Hang Lung Mathematics Awards
Vol. 7 (2016)

Hang Lung Mathematics Awards 2016

Honorable Mention

On Hilbert Functions and Positive-definite Quadratic Forms

Team member: Chak Him Au
Teacher: Mr. Yan Ching Chan
School: Po Leung Kuk Centenary Li Shiu Chung Memorial College

Hang Lung Mathematics Awards
Vol. 7 (2016), pp. 145–160

ON HILBERT FUNCTIONS AND POSITIVE-DEFINITE QUADRATIC FORMS

TEAM MEMBER
Chak Him Au

TEACHER
Mr. Yan Ching Chan

SCHOOL
Po Leung Kuk Centenary Li Shiu Chung Memorial College

Abstract. In this project, we give an explicit construction of positive definite quadratic forms of arbitrary dimension by using a family of real analytic functions whose coefficients in their Taylor expansions are strictly positive. We also prove a variant result that allows the construction if the number of positive coefficients has a positive upper density.

1. Acknowledgement

In this project, there are several people giving assistance to my work. Without their contributions and efforts, this project will never be completed so successfully. Now I would like to express my sincere gratitude to them:

Firstly, my teacher advisor, Mr. Yan Ching Chan, plays an important role in the whole project. Not only did he encourage me to participate in this biannual event, he also held several meetings with me to discuss the content and directions of the project. At the meetings, I usually presented my work to him and he asked me some questions in order to strengthen our understandings to the content as well as giving me some useful feedback.

In addition to Mr.Chan's contributions, what is equally noteworthy is the assistance from Mr. Kwok Wing Tsoi. He gave me an inspiring initial problem and edited the project in latex so that it looks more decent and clear. What's more, he spotted two mathematical errors in the project when proofreading it. Also, his mathematical expertise is of great help to me when I faced with difficulties.

Moreover, Mr. Hok Kan Yu, also lent me a helping hand in the project. He was the one who taught me how to use latex. Without his patience and clear illustration, I can never master latex within such a short period of time.

2. Introduction and Main Results

2.1. Motivations: Quadratic Functions in One Variable

Let us first recall some elementary theory of quadratic functions. Suppose we are given a function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = ax^2 + bx + c$$

for some fixed real numbers a, b, c with $a > 0$. It is known that $f(x)$ always non-negative (indeed almost always positive) if its **discriminant** $\Delta(f) = b^2 - 4ac < 0$. The underlying principle of discriminant is the trick of **completing-the-square**. i.e. we can rewrite

$$f(x) = a\Big(x - \frac{b}{2}\Big)^2 + \Big(-\frac{\Delta(f)}{4a}\Big).$$

[See reviewer's comment (2)]

If we put $x = X_1/X_2$ and define $F : \mathbb{R}^2 \to \mathbb{R}$ by

$$F(X_1, X_2) = X_2^2 f(x) = aX_1^2 + bX_1X_2 + cX_2^2,$$

then the theory of quadratic functions immediately tells us that

Lemma 1. *If $a > 0$ and $\Delta(f) = b^2 - 4ac < 0$, then*

1 $F(X_1, X_2) \geq 0$ *for all* $X_1, X_2 \in \mathbb{R}$

2 $F(X_1, X_2) = 0$ *if and only if* $X_1 = X_2 = 0$.

For functions that satisfy the above properties in the above lemma, they are called positive-definite. In other words, the theory of quadratic functions provides us an easy criterion to choose $a, b, c \in \mathbb{R}$ such that the associated function $F(X_1, X_2) = aX_1^2 + bX_1X_2 + cX_2^2$ is positive-definite.

The analysis above can be regarded as a two-dimensional story of quadratic forms. In general, if we fix a positive integer n and consider a quadratic form (the terminology will be made precise in the following chapters) $Q : \mathbb{R}^n \to \mathbb{R}$ defined by

$$Q(x_1, x_2, \dots, x_n) \sum_{1\leq i\leq n} \sum_{1\leq j\leq n} a_{ij}X_iX_j$$

[See reviewer's comment (3)]

for some fixed $a_{ij} \in \mathbb{R}$, basic linear algebra allows us to generalise the trick of completing square and determine whether Q is positive-definite, as in the two-dimensional case. In other words, once we have fixed some a_{ij}, it is known how to determine whether Q is positive-definite. However, one may ask if there is a converse construction. More explicitly,

> Is there a systematic way to choose the coefficients a_{ij} such that Q is positive-definite?

In this project, we are going to discuss the above question and ultimately provide a concrete construction of a family of such quadratic forms. We will also see that such construction allows us to prove a family of non-trivial inequalities.

2.2. Main Results

In this section, we record the main results of this project.

Theorem 2. *Fix any positive integer n. Let $K(x) = \sum_{k=0}^{\infty} a_k x^k$ be a real analytic function with all $a_i > 0$ and radius of convergence R. If $\alpha_1, \alpha_2, \ldots, \alpha_n$ are pairwise distinct real numbers in the range $(-\sqrt{R}, \sqrt{R})$, then the quadratic form $Q_K : \mathbb{R}^n \to \mathbb{R}$ [See reviewer's comment (4)]*

$$Q_K(X_1, X_2, \ldots, X_n) = \sum_{1 \le i \le n} \sum_{1 \le j \le n} K(\alpha_i \alpha_j) X_i X_j$$

is positive-definite; in other words, we have

$$Q_K(X_1, X_2, \ldots, X_n) > 0$$

unless $X_1 = X_2 = \ldots = X_n = 0$.

Example 3. *Consider the function $K : (-1, 1) \to \mathbb{R}$ given by*

$$K(x) = \frac{1}{1-x} = \sum_{k=0}^{\infty} x^k.$$

This has radius of convergence 1. Then Theorem 2 asserts that for any fixed positive integer n and $\alpha_1, \alpha_2, \ldots, \alpha_n \in (-1, 1)$ which are pairwise distinct, we have

$$\sum_{1 \le i \le n} \sum_{1 \le j \le n} \frac{1}{1 - \alpha_i \alpha_j} X_i X_j > 0$$

for any $X_1, X_2, \ldots, X_n \in \mathbb{R}$ unless $X_1 = X_2 = \ldots = X_n = 0$.

Example 4. *Consider the exponential function $\exp : \mathbb{R} \to \mathbb{R}$ whose radius of convergence is ∞. Recall that*

$$\exp(x) = \sum_{k=0}^{\infty} \frac{1}{k!} x^k.$$

Then Theorem 2 asserts that for any fixed positive integer n and $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$ which are pairwise distinct, we have

$$\sum_{1 \le i \le n} \sum_{1 \le j \le n} \exp(\alpha_i \alpha_j) X_i X_j > 0$$

for any $X_1, X_2, \ldots, X_n \in \mathbb{R}$ *unless* $X_1 = X_2 = \ldots = X_n = 0$.

In addition to Theorem 2, we also prove the following variant version of it in Chapter 5 which allows construction of positive definite quadratic forms via a wider family of analytic functions with humble additional hypothesis on the choice of α_i's.

Theorem 5. *Fix any positive integer* n. *Let* $K(x) = \sum_{k=0}^{\infty} a_k x^k$ *be a real analytic function with* $a_k \geq 0$ *for all* k *and radius of convergence* R. *Define the set*

$$H_K = \{k \in \mathbb{N} \cup \{0\} : a_k > 0\}.$$

If H_K *has positive upper density, then for any* $\alpha_1, \alpha_2, \ldots, \alpha_n \in (-\sqrt{R}, \sqrt{R})$ *such that*

1 $\alpha_i \neq 0$ *for all* $i = 1, 2, \ldots, n$ *and*

2 $|\alpha_i| \neq |\alpha_j|$ *for any* $i \neq j$,

then the quadratic form $Q_K : \mathbb{R}^n \to \mathbb{R}$ *given by*

$$Q_K(X_1, X_2, \ldots, X_n) = \sum_{1 \leq i \leq n} \sum_{1 \leq j \leq n} K(\alpha_i \alpha_j) X_i X_j$$

is positive definite.

We will explain the terminology of Theorem 5 in Chapter 5 of this report. Roughly speaking, it says that if there are "sufficiently many" non-zero coefficients in the Taylor expansion of $K(x)$, then one can build a positive definite quadratic form in almost the same spirit of Theorem 2.

2.3. Outline of the Report

In this chapter, we have recorded the main results of the project. In order to make this report self-contained, we, in Chapter 3, are going to review the classical results from linear algebra and analysis that are essential in the proof of our main results. In particular, we will cite all the necessary results from quadratic forms and Hilbert spaces. After reviewing all the necessary background results, we are going to prove our main result (Theorem 2) in Chapter 4. In Chapter 5, we discuss a possible variant of our main result (Theorem 5) and give a proof of it using some classical results from additive combinatorics.

3. Background in Algebra and Analysis

In this chapter, we are going to give all the necessary background from algebra and analysis. We will state most of these standard results without proofs but illustrate them with explicit examples. The proofs of these results can be found in [1] or [3].

3.1. Quadratic Forms and Symmetric Matrices

Definition 6. *Let $n \geq 1$ be an integer. A quadratic form of dimension n is a function $Q : \mathbb{R}^n \to \mathbb{R}$ of the form*

$$Q(X_1, X_2, \ldots, X_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} X_i X_j$$

for some real numbers a_{ij}. In other words, it is a homogeneous polynomial of degree 2 in n variables.

Example 7. *$Q_1(x) = 5x^2$ is a unary quadratic form whereas $Q_2(x, y) = 6x^2 + 4xy - y^2$ is a binary quadratic form.*

Definition 8. *Let $Q : \mathbb{R}^n \to \mathbb{R}$ be a quadratic form.*

1. *Q is called positive semi-definite if $Q(X_1, X_2, \ldots, X_n) \geq 0$ for any X_1,X_2,..., $X_n \in \mathbb{R}$.*
2. *Q is called positive definite if it is positive semi-definite and $Q(X_1, X_2, \ldots, X_n) = 0$ if and only if $X_1 = X_2 = \ldots = X_n = 0$.*

Example 9. *The binary quadratic form $Q_1(x, y) = x^2 + y^2$ is clearly positive definite. However, the quadratic form $Q_2(x, y) = x^2$ is positive semi-definite but not positive definite; for example, we have $Q_2(0, 1) = 0$.*

Recall that a square matrix A is called symmetric if $A^T = A$ where A^T is the transpose of A. For any quadratic form $Q : \mathbb{R}^n \to \mathbb{R}$, we can associate a symmetric matrix with it as follows;

$$Q(X_1, X_2, \ldots, X_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} X_i X_j \longmapsto A_Q = (\alpha_{ij})_{1 \leq i,j \leq n},$$

where $\alpha_{ij} = \frac{a_{ij} + a_{ji}}{2}$ for all $1 \leq i, j \leq n$. Conversely, given a n-by-n symmetric matrix A, we can obtain a quadratic form Q_A of dimension n by defining

$$Q_A(X_1, X_2, \ldots, X_n) = x^T A x,$$

where $x^T = (X_1, X_2, \ldots, X_n)$.

Example 10. *For $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, its associated quadratic form is $Q_A(x, y) = 2x^2 + 2xy + 2y^2$. On the other hand, the symmetric matrix associated with the quadratic form $Q(x, y) = x^2 + 4xy + 2y^2$ is $\begin{pmatrix} 1 & 2 \\ 2 & 2 \end{pmatrix}$.*

We say that a symmetric matrix A positive (semi)-definite if its associated quadratic form is also positive (semi)-definite. For example, the identity matrix $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is positive-definite whereas $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ is only positive semi-definite.

Let A be a n-by-n square matrix. Recall that $\lambda \in \mathbb{C}$ is called an eigenvalue of A if there exists a non-zero $\mathbf{v} \in \mathbb{C}^n$ such that $A\mathbf{v} = \lambda\mathbf{v}$. Rearranging the equation gives $(A - \lambda I_n)\mathbf{v} = 0$. Since $\mathbf{v} \neq 0$, we observe that $\lambda \in \mathbb{C}$ is an eigenvalue of A if and only if λ satisfies $\det(A - \lambda I_n) = 0$.

Example 11. *Let* $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. *Since* $\det(A - xI_2) = x^2 - 4x + 3 = (x-1)(x-3)$. *The eigenvalues of A are 1 and 3.*

We record here a central result about real symmetric matrices.

Theorem 12. *Let A be a symmetric matrix whose entries are real numbers. Then*

1. *All eigenvalues of A are real.*
2. *There exists a real invertible matrix P such that $P^{-1}AP$ is a diagonal matrix whose diagonal entries consist of all the eigenvalues of A.*

Proof. See [1]. □

Example 13. *Let* $P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. *Since its determinant is -2, it is invertible and* $P^{-1}AP = \frac{1}{-2}\begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$, *which is a diagonal matrix whose diagonal entries consist of all the eigenvalues of A.*

Indeed, the second part of **Theorem 12** can be regarded as a generalisation of the trick of completing-the-square for binary quadratic forms. One immediate corollary is that one can determine the definiteness of a quadratic form completely from the signs of the eigenvalues of its associated symmetric matrix.

Corollary 14. *Let A be a real symmetric matrix. Then*

1. *A is positive semi-definite if and only if all its eigenvalues are non-negative.*
2. *A is positive definite if and only if all its eigenvalues are positive.*

Example 15. *Let* $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. *Since the eigenvalues of A are 1 and 3 which are both positive, the associated quadratic form* $Q_A(x,y) = 2x^2 + 2xy + 2y^2$ *is positive definite.*

Since eigenvalues of a matrix A are roots of the polynomial $f_A(x) = \det(A - xI)$, the product of all eigenvalues is given by the constant term of this polynomial $f_A(0) = \det(A)$. Therefore, combining **Corollary 14** with this observation, we have the following.

Corollary 16. *Let A be a real symmetric matrix. Then A is positive definite if and only if A is positive semi-definite and invertible.*

3.2. Inner Products and Hilbert Spaces

In this section, we let V be a real vector space.

Definition 17. *An inner product on the vector space V is a function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ such that for any $x, y, z \in V$, we have*

1. $\langle x, y \rangle = \langle y, x \rangle$,
2. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$,
3. $\langle x, x \rangle \geq 0$. *Moreover,* $\langle x, x \rangle = 0$ *if and only if* $x = 0$.

Example 18. *On the Euclidean space $\mathbb{R}^n$, an inner product is given by the standard scalar product, namely*

$$\langle (x_1, x_2, \cdots, x_n), (y_1, y_2, \cdots, y_n) \rangle = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

Example 19 (l^2-space)**.** *Define the space of square-summable sequences*

$$l^2(\mathbb{R}) = \left\{ (a_i)_{i=0}^{\infty} : a_i \in \mathbb{R} \text{ and } \sum_{i=1}^{\infty} a_i^2 \text{ converges} \right\}.$$

This is a real vector space with an inner product given by

$$\langle (a_i)_{i=0}^{\infty}, (b_i)_{i=0}^{\infty} \rangle = \sum_{i=0}^{\infty} a_i b_i$$

for any $(a_i)_{i=0}^{\infty}, (b_i)_{i=0}^{\infty} \in l^2(\mathbb{R})$. Note that the infinite summation on the right hand side converges because the AM-GM inequality says $a_i b_i \leq \dfrac{a_i^2 + b_i^2}{2}$ for all $i = 0, 1, 2, \ldots$ and the sequences $(a_i)_{i=0}^{\infty}, (b_i)_{i=0}^{\infty} \in l^2(\mathbb{R})$. [See reviewer's comment (5)]

For a real vector space V with an inner product $\langle \cdot, \cdot \rangle$, one can associate a distance function with it by defining

$$d(x, y) = \sqrt{\langle x - y, x - y \rangle}.$$

For example, on the Euclidean space $\mathbb{R}^n$ with the standard scalar product, the associated distance function is the usual Euclidean distance given by

$$((x_1, x_2, \cdots, x_n), (y_1, y_2, \cdots, y_n)) = \sqrt{\sum_{i=0}^{n} (x_i - y_i)^2}.$$

[See reviewer's comment (6)] A so-called real Hilbert space is a real vector space with an inner product which satisfies a strong analytic property with respect to its associated distance function: any Cauchy sequence converges in the space with respect to the associated distance function. The slogan is

A Hilbert Space is an inner product space in which all sequences that should converge indeed converge.

We are not going to give a precise definition of a Hilbert space but we should record the following standard fact from real analysis.

Theorem 20. *Both the Euclidean n-space $\mathbb{R}^n$ and $l^2(\mathbb{R})$ are real Hilbert spaces.*

Proof. See [3] $\square$

Since in the rest of this report, any Hilbert space would either refers to the Euclidean space $\mathbb{R}^n$ or the sequence space $l^2(\mathbb{R})$, it is safe to bear no more than these two examples in mind to read the rest of this report. We end this section with a main theorem that is useful to us concerning Hilbert spaces.

Definition 21. *Fix any n and $X \subset \mathbb{R}^n$ be closed. A function $K : X \times X \to \mathbb{R}$ is called a Hilbert function if there exists a function $\phi : X \to \mathcal{W}$ where $\mathcal{W}$ is a real Hilbert space such that*

$$K(x_1, x_2) = \langle \phi(x_1), \phi(x_2) \rangle_{\mathcal{W}},$$

where $\langle \cdot, \cdot \rangle_{\mathcal{W}}$ is the inner product endowed by $\mathcal{W}$.

Example 22. *Let $\mathcal{W} = \mathbb{R}^2$ with the standard inner product. The function $K : \mathbb{R}^2 \to \mathbb{R}$*

$$K((x_1, x_2), (y_1, y_2)) = x_1 y_1 + x_2 y_2 = \langle (x_1, y_1), (x_2, y_2) \rangle_{\mathcal{W}}$$

is a Hilbert function by taking $\phi : \mathbb{R}^2 \to \mathcal{W}$ to be the identity map.

Theorem 23 (Mercer)**.** *Fix any n and $X \subset \mathbb{R}^n$ be closed. A function $K : X \times X \to \mathbb{R}$ is a Hilbert function if and only if for any positive integer m and arbitrary real numbers $\alpha_1, \alpha_2, \ldots, \alpha_m$, the matrix*

$$\begin{pmatrix} K(\alpha_1, \alpha_1) & K(\alpha_1, \alpha_2) & \cdots & K(\alpha_1, \alpha_m) \\ K(\alpha_2, \alpha_1) & K(\alpha_2, \alpha_2) & \cdots & K(\alpha_2, \alpha_m) \\ \vdots & \vdots & \ddots & \vdots \\ K(\alpha_m, \alpha_1) & K(\alpha_m, \alpha_2) & \cdots & K(\alpha_m, \alpha_m) \end{pmatrix}$$

is positive semi-definite.

3.3. Special Matrices and Properties

In this last section, we cite two results from linear algebra concerning matrices of special forms.

Theorem 24 (Vandermonde)**.** *Let $a_1, a_2, \ldots, a_n$ be complex numbers. Then we have*

$$\det \begin{pmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots & \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{pmatrix} = \prod_{1 \le i < j \le n} (x_j - x_i).$$

[See reviewer's comment (7)]

Proof. See [1] □

Theorem 25. *Let V be a real vector space with an inner product $\langle\cdot,\cdot\rangle$. For any fixed n and $v_1, v_2, \ldots, v_n \in V$, the matrix*

$$\begin{pmatrix} \langle v_1, v_1\rangle & \langle v_1, v_2\rangle & \cdots & \langle v_1, v_n\rangle \\ \langle v_2, v_1\rangle & \langle v_2, v_2\rangle & \ldots & \langle v_1, v_1\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle v_n, v_1\rangle & \langle v_n, v_2\rangle & \cdots & \langle v_n, v_n\rangle \end{pmatrix}$$

is invertible if and only if $v_1, v_2, \ldots, v_n$ are linearly independent in V.

[See reviewer's comment (8)]

4. Proof of the Main Result

We recall the main theorem of this report.

Theorem 26 ((Theorem 2))**.** *Fix any positive integer n. Let $K(x) = \sum_{k=0}^{\infty} a_k x^k$ be a real analytic function with all $a_k > 0$ and radius of convergence R. If α_1, α_2,...,α_n are pairwisely distinct real numbers in the range $(-\sqrt{R}, \sqrt{R})$, then the quadratic form $Q_K : \mathbb{R}^n \to \mathbb{R}$*

$$Q_K(X_1, X_2, \ldots, X_n) = \sum_{1\le i\le n} \sum_{1\le j\le n} K(\alpha_i\alpha_j) X_i X_j$$

is positive-definite; in other words, we have

$$Q_K(X_1, X_2, \ldots, X_n) > 0$$

unless $X_1 = X_2 = \ldots = X_n = 0$.

The proof occupies the rest of this chapter.

4.1. Proof of Theorem 2

Let $K(x) = \sum_{k=0}^{\infty} a_k x^k$ be a real analytic function with all $a_k > 0$ whose radius of convergence equals to R (possibly infinite). Fix a positive integer n and $\alpha_1, \alpha_2, \ldots, \alpha_n$ be distinct real numbers in $(-\sqrt{R}, \sqrt{R})$.

First we observe that since by hypothesis $a_k > 0$ for all $k = 0, 1, 2, \ldots$, for any $1 \le i, j \le n$, we can rewrite

$$K(\alpha_i\alpha_j) = \sum_{k=0}^{\infty} a_k (\alpha_i\alpha_j)^k = \sum_{k=0}^{\infty} (\sqrt{a_k}\alpha_i)^k (\sqrt{a_k}\alpha_j)^k. \tag{1}$$

[See reviewer's comment (9)]
We also note that for any $1 \le i \le n$, the infinite sum

$$\sum_{k=0}^{\infty}((\sqrt{a_k}\alpha_i)^k)^2 = \sum_{k=0}^{\infty} a_k(\alpha_i^2)^k = K(\alpha_i^2) \tag{2}$$

[See reviewer's comment (10)]
converges because $\alpha_i^2 \in (-R, R)$ and the radius of convergence of $K(x)$ equals to R. Write

$$\mathcal{W} = l^2(\mathbb{R}) = \left\{(a_i)_{i=0}^{\infty} : a_i \in \mathbb{R} \text{ and } \sum_{i=1}^{\infty} a_i^2 \text{ converges}\right\}.$$

be the Hilbert space of square-summable sequences (See Example 19). Define $\epsilon > 0$ to be any positive real numbers small enough such that all α_i's are contained in the closed interval $[-\sqrt{R}+\epsilon, \sqrt{R}-\epsilon]$. Now we define a function $\phi : [-\sqrt{R}+\epsilon, \sqrt{R}-\epsilon] \to l^2(\mathbb{R})$ by

$$x \longmapsto (\sqrt{a_0}, \sqrt{a_1}x, \sqrt{a_2}x^2, \ldots, \sqrt{a_k}x^k, \ldots)$$

This function is well-defined (i.e. the image is indeed square-summable) by the convergence analysis in (2) above. As a result, combining with (1), we can rewrite

$$K(\alpha_i\alpha_j) = \sum_{k=0}^{\infty} a_k(\alpha_i\alpha_j)^k = \langle \phi(\alpha_i), \phi(\alpha_j)\rangle_{\mathcal{W}} \tag{3}$$

for any $1 \le i, j \le n$. Now we consider the quadratic form

$$Q_K(X_1, X_2, \ldots, X_n) = \sum_{1\le i\le n}\sum_{1\le j\le n} K(\alpha_i\alpha_j)X_iX_j.$$

The symmetric matrix associated with the quadratic form is

$$A_K = \begin{pmatrix} K(\alpha_1^2) & K(\alpha_1\alpha_2) & \cdots & K(\alpha_1\alpha_n) \\ K(\alpha_2\alpha_1) & K(\alpha_2^2) & \cdots & K(\alpha_2\alpha_n) \\ \vdots & \vdots & \ddots & \vdots \\ K(\alpha_n\alpha_1) & K(\alpha_n\alpha_2) & \cdots & K(\alpha_n^2) \end{pmatrix}$$

Now, using (3), we can rewrite this matrix as

$$A_K = \begin{pmatrix} \langle \phi(\alpha_1), \phi(\alpha_1)\rangle_{\mathcal{W}} & \langle \phi(\alpha_1), \phi(\alpha_2)\rangle_{\mathcal{W}} & \cdots & \langle \phi(\alpha_1), \phi(\alpha_n)\rangle_{\mathcal{W}} \\ \langle \phi(\alpha_2), \phi(\alpha_1)\rangle_{\mathcal{W}} & \langle \phi(\alpha_2), \phi(\alpha_2)\rangle_{\mathcal{W}} & \cdots & \langle \phi(\alpha_2), \phi(\alpha_n)\rangle_{\mathcal{W}} \\ \vdots & \vdots & \ddots & \vdots \\ \langle \phi(\alpha_n), \phi(\alpha_1)\rangle_{\mathcal{W}} & \langle \phi(\alpha_n), \phi(\alpha_2)\rangle_{\mathcal{W}} & \cdots & \langle \phi(\alpha_n), \phi(\alpha_n)\rangle_{\mathcal{W}} \end{pmatrix}. \tag{4}$$

Therefore, by Mercer's Theorem (Theorem 23), we deduce that the matrix A_K is positive semi-definite. Last but not least, we are going to prove the following.

Claim 27. *The matrix A_K is invertible.*

Proof of Claim. By using (4) and Theorem 25, the claim is equivalent to proving that $\phi(\alpha_1), \phi(\alpha_2), \ldots, \phi(\alpha_n)$ are linearly independent in the space $l^2(\mathbb{R})$.

Suppose $\lambda_1, \lambda_2, \ldots, \lambda_n$ are real numbers such that

$$\lambda_1\phi(\alpha_1) + \lambda_2\phi(\alpha_2) + \ldots + \lambda_n\phi(\alpha_n) = 0 \tag{5}$$

where $0 \in l^2(\mathbb{R})$ denotes the zero sequence $(0, 0, 0, \ldots)$. Recall that the function ϕ is given by

$$x \longmapsto (\sqrt{a_0}, \sqrt{a_1}x, \sqrt{a_2}x^2, \ldots, \sqrt{a_k}x^k, \ldots)$$

If we expand the left hand side of (5),we have

$$\left(\sqrt{a_0}\sum_{i=1}^{n}\lambda_i, \sqrt{a_1}\sum_{i=1}^{n}\lambda_i\alpha_i, \ldots, \sqrt{a_k}\sum_{i=1}^{n}\lambda_i\alpha_i^k, \ldots\right) = 0. \tag{6}$$

By equating the first n terms of (6), we have

$$\sqrt{a_0}\sum_{i=1}^{n}\lambda_i = \sqrt{a_1}\sum_{i=1}^{n}\lambda_i\alpha_i = \ldots = \sqrt{a_{n-1}}\sum_{i=1}^{n}\lambda_i\alpha_i^{n-1} = 0.$$

Since, by hypothesis, we have $a_0, a_1, \ldots, a_{n-1}$ are all positive. This reduces to a system of linear equations in λ_i's as follow

$$\sum_{i=1}^{n}\lambda_i = \sum_{i=1}^{n}\lambda_i\alpha_i = \ldots = \sum_{i=1}^{n}\lambda_i\alpha_i^{n-1} = 0.$$

We can rewrite this system of equations in the matrix form as follows.

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_n \\ \alpha_1^2 & \alpha_2^2 & \alpha_3^2 & \cdots & \alpha_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{n-1} & \alpha_2^{n-1} & \alpha_3^{n-1} & \cdots & \alpha_n^{n-1} \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{7}$$

By Theorem 24 (Vandermonde), the square matrix in the left hand side of (7) had determinant

$$\Delta = \prod_{1 \le i < j \le n} (\alpha_j - \alpha_i).$$

In particular, by hypothesis, the α_i's are chosen to be pairwise distinct. As a result, the determinant Δ is non-zero. Hence, the homogeneous system (7) has only trivial solution, namely,

$$\lambda_1 = \lambda_2 = \lambda_3 = \ldots = \lambda_n = 0.$$

This proves that $\phi(\alpha_1), \phi(\alpha_2), \ldots, \phi(\alpha_n)$ are linearly independent in $l^2(\mathbb{R})$ and hence the result follows by Theorem 25. $\square$

Since the matrix A_K is

- positive semi-definite by Mercer's Theorem (Theorem 23) and
- invertible (by Claim),

we conclude that A_K is positive definite by Corollary 16. This finishes the proof of Theorem 2.

5. Variants of the Main Result

In this last chapter, we are going for possible generalisations of our main result. Our main result (Theorem 2) requires the Taylor series $K(x)$ to satisfy a rather strong hypothesis: we require all the coefficients in the Taylor expansion of $K(x)$ to be positive. For example, our main result is not applicable to the function $K : \mathbb{R} \to \mathbb{R}$ given by

$$K(x) = x\exp(x) = \sum_{k=1}^{\infty} \frac{1}{(k-1)!} x^k.$$

However, we shall see that for any fixed positive integer n, if we choose any **non-zero** pairwise distinct real numbers $\alpha_1, \alpha_2, \ldots, \alpha_n$, the quadratic form

$$Q_K(X_1, X_2, \ldots, X_n) = \sum_{1\le i\le n} \sum_{1\le j\le n} K(\alpha_i\alpha_j) X_i X_j$$

is indeed positive definite. More generally, we shall see that by imposing very mild hypothesis on the α_i's, we can construct positive definite quadratic form from a more general family of real analytic functions $K(x)$.

5.1. A Variant Version of Theorem 2

Let us look at the proof of Theorem 2 (See Section 4.1) again. While we were trying to establish the linear independence of the elements $\phi(\alpha_1), \phi(\alpha_2), \ldots, \phi(\alpha_n)$, we show this by **manually** chosen to equate the first n-terms of (6). Therefore, it is rather natural to ask whether we can instead choose any other n terms to equate to recover the linear independence. It turns out that we can achieve it by sacrificing some freedom in choosing the terms α_i's.

In particular, we shall prove the following variant of our main result.

Theorem 28 (Theorem 5)**.** *Fix any positive integer n. Let $K(x) = \sum_{k=0}^{\infty} a_k x^k$ be a real analytic function with $a_k \ge 0$ for all k and radius of convergence R. Define the set*

$$H_K = \{k \in \mathbb{N} \cup \{0\} : a_k > 0\}.$$

If H_K has positive upper density, then for any $\alpha_1, \alpha_2, \ldots, \alpha_n \in (-\sqrt{R}, \sqrt{R})$ such that

1. *$\alpha_i \ne 0$ for all $i = 1, 2, \ldots, n$ and*
2. *$|\alpha_i| \ne |\alpha_j|$ for any $i \ne j$,*

then the quadratic form $Q_K : \mathbb{R}^n \to \mathbb{R}$ given by

$$Q_K(X_1, X_2, \ldots, X_n) = \sum_{1\le i\le n} \sum_{1\le j\le n} K(\alpha_i\alpha_j) X_i X_j$$

is positive definite.

We will explain the terminology of this theorem in the following sections.

5.2. Results from Additive Combinatorics

To prove Theorem 28, we need to employ some classical results from additive combinatorics. First we shall introduce the notion of density in natural numbers.

Definition 29. *Let A be a subset of $\mathbb{N} \cup \{0\}$. The upper density of A is given by*

$$d(A) = \limsup_{n\to\infty} \frac{|A \cap \{0, 1, 2, 3, \ldots, n\}|}{n}$$

Example 30. *Any finite subset of $\mathbb{N} \cup \{0\}$ has upper density zero.*

Example 31. *By the Prime Number Theorem, the set of prime numbers also has upper density zero.*

Example 32. *Fix a pair of positive integers a and k. Consider the set*

$$A = \{k + na : n \in \mathbb{N} \cup \{0\}\}$$

Then it is easy to show that

$$d(A) = \frac{1}{a} > 0$$

so A has a positive upper density.

The following is a classical result in additive combinatorics regarding behaviour of positively dense subsets of natural numbers.

Theorem 33 (Szemerédi)**.** *Let A be a subset of $\mathbb{N} \cup \{0\}$. If A has a positive upper density, then A contains an arithmetic progression of arbitrary length.*

Roughly speaking, Szemerédi's Theorem says that if a subset of natural numbers is "large enough", it should contain an arithmetic progression of any specified length. (indeed, Szemerédi's Theorem says that such subset would contain infinitely many arithmetic progressions of any specified length)

5.3. Proof of Theorem 5

Fix a positive integer n. Let $K(x) = \sum_{k=0}^{\infty} a_k x^k$ and $\alpha_1, \alpha_2, \ldots, \alpha_n$ be as claimed in the hypothesis of Theorem 28. Following the proof of Theorem 2, if we put $\mathcal{W} = l^2(\mathbb{R})$ and define ϕ as in (3), we can rewrite the matrix associated with the quadratic form Q_K as follows.

$$A_k = \begin{pmatrix} \langle\phi(\alpha_1), \phi(\alpha_1)\rangle_{\mathcal{W}} & \langle\phi(\alpha_1), \phi(\alpha_2)\rangle_{\mathcal{W}} & \cdots & \langle\phi(\alpha_1), \phi(\alpha_n)\rangle_{\mathcal{W}} \\ \langle\phi(\alpha_2), \phi(\alpha_1)\rangle_{\mathcal{W}} & \langle\phi(\alpha_2), \phi(\alpha_2)\rangle_{\mathcal{W}} & \cdots & \langle\phi(\alpha_2), \phi(\alpha_n)\rangle_{\mathcal{W}} \\ \vdots & \vdots & \ddots & \vdots \\ \langle\phi(\alpha_n), \phi(\alpha_1)\rangle_{\mathcal{W}} & \langle\phi(\alpha_n), \phi(\alpha_2)\rangle_{\mathcal{W}} & \cdots & \langle\phi(\alpha_n), \phi(\alpha_n)\rangle_{\mathcal{W}} \end{pmatrix}. \tag{8}$$

Therefore, Mercer's Theorem (Theorem 23) says that A_K is positive semi-definite and by Corollary 16, it is now sufficient to prove that A_K is invertible. In addition, by Theorem 25, it is equivalent to prove that $\phi(\alpha_1), \phi(\alpha_2), \ldots, \phi(\alpha_n)$ are linearly independent in $l^2(\mathbb{R})$.

Suppose $\lambda_1, \lambda_2, \ldots, \lambda_n$ are real numbers such that

$$\lambda_1\phi(\alpha_1) + \lambda_2\phi(\alpha_2) + \ldots + \lambda_n\phi(\alpha_n) = 0.$$

Equivalently, it says

$$\left(\sqrt{a_0}\sum_{i=1}^{n}\lambda_i, \sqrt{a_1}\sum_{i=1}^{n}\lambda_i\alpha_i, \ldots, \sqrt{a_k}\sum_{i=1}^{n}\lambda_i\alpha_i^k, \ldots\right) = 0. \tag{9}$$

By hypothesis, the set

$$H_K = \{k \in \mathbb{N} \cup \{0\} : a_k > 0\}.$$

has positive upper density. Therefore, using Szemerédi's Theorem (Theorem 33), H_K contains an arithmetic progression of length n. We write

$$A = \{aN + k : N = 0, 1, 2, \ldots, n-1\} \subseteq H_K$$

to be such arithmetic progression. By equating the l-th terms of the equation (10) for all $l \in A$, we obtain a system of linear equations in λ_i's as follows.

$$\sum_{i=1}^{n}\lambda_i\alpha_i^l \quad \text{for all } l \in A$$

We can rewrite this system of linear equations in matrix form.

$$\begin{pmatrix} \alpha_1^k & \alpha_2^k & \alpha_3^k & \cdots & \alpha_n^k \\ \alpha_1^{k+a} & \alpha_2^{k+a} & \alpha_3^{k+a} & \cdots & \alpha_n^{k+a} \\ \alpha_1^{k+2a} & \alpha_2^{k+2a} & \alpha_3^{k+2a} & \cdots & \alpha_n^{k+2a} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{k+(n-1)a} & \alpha_2^{k+(n-1)a} & \alpha_3^{k+(n-1)a} & \cdots & \alpha_n^{k+(n-1)a} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{10}$$

Write M_K to be the square matrix that appears on the left of (10). Then by using Theorem 24 (Vandermonde), we have

$$\begin{aligned} \det(M_K) &= \prod_{i=1}^{n}\alpha_i^k \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ \alpha_1^a & \alpha_2^a & \alpha_3^a & \cdots & \alpha_n^a \\ \alpha_1^{2a} & \alpha_2^{2a} & \alpha_3^{2a} & \cdots & \alpha_n^{2a} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{(n-1)a} & \alpha_2^{(n-1)a} & \alpha_3^{(n-1)a} & \cdots & \alpha_n^{(n-1)a} \end{vmatrix} \\ &= \prod_{i=1}^{n}(\alpha_i)^k \prod_{1\le i<j\le n}(\alpha_j^a - \alpha_i^a) \end{aligned}$$

Now we consider two cases.

- If a is odd, then we rewrite

$$\det(M_K) = \prod_{i=1}^{n} \alpha_i^k \prod_{1\le i<j\le n} \alpha_i^n \left(\left(\frac{\alpha_j}{\alpha_i} \right)^a - 1 \right). \tag{11}$$

  Recall that the polynomial $x^a - 1$ has a distinct roots

$$\{\exp(2\pi i m/a) : m = 0, 1, 2, \ldots, a-1\}.$$

  It is clear that when a is odd, its only real root is 1. Hence the second product on the right hand side of (11) is non-zero. Moreover, by hypothesis, none of the α_i's are zeroes. We can conclude that $\det(M_K) \neq 0$ and thus the matrix M_K is invertible in this case.
- If a is even, then we can rewrite

$$\det(M_K) = \prod_{i=1}^{n} \alpha_i^k \prod_{1\le i<j\le n} (\alpha_j^2 - \alpha_i^2)(\alpha_j^{a-2} + \alpha_j^{a-4}\alpha_i^2 + \cdots + \alpha_i^{a-2}).$$

  Since a is even and none of the α_i's are zero, for any fixed $1 \le i < j \le n$, the factor

$$\alpha_j^{a-2} + \alpha_j^{a-4}\alpha_i^2 + \cdots + \alpha_i^{a-2}$$

  is always positive. In addition, we also have, by hypothesis, that $|\alpha_i| \neq |\alpha_j|$ for any $i \neq j$. This asserts that $\det(M_K) \neq 0$ and thus the matrix M_K is invertible in this case.

In either case, the matrix M_K is invertible and thus the only solution to the homogeneous system (10) is the trivial one. i.e.

$$\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0.$$

Hence, we have established the linear independence of $\phi(\alpha_1), \phi(\alpha_2), \ldots, \phi(\alpha_n)$ and we are done.

REFERENCES

[1] Lay, D. C., *Linear algebra and its applications*, 4th ed., Pearson Education, 2010.

[2] Cristianini, N., Shawe-Taylor, J. and Christianini, N., *An introduction to support vector machines: And other kernel-based learning methods*, Cambridge University Press, 2000.

[3] Rudin, W., *Principles of mathematical analysis*, 3rd ed., McGraw-Hill Higher Education, New York, 1976.

[4] Szemerédi, E., *On sets of integers containing no k elements in arithmetic progression*, Acta Mathematica **27** (1975), 199-245.

[5] Stein, E. M., Shakarchi, R., *Functional analysis: Introduction to further topics in analysis*, Princeton University Press, 2008.

[6] Golan, J. S., *The linear algebra a beginning graduate student ought to know*, 3rd ed., Springer, Dordrecht, 2012.

Reviewer's Comments

Grammatical mistakes and typos

1 The reviewer has comments on the wordings, which have been amended in this paper.
2 $f(x) = a\left(x + \frac{b}{2a}\right)^2 + \left(-\frac{\Delta(f)}{4a}\right)$
3 Left hand side should be $Q(X_1, X_2, \ldots, X_n)$
4 radius of convergence R (possibly $+\infty$)
5 $a_i b_i \to |a_i b_i|$
6 Left hand side should be $d((x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n))$
7 Right hand side should be $\prod\limits_{1 \leq i < j \leq n} (a_j - a_i)$
8 The last entry on the second row of the matrix should be $\langle v_2, v_n \rangle$.
9 Right hand side should be $\sum\limits_{k=0}^{\infty} (\sqrt{a_k}\alpha_i^k)(\sqrt{a_k}\alpha_j^k)$.
10 Left hand side should be $\sum\limits_{k=0}^{\infty} (\sqrt{a_k}\alpha_i^k)^2$.

Comments

The paper is about constructing positive-definite quadratic forms using real analytic functions with positive coefficients. The method in the paper is interesting. The author made use of such a real analytic function to construct an embedding of an Euclidean space $\mathbb{R}^n$ into the Hilbert space l^2. The quadratic form is then given by the pullback of the inner product on l^2 under that embedding. The author also proved a variant result on a wider class of analytic functions using results from additive combinatorics.

The paper is well-organized with a summary of results, background materials, examples and proofs of theorems. The last part of the proof of Theorem 5 can be simplified. It is pretty clear that $\det(M_K)$ is non-zero from the expression of it obtained from Theorem 24. There is no need to consider the cases whether a is even or odd separately.