{bmatrix} 3/2 \\ 3/2 \\ 1/2 \\ 1/2 \end{bmatrix} \tag{79}$$

Let us use a scale $\bar{v}_2' = 2\bar{v}_2$

$$\bar{v}_3 = \bar{x}_3 - \left(\frac{\bar{v}_1 \cdot \bar{x}_3}{\bar{v}_1 \cdot \bar{v}_1}\right)\bar{v}_1 - \left(\frac{\bar{v}_2' \cdot \bar{x}_3}{\bar{v}_2' \cdot \bar{v}_2'}\right)\bar{v}_2' = \begin{bmatrix} -1/2 \\ 0 \\ 1/2 \\ 1 \end{bmatrix} \tag{80}$$

Let us use a scale $\bar{v}_3' = 2\bar{v}_3$.
Then, an orthogonal basis for W is $\{\bar{v}_1, \bar{v}_2', \bar{v}_3'\}$.

- Finally we normalize the basis vectors. We should get

$$\bar{q}_1 = \frac{\bar{v}_1}{||\bar{v}_1||} = \begin{bmatrix} 1/2 \\ -1/2 \\ -1/2 \\ 1/2 \end{bmatrix} \tag{81}$$

$$\bar{q}_2 = \frac{\bar{v}_2'}{||\bar{v}_2'||} = \begin{bmatrix} 3\sqrt{5}/10 \\ 3\sqrt{5}/10 \\ \sqrt{5}/10 \\ \sqrt{5}/10 \end{bmatrix} \tag{82}$$

$$\bar{q}_3 = \frac{\bar{v}_3'}{||\bar{v}_3'||} = \begin{bmatrix} -\sqrt{6}/6 \\ 0 \\ \sqrt{6}/6 \\ \sqrt{6}/3 \end{bmatrix} \tag{83}$$

The QR Factorization

If A is an $m \times n$ $(m \geq n)$ matrix with linearly independent columns $\bar{a}_i$, with $i = 1, \cdots, n$, then applying the Gram-Schmidt Process to theses columns yields orthonormal vectors $\bar{q}_i$, with $i = 1, \cdots, n$. From Theorem 5.15, for each $i = 1, \cdots, n$,

$$W_i = span(\bar{a}_1, \cdots, \bar{a}_i) = span(\bar{q}_1, \cdots, \bar{q}_i) \tag{84}$$

Therefore, there are scalars $r_{1i}, \cdots, r_{ii}$ such that

$$\bar{a}_i = r_{1i}\bar{q}_1 + \cdots + r_{ii}\bar{q}_i \tag{85}$$

for each $i = 1, \cdots, n$. That is,

$$\bar{a}_1 = r_{11}\bar{q}_1 \tag{86}$$
$$\bar{a}_2 = r_{12}\bar{q}_1 + r_{22}\bar{q}_2 \tag{87}$$
$$\vdots \tag{88}$$
$$\bar{a}_n = r_{1n}\bar{q}_1 + \cdots + r_{nn}\bar{q}_n \tag{89}$$

in matrix form it read,

$$A = [\bar{a}_1 \ \cdots \ \bar{a}_n] = [\bar{q}_1 \ \cdots \ \bar{q}_n] \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{bmatrix} = QR \tag{90}$$

where Q is a matrix $m \times n$ which has orthonormal columns

$$Q = [\bar{q}_1|\bar{q}_2|\cdots|\bar{q}_n] \tag{91}$$

and R is an invertible (see Exercise 23, pag. 392 in the book) upper triangular matrix $n \times n$ with non zero diagonal entries,

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{bmatrix} \tag{92}$$

Theorem 5.16: The QR Factorization. Let A be 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.

Exercise for the student in class (Example 5.15): Find a QR factorization of

$$A = \begin{bmatrix} 1 & 2 & 2 \\ -1 & 1 & 2 \\ -1 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix} \tag{93}$$

Solution:

- First we identify the columns of A with vector

$$\bar{x}_1 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \quad \bar{x}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \quad \bar{x}_3 = \begin{bmatrix} 2 \\ 2 \\ 1 \\ 2 \end{bmatrix} \tag{94}$$

- Next we check if the are LI

- Next we orthonormalized these vectors. I was done in the previous example:

$$\bar{q}_1 = \frac{\bar{v}_1}{||\bar{v}_1||} = \begin{bmatrix} 1/2 \\ -1/2 \\ -1/2 \\ 1/2 \end{bmatrix} \tag{95}$$

$$\bar{q}_2 = \frac{\bar{v}_2'}{||\bar{v}_2'||} = \begin{bmatrix} 3\sqrt{5}/10 \\ 3\sqrt{5}/10 \\ \sqrt{5}/10 \\ \sqrt{5}/10 \end{bmatrix} \tag{96}$$

$$\bar{q}_3 = \frac{\bar{v}_3'}{||\bar{v}_3'||} = \begin{bmatrix} -\sqrt{6}/6 \\ 0 \\ \sqrt{6}/6 \\ \sqrt{6}/3 \end{bmatrix} \tag{97}$$

- Next, we build the matrix Q with these orthonormal vectors.

$$Q = [\bar{q}_1|\bar{q}_2|\bar{q}_3] \tag{98}$$

$$= \begin{bmatrix} 1/2 & 3\sqrt{5}/10 & -\sqrt{6}/6 \\ -1/2 & 3\sqrt{5}/10 & 0 \\ -1/2 & \sqrt{5}/10 & \sqrt{6}/6 \\ 1/2 & \sqrt{5}/10 & \sqrt{6}/6 \end{bmatrix} \tag{99}$$

- Finally, we calculate the R matrix from

$$A = QR \Rightarrow Q^T A = R \tag{100}$$

It should gives,

$$R = \begin{bmatrix} 2 & 1 & 1/2 \\ 0 & \sqrt{5} & 3\sqrt{5}/2 \\ 0 & 0 & \sqrt{6}/2 \end{bmatrix} \tag{101}$$

Orthogonal Diagonalization of Symmetric Matrices

A square matrix with real entries will not necessarily have real eigenvalues. We also found that not all square matrices are diagonalizable. The situation changes dramatically if we restrict our attention to real symmetric matrices ($A = A^T$).

Orthogonally Diagonalizable Matrix: A square matrix A is orthogonally diagonalizable if there exist an orthogonal matrix Q and a diagonal matrix D such that $Q^T AQ = D$.

Theorem 5.17: If A is orthogonally diagonalizable, then A is symmetric.
Proof: If A is orthogonally diagonalizable, then there exist an orthogonal matrix Q and a diagonal matrix D such that $Q^T AQ = D$. Since $Q^{-1} = Q^T$, we have $Q^T Q = I = QQ^T$, so

$$QDQ^T = QQ^T AQQ^T = IAI = A \tag{102}$$

But then

$$A^T = (QDQ^T)^T = QDQ^T = A \Rightarrow A^T = A \tag{103}$$

Theorem 5.18: If A is a real symmetric matrix, then the eigenvalues of A are real.
Proof: See book, pag. 398.

Theorem 5.19: If A is a symmetric matrix, then any two eigenvectors corresponding to distinct eigenvalues of A are orthogonal.
Proof: See book, pag. 399.

Theorem 5.20: The (real) Spectral Theorem. Let A be an $n \times n$ real matrix. Then A is symmetric if and only if it is orthogonally diagonalizable.
Proof: See book, pag. 399.

Exercise for the student in class (Example 5.18): Orthogonally diagonalize the matrix
$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$

Solution:

- First we calculate the eigenvalue from $det(A - \lambda I)$. We should get $\lambda_1 = 4$ and $\lambda_2 = 1$
- Next we calculate the eigenvectors for each eigenvalue from the homogeneous solution $(A - \lambda_i I)\bar{v} = 0$, for $i = 1, 2$. We should get

$$E_4 = span\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right) \tag{104}$$

$$E_1 = span\left(\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}\right) \tag{105}$$

- Since the two eigenvectors of $\lambda = 1$ are not orthogonal we apply the Gram-Schmidt method starting from the first eigenvector. We should get

$$\begin{bmatrix} -1/2 \\ 1 \\ -1/2 \end{bmatrix} \tag{106}$$

- Next we normalized the three orthogonal eigenvectors and build the orthogonal matrix Q. We should get

$$Q = \begin{bmatrix} 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \end{bmatrix} \tag{107}$$

- Finally we make the matrix products. We should get

$$Q^T A Q = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{108}$$

which is the usual matrix D build from the eigenvalues!.

Spectral Decomposition. The Spectral Theorem allows us to write a real symmetric matrix A in the form

$$A = QDQ^T \tag{109}$$

where Q is orthogonal form by the eigenvectors of A as columns and D is diagonal form by the eigenvalues of A in the same order as the eigenvectors in Q,

$$\begin{aligned} A &= QDQ^T && (110) \\ &= [\bar{q}_1 \ \cdots \ \bar{q}_n] \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} \bar{q}_1^T \\ \vdots \\ \bar{q}_n^T \end{bmatrix} && (111) \\ &= [\lambda_1\bar{q}_1 \ \cdots \ \lambda_n\bar{q}_n] \begin{bmatrix} \bar{q}_1^T \\ \vdots \\ \bar{q}_n^T \end{bmatrix} && (112) \\ &= \lambda_1\bar{q}_1\bar{q}_1^T + \cdots + \lambda_n\bar{q}_n\bar{q}_n^T && (113) \end{aligned}$$

This is called the spectral decomposition of A. Each of the terms $\lambda_i\bar{q}_i\bar{q}_i^T$ is a rank 1 matrix, by Exercise 56 in Section 3.5, and $\bar{q}_i\bar{q}_i^T$ is the matrix of the projection onto the subspace spanned by $\bar{q}_i$. (See Exercise 25, pag. 405). For this reason, the spectral decomposition

$$A = \lambda_1\bar{q}_1\bar{q}_1^T + \cdots + \lambda_n\bar{q}_n\bar{q}_n^T \tag{114}$$

is sometimes referred to as the **projection form of the Spectral Theorem.**

Application: Example 5.20. Find a 2×2 matrix with eigenvalues $\lambda_1 = 3$ and $\lambda_2 = -2$ and corresponding eigenvalues

$$\bar{v}_1 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \quad \bar{v}_2 = \begin{bmatrix} -4 \\ 3 \end{bmatrix} \tag{115}$$

Solution

- Firs we normalized the eigenvectors

$$\bar{q}_1 = \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix}, \quad \bar{q}_2 = \begin{bmatrix} -4/5 \\ 3/5 \end{bmatrix} \tag{116}$$

- Next we compute $\bar{q}_i\bar{q}_i^T$ for $i = 1, 2$. We should get

$$\begin{aligned} \bar{q}_1\bar{q}_1^T &= \begin{bmatrix} 9/25 & 12/25 \\ 12/25 & 16/25 \end{bmatrix} && (117) \\ \bar{q}_2\bar{q}_2^T &= \begin{bmatrix} 16/25 & -12/25 \\ -12/25 & 9/25 \end{bmatrix} && (118) \end{aligned}$$

- Finally we compute the matrix A in its spectral decomposition. We should get

$$\begin{aligned} A &= \lambda_1\bar{q}_1\bar{q}_1^T + \lambda_2\bar{q}_2\bar{q}_2^T && (119) \\ &= \begin{bmatrix} -1/5 & 12/5 \\ 12/5 & 6/5 \end{bmatrix} && (120) \end{aligned}$$

Applications

Quadratic Forms

A quadratic from is x, y (and z) is a sum of terms, each of which has total degree two in the variables, i.e. $ax^2 + by^2(+cz^2) + dxy(+exz + fyz)$ in 2 (or 3) variables.

Quadratic forms can be represented using matrices

$$ax^2 + by^2 + cxy = [x\ y] \begin{bmatrix} a & c/2 \\ c/2 & b \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{121}$$

and

$$ax^2 + by^2 + cz^2 + dxy + exz + fyz = [x\ y\ z] \begin{bmatrix} a & d/2 & e/2 \\ d/2 & b & f/2 \\ e/2 & f/2 & c \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{122}$$

Quadratic Form and its associated matrix: A quadratic form in n variables is a function $f : \mathbb{R}^n \to \mathbb{R}$ of the form

$$f(\bar{x}) = \bar{x}^T A \bar{x} \tag{123}$$

where A is a symmetric $n \times n$ matrix and $\bar{x}$ is in $\mathbb{R}^n$. We refer to A as the matrix associated with f.

Quadratic Form Expansion: We can expand a quadratic form in n variables $\bar{x}^T A \bar{x}$ as follows:

$$\bar{x}^T A \bar{x} = \sum_{i=1}^{n} a_{ii} x_i^2 + \sum_{i<j}^{n} 2a_{ij} x_i x_j \tag{124}$$

About diagonalizing the associated matrix: In general, the matrix of a quadratic form is a symmetric matrix and that such matrices can always be diagonalized. We will now use this fact to show that, for every quadratic from, we can eliminate the cross-product terms by means of a suitable change of variable. Let $f(\bar{x}) = \bar{x}^T A \bar{x}$ be a quadratic form in n variables, with A a symmetric $n \times n$ matrix. By the Spectral Theorem, there is an orthogonal matrix Q that diagonalizes A; that is, $Q^T A Q = D$, where D is the diagonal matrix displaying the eigenvalues of A. We now set

$$\bar{x} = Q\bar{y} \Rightarrow \bar{y} = Q^{-1}\bar{x} = Q^T \bar{x} \tag{125}$$

Then

$$\begin{aligned} \bar{x}^T A \bar{x} &= (Q\bar{y})^T A (Q\bar{y}) && (126) \\ &= \bar{y}^T Q^T A Q \bar{y} && (127) \\ &= \bar{y}^T D \bar{y} && (128) \\ &= \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 && (129) \\ &= \sum_{i=1}^{n} \lambda_i y_i^2 && (130) \end{aligned}$$

This process is called diagonalizing a quadratic form.

Theorem 5.23: The Principal Axes Theorem. Every quadratic form can be diagonalized. Specifically, if A is the $n \times n$ symmetric matrix associated with the quadratic form $\bar{x}^T A \bar{x}$ and if Q is an orthogonal matrix such that $Q^T A Q = D$ is a diagonal matrix, then the change of variable $\bar{x} = Q\bar{y}$ transform the quadratic form $\bar{x}^T A \bar{x}$ into the quadratic form $\bar{y}^T D \bar{y}$, which has no cross-product terms. If the eigenvalues of A are $\lambda_1, \cdots, \lambda_n$ and $\bar{y} = [y_1 \ \cdots \ y_n]^T$, then

$$\bar{x}^T A \bar{x} = \bar{y}^T D \bar{y} = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 \tag{131}$$

Classification of the quadratic forms: A quadratic form $f(\bar{x}) = \bar{x}^T A \bar{x}$ is classified as one of the following:

1. positive defined if $f(\bar{x}) > 0$ for all $\bar{x} \neq 0$.
2. positive semidefined if $f(\bar{x}) \geq 0$ for all $\bar{x}$.
3. negative defined if $f(\bar{x}) < 0$ for all $\bar{x} \neq 0$.
4. negative semidefined if $f(\bar{x}) \leq 0$ for all $\bar{x}$.
5. indefinite if $f(\bar{x})$ takes on both positive and negative values.

Classification of symmetric matrices: A symmetric matrix A is called positive defined, positive semidefined, negative define, negative semidefined, or indefinite if the associated quadratic form $f(\bar{x}) = \bar{x}^T A \bar{x}$ has the corresponding property.

Theorem 5.24: Let A be an $n \times n$ symmetric matrix. The quadratic form $f(\bar{x}) = \bar{x}^T A \bar{x}$ is

a. positive definite if and only if all of the eigenvalues of A are positive.

b. positive semidefinite if and only if all of the eigenvalues of A are nonnegative.

c. negative definite if and only if all of the eigenvalues of A are negative.

d. negative semidefinite if and only if all of the eigenvalues of A are nonpositive.

e. indefinite if and only if A has both positive and negative eigenvalues.

Theorem 5.25: Let $f(\bar{x}) = \bar{x}^T A \bar{x}$ be a quadratic form with associated $n \times n$ symmetric matrix A. Let the eigenvalues of A be $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Then the following are true, subject to the constrain $||\bar{x}|| = 1$:

a. $\lambda_1 \geq f(\bar{x}) \geq \lambda_n$

b. The maximum value of $f(\bar{x})$ is λ_1, and it occurs when $\bar{x}$ is a unit eigenvector corresponding to λ_1

c. The minimum value of $f(\bar{x})$ is λ_n, and it occurs when $\bar{x}$ is a unit eigenvector corresponding to λ_n

Graphing Quadratic Equations

Conic sections

The general form of a quadratic equation in two variables is

$$ax^2 + by^2 + cxy + dx + ey + f = 0 \tag{132}$$

where at least one of a, b, and c is nonzero. The graphs of such quadratic equations are called conic sections. They can be obtain by taking cross sections of a double cone. The most important of the conic sections are the ellipses (with circles as a special case), hyperbolas, and parabolas. These are called nondegenerate conic. It is also possible for a cross section of a cone to result in a single point, a straight line, or a pair of lines. These are called degenerate conics.

The graph of a nondegenerate conic is said to be in standard position relative to the coordinate axes if its equation can be expressed in one of the following forms:

- Ellipse or circle: $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$; $a, b > 0$
- Hyperbola: $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$; $a, b > 0$ or $-\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$; $a, b > 0$
- Parabola: $y = ax^2$, $a \gtrless 0$, $x = ay^2$, $a \gtrless 0$.

Other cases

- If a quadratic equation contains too many terms to be written in one of the above forms, then its graph is not in standard position. When there are additional terms but no xy term, the graphs of the conic has been translated out of standard position. It can be identify by completing the squares.
- If a quadratic equation contains a cross-product term, then it represents a conic that has been rotated.

Quadratic surface

The general form of a quadratic equation in three variables is

$$ax^2 + by^2 + cz^2 + dxy + exz + fyz + gx + hy + iz + j = 0 \tag{133}$$

where at least one of $a, b, \cdots, f$ is nonzero. The graphs of such quadratic equation is called a quadratic surface.

Some quadratic in standard position are

- Ellipsoid: $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$
- Hyperboloid: $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$
- Hyperboloid of two sheets: $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1$
- Elliptic cone: $z^2 = \frac{x^2}{a^2} + \frac{y^2}{b^2}$
- Elliptic paraboloid: $z = \frac{x^2}{a^2} + \frac{y^2}{b^2}$
- Hyperbolic paraboloid: $z = \frac{x^2}{a^2} - \frac{y^2}{b^2}$

A quadratic graph that have been translated out of standard position can be identify using the complete-the-squares method.