Chapter 11

APPLICATIONS OF REAL INNER PRODUCT SPACES

11.1. Least Squares Approximation

Given a continuous function $f:[a,b]\to\mathbb{R}$, we wish to approximate f by a polynomial $g:[a,b]\to\mathbb{R}$ of degree at most k, such that the error

$$\int_a^b |f(x)-g(x)|^2\,\mathrm{d}x$$

is minimized. The purpose of this section is to study this problem using the theory of real inner product spaces. Our argument is underpinned by the following simple result in the theory.

PROPOSITION 11A. *Suppose that V is a real inner product space, and that W is a finite-dimensional subspace of V. Given any $\mathbf{u}\in V$, the inequality*

$$\|\mathbf{u}-\mathrm{proj}_W\mathbf{u}\|\le\|\mathbf{u}-\mathbf{w}\|$$

holds for every $\mathbf{w}\in W$.

In other words, the distance from $\mathbf{u}$ to any $\mathbf{w}\in W$ is minimized by the choice $\mathbf{w}=\mathrm{proj}_W\mathbf{u}$, the orthogonal projection of $\mathbf{u}$ on the subspace W. Alternatively, $\mathrm{proj}_W\mathbf{u}$ can be thought of as the vector in W closest to $\mathbf{u}$.

Proof of Proposition 11A. Note that

$$\mathbf{u}-\mathrm{proj}_W\mathbf{u}\in W^\perp \qquad\text{and}\qquad \mathrm{proj}_W\mathbf{u}-\mathbf{w}\in W.$$

It follows from Pythagoras's theorem that

$$\|\mathbf{u}-\mathbf{w}\|^2 = \|(\mathbf{u}-\text{proj}_W\mathbf{u})+(\text{proj}_W\mathbf{u}-\mathbf{w})\|^2 = \|\mathbf{u}-\text{proj}_W\mathbf{u}\|^2 + \|\text{proj}_W\mathbf{u}-\mathbf{w}\|^2,$$

so that

$$\|\mathbf{u}-\mathbf{w}\|^2 - \|\mathbf{u}-\text{proj}_W\mathbf{u}\|^2 = \|\text{proj}_W\mathbf{u}-\mathbf{w}\|^2 \geq 0.$$

The result follows immediately. ○

Let V denote the vector space $C[a,b]$ of all continuous real valued functions on the closed interval $[a,b]$, with inner product

$$\langle f,g\rangle = \int_a^b f(x)g(x)\,\mathrm{d}x.$$

Then

$$\int_a^b |f(x)-g(x)|^2\,\mathrm{d}x = \langle f-g, f-g\rangle = \|f-g\|^2.$$

It follows that the least squares approximation problem is reduced to one of finding a suitable polynomial g to minimize the norm $\|f-g\|$.

Now let $W = P_k[a,b]$ be the collection of all polynomials $g : [a,b] \to \mathbb{R}$ with real coefficients and of degree at most k. Note that W is essentially P_k, although the variable is restricted to the closed interval $[a,b]$. It is easy to show that W is a subspace of V. In view of Proposition 11A, we conclude that

$$g = \text{proj}_W f$$

gives the best least squares approximation among polynomials in $W = P_k[a,b]$. This subspace is of dimension $k+1$. Suppose that $\{v_0, v_1, \ldots, v_k\}$ is an orthogonal basis of $W = P_k[a,b]$. Then by Proposition 9L, we have

$$g = \frac{\langle f,v_0\rangle}{\|v_0\|^2}v_0 + \frac{\langle f,v_1\rangle}{\|v_1\|^2}v_1 + \ldots + \frac{\langle f,v_k\rangle}{\|v_k\|^2}v_k.$$

EXAMPLE 11.1.1. Consider the function $f(x) = x^2$ in the interval $[0,2]$. Suppose that we wish to find a least squares approximation by a polynomial of degree at most 1. In this case, we can take $V = C[0,2]$, with inner product

$$\langle f,g\rangle = \int_0^2 f(x)g(x)\,\mathrm{d}x,$$

and $W = P_1[0,2]$, with basis $\{1,x\}$. We now apply the Gram-Schmidt orthogonalization process to this basis to obtain an orthogonal basis $\{1, x-1\}$ of W, and take

$$g = \frac{\langle x^2,1\rangle}{\|1\|^2}1 + \frac{\langle x^2,x-1\rangle}{\|x-1\|^2}(x-1).$$

Now

$$\langle x^2,1\rangle = \int_0^2 x^2\,\mathrm{d}x = \frac{8}{3} \qquad \text{and} \qquad \|1\|^2 = \langle 1,1\rangle = \int_0^2 \mathrm{d}x = 2,$$

while

$$\langle x^2, x-1\rangle = \int_0^2 x^2(x-1)\,\mathrm{d}x = \frac{4}{3} \qquad \text{and} \qquad \|x-1\|^2 = \langle x-1, x-1\rangle = \int_0^2 (x-1)^2\,\mathrm{d}x = \frac{2}{3}.$$

It follows that

$$g = \frac{4}{3} + 2(x-1) = 2x - \frac{2}{3}.$$

EXAMPLE 11.1.2. Consider the function $f(x) = \mathrm{e}^x$ in the interval $[0,1]$. Suppose that we wish to find a least squares approximation by a polynomial of degree at most 1. In this case, we can take $V = C[0,1]$, with inner product

$$\langle f, g\rangle = \int_0^1 f(x)g(x)\,\mathrm{d}x,$$

and $W = P_1[0,1]$, with basis $\{1, x\}$. We now apply the Gram-Schmidt orthogonalization process to this basis to obtain an orthogonal basis $\{1, x-1/2\}$ of W, and take

$$g = \frac{\langle \mathrm{e}^x, 1\rangle}{\|1\|^2}1 + \frac{\langle \mathrm{e}^x, x-1/2\rangle}{\|x-1/2\|^2}\left(x - \frac{1}{2}\right).$$

Now

$$\langle \mathrm{e}^x, 1\rangle = \int_0^1 \mathrm{e}^x\,\mathrm{d}x = \mathrm{e} - 1 \qquad \text{and} \qquad \langle \mathrm{e}^x, x\rangle = \int_0^1 \mathrm{e}^x x\,\mathrm{d}x = 1,$$

so that

$$\left\langle \mathrm{e}^x, x - \frac{1}{2}\right\rangle = \langle \mathrm{e}^x, x\rangle - \frac{1}{2}\langle \mathrm{e}^x, 1\rangle = \frac{3}{2} - \frac{\mathrm{e}}{2}.$$

Also

$$\|1\|^2 = \langle 1, 1\rangle = \int_0^1 \mathrm{d}x = 1 \qquad \text{and} \qquad \left\|x - \frac{1}{2}\right\|^2 = \left\langle x - \frac{1}{2}, x - \frac{1}{2}\right\rangle = \int_0^1 \left(x - \frac{1}{2}\right)^2 \mathrm{d}x = \frac{1}{12}.$$

It follows that

$$g = (\mathrm{e}-1) + (18 - 6\mathrm{e})\left(x - \frac{1}{2}\right) = (18 - 6\mathrm{e})x + (4\mathrm{e} - 10).$$

REMARK. From the proof of Proposition 11A, it is clear that $\|\mathbf{u} - \mathbf{w}\|$ is minimized by the unique choice $\mathbf{w} = \mathrm{proj}_W \mathbf{u}$. It follows that the least squares approximation problem posed here has a unique solution.

11.2. Quadratic Forms

A real quadratic form in n variables $x_1, \ldots, x_n$ is an expression of the form

$$\sum_{\substack{i=1\\ i\le j}}^{n} \sum_{j=1}^{n} c_{ij}x_i x_j, \tag{1}$$

where $c_{ij} \in \mathbb{R}$ for every $i, j = 1, \ldots, n$ satisfying $i \le j$.

EXAMPLE 11.2.1. The expression $5x_1^2 + 6x_1x_2 + 7x_2^2$ is a quadratic form in two variables x_1 and x_2. It can be written in the form

$$5x_1^2 + 6x_1x_2 + 7x_2^2 = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 5 & 3 \\ 3 & 7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

EXAMPLE 11.2.2. The expression $4x_1^2 + 5x_2^2 + 3x_3^2 + 2x_1x_2 + 4x_1x_3 + 6x_2x_3$ is a quadratic form in three variables x_1, x_2 and x_3. It can be written in the form

$$4x_1^2 + 5x_2^2 + 3x_3^2 + 2x_1x_2 + 4x_1x_3 + 6x_2x_3 = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} 4 & 1 & 2 \\ 1 & 5 & 3 \\ 2 & 3 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Note that in both examples, the quadratic form can be described in terms of a real symmetric matrix. In fact, this is always possible. To see this, note that given any quadratic form (1), we can write, for every $i, j = 1, \ldots, n$,

$$a_{ij} = \begin{cases} c_{ij} & \text{if } i = j, \\ \frac{1}{2}c_{ij} & \text{if } i < j, \\ \frac{1}{2}c_{ji} & \text{if } i > j. \end{cases} \tag{2}$$

Then

$$\sum_{\substack{i=1 \\ i \le j}}^{n} \sum_{j=1}^{n} c_{ij}x_ix_j = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}x_ix_j = \begin{pmatrix} x_1 & \ldots & x_n \end{pmatrix} \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

The matrix

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}$$

is clearly symmetric, in view of (2).

We are interested in the case when $x_1, \ldots, x_n$ take real values. In this case, we can write

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

It follows that a quadratic form can be written as

$$\mathbf{x}^t A \mathbf{x},$$

where A is an $n \times n$ real symmetric matrix and $\mathbf{x}$ takes values in $\mathbb{R}^n$.

Many problems in mathematics can be studied using quadratic forms. Here we shall restrict our attention to two fundamental problems which are in fact related. The first is the question of what conditions the matrix A must satisfy in order that the inequality

$$\mathbf{x}^t A \mathbf{x} > 0$$

holds for every non-zero $\mathbf{x} \in \mathbb{R}^n$. The second is the question of whether it is possible to have a change of variables of the type $\mathbf{x} = P\mathbf{y}$, where P is an invertible matrix, such that the quadratic form $\mathbf{x}^t A \mathbf{x}$ can be represented in the alternative form $\mathbf{y}^t D \mathbf{y}$, where D is a diagonal matrix with real entries.

Definition. A quadratic form $\mathbf{x}^t A\mathbf{x}$ is said to be positive definite if $\mathbf{x}^t A\mathbf{x} > 0$ for every non-zero $\mathbf{x} \in \mathbb{R}^n$. In this case, we say that the symmetric matrix A is a positive definite matrix.

To answer our first question, we shall prove the following result.

PROPOSITION 11B. *A quadratic form $\mathbf{x}^t A\mathbf{x}$ is positive definite if and only if all the eigenvalues of the symmetric matrix A are positive.*

Our strategy here is to prove Proposition 11B by first studying our second question. Since the matrix A is real and symmetric, it follows from Proposition 10E that it is orthogonally diagonalizable. In other words, there exists an orthogonal matrix P and a diagonal matrix D such that $P^t AP = D$, and so $A = PDP^t$. It follows that

$$\mathbf{x}^t A\mathbf{x} = \mathbf{x}^t PDP^t\mathbf{x},$$

and so, writing

$$\mathbf{y} = P^t\mathbf{x},$$

we have

$$\mathbf{x}^t A\mathbf{x} = \mathbf{y}^t D\mathbf{y}.$$

Also, since P is an orthogonal matrix, we also have $\mathbf{x} = P\mathbf{y}$. This answers our second question.

Furthermore, in view of the Orthogonal diagonalization process, the diagonal entries in the matrix D can be taken to be the eigenvalues of A, so that

$$D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

where $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ are the eigenvalues of A. Writing

$$\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

we have

$$\mathbf{x}^t A\mathbf{x} = \mathbf{y}^t D\mathbf{y} = \lambda_1 y_1^2 + \ldots + \lambda_n y_n^2. \tag{3}$$

Note now that $\mathbf{x} = \mathbf{0}$ if and only if $\mathbf{y} = \mathbf{0}$, since P is an invertible matrix. Proposition 11B now follows immediately from (3).

Example 11.2.3. Consider the quadratic form $2x_1^2 + 5x_2^2 + 2x_3^2 + 4x_1x_2 + 2x_1x_3 + 4x_2x_3$. This can be written in the form $\mathbf{x}^t A\mathbf{x}$, where

$$A = \begin{pmatrix} 2 & 2 & 1 \\ 2 & 5 & 2 \\ 1 & 2 & 2 \end{pmatrix} \quad \text{and} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

The matrix A has eigenvalues $\lambda_1 = 7$ and (double root) $\lambda_2 = \lambda_3 = 1$; see Example 10.3.1. Furthermore, we have $P^t AP = D$, where

$$P = \begin{pmatrix} 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \\ 2/\sqrt{6} & 0 & -1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \end{pmatrix} \quad \text{and} \quad D = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Writing $\mathbf{y} = P^t\mathbf{x}$, the quadratic form becomes $7y_1^2 + y_2^2 + y_3^2$ which is clearly positive definite.

EXAMPLE 11.2.4. Consider the quadratic form $5x_1^2 + 6x_2^2 + 7x_3^2 - 4x_1x_2 + 4x_2x_3$. This can be written in the form $\mathbf{x}^t A\mathbf{x}$, where

$$A = \begin{pmatrix} 5 & -2 & 0 \\ -2 & 6 & 2 \\ 0 & 2 & 7 \end{pmatrix} \qquad \text{and} \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

The matrix A has eigenvalues $\lambda_1 = 3$, $\lambda_2 = 6$ and $\lambda_3 = 9$; see Example 10.3.3. Furthermore, we have $P^t AP = D$, where

$$P = \begin{pmatrix} 2/3 & 2/3 & -1/3 \\ 2/3 & -1/3 & 2/3 \\ -1/3 & 2/3 & 2/3 \end{pmatrix} \qquad \text{and} \qquad D = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 9 \end{pmatrix}.$$

Writing $\mathbf{y} = P^t\mathbf{x}$, the quadratic form becomes $3y_1^2 + 6y_2^2 + 9y_3^2$ which is clearly positive definite.

EXAMPLE 11.2.5. Consider the quadratic form $x_1^2 + x_2^2 + 2x_1x_2$. Clearly this is equal to $(x_1 + x_2)^2$ and is therefore not positive definite. The quadratic form can be written in the form $\mathbf{x}^t A\mathbf{x}$, where

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \qquad \text{and} \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

It follows from Proposition 11B that the eigenvalues of A are not all positive. Indeed, the matrix A has eigenvalues $\lambda_1 = 2$ and $\lambda_2 = 0$, with corresponding eigenvectors

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Hence we may take

$$P = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \qquad \text{and} \qquad D = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}.$$

Writing $\mathbf{y} = P^t\mathbf{x}$, the quadratic form becomes $2y_1^2$ which is not positive definite.

Problems for Chapter 11

1. Consider the function $f : [-1, 1] \to \mathbb{R} : x \mapsto x^3$. We wish to find a polynomial $g(x) = ax + b$ which minimizes the error

$$\int_{-1}^{1} |f(x) - g(x)|^2 \, \mathrm{d}x.$$

Follow the steps below to find this polynomial g:

a) Consider the real vector space $C[-1, 1]$. Write down a suitable real inner product on $C[-1, 1]$ for this problem, explaining carefully the steps that you take.

b) Consider now the subspace $P_1[-1, 1]$ of all polynomials of degree at most 1. Describe the polynomial g in terms of f and orthogonal projection with respect to the inner product in part (a). Give a brief explanation for your choice.

c) Write down a basis of $P_1[-1, 1]$.

d) Apply the Gram-Schmidt process to your basis in part (c) to obtain an orthogonal basis of $P_1[-1, 1]$.

e) Describe your polynomial in part (b) as a linear combination of the elements of your basis in part (d), and find the precise values of the coefficients.

2. For each of the following functions, find the best least squares approximation by linear polynomials of the form $ax + b$, where $a, b \in \mathbb{R}$:

a) $f : [0, \pi/2] \to \mathbb{R} : x \mapsto \sin x$

b) $f : [0, 1] \to \mathbb{R} : x \mapsto x^3$

c) $f : [0, 2] \to \mathbb{R} : x \mapsto \mathrm{e}^x$

3. Consider the quadratic form $2x_1^2 + x_2^2 + x_3^2 + 2x_1x_2 + 2x_1x_3$ in three variables x_1, x_2, x_3.

a) Write the quadratic form in the form $\mathbf{x}^t A\mathbf{x}$, where

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

and where A is a symmetric matrix with real entries.

b) Apply the Orthogonal diagonalization process to the matrix A.

c) Find a transformation of the type $\mathbf{x} = P\mathbf{y}$, where P is an invertible matrix, so that the quadratic form can be written as $\mathbf{y}^t D\mathbf{y}$, where

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

and where D is a diagonal matrix with real entries. You should give the matrices P and D explicitly.

d) Is the quadratic form positive definite? Justify your assertion both in terms of the eigenvalues of A and in terms of your solution to part (c).

4. For each of the following quadratic forms in three variables, write it in the form $\mathbf{x}^t A\mathbf{x}$, find a substitution $\mathbf{x} = P\mathbf{y}$ so that it can be written as a diagonal form in the variables y_1, y_2, y_3, and determine whether the quadratic form is positive definite:

a) $x_1^2 + x_2^2 + 2x_3^2 - 2x_1x_2 + 4x_1x_3 + 4x_2x_3$

b) $3x_1^2 + 2x_2^2 + 3x_3^2 + 2x_1x_3$

c) $3x_1^2 + 5x_2^2 + 4x_3^2 + 4x_1x_3 - 4x_2x_3$

d) $5x_1^2 + 2x_2^2 + 5x_3^2 + 4x_1x_2 - 8x_1x_3 - 4x_2x_3$

e) $x_1^2 - 5x_2^2 - x_3^2 + 4x_1x_2 + 6x_2x_3$

5. Determine which of the following matrices are positive definite:

a) $\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ b) $\begin{pmatrix} 3 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 1 \end{pmatrix}$ c) $\begin{pmatrix} 6 & 1 & 7 \\ 1 & 1 & 2 \\ 7 & 2 & 9 \end{pmatrix}$

d) $\begin{pmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{pmatrix}$ e) $\begin{pmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{pmatrix}$ f) $\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$