Chapter 4

Bilinear and quadratic forms

This chapter mainly focus on the following objectives:

- *define the concept of a bilinear form on a vector space.*
- *explain the equivalence of bilinear forms with matrices.*
- *represent a bilinear form on a vector space as a square matrix.*
- *define the concept of a quadratic form on* $\mathbb{R}^n$.
- *explain the notion of a Hermitian form on a vector space.*

Introduction

Bilinear forms occupy a unique place in all of mathematics. The study of linear transformations alone is incapable of handling the notions of orthogonality in geometry, optimization in many variables, Fourier series and so on forth. in optimization theory, the relevance of quadratic forms is all the more. The concept of dot product is a particular instance of a bilinear form. Quadratic forms, in particular, play an all important role in deciding the minima and maxima of functions of several variables. Hermitian forms appear naturally in harmonic analysis, communication systems and representation theory. The theory of quadratic forms derives much motivation from number theory. in short, there are enough reasons to undertake a basic study of bilinear and quadratic forms.

We first remark that in this lesson, we shall deal with the fields $\mathbb{F} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$ *only. Now, let us begin with definitions and examples.*

4.1 Bilinear forms

4.1.1 Definition of a bilinear form on a vector space:

We know that a linear functional if a scalar-valued linear transformation on a vector space. In a similar spirit, a bilinear form on a vector space is also a scalar-valued mapping of vector space. The difference lies in the fact while a linear functional is a function of single variable, a bilinear form is a function of two vector variables. In other words, while a linear functional on a linear functional on a vector space V *has the domain set* V*, a bilinear form on* V *has the domain set the Cartesian product* $V \times V$*. A bilinear form is linear in both the variables. hence, the name bears the adjective 'bilinear'.*

Definition 4.1.1.1. *: Let V be a vector space over a field $\mathbb{F}$. A bilinear form on V is a mapping $f : V \times V \to \mathbb{F}$ such that it is linear in both coordinates.That is for all $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$ and $\lambda \in \mathbb{F}$, we have,*

1. $f(a\boldsymbol{u} + \lambda\boldsymbol{v}, \boldsymbol{w}) = af(\boldsymbol{u}, \boldsymbol{w}) + \lambda f(\boldsymbol{v}, \boldsymbol{w})$
2. $f(\boldsymbol{u}, a\boldsymbol{v} + \lambda\boldsymbol{w}) = af(\boldsymbol{u}, \boldsymbol{v}) + \lambda f(\boldsymbol{u}, \boldsymbol{w})$
3. $f(\boldsymbol{u}, a\boldsymbol{v}) = af(\boldsymbol{u}, \boldsymbol{v})$
4. $f(a\boldsymbol{u}, \boldsymbol{v}) = af(\boldsymbol{u}, \boldsymbol{v})$

Thus, a bilinear form on a vector space V is a function on $V \times V$ such that it is linear in both coordinates.

Example 4.1.1.1. *Every square matrix, having entries from a field $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, gives rise to a bilinear form. Let A be an $n \times n$ matrix over a field $\mathbb{F}$. Then, the function $f : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ defined by $f(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}^t A\boldsymbol{y}$ is a bilinear form, on the vector space $V = \mathbb{F}^n$, for every pair of vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{F}^n$.*
One can easily verify the bilinearity of the mapping f using simple properties of matrix addition, matrix multiplication and matrix transpose and demonstrates that every square matrix over a field produces a bilinear form.

4.1.2 Special types of Bilinear forms

Bilinear forms of significant importance include: symmetric, skew-symmetric, and alternating bilinear forms. The forms are conceptually inter-linked. We begin with their definitions.

Definition 4.1.2.1. *A bilinear form $f : V \times V \to \mathbb{F}$ is*

1. ***symmetric*** *if $f(\mathbf{u}, \mathbf{v}) = f(\mathbf{v}, \mathbf{u})$, $\forall \mathbf{u}, \ \mathbf{v} \in V$;*
2. ***Skew-symmetric*** *if $f(\boldsymbol{u}, \boldsymbol{v}) = -f(\boldsymbol{v}, \boldsymbol{u}), \forall \boldsymbol{u}, \boldsymbol{v} \in V$;*
3. ***alternating*** *if $f(\boldsymbol{v}, \boldsymbol{v}) = 0, \forall \boldsymbol{v} \in V$ and the vector $\boldsymbol{v}$ is called* ***an isotrophic vector****;*
4. ***non-degenerate*** *if $\forall \boldsymbol{v} \in V$, there exist $\boldsymbol{w} \in V$, such that $f(\boldsymbol{w}, \boldsymbol{v}) \neq 0$;*
5. *positive definite if $f(\boldsymbol{v}, \boldsymbol{v}) \forall \boldsymbol{v} \in V$ and the equality hold if $\boldsymbol{v} = 0$.*

Example 4.1.2.1. *:*

1. *The usual dot product of vectors in $\mathbb{R}^n$ defines a symmetric bilinear form on $\mathbb{R}^n$. The mapping*

$$f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$

defined by

$$f(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{xy} = x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

is a bilinear form on $\mathbb{R}^n$ for every vector $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$ and $\boldsymbol{y} = (y_1, y_2, \cdots, y_n)$ in $\mathbb{R}^n$ and satisfies the symmetry property $f(\boldsymbol{x}, \boldsymbol{y}) = f(\boldsymbol{y}, \boldsymbol{x})$, since he dot product is commutative.

2. *Let $V = \mathbb{R}^2$ and its elements be viewed as a column vectors. Then, the determinant map*

$$det : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$$

given by

$$det \left[\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix} \right] = det \begin{pmatrix} a & c \\ b & d \end{pmatrix} = ad - bc$$

is a skew-symmetric and alternating bilinear form on $\mathbb{R}^2$. Skew-symmetric follows because interchanging the columns of the matrix changes the sign of the determinant; and it is alternating because whenever the two columns are identical, the deretminant is zero.

3. *The bilinear form $f : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ given by $f(\boldsymbol{z}, \boldsymbol{w}) = Im(\boldsymbol{z}\overline{\boldsymbol{w}}), \forall \boldsymbol{z}, \boldsymbol{w} \in \mathbb{C}$ is a skew-symmetric bilinear form on $\mathbb{C}$ because $\forall \boldsymbol{z}, \boldsymbol{w} \in \mathbb{C}$, we have $f(\boldsymbol{z}, \boldsymbol{w}) = Im(\boldsymbol{z}\overline{\boldsymbol{w}}) = -Im(\boldsymbol{w}\overline{\boldsymbol{z}}) = -f(\boldsymbol{w}, \boldsymbol{z})$.*

4. *The bilinear form $f : \mathbb{Q}^3 \times \mathbb{Q}^3 \to \mathbb{Q}$ (in 1) given by $f(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{ur} + 3\boldsymbol{vt} - \boldsymbol{wr} + 2\boldsymbol{ut}$ for all $\boldsymbol{x} = \begin{pmatrix} u \\ v \\ w \end{pmatrix} \in \mathbb{Q}^3$ and $\boldsymbol{y} = \begin{pmatrix} r \\ s \\ t \end{pmatrix} \in Q^3$ is neither symmetric nor alternating on $\mathbb{Q}^3$.*

5. $A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 0 & 3 \\ -1 & 0 & 0 \end{pmatrix}$. *Construct the corresponding $\mathbb{R}$-bilinear form on $\mathbb{R}^3$.*

   ***Solution**:*
   The desired bilinear form $f : \mathbb{R}^3 \to \mathbb{R}^3$ is

$$\begin{aligned} f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^t A \mathbf{y} &= (u \ v \ w) \begin{pmatrix} 1 & 0 & 2 \\ 0 & 0 & 3 \\ -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} r \\ s \\ t \end{pmatrix} \\ &= ur + 3vt - wr + 2ut; \end{aligned}$$

   for all $\mathbf{x} = \begin{pmatrix} u \\ v \\ w \end{pmatrix} \in \mathbb{R}^3$ and $\mathbf{y} = \begin{pmatrix} r \\ s \\ t \end{pmatrix} \in \mathbb{R}^3$.

Remarks

- *In all characteristics, an alternating bilinear form is skew-symmetric. A bilinear form is skew-symmetric if and only if it is alternating.*

- *In charactristic not 2, every bilinear form $f(\boldsymbol{u}, \boldsymbol{v})$ on V can be written uniquely as a sum of a symmetric and an alternating bilinear form as for every $\boldsymbol{u}, \boldsymbol{v} \in V$, we have*

$$f(\boldsymbol{u}, \boldsymbol{v}) = \frac{f(\boldsymbol{u}, \boldsymbol{v}) + f(\boldsymbol{v}, \boldsymbol{w}}{2} + \frac{f(\boldsymbol{u}, \boldsymbol{v}) - f(\boldsymbol{v}, \boldsymbol{w})}{2}.$$

- *In all characteristic not 2, every symmetric biinear form $f(\boldsymbol{u}, \boldsymbol{v})$ on an $n-$dimensional vector space V is completely determined by its values $f(\boldsymbol{u}, \boldsymbol{v})$ on the diagonal, as for every $\boldsymbol{u}, \boldsymbol{v} \in V$, we have using the symmetry of the bilinear form f,*

$$f(\boldsymbol{u}, \boldsymbol{v}) = \frac{f(\boldsymbol{u}+\boldsymbol{v}, \boldsymbol{u}+\boldsymbol{v}) - f(\boldsymbol{v}, \boldsymbol{v}) - f(\boldsymbol{u}, \boldsymbol{u})}{2} = \frac{1}{2}(f(\boldsymbol{u}, \boldsymbol{v}) + f(\boldsymbol{v}, \boldsymbol{u})) = f(\boldsymbol{u}, \boldsymbol{v}).$$

  This is called polarization identity.

Remark 4.1.1. *A matrix A is symmetric iff*

$$A = \frac{A + A^T}{2} + \frac{A - A^T}{2}.$$

4.1.3 Matrix representation of a Bilinear forms:

Let $V = \mathbb{F}^n$, then every $n \times n$ matrix A gives rise to a bilinear form by the formula $f_A(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{u}^t A \boldsymbol{v}$ that is if $f : V \times W \to \mathbb{F}$ is bilinear form over a field $\mathbb{F}$, $B = \{v_1, v_2, \cdots, v_n\}$ and $B' = \{w_1, w_2, \cdots, w_n\}$ being basis of vector space V and W respectively. For any $\boldsymbol{v} \in V$, $\boldsymbol{w} \in W$ can be written as $v = \sum_{i=1}^{m}(\alpha_i v_i)$ and $w = \sum_{j=1}^{n}(\beta_j v_j)$. Therefore,

$$f(\boldsymbol{v}, \boldsymbol{w}) = f(\sum_{i=1}^{m}(\alpha_i v_i), \sum_{j=1}^{n}(\beta_j v_j)) = \sum_{i=1}^{m}(f(\alpha_i v_i, \sum_{j=1}^{n}(\beta_j v_j)))$$

(Since f is bilinear map.)

$$= \sum_{i=1}^{m}(\alpha_i f(v_i, \sum_{j=1}^{n}(\beta_j v_j)))$$

$$= \sum_{i=1}^{m}\alpha_i(\sum_{j=1}^{n}(\beta_j f(v_i, v_j)))$$

(f is linear with the 2nd variable.)

$$= \sum_{i=1}^{m}\alpha_i \sum_{j=1}^{n}\beta_j f(v_i, v_j)$$

*. We have $m \times n$ scalars. The mn scalars $f(v_i, v_j)$ completely determines the values of f. Let $a_{ij} = f(v_i, v_j)$ consider the matrix $A = (a_{ij})_{m\times n} = f(v_i)f(v_j)$, then $f(x, y) = x^t A y$ where $x = (x_1, x_2, \ldots, x_m)$ and $y = (y_1, y_2, ..., y_n)$. A is called the **matrix representation** of the bilinear form $f : V \times W \to \mathbb{F}$.*
Here

$$A = [a_{ij}]_{n\times n}, \ a_{ij} = f(\mathbf{e_i}, \mathbf{e_j}), \ \forall i, j = 1 : n.$$

Example 4.1.3.1. *The bilinear form $f(\boldsymbol{x}, \boldsymbol{y}) = x_1y_1 + 2x_1y_2 + 3x_2y_1 + 4x_2y_2$ has matrix $\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}$ with respect to standard basis; $f((x_1, x_2), (y_1, y_2)) = (x_1, x_2)\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}\begin{pmatrix}y_1\\ y_2\end{pmatrix}$*

Example 4.1.3.2. *Take $V = \mathbb{R}^2$. The following matrices are bilinear forms $\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$, $\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$, $\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}$.*

Example 4.1.3.3. *If $A = \begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}$ then the corresponding bilinear form is*

$$f((x_1, x_2), (y_1, y_2)) = \begin{pmatrix}x_1 & y_1\end{pmatrix}\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}\begin{pmatrix}x_2\\ y_2\end{pmatrix} = x_1x_2 + 2x_1y_2 + 3y_1x_2 + 4y_1y_2.$$

Example 4.1.3.4. *Calculate the matrix of the bilinear form det on $\mathbb{R}^2$ relative to the standard basis for $\mathbb{R}^2$.*

Solution *The bilinear form $f = det$ on $\mathbb{R}^2$ is given by*

$$f\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix}\right) = det\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix}\right) = \det\begin{pmatrix} 1 & c \\ b & d \end{pmatrix} = ad - bc;$$

so that with respect to the standard basis $\mathcal{B} = \{\mathbf{e_1} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \ \mathbf{e_2} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, *we have,*

$$a_{11} = f(\mathbf{e_1}, \mathbf{e_2}) = det\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = 0; \ a_{12} = f(\mathbf{e_1}, \mathbf{e_2}) = det\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1;$$
$$a_{21} = f(\mathbf{e_2}, \mathbf{e_1}) = det\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = -1, \ a_{22} = f(\mathbf{e_2}, \mathbf{e_2}) = det\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = 0;$$

and therefore, the matrix of this bilinear form det is $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Exercise 4.1.3.1. *:*
Compute the matrix of the

a. *bilinear form* $f : \mathbb{Q}^3 \to \mathbb{Q}$ *given by* $f(\mathbf{x}, mathbfy) = x_1y_2 + x_3y_2 + x_2y_1, \ \forall \mathbf{x} = (x_1, x_2, x_3), \ \mathbf{y} = (y_1, y_2, y_3) \in \mathbb{Q}^3$ *relative to the basis for* $\mathbb{Q}^3$.

b. $f(\mathbf{x}, mathbfy) = 2x_1y_1 - 3x_2y_2 + x_3y_3, \ \forall (x_1, x_2, x_3), \ (y_1, y_2, y_3) \in \mathbb{R}^3$ *relative to the ordered basis* $\mathcal{B} = (1, 1, 1), \ (-2, 1, 1), \ (2, 1, 0)$.

4.2 Quadratic forms

In this section we will use matrix methods to study real-valued functions of several variables in which each term is either the square of a variable or the product of two variables. Symmetric bilinear forms give rise to what are known as quadratic forms. In optimization theory, machine learning, applied probability, geometry, vibrations of mechanical systems, statistics, and electrical engineering. and above all, in number theory, quadratic forms have deep and serious applications. There are many open problems regarding quadratic forms.
Let V be a finite dimensional vector space over a field $\mathbb{F}$.
A certain bilinear form is symmetric if it's associated matrix is symmetric.

4.2.1 Quadrics

Let A denote an $n \times n$ symmetric matrix with real entries and $\mathbf{x}$ denote an $n \times 1$ column vector.

Definition 4.2.1.1. *Given a symmetric bilinear form f on V, the associated quadratic form is the function $q(v) = f(\mathbf{x}, \mathbf{x}) = \mathbf{x}^T A\mathbf{x}$. The quadratic form q has the property $q(\lambda x) = \lambda^2 q(x)$.*

Note that

$$Q = \mathbf{x}^{'}A\mathbf{x} = (x_1,\ x_2, ...,\ x_n)\begin{pmatrix} a_{11} & ... & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$
$$= (x_1,\ x_2, ...,\ x_n)\begin{pmatrix} \sum_{i=1}^{n} a_{a1i}x_i \\ \vdots \sum_{i=1}^{n} a_{ni}x_i \end{pmatrix}$$
$$= (a_{11}x_1^2 + a_{12}x_1x_2 + ... + a_{1n}x_1x_n) + ... + (a_{n1}x_nx_1 + a_{22}x_nx_2 + ... + a_{nn}x_n^2)$$
$$= \sum_{i\leq j} a_{ij}x_ix_j.$$

Example 4.2.1.1. *The corresponding quadratic form of the bilinear form f defined by $\begin{pmatrix} 6 & 0 \\ 0 & 5 \end{pmatrix}$ is $q(x,y) = f((x,y),(x,y)) = (x,y)\begin{pmatrix} 6 & 0 \\ 0 & 5 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 6x^2 + 5y^2.$*

Proposition 4.2.1.1. *Let f be a bilinear form on V and B be a basis for V. Then f is a symmetric bilinear form iff the matrix representation of f with respect to the basis B is a symmetric bilinear form.*

Theorem 4.2.1.1. *(polarization theorem) If $1 + 1 \neq 0$ in $\mathbb{F}$ then for any quadratic form q the underlying symmetric bilinear form is unique.*

Proof. If $\mathbf{u}, \mathbf{v} \in V$ then

$$q(u+v) = f(u+v, u+v)$$
$$= f(u,u) + 2f(u,v) + f(v,v)$$
$$= q(u) + q(v) + 2f(u,v)$$

So, $f(u,v) = \frac{1}{2}(q(u+v) - q(u) - q(v))$. □

Remark 4.2.1. *:*
A bilinear form is symmetric if, and only if, its matrix relative to some basis is symmetric.
A bilinear form is alternating if, and only if, its matrix relative to some basis is skew-symmetric.

Example 4.2.1.2. *: Consider $A = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}$ it is a symmetric matrix.*

Let f be the corresponding bilinear form. We have

$$f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = \begin{pmatrix} x_1 & y_1 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = 2x_1x_2 + x_1y_2 + x_2y_1,$$

and

$$q(x,y) = 2x^2 + 2xy = f((x,y),(x,y)).$$

Let $\mathbf{u} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$.
Then

$$\frac{1}{2}(q(\mathbf{u}+\mathbf{v}) - q(\mathbf{u}) - q(\mathbf{v})) = \frac{1}{2}(2(x_1+x_2)^2 + 2(x_1+x_2)(y_1+y_2) - (x_1)^2 - 2x_1y_1 - 2(x_2)^2 - 2x_1y_2)$$
$$= \frac{1}{2}(4x_1x_2 + 2(x_1y_2 + x_2y_1)) = f\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right).$$

If $A = (a_{ij})$ is a symmetric matrix,then the corresponding bilinear form is

$$f(x,y) = \sum_{i}(a_{ij})(x_i y_i) + \sum_{i<j}(a_{ij})(x_i y_j + x_j y_i)$$

and the corresponding quadratic form is $q(x) = \sum a_{ij}(x_i)^2$.

Definition 4.2.1.2. *We say that a quadratic form f on a vector space dimension has signature (a,b,c) if there exists a basis B such that $[f]_B = D(a,b,c)$.*

Example 4.2.1.3. *Consider the matrix*

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

and the vector $\mathbf{x}$, *Q is given by*

$$\begin{aligned} Q = \mathbf{x}^{'}A\mathbf{x} &= (x_1\ x_2)\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= (x_1 + 2x_2 2x_1 + x_2)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= x_1^2 + 2x_1x_2 + 2x_1x_2 + x_2^2 \\ &= x_1^2 + 4x_1x_2 + x_2^2. \end{aligned}$$

Definition 4.2.1.3. *For any $n \times n$ symmetric real matrix A, a real row vector $\mathbf{b}$ and real number γ, the set $x \in \mathbb{R}^n : x^T Ax + bx + \gamma = 0$ is called* ***quadric***. *If $n = 2$ is called a conic.*

Example 4.2.1.4. *The set $(x,y,z) \in \mathbb{R}^3 : x^2 - y^2 - z + 1 = 0$: Consider a quadric $V = \{x \in \mathbb{R}^n : x^T Ax + bx + \gamma = 0\}$. We have $A = Q^T DQ$ for an orthogonal Q and a diagonal matrix D.*

Classification of the quadratic form :
$Q = \mathbf{x}^{'}A\mathbf{x}$; *A quadratic form is said to be:*

a: *negative definite: $Q < 0$ when $x \neq 0$.*

b: *negative semidefinite: $Q \leq 0$ for all $\mathbf{x}$ and $Q = 0$ for some $\mathbf{x} \neq 0$.*

c: *positive definite: $Q > 0$ when $\mathbf{x} \neq 0$*

d: *Positive semidefinte: $Q \geq 0$ for all $\mathbf{x}$ and $Q = 0$ for some $x \neq 0$.*

e: *indefinite: $Q > 0$ for some $\mathbf{x}$ and $Q < 0$ for some other $\mathbf{x}$.*

Example 4.2.1.5. *The quadratic forms $Q = x_1^2 + 2x_2^2 + 4x_3^2$ and $Q = 3x_1 + 3x_2^2$ are positive definite. as shown below*

Example 4.2.1.6. *classification of conics: $p = 2 : a_1x_1^2 + a_2x_2^2 = \gamma$ for $a_1, a_2 > 0$ empty if $\gamma < 0$, the point $(0,0)$ if $\gamma = 0$, an ellipse with axis $\sqrt{\frac{\gamma}{a_1}}$ and $\sqrt{\frac{\gamma}{a_2}}$, if $\gamma > 0$.*

Example 4.2.1.7. *The dot product $\mathbf{u}.\mathbf{v}$ on $\mathbb{R}^n$ is symmetric bilinear form.*

Matrix representation of quadratic forms *Let Q be a quadratic form. The matrix representation of Q is A for which its diagonal entries are the coefficient of the squared terms, and off-diagonals are half of the coefficients of the cross product terms. Here it's associated matrix is always symmetric.*

Example 4.2.1.8. *Find the associated matrix of the quadratic forms*

a. $2x^2 + 6xy - 5y^2$ *b.* $x^2 + yx_3^2 + 4x_1x_2 - 2x_1x_3 + 8x_2x_3$

4.2.2 Conic Sections

*Recall that a **conic section** or **conic** is a curve that results by cutting a double-napped cone with a plane.*
The most important conic sections are ellipses, hyperbolas, and parabolas, which result when the cutting plane does not pass through the vertex.
Circles are special cases of ellipses that result when the cutting plane is perpendicular to the axis of symmetry of the cone. If the cutting plane passes through the vertex, then the resulting intersection is called a degenerate conic. The possibilities are a point, a pair of intersecting lines, or a single line.
Quadratic forms in $\mathbb{R}^2$ arise naturally in the study of conic sections. For example, it is shown in analytic geometry that an equation of the form

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0$$

in which a, b, and c are not all zero, represents a conic section.
If $d = e = 0$ in, then there are no linear terms, so the equation becomes

$$ax^2 + 2bxy + cy^2 + f = 0 \tag{4.1}$$

*and is said to represent a **central conic**. These include **circles, ellipses, and hyperbolas**, but not parabolas. Furthermore, if $b = 0$, then there is no cross product term (i.e., term involving xy), and the equation*

$$ax2 + cy2 + f = 0 \tag{4.2}$$

is said to represent a central conic in standard position.

Theorem 4.2.2.1. *If A is a symmetric matrix, then:*

(a) $\mathbf{x}^T A\mathbf{x}$ *is positive definite if and only if all eigenvalues of A are positive.*

(b) $\mathbf{x}^T A\mathbf{x}$ *is negative definite if and only if all eigenvalues of A are negative.*

(c) $\mathbf{x}^T A\mathbf{x}$ *is indefinite if and only if A has at least one positive eigenvalue and at least one negative eigenvalue.*

Classification of Conic sections *:*
$\mathbf{x}^T B\mathbf{x} = k$ *is the equation of a conic, and if $k \neq 0$, then we can divide through by k and rewrite the equation in the form*

$$\mathbf{x}^T A\mathbf{x} = 1$$

where $A = \frac{1}{k}B$. If we now rotate the coordinate axes to eliminate the cross product term(if any) in this equation, then the equation of the conic in the new coordinate system will be of the form

$$\lambda_1 x_1^2 + \lambda_n x_2^2 + ... + \lambda x_n^2 = 1 \tag{4.3}$$

in which $\lambda_i{}'s$ are eigenvalues of A. The particular type of conic represented by this equation will depend on the signs of the eigenvalues $\lambda_i{}'s$.
$\mathbf{x}^T A\mathbf{x} = 1$ *represents an ellipse if all eigenvalues are positive.*
$\mathbf{x}^T A\mathbf{x} = 1$ *has no graph if all eigenvalues are negative.*
$\mathbf{x}^T A\mathbf{x} = 1$ *represents a hyperbola if the sign of the eigenvalues are opposite.*

Theorem 4.2.2.2. *If A is a symmetric matrix, then:*

(a) *A is positive definite if and only if the determinant of every principal submatrix is positive.*

(b) *A is negative definite if and only if the determinants of the principal submatrices alternate between negative and positive values starting with a negative value for the determinant of the first principal submatrix.*

(c) *A is indefinite if and only if it is neither positive definite nor negative definite and at least one principal submatrix has a positive determinant and at least one has a negative determinant.*

4.3 Application of Bilinear and quadratic forms

4.3.1 Optimization Using Quadratic Forms

Our first goal in this section is to consider the problem of finding the maximum and minimum values of a quadratic form $\mathbf{x}^T A\mathbf{x}$ subject to the constraint $||\mathbf{x}|| = 1$. Problems of this type arise in a wide variety of applications.
To visualize this problem geometrically in the case where $\mathbf{x}^T A\mathbf{x}$ is a quadratic form on $\mathbb{R}^2$, view $z = \mathbf{x}^T A\mathbf{x}$ as the equation of some surface in a rectangular $xyz-$coordinate system and view $||x|| = 1$ as the unit circle centered at the origin of the $xy-$plane. Geometrically, the problem of finding the maximum and minimum values of $\mathbf{x}^T A\mathbf{x}$ subject to the requirement $||x|| = 1$ amounts to finding the highest and lowest points on the intersection of the surface with the right circular cylinder determined by the circle

Theorem 4.3.1.1. *Constrained Extremum Theorem*:
Let A be a symmetric $n \times n$ matrix whose eigenvalues in order of decreasing size are

$$\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n.$$

Then:

(a) *The quadratic form $\mathbf{x}^T A\mathbf{x}$ attains a maximum value and a minimum value on the set of vectors for which $||\mathbf{x}|| = 1$.*

(b) *The maximum value attained in part (a) occurs at a vector corresponding to the eigenvalue λ_1.*

(c) *The minimum value attained in part (a) occurs at a vector corresponding to the eigenvalue λ_n*

The condition $||\mathbf{x}|| = 1$ in this theorem is called a ***constraint****, and the maximum or minimum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraint is called a constrained extremum. This constraint can also be expressed as $\mathbf{x}^T\mathbf{x} = 1$ or as $x_1^2 + x_2^2 + ... + x_n^2 = 1$, when convenient.*

Example 4.3.1.1. *Finding Constrained Exxtrema*
Find the maximum and minimum values of the quadratic form

$$z = 5x^2 + 5y^2 + 4xy$$

subject to the constraint $x^2 + y^2 = 1$.
Solution*:*
The quadratic form can be expressed in matrix notation as

$$z = 5x^2 + 5y^2 + 4xy = \mathbf{x}^T A\mathbf{x} = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

We leave it for you to show that the eigenvalues of A are $\lambda_1 = 7$ and $\lambda_2 = 3$ and that corresponding eigenvectors are

$$\lambda_1 = 7 : \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \ \lambda_2 = 3 : \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Normalizing these eigenvectors yields

$$\lambda_1 = 7 : \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}, \ \lambda_2 = 3 : \begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}. \tag{4.4}$$

Thus, the constrained extrema are
Constrained maximum: $z = 7$ at $(x, y) = \left(\frac{1}{\sqrt{2}}\right)$
constrained minumu: $z = 3$ at $\left(\frac{1}{\sqrt{2}}\right)$. Since the negatives of the eigenvectors above are also unit eigenvectors, they too produce the maximum and minimum values of z; that is, the constrained maximum $z = 7$ also occurs at the point $(x, y) = \left(\frac{-1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right)$ and the constrained minimum $z = 3$ at $(x, y) = \left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right)$.

4.3.2 Constrained Extrema and Level Curves

A useful way of visualizing the behavior of a function $f(x, y)$ of two variables is to consider the curves in the $xy-$plane along which $f(x, y)$ is constant. These curves have equations of the form

$$f(x, y) = k$$

and are called the ***level curves*** *of f. In particular, the level curves of a quadratic form $\mathbf{x}^T A\mathbf{x}$ on $\mathbb{R}^2$ have equations of the form*

$$\mathbf{x}^T A\mathbf{x} = k \tag{4.5}$$

so the maximum and minimum values of $\mathbf{x}^T A\mathbf{x}$ subject to the constraint $||\mathbf{x}|| = 1$ are the largest and smallest values of k for which the graph of the above equation intersects the unit circle. Typically, such values of k produce level curves that just touch the unit circle and the coordinates of the points where the level curves just touch produce the vectors that maximize or minimize $\mathbf{x}^T A\mathbf{x}$ subject to the constraint $||\mathbf{x}|| = 1$.

4.3.3 Relative Extrema of Function of Two Variables using Quadratic forms

We will conclude this section by showing how quadratic forms can be used to study characteristics of real-valued functions of two variables. Recall that if a function $f(x, y)$ has first-order partial derivatives, then its relative maxima and minima, if any, occur at points where the condition

$$f_x(x, y) = 0 \text{ and } f_y(x, y) = 0$$

are both true. These are called ***critical points*** *of f. The specific behavior of f at a critical point (x_0, y_0) is determined by the sign of*

$$D(x, y) = f(x, y) - f(x_0, y_0) \tag{4.6}$$

at points (x, y) that are close to, but different from, (x_0, y_0):

- *If $D(x, y) > 0$ at points (x, y) that are sufficiently close to, but different from, (x_0, y_0), then $f(x_0, y_0) < f(x, y)$ at such points and f is said to have a* ***relative minimum*** *at (x_0, y_0).*
- *If $D(x, y) < 0$ at points (x, y) that are sufficiently close to, but different from, (x_0, y_0), then $f(x_0, y_0) > f(x, y)$ at such points and f is said to have a* ***relative maximum*** *at (x_0, y_0).*
- *If $D(x, y)$ has both positive and negative values inside every circle centered at (x_0, y_0), then there are points (x, y) that are arbitrarily close to (x_0, y_0) at which $f(x_0, y_0) < f(x, y)$ and points (x, y) that are arbitrarily close to (x_0, y_0) at which $f(x_0, y_0) > f(x, y)$. In this case we say that f has a* ***saddle point*** *at (x_0, y_0)*

Theorem 4.3.3.1. *(Second Derivative Test):*
Suppose that (x_0, y_0) is a critical point of $f(x, y)$ and that f has continuous second order partial derivatives in some circular region centered at (x_0, y_0). Then:

(a) f has a relative minimum at (x_0, y_0) if

$$f_{xx}(x_0, y_0)f_{yy}(x_0, y_0) - f_{xy}^2(x_0, y_0) > 0 \text{ and } f_{xx}(x_0, y_0) > 0.$$

(b) f has a relative maximum at (x_0, y_0) if

$$f_{xx}(x_0, y_0)f_{yy}(x_0, y_0) - f_{xy}^2(x_0, y_0) > 0 \text{ and } f_{xx}(x_0, y_0) < 0.$$

(c) f has a saddle point at (x_0, y_0) if

$$f_{xx}(x_0, y_0)f_{yy}(x_0, y_0) - f_{xy}^2(x_0, y_0) < 0.$$

(d) The test is inconclusive if

$$f_{xx}(x_0, y_0)f_{yy}(x_0, y_0) - f_{xy}^2(x_0, y_0) = 0.$$

Our interest here is in showing how to reformulate this theorem using properties of symmetric matrices. For this purpose we consider the symmetric matrix

$$H(x, y) = \begin{bmatrix} f_{xx}(x, y) & f_{xy}(x, y) \\ f_{xy}(x, y) & f_{yy}(x, y) \end{bmatrix}$$

which is called the Hessian or Hessian matrix of f in honor of the German mathematician and scientist Ludwig Otto Hesse (1811–1874). The notation$H(x, y)$ emphasizes that the entries in the matrix depend on x and y. The Hessian is of interest because

$$det[H(x_0, y_0)] = \begin{vmatrix} f_{xx}(x,y) & f_{xy}(x,y) \\ f_{xy}(x,y) & f_{yy}(x,y) \end{vmatrix} =$$

is the expression that appears in above Theorem. We can now reformulate the second derivative test as follows.

Theorem 4.3.3.2. *(Hessian Form of the Second Derivative Test):*
Suppose that (x_0, y_0) is a critical point of $f(x, y)$ and that f has continuous second order partial derivatives in some circular region centered at (x_0, y_0). If $H(x_0, y_0)$ is the Hessian of f at (x_0, y_0), then:

(a) f has a relative minimum at (x_0, y_0) if $H(x_0, y_0)$ is positive definite.

(b) f has a relative maximum at (x_0, y_0) if $H(x_0, y_0)$ is negative definite.

(c) f has a saddle point at (x_0, y_0) if $H(x_0, y_0)$ is indefinite.

(d) The test is inconclusive otherwise.

Example 4.3.3.1. *Using the Hessian to Classify Relative Extrema*
Find the critical points of the function

$$f(x, y) = \frac{1}{3}x^3 + xy^2 - 8xy + 3$$

and use the eigenvalues of the Hessian matrix at those points to determine which of them, if any, are relative maxima, relative minima, or saddle points.
Solution:
To find both the critical points and the Hessian matrix we will need to calculate the first and second partial derivatives of f. These derivatives are

$$f_x(x, y) = x^2 + y^2 - 8y,\ f_y(x, y) = 2xy - 8x,\ f_{xy}(x, y) = 2y - 8f_{xx}(x, y) = 2x,\ f_{yy}(x, y) = 2x.$$

Thus, the Hessian matrix is

$$H(x, y) = \begin{bmatrix} f_{xx}(x,y) & f_{xy}(x,y) \\ f_{xy}(x,y) & f_{yy}(x,y \end{bmatrix} = \begin{bmatrix} 2x & 2y-8 \\ 2y-8 & 2x \end{bmatrix}$$

To find the critical points we set f_x and f_y equal to zero. This yields the equations

$$f_x(x, y) = x^2 + y^2 - 8y = 0 \text{ and } f_y(x, y) = 2xy - 8x = 2x(y - 4) = 0$$

Solving the second equation yields $x = 0$ or $y = 4$. Substituting $x = 0$ in the first equation and solving for y yields $y = 0$ or $y = 8$; and substituting $y = 4$ into the first equation and solving for x yields $x = 4$ or $x = -4$. Thus, we have four critical points:

$$(0, 0),\ (0, 8),\ (4, 4),\ (-4, 4).$$

Evaluating the Hessian matrix at these points yields

$$H(0,0) = \begin{bmatrix} 0 & -8 \\ -8 & 0 \end{bmatrix},\ H(0,8) = \begin{bmatrix} 0 & 8 \\ 8 & 0 \end{bmatrix},\ H(4,4) = \begin{bmatrix} 8 & 0 \\ 0 & 8 \end{bmatrix},\ H(-4,4) = \begin{bmatrix} -8 & 0 \\ 0 & 8 \end{bmatrix}.$$

We leave it for you to find the eigenvalues of these matrices and deduce the following classifications of the stationary points:

Critical point (x_0, y_0)	λ_1	λ_2	*Classification*
(0,0)	*8*	*-8*	*saddle point*
(0,8)	*8*	*-8*	*saddle point*
(4,4)	*8*	*8*	*Relative minimum*
(-4,4)	*-8*	*-8*	*Relative maximum*