NTNU
Norwegian University of Science and Technology

Inequalities in Hilbert Spaces

Jan Wigestrand

Master of Science in Mathematics
Submission date: March 2008
Supervisor: Eugenia Malinnikova, MATH

Norwegian University of Science and Technology
Department of Mathematical Sciences

NTNU Norwegian University of Science and Technology>
Faculty of Information Technology, Mathematics and
Electrical Engineering>
Department of Mathematical Sciences.

Master Thesis

Inequalities in Hilbert Spaces

by

Jan Wigestrand

Supervisor: Eugenia Malinnikova

Trondheim, March 25, 2008.

Contents

Acknowledgments

I would like to thank Eugenia Malinnikova, for her most valuable help in writing this thesis.

The start date is March 26, 2007.
The deadline is March 26, 2008.

Jan Wigestrand,
Trondheim, March 25, 2008.

Introduction

The main result in this thesis is a new generalization of Selberg's inequality in Hilbert spaces with a proof, see page 25.

In Chapter 1 we define Hilbert spaces and give a proof of the Cauchy-Schwarz inequality and the Bessel inequality. As an example of application of the Cauchy-Schwarz inequality and the Bessel inequality, we give an estimate *for the dimension of an eigenspace* of an integral operator.
Next we give a proof of Selberg's inequality including the equality conditions following [Furuta].

In Chapter 2 we give selected facts on positive semidefinite matrices with proofs or references.
Then we use this theory for positive semidefinite matrices to study inequalities.
First we give a proof of a generalized Bessel inequality following [Akhiezer,Glazman], then we use the same technique to give a new proof of Selberg's inequality.
We conclude with a new generalization of Selberg's inequality with a proof.
In the last section of Chapter 2 we show how the matrix approach developed in Chapter 2.1 and Chapter 2.2 can be used to obtain optimal frame bounds.

We introduce a new notation for frame bounds, see page vii.

Notation

a b	frame bounds `( \text{\LARGE{$a$}} \text{\Large{$b$}} )`.
$\mathbb{P}$	the set $\{1, 2, 3, \ldots\}$ of all positive integers.
$\mathbb{N}$	the set $\{0, 1, 2, 3, \ldots\}$ of all nonnegative integers.
$\mathbb{R}$	the set of all real numbers.
$\mathbb{C}$	the set of all complex numbers $z = a + ib$ ($a \in \mathbb{R}$, $b \in \mathbb{R}$, $i^2 = -1$).
$\mathcal{H}$	Hilbert space.
A^*	complex conjugate transpose matrix of A, $A^* = \overline{A}^T$.
$A \geq 0$	A is positive semidefinite.
$A > 0$	A is positive definite.
$A^{\frac{1}{2}}$	is the square root of a positive semidefinite matrix A.
I	identity matrix.
U	unitary matrix, $UU^* = U^*U = I$.
$u \perp v$	u and v are orthogonal vectors.
λ	eigenvalue.
diag$(\lambda_1, \ldots, \lambda_n)$	diagonal matrix.
$\lambda_{\max}(A)$	largest eigenvalue of matrix A.
$\lambda_{\min>0}(A)$	smallest positive eigenvalue of matrix A.

Chapter 1

Classical inequalities

1.1 Hilbert Spaces

We will study inequalities in Hilbert spaces and in this section we give the definitions and examples of Hilbert spaces.

1.1.1 Hilbert spaces

Definition 1. A vector space $\mathcal{H}$ with a map $\langle .,. \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$
(for real vector spaces $\langle .,. \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$) is called an
inner product space if the following properties are satisfied:
(I1) $\langle x, x \rangle = 0 \Leftrightarrow x = 0$,
(I2) $\langle x, x \rangle \geq 0 \quad \forall x \in \mathcal{H}$,
(I3) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle \quad \forall x \in \mathcal{H} \, \forall y \in \mathcal{H} \, \forall z \in \mathcal{H}$,
(I4) $\langle \alpha x, x \rangle = \alpha \langle x, x \rangle \quad \forall x \in \mathcal{H} \, \forall \alpha \in \mathbb{C}$ (for real vector spaces $\alpha \in \mathbb{R}$),
(I5) $\langle x, y \rangle = \overline{\langle y, x \rangle} \quad \forall x \in \mathcal{H} \, \forall y \in \mathcal{H}$ (the bar denotes complex conjugation).

If in addition $\mathcal{H}$ is complete, that is
(I6) $\left(\lim_{n,m \to \infty} \langle x_n - x_m, x_n - x_m \rangle = 0 \quad x_n \in \mathcal{H} \, \forall n \in \mathbb{P} \, \forall m \in \mathbb{P} \right) \Rightarrow$
$\left(\exists x \in \mathcal{H} \quad \lim_{n \to \infty} \langle x - x_n, x - x_n \rangle = 0 \right)$,
then $\mathcal{H}$ is called a **Hilbert space**.

From now on $\mathcal{H}$ will denote a Hilbert space.

The norm in $\mathcal{H}$ is defined by

(I7) $\|x\| = \sqrt{\langle x, x\rangle} \quad \forall x \in \mathcal{H}$.

(I8) Every Hilbert space has an orthonormal basis, see [Folland,p176].

It means that there exists a system $\{e_\alpha\}_{\alpha\in Å}$ of elements in $\mathcal{H}$ that is linearly independent, that is $\langle e_\alpha, e_\beta\rangle = 0$ if $\alpha \neq \beta$ and $\|e_\alpha\| = \sqrt{\langle e_\alpha, e_\alpha\rangle} = 1$ for each α and for each $x \in \mathcal{H}$ we have $x = \sum_{\alpha\in Å} \langle x, e_\alpha\rangle e_\alpha$ (the series converges in $\mathcal{H}$).
If we have a separable Hilbert space we can replace $\{e_\alpha\}_{\alpha\in Å}$ by $\{e_j\}_{j\geq 1}$ and the sentence above can be reformulated in the following way:
It means that there exists a system $\{e_j\}_{j\geq 1}$ of elements in $\mathcal{H}$ that is linearly independent, that is $\langle e_j, e_k\rangle = 0$ if $j \neq k$ and $\|e_j\| = \sqrt{\langle e_j, e_j\rangle} = 1$ for each j and for each $x \in \mathcal{H}$ we have $x = \sum_{j\geq 1} \langle x, e_j\rangle e_j$ (the series converges in $\mathcal{H}$).

From now on a Hilbert space will be synonymous with a separable Hilbert space unless otherwise specified.

When the basis of $\mathcal{H}$ is finite we say that $\mathcal{H}$ is finite dimensional otherwise we say that $\mathcal{H}$ has infinite dimension.

1.1.2 Examples of Hilbert spaces

(a) $\mathbb{R}^3$ (real vector space) is a three-dimensional Hilbert space.

$\langle x, y\rangle = x_1y_1 + x_2y_2 + x_3y_3 \quad x,\ y$ are vectors in $\mathbb{R}^3$.

$\langle x, y\rangle = x^T y \quad x,\ y$ are vectors in $\mathbb{R}^3$. $x^T = [x_1\ x_2\ x_3]$, $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$.

$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ is an orthonormal basis for $\mathbb{R}^3$.

(b) $\mathbb{R}^n$ (real vector space) is an n-dimensional Hilbert space.

$\langle x, y\rangle = \sum_{j=1}^{n} x_j y_j \quad x,\ y$ are vectors in $\mathbb{R}^n$.

$\langle x, y\rangle = x^T y \quad x,\ y$ are vectors in $\mathbb{R}^n$, $x^T = [x_1\ x_2\ \dots\ x_n]$, $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$.

An orthonormal basis for $\mathbb{R}^n$ consists of n vectors each of dimension n.

$$\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}, \begin{bmatrix}0\\1\\\vdots\\0\end{bmatrix}, \ldots, \begin{bmatrix}0\\0\\\vdots\\1\end{bmatrix}$$ is an orthonormal basis for $\mathbb{R}^n$.

(c) $\langle x, y\rangle_A = (Ax)^T y = x^T A^T y = x^T A y \quad x, y$ are vectors in $\mathbb{R}^n$, $A \in \mathbb{R}^{n\times n}$,
A is positive definite ($x^T A x > 0 \; x \neq 0$, for more details see Chapter 2.1.1).
(I3), (I4) are clearly satisfied.
We see that if A is positive definite, then (I1) and (I2) are satisfied.
A is positive definite implies that A is symmetric ($A^T = A$).
We use the property that A is symmetric to show that (I5) is satisfied.

$$\langle x, y\rangle_A = (Ax)^T y = x^T A^T y = (x^T A^T y)^T = y^T A x = (A^T y)^T x$$
$$= \langle y, x\rangle_{A^T} = \langle y, x\rangle_A \quad x, y \text{ are vectors in } \mathbb{R}^n,\ A \in \mathbb{R}^{n\times n}.$$

It follows that (I5) is satisfied and that $\langle x, y\rangle_A$ is an inner product.
Since A is symmetric the eigenvectors from different eigenspaces are orthogonal. We can find an orthonormal basis for A by first finding a basis for each eigenspace of A, then apply the Gram-Schmidt process to each of these bases.

(d) $\mathbb{C}^n$ (complex vector space) is an n-dimensional Hilbert space.

$$\langle x, y\rangle = \sum_{j=1}^{n} x_j \overline{y_j} \quad x, y \text{ are vectors in } \mathbb{C}^n.$$

$$\langle x, y\rangle = x^T \overline{y} \quad x, y \text{ are vectors in } \mathbb{C}^n,\ x^T = [x_1\ x_2\ \ldots\ x_n],\ x = \begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}.$$

An orthonormal basis for $\mathbb{C}^n$ consists of n vectors each of dimension n.

$$\begin{bmatrix}\frac{1}{\sqrt{2}}(1+i)\\0\\\vdots\\0\end{bmatrix}, \begin{bmatrix}0\\\frac{1}{\sqrt{2}}(1+i)\\\vdots\\0\end{bmatrix}, \ldots, \begin{bmatrix}0\\0\\\vdots\\\frac{1}{\sqrt{2}}(1+i)\end{bmatrix}$$ is an orthonormal basis for $\mathbb{C}^n$.

(e) $\ell^2 = \left\{ x = \{\xi_1, \xi_2, \ldots\} \,:\, \sum_{j=1}^{\infty} |\xi_j|^2 < \infty \quad \xi_j \in \mathbb{C}\ \forall j \in \mathbb{P} \right\}$.
ℓ^2 is an infinite-dimensional Hilbert space.

$$\langle x, y\rangle = \sum_{j=1}^{\infty} \xi_j \overline{\eta_j} \quad x = \{\xi_1, \xi_2, \ldots\},\ y = \{\eta_1, \eta_2, \ldots\}\ \xi_j \in \mathbb{C}\ \eta_j \in \mathbb{C}\ \forall j \in \mathbb{P}.$$

An orthonormal basis for ℓ^2 consists of infinitely many vectors each of infinite dimension.

$$\begin{bmatrix}1\\0\\\vdots\\ \end{bmatrix}, \begin{bmatrix}0\\1\\0\\\vdots\end{bmatrix}, \ldots \text{ is an orthonormal basis for } \ell^2.$$

(f) $L^2(X, \mathcal{M}, \mu) = \left\{ f : X \to \mathbb{C} \,:\, f \text{ is measurable and} \left(\int |f|^2 d\mu\right)^{1/2} < \infty \right\}$

where $(X, \mathcal{M}, \mu)$ is a measure space and f is a measurable function on X. $L^2(X, \mathcal{M}, \mu)$ is an infinite-dimensional Hilbert space.

$$\langle x, y\rangle = \int x(t)\overline{y(t)}\, d\mu(t) \quad \forall x(t) \in L^2(\mu)\ \forall y(t) \in L^2(\mu).$$

(g) An orthonormal basis for $L^2[0, 2\pi]$ is

$$\frac{1}{\sqrt{2\pi}},\ \frac{1}{\sqrt{\pi}}\cos t,\ \frac{1}{\sqrt{\pi}}\sin t,\ \frac{1}{\sqrt{\pi}}\cos 2t,\ \frac{1}{\sqrt{\pi}}\sin 2t,\ \ldots\ .$$

1.1.3 Riesz's representation theorem

We will use the Riesz representation theorem in Chapter 1.3.3 example (b).

Let $\mathcal{H}^*$ be the set of all bounded linear functionals on a Hilbert space $\mathcal{H}$.
For all $F \in \mathcal{H}^*$ we define $||F||_{\mathcal{H}^*} = \sup_{x \in \mathcal{H},\ ||x||=1} |F(x)|$.

Theorem 1 (Riesz's representation theorem).

$$\forall F \in \mathcal{H}^* \text{ there exists a unique } y \in \mathcal{H} \text{ such that } F(x) = \langle x, y\rangle \quad \forall x \in \mathcal{H} \tag{1.1}$$

Moreover, we have $||y|| = ||F||_{\mathcal{H}^*}$.

A proof for Theorem 1 can be found in [Schechter,p30].

1.2 Cauchy-Schwarz inequality

The Cauchy-Schwarz inequality is one of the most used inequalities in mathematics. Probably the most used inequality in advanced mathematical analysis. The inequality is often used without explicit referring to it. See Chapter 1.3.3 (c) for an example where the Cauchy-Schwarz inequality is used.

1.2.1 Cauchy-Schwarz inequality

Theorem 2 (Cauchy-Schwarz inequality).
In an inner product space X,

$$|\langle x, y\rangle| \leq ||x||\,||y|| \quad \forall x \in X\, \forall y \in X \tag{1.2}$$

The equality (1.2) holds if and only if x and y are linearly dependent.

Proof. $y = 0$ is trivial.
Let $y \neq 0$, then for any $\alpha \in \mathbb{C}$ we have
$0 \leq ||x - \alpha y||^2 = \langle x - \alpha y, x - \alpha y\rangle = ||x||^2 - \alpha\langle y, x\rangle - \overline{\alpha}\langle x, y\rangle + |\alpha|^2||y||^2$.
Choose $\alpha = \dfrac{\langle x, y\rangle}{||y||^2}$ and we have

$$0 \leq ||x||^2 - \frac{|\langle x, y\rangle|^2}{||y||^2} - \frac{|\langle x, y\rangle|^2}{||y||^2} + \frac{|\langle x, y\rangle|^2}{||y||^4}||y||^2 = ||x||^2 - \frac{|\langle x, y\rangle|^2}{||y||^2}\,,$$

and $|\langle x, y\rangle| \leq ||x||\,||y||$ follows.

If $|\langle x, y\rangle| = ||x||\,||y||$, then
we can choose $\alpha \in \mathbb{C}$ $|\alpha| = 1$, such that $\alpha\langle x, y\rangle = |\langle x, y\rangle|$, and we have

$$\begin{aligned}\Big\|\,||x||y - \alpha||y||x\Big\|^2 &= \langle ||x||y - \alpha||y||x, ||x||y - \alpha||y||x\rangle \\ &= ||x||\,||x||\langle y, y\rangle - \alpha||y||\,||x||\langle x, y\rangle - \overline{\alpha}||x||\,||y||\langle y, x\rangle + \alpha\overline{\alpha}||y||\,||y||\langle x, x\rangle \\ &= ||x||\,||x||\,||y||\,||y|| - ||y||\,||x||\,||x||\,||y|| - ||x||\,||y||\,||x||\,||y|| + ||y||\,||y||\,||x||\,||x|| \\ &= 0\end{aligned}$$

According to (I1), we must have $||x||y = \alpha||y||x$, so x and y are linearly dependent.

If $y = \beta x$, $\beta \in \mathbb{C}$, then
$|\langle x, \beta x\rangle|^2 = \langle x, \beta x\rangle\,\overline{\langle x, \beta x\rangle} = \langle x, \beta x\rangle\,\langle \beta x, x\rangle = \beta\langle x, \beta x\rangle\,\langle x, x\rangle = ||x||^2\,||\beta x||^2$,
and we have $|\langle x, \beta x\rangle| = ||x||\,||\beta x||$. ∎

1.2.2 Examples of Cauchy-Schwarz inequality

(a) In $\mathbb{R}^3$ we have $|\langle x, y\rangle| \leq ||x||\,||y|| \quad x,\, y$ are vectors in $\mathbb{R}^3$.
We have equality if x and y are linearly dependent. This can be seen from Lagrange identity which gives us $\langle x, y\rangle^2 = ||x||^2\,||y||^2 - |x \times y|^2$,
$x \times y$ is the vector product, $x,\, y$ are vectors in $\mathbb{R}^3$.

(b) In $\mathbb{R}^n$ we have $\left|\sum_{j=1}^{n} x_j y_j\right| \leq \sqrt{\sum_{j=1}^{n} x_j^2}\,\sqrt{\sum_{j=1}^{n} y_j^2} \quad x,\, y$ are vectors in $\mathbb{R}^n$.

$\left|x^T y\right| \leq ||x||\,||y|| \quad x,\, y$ are vectors in $\mathbb{R}^n$, $x^T = [x_1\ x_2\ \ldots\ x_n]$, $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$.

(c) $\left|x^T A y\right| \leq \sqrt{x^T A x}\,\sqrt{y^T A y} \quad x,\, y$ are vectors in $\mathbb{R}^n$.
A is positive definite, $A \in \mathbb{R}^{n\times n}$.

(d) In $\mathbb{C}^n$ we have $\left|\sum_{j=1}^{n} x_j \overline{y_j}\right| \leq \sqrt{\sum_{j=1}^{n} |x_j|^2}\,\sqrt{\sum_{j=1}^{n} |y_j|^2} \quad x,\, y$ are vectors in $\mathbb{C}^n$.

(e) In ℓ^2 we have $\left|\sum_{j=1}^{\infty} \xi_j \overline{\eta_j}\right| \leq \sqrt{\sum_{j=1}^{\infty} |\xi_j|^2}\,\sqrt{\sum_{j=1}^{\infty} |\eta_j|^2}$
$x = \{\xi_1, \xi_2, \ldots\},\ y = \{\eta_1, \eta_2, \ldots\}\ \xi_j \in \mathbb{C}\ \eta_j \in \mathbb{C}\ \forall j \in \mathbb{P}$.

(f) In $L^2(\mu)$ we have $\left|\int x(t)\overline{y(t)}\, d\mu(t)\right| \leq \sqrt{\int |x(t)|^2 d\mu(t)}\,\sqrt{\int |y(t)|^2 d\mu(t)}$
$\forall x(t) \in L^2(\mu)\ \forall y(t) \in L^2(\mu)$.

(g) In $L^2(\mu)$, Assume that $\mu < +\infty$ and $g \equiv 1$ and $f \in L^2(\mu)$, then
the Cauchy-Schwarz inequality implies $\int |f| d\mu \leq \sqrt{\int |f|^2 d\mu}\,\sqrt{\mu(X)}$.
If μ is a probability measure, then $\mu(X) = 1$.

1.3 Bessel's inequality

Another widely used inequality for vectors in inner product spaces is the Bessel inequality.
The Cauchy-Schwarz inequality follows from the Bessel inequality.
In this section we prove the inequality and use it to give an estimate for the dimension of an eigenspace of an integral operator.

1.3.1 Bessel's inequality

Theorem 3 (Bessel's inequality).
Let $\{e_j\}_{j\geq 1}$ be an orthonormal system in a Hilbert space $\mathcal{H}$. Then

$$\sum_{j\geq 1} |\langle x, e_j\rangle|^2 \leq \|x\|^2 \quad \forall x \in \mathcal{H} \tag{1.3}$$

Proof. Let $\alpha_k = \langle x, e_k\rangle$, then for any $n \in \mathbb{P}$ we have

$$\begin{aligned}
\left\| x - \sum_{k=1}^{n} \alpha_k e_k \right\|^2 &= \left\langle x - \sum_{k=1}^{n} \alpha_k e_k, x - \sum_{k=1}^{n} \alpha_k e_k \right\rangle \\
&= \|x\|^2 - \left\langle \sum_{k=1}^{n} \alpha_k e_k, x \right\rangle - \left\langle x, \sum_{k=1}^{n} \alpha_k e_k \right\rangle + \sum_{k=1}^{n} |\alpha_k|^2 \\
&= \|x\|^2 - \sum_{k=1}^{n} \alpha_k \overline{\langle x, e_k\rangle} - \sum_{k=1}^{n} \overline{\alpha_k} \langle x, e_k\rangle + \sum_{k=1}^{n} |\alpha_k|^2 \\
&= \|x\|^2 - \sum_{k=1}^{n} |\langle x, e_k\rangle|^2 + \sum_{k=1}^{n} |\langle x, e_k\rangle - \alpha_k|^2 \\
&= \|x\|^2 - \sum_{k=1}^{n} |\langle x, e_k\rangle|^2.
\end{aligned}$$

We have $\displaystyle\sum_{k=1}^{n} |\langle x, e_k\rangle|^2 = \|x\|^2 - \left\| x - \sum_{k=1}^{n} \alpha_k e_k \right\|^2 \leq \|x\|^2.$

Let $n \to \infty$ in the last inequality. We have a sequence of nonnegative numbers, where the sum of the numbers is bounded from above. Hence (1.3) follows. ∎

The inner products $\langle x, e_j\rangle$ in (1.3) are called the **Fourier coefficients** of x with respect to the orthonormal system $\{e_j\}_{j\geq 1}$.

Remark 1. We will look at a more general system later.

Theorem 4.
Let $\{e_j\}_{j\geq 1}$ be an orthonormal system in a Hilbert space $\mathcal{H}$.
Then $\{e_j\}_{j\geq 1}$ is an orthonormal basis if and only if for all $x\in\mathcal{H}$ we have

$$||x||^2 = \sum_{j\geq 1} |\langle x, e_j\rangle|^2 \quad \textbf{(Parseval's identity)} \tag{1.4}$$

A proof for Theorem 4 can be found in [Weidmann,p39].

If we have an orthonormal system with only one element $\left(e_1 = \frac{y}{||y||}\right)$, then the Bessel inequality becomes the Cauchy-Schwarz inequality.

1.3.2 Examples of Bessel's inequality

(a) In $\mathbb{R}^3$ we have $||x||^2 = \sum_{j=1}^{3} |\langle x, e_j\rangle|^2 = \sum_{j=1}^{3} x_j^2$ where $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ is a vector in $\mathbb{R}^3$

and $e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ is an orthonormal basis for $\mathbb{R}^3$.

(b) In $\mathbb{R}^n$ we have $||x||^2 = \sum_{j=1}^{n} |\langle x, e_j\rangle|^2 = \sum_{j=1}^{n} x_j^2$ where $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ is a vector

in $\mathbb{R}^n$ and $e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, $e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}$, $\ldots$, $e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$ is an orthonormal basis for $\mathbb{R}^n$.

(c) In $\mathbb{C}^n$ we have $||x||^2 = \sum_{j=1}^{n} |\langle x, e_j\rangle|^2 = \sum_{j=1}^{n} |x_j|^2$ where $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ is a

vector in $\mathbb{C}^n$ and $e_1 = \begin{bmatrix} \frac{1}{\sqrt{2}}(1+i) \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, $e_2 = \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}}(1+i) \\ \vdots \\ 0 \end{bmatrix}$, $\ldots$, $e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \frac{1}{\sqrt{2}}(1+i) \end{bmatrix}$ is an orthonormal basis for $\mathbb{C}^n$.

(d) In ℓ^2 we have $||x||^2 = \sum_{j=1}^{\infty} |\xi_j|^2 \quad \forall x \in \ell^2$ where
$x = \{\xi_1, \xi_2, \ldots\} \quad \xi_j \in \mathbb{C} \ \forall j \in \mathbb{P}$.

(e) In $L^2[0, 2\pi]$ we have $||x||^2 = \frac{a_0^2}{2} + \sum_{n=1}^{\infty} a_n^2 + \sum_{n=1}^{\infty} b_n^2 \quad \forall x \in L^2[0, 2\pi]$

where $a_0 = \langle x, e_0 \rangle = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} x\, dt$, $a_0 \in \mathbb{R}$ and

$a_n = \langle x, e_n \rangle = \frac{1}{\sqrt{\pi}} \int_0^{2\pi} x \cos nt\, dt$, $a_n \in \mathbb{R} \ \forall n \in \mathbb{P}$ and

$b_n = \langle x, e_n \rangle = \frac{1}{\sqrt{\pi}} \int_0^{2\pi} x \sin nt\, dt$, $b_n \in \mathbb{R} \ \forall n \in \mathbb{P}$.

$e_0 = \frac{1}{\sqrt{2\pi}}$, $e_1 = \frac{1}{\sqrt{\pi}} \cos t$, $e_2 = \frac{1}{\sqrt{\pi}} \sin t$, $e_3 = \frac{1}{\sqrt{\pi}} \cos 2t$, $e_4 = \frac{1}{\sqrt{\pi}} \sin 2t, \ldots$ is an orthonormal basis for $L^2[0, 2\pi]$.

$\frac{a_0}{\sqrt{2\pi}} + \sum_{n=1}^{\infty} \frac{a_n}{\sqrt{\pi}} \cos nt + \sum_{n=1}^{\infty} \frac{b_n}{\sqrt{\pi}} \sin nt$ is the Fourier series of x.

(f) If e_2 is not included in example (a),(b),(c),(e) above, then we do not have an orthonormal basis and we have inequality instead of equality ($||x||^2 \geq$ instead of $||x||^2 =$).

1.3.3 Application of Bessel's inequality

(a) Let $S = \sum_{n=1}^{\infty} \frac{\sin nx}{n^a} \quad 0 < a \leq \frac{1}{2}$. S converges.
We will show that S can not be a Fourier series of a Riemann integrable function $f(x)$.

From the Bessel inequality we have $\sum_{n=1}^{\infty} \frac{1}{n^{2a}} \leq \frac{1}{\pi} \int_0^{2\pi} f^2(x)\, dx$ where $\frac{1}{n^{2a}}$ are the Fourier coefficients. This is impossible since $\sum_{n=1}^{\infty} \frac{1}{n^{2a}} = \infty$ when $0 \leq a \leq \frac{1}{2}$. Hence S can not be a Fourier series, see [Gelbaum,Olmsted,p70].

(b) Let $\mathcal{H}$ be a closed subspace of $L^2[0, 1]$ that is contained in $C[0, 1]$, where $C[0, 1]$ is defined as the space of continuous functions on $[0, 1]$. We will show that $\mathcal{H}$ is finite dimensional, see [Folland,p178 (ex.66)].

$\mathcal{H}$ is a Hilbert space since $\mathcal{H}$ is a closed subspace of $L^2[0, 1]$.
Both $\mathcal{H}$ with L^2-norm and $C[0, 1]$ with $||f||_{[0,1]} = \sup\left\{|f(x)| : x \in [0, 1]\right\}$ are Banach spaces, see [Griffel,p108].
Consider the inclusion $\mathcal{H} \to C[0, 1]$ as a linear map of Banach spaces.

This map $f \mapsto f$ is closed.
We have to check that $\left\{(f,f)\in\mathcal{H}\times C[0,1] \ : \ f\in\mathcal{H}\right\}$ is a closed subset of $\mathcal{H}\times C[0,1]$.
Suppose that (f_n, f_n) is a Cauchy sequence in $\mathcal{H}\times C[0,1]$, then f_n is a Cauchy sequence in $C[0,1]$ and there exists an f such that $f_n \rightrightarrows f$ (uniformly on $[0,1]$). Then $f_n \to f$ in $\mathcal{H}$, that is $||f_n - f||_2 \to 0$ when $n \to \infty$ since $||f_n - f||_2^2 = \int_0^1 |f_n - f|^2 dt \le ||f_n - f||_{[0,1]}^2$.
f_n is a Cauchy sequence in $\mathcal{H}$ implies that f_n is a Cauchy sequence in $C[0,1]$ since $\mathcal{H} \subset C[0,1]$.
By the closed graph theorem the inclusion is bounded, thus there exists a C such that $||f||_{[0,1]} \le C||f||_2$ for any $f\in\mathcal{H}$ where $||f||_{[0,1]} = \sup\left\{|f(x)| : x\in[0,1]\right\}$ and $||f||_2 = \left(\int_0^1 |f|^2 \, d\mu\right)^{1/2}$.

Let $x\in[0,1]$ and consider a linear functional $F_x : \mathcal{H} \to \mathcal{H}$ where $F_x(f) = f(x)$. It is bounded since $|F_x(f)| = |f(x)| \le ||f||_{[0,1]} \le C||f||_2$ for all $f\in\mathcal{H}$.
Hence it is a continuous linear functional on $\mathcal{H}$.

By Riesz representation theorem there exists a unique $g_x\in\mathcal{H}$ such that $f(x) = \langle f, g_x\rangle = \int_0^1 f(t)g_x(t)\,dt$ for all $f\in\mathcal{H}$.

Further

$$
\begin{aligned}
|f(x)| = |\langle f, g_x\rangle| &\le C||f||_2 \quad \forall f\in\mathcal{H}\\
&\Downarrow\\
|\langle g_x, g_x\rangle| &\le C||g_x||_2\\
&\Downarrow\\
||g_x||_2^2 &\le C||g_x||_2\\
&\Downarrow\\
||g_x||_2 &\le C.
\end{aligned}
$$

Let $\{f_j\}_{j=1}^n$ be an orthonormal system of functions in $\mathcal{H}$.
Then by using Riesz representation theorem and Bessel's inequality and $||g_x||_2 \le C$ from previous result we have

$$\sum_{j=1}^n |f_j(x)|^2 = \sum_{j=1}^n |\langle f_j, g_x\rangle|^2 = \sum_{j=1}^n |\langle g_x, f_j\rangle|^2 \le ||g_x||_2^2 \le C^2 \quad x\in[0,1]. \qquad (*)$$

$$\Downarrow$$

$$n = \int_0^1 \sum_{j=1}^n |f_j(x)|^2 dx \le \int_0^1 C^2 dx = C^2. \qquad (**)$$

$$\Downarrow$$

$$n \le C^2.$$

Thus dim $\mathcal{H} \le C^2$ and $\mathcal{H}$ is finite dimensional.

Remark 2. $C[0,1]$ contains a subspace of polynomials where $1, x, x^2, \dots$ are linearly independent. It is infinite dimensional and is contained in $L^2[0,1]$, but not in $\mathcal{H}$ which is a closed subspace of $L^2[0,1]$, that is contained in $C[0,1]$. The closure ($L^2[0,1]$) of this subspace is not contained in $C[0,1]$.

(c) Let $K(x,y)$ be a continuous function on $[a,b]\times[a,b]$.
A continuous function f on $[a,b]$ is called an eigenfunction for K with respect to a real eigenvalue r if $f(y) = r\int_a^b K(x,y)f(x)\,dx$.
We will without loss of generality use $[0,1]$ instead of $[a,b]$.
Let $E_r = \left\{ f\in C[0,1] : f(y) = r\int_0^1 K(x,y)f(x)\,dx \right\}$.
We will give an estimate for the dimension of E_r, see [Lang,p108 (ex.7)].

Let $\mathcal{H}_r = \left\{ f\in L^2[0,1] : \; f(y) = r\int_0^1 K(x,y)f(x)\,dx \right\}$.
Let $h\in\mathcal{H}_r$. We want to show that h is continuous.

$$|h(y) - h(y_0)| \le |r|\int_0^1 |K(x,y) - K(x,y_0)|\,|h(x)|\,dx$$
$$\Downarrow$$
$$\le |r| \left(\int_0^1 |K(x,y) - K(x,y_0)|^2\,dx\right)^{\frac{1}{2}} \left(\int_0^1 |h(x)|^2\,dx\right)^{\frac{1}{2}} \to 0 \text{ when } y \to y_0$$

since $K(x,y)$ is uniformly continuous and $h\in L^2[0,1]$.
So we have that $\mathcal{H}_r \subseteq E_r \subset C[0,1]$ since all functions in $\mathcal{H}_r$ are continuous as we shown above.
Clearly $E_r \subseteq \mathcal{H}_r \subset L^2[0,1]$. Hence $E_r = \mathcal{H}_r$.

If we can show that $\mathcal{H}_r$ is a closed subspace of $L^2[0,1]$, then we can use results from example (b) above to give an estimate for the dimension of E_r.
Let $f_n \to f$ in $L^2[0,1]$ and $f_n\in\mathcal{H}_r$ and $g\in\mathcal{H}_r$. Then we have

$$|f_n(y) - g(y)| \le |r|\int_0^1 |K(x,y)|\,|f_n(x) - g(x)|\,dx$$
$$\Downarrow$$
$$\le |r| \left(\int_0^1 |K(x,y)|^2\,dx\right)^{\frac{1}{2}} \left(\int_0^1 |f_n(x) - g(x)|^2\,dx\right)^{\frac{1}{2}} \to 0 \text{ when } f_n \to g.$$

g is continuous and $f_n \rightrightarrows g$ on [0,1].

Then $f = g$ almost everywhere and $f\in\mathcal{H}_r$.

Hence $\mathcal{H}_r$ is a closed subspace of $L^2[0,1]$.

Assume that dim $E_r \ge n$ where $n\in\mathbb{P}$, then there exists an orthonormal

system $\{f_j\}_{j=1}^n$ in $\mathcal{H}$. From $(*)$ and $(**)$ in example (b) above we have that
$n = \int_0^1 \sum_{j=1}^n |f_j(x)|^2 dx \leq \int_0^1 \|rK(x,y)\|_2^2 \, dx.$

$$\Downarrow$$

$$n \leq r^2 \int_0^1 \int_0^1 K^2(x,y) \, dydx.$$

Thus dim $E_r \leq r^2 \int_0^1 \int_0^1 K^2(x,y) \, dydx$.

1.4 Selberg's inequality

Selberg's inequality is not so well known as the Cauchy-Schwarz and the Bessel inequality. It is an interesting inequality and it is also a generalization of the Cauchy-Schwarz and the Bessel inequality.

1.4.1 Selberg's inequality

Theorem 5 (Selberg's inequality).
In a Hilbert space $\mathcal{H}$,

$$\sum_{j=1}^{n} \frac{|\langle x, y_j\rangle|^2}{\sum_{k=1}^{n}|\langle y_j, y_k\rangle|} \leq \|x\|^2 \quad \forall x \in \mathcal{H} \ y_j \neq 0 \ y_j \in \mathcal{H} \tag{1.5}$$

The equality (1.5) holds if and only if

(C) $x = \sum_{j=1}^{n} \alpha_j y_j$, $\alpha_j \in \mathbb{C}$, and for each pair (j, k), $j \neq k$,

(C1) $\langle y_j, y_k\rangle = 0$,
or
(C2) $|\alpha_j| = |\alpha_k|$ and $\langle \alpha_j y_j, \alpha_k y_k\rangle \geq 0$.

See [Furuta,p218].

Proof. For any $\alpha_j \in \mathbb{C}$ we have

$$0 \leq \left\| x - \sum_{j=1}^{n} \alpha_j y_j \right\|^2 = \left\langle x - \sum_{j=1}^{n} \alpha_j y_j, x - \sum_{j=1}^{n} \alpha_j y_j \right\rangle.$$

$$= \|x\|^2 - \left\langle \sum_{j=1}^{n} \alpha_j y_j, x \right\rangle - \left\langle x, \sum_{j=1}^{n} \alpha_j y_j \right\rangle + \left\langle \sum_{j=1}^{n} \alpha_j y_j, \sum_{j=1}^{n} \alpha_j y_j \right\rangle.$$

$$= \|x\|^2 - \sum_{j=1}^{n} \alpha_j \langle y_j, x\rangle - \sum_{j=1}^{n} \overline{\alpha_j} \langle x, y_j\rangle + \sum_{j=1}^{n} \sum_{k=1}^{n} \alpha_j \overline{\alpha_k} \langle y_j, y_k\rangle.$$

From $0 \leq (|\alpha_j| - |\alpha_k|)^2$ we have $|\alpha_j \overline{\alpha_k}| \leq \frac{1}{2}|\alpha_j|^2 + \frac{1}{2}|\alpha_k|^2$, and we have

$$\leq \|x\|^2 - \sum_{j=1}^{n} \alpha_j \overline{\langle x, y_j\rangle} - \sum_{j=1}^{n} \overline{\alpha_j} \langle x, y_j\rangle + \frac{1}{2} \sum_{j=1}^{n} \sum_{k=1}^{n} |\alpha_j|^2 |\langle y_j, y_k\rangle| + \frac{1}{2} \sum_{j=1}^{n} \sum_{k=1}^{n} |\alpha_k|^2 |\langle y_j, y_k\rangle|.$$

We can choose $\alpha_j = \dfrac{\langle x, y_j\rangle}{\sum_{k=1}^n |\langle y_j, y_k\rangle|}$, and we have

$$= \|x\|^2 - \sum_{j=1}^n \frac{\langle x, y_j\rangle \overline{\langle x, y_j\rangle}}{\sum_{k=1}^n |\langle y_j, y_k\rangle|} - \sum_{j=1}^n \frac{\overline{\langle x, y_j\rangle} \langle x, y_j\rangle}{\sum_{k=1}^n |\langle y_j, y_k\rangle|} + \frac{1}{2} \sum_{j=1}^n \sum_{k=1}^n \frac{|\langle x, y_j\rangle|^2 \, |\langle y_j, y_k\rangle|}{\left(\sum_{k=1}^n |\langle y_j, y_k\rangle|\right)^2} +$$

$$\frac{1}{2} \sum_{j=1}^n \sum_{k=1}^n \frac{|\langle x, y_k\rangle|^2 \, |\langle y_j, y_k\rangle|}{\left(\sum_{j=1}^n |\langle y_k, y_j\rangle|\right)^2} .$$

$$= \|x\|^2 - \sum_{j=1}^n \frac{|\langle x, y_j\rangle|^2}{\sum_{k=1}^n |\langle y_j, y_k\rangle|} - \sum_{j=1}^n \frac{|\langle x, y_j\rangle|^2}{\sum_{k=1}^n |\langle y_j, y_k\rangle|} + \frac{1}{2} \sum_{j=1}^n |\langle x, y_j\rangle|^2 \frac{\sum_{k=1}^n |\langle y_j, y_k\rangle|}{\left(\sum_{k=1}^n |\langle y_j, y_k\rangle|\right)^2} +$$

$$\frac{1}{2} \sum_{k=1}^n |\langle x, y_k\rangle|^2 \frac{\sum_{j=1}^n |\langle y_j, y_k\rangle|}{\left(\sum_{j=1}^n |\langle y_j, y_k\rangle|\right)^2} .$$

$$= \|x\|^2 - 2 \sum_{j=1}^n \frac{|\langle x, y_j\rangle|^2}{\sum_{k=1}^n |\langle y_j, y_k\rangle|} + \frac{1}{2} \sum_{j=1}^n \frac{|\langle x, y_j\rangle|^2}{\sum_{k=1}^n |\langle y_j, y_k\rangle|} + \frac{1}{2} \sum_{k=1}^n \frac{|\langle x, y_k\rangle|^2}{\sum_{j=1}^n |\langle y_j, y_k\rangle|} .$$

$$= \|x\|^2 - 2 \sum_{j=1}^n \frac{|\langle x, y_j\rangle|^2}{\sum_{k=1}^n |\langle y_j, y_k\rangle|} + \frac{1}{2} \sum_{j=1}^n \frac{|\langle x, y_j\rangle|^2}{\sum_{k=1}^n |\langle y_j, y_k\rangle|} + \frac{1}{2} \sum_{j=1}^n \frac{|\langle x, y_j\rangle|^2}{\sum_{k=1}^n |\langle y_j, y_k\rangle|} ,$$

and $\displaystyle\sum_{j=1}^n \frac{|\langle x, y_j\rangle|^2}{\sum_{k=1}^n |\langle y_j, y_k\rangle|} \le \|x\|^2$ follows.

We will show that

(C)

$$\Downarrow$$

$$\sum_{j=1}^n \frac{|\langle x, y_j\rangle|^2}{\sum_{k=1}^n |\langle y_j, y_k\rangle|} = \|x\|^2$$

$$\Downarrow$$

$$x = \sum_{j=1}^n \alpha_j y_j \ \wedge \ 2\alpha_j \overline{\alpha_k} \langle y_j, y_k\rangle = |\alpha_j|^2 \, |\langle y_j, y_k\rangle| \ + \ |\alpha_k|^2 \, |\langle y_j, y_k\rangle| \qquad (*)$$

$$\Downarrow$$

(C).

If (C), then for each pair (j,k), $j \neq k$, where (C1) is true, we have

$\langle \alpha_k y_k, \alpha_j y_j\rangle = |\alpha_j|^2 \, |\langle y_k, y_j\rangle| \quad (**)$

And for each pair (j,k), $j \neq k$, where (C2) is true, we have $(**)$.

$$\sum_{j=1}^{n} \frac{\left|\left\langle \sum_{k=1}^{n} \alpha_k y_k, y_j\right\rangle\right|^2}{\sum_{k=1}^{n}|\langle y_j, y_k\rangle|} = \sum_{j=1}^{n} \frac{\left|\sum_{k=1}^{n} \alpha_k \langle y_k, y_j\rangle\right|^2}{\sum_{k=1}^{n}|\langle y_j, y_k\rangle|} = \sum_{j=1}^{n} \frac{\left|\sum_{k=1}^{n} \alpha_k \langle y_k, y_j\rangle\right|^2 |\alpha_j|^2}{\sum_{k=1}^{n}|\langle y_j, y_k\rangle| \, |\alpha_j|^2}$$

$$= \sum_{j=1}^{n} \frac{\left(\sum_{k=1}^{n} \alpha_k \langle y_k, y_j\rangle\right) \sum_{k=1}^{n} \overline{\alpha_k} \langle y_j, y_k\rangle \, \alpha_j \overline{\alpha_j}}{\sum_{k=1}^{n}|\langle y_j, y_k\rangle| \, |\alpha_j|^2} = \sum_{j=1}^{n} \frac{\left(\sum_{k=1}^{n} \langle \alpha_k y_k, \alpha_j y_j\rangle\right) \sum_{k=1}^{n} \alpha_j \overline{\alpha_k} \langle y_j, y_k\rangle}{\sum_{k=1}^{n}|\langle y_k, y_j\rangle| \, |\alpha_j|^2}.$$

We use $(**)$, and we have

$$= \sum_{j=1}^{n} \frac{\left(\sum_{k=1}^{n} \langle \alpha_k y_k, \alpha_j y_j\rangle\right) \sum_{k=1}^{n} \alpha_j \overline{\alpha_k} \langle y_j, y_k\rangle}{\sum_{k=1}^{n}\langle \alpha_k y_k, \alpha_j y_j\rangle} = \sum_{j=1}^{n}\sum_{k=1}^{n} \alpha_j \overline{\alpha_k} \langle y_j, y_k\rangle, \text{ and}$$

$$\|x\|^2 = \left\langle \sum_{j=1}^{n} \alpha_j y_j, \sum_{k=1}^{n} \alpha_k y_k \right\rangle = \sum_{j=1}^{n}\sum_{k=1}^{n} \alpha_j \overline{\alpha_k} \langle y_j, y_k\rangle.$$

Hence (C) implies $\displaystyle\sum_{j=1}^{n} \frac{|\langle x, y_j\rangle|^2}{\sum_{k=1}^{n}|\langle y_j, y_k\rangle|} = \|x\|^2$.

If $\displaystyle\sum_{j=1}^{n} \frac{|\langle x, y_j\rangle|^2}{\sum_{k=1}^{n}|\langle y_j, y_k\rangle|} = \|x\|^2$, then choose $\alpha_j = \dfrac{\langle x, y_j\rangle}{\sum_{k=1}^{n}|\langle y_j, y_k\rangle|}$.

From the proof of the inequality (1.5) we have that equality (1.5) holds when we have

$$0 = \left\|x - \sum_{j=1}^{n} \alpha_j y_j\right\|^2 \text{ and } \sum_{j=1}^{n}\sum_{k=1}^{n} \alpha_j \overline{\alpha_k} \langle y_j, y_k\rangle = \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} |\alpha_j|^2 \, |\langle y_j, y_k\rangle| + \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} |\alpha_k|^2 \, |\langle y_j, y_k\rangle|.$$

For each pair (j,k), $j \neq k$ we have $\frac{1}{2}|\alpha_j|^2 \, |\langle y_j, y_k\rangle| + \frac{1}{2}|\alpha_k|^2 \, |\langle y_j, y_k\rangle| \geq 0$ and

$$\left|\alpha_j \overline{\alpha_k} \langle y_j, y_k\rangle\right| \leq \frac{1}{2}|\alpha_j|^2 \, |\langle y_j, y_k\rangle| + \frac{1}{2}|\alpha_k|^2 \, |\langle y_j, y_k\rangle|.$$

Hence $\displaystyle\sum_{j=1}^{n} \frac{|\langle x, y_j\rangle|^2}{\sum_{k=1}^{n}|\langle y_j, y_k\rangle|} = \|x\|^2$ implies $(*)$.

If $(*)$, then for each pair (j,k), $j \neq k$, assume that (C1) is not true.

Then $\langle \alpha_j y_j, \alpha_k y_k\rangle \geq 0$, and

$$\frac{2\alpha_j\overline{\alpha_k}\,\langle y_j, y_k\rangle}{|\langle y_j, y_k\rangle|} = |\alpha_j|^2 + |\alpha_k|^2$$

$$\Downarrow$$

$$\frac{\left|2\alpha_j\overline{\alpha_k}\,\langle y_j, y_k\rangle\right|}{|\langle y_j, y_k\rangle|} = |\alpha_j|^2 + |\alpha_k|^2$$

$$\Downarrow$$

$$2|\alpha_j|\,|\alpha_k| = |\alpha_j|^2 + |\alpha_k|^2$$

$$\Updownarrow$$

$$|\alpha_j| = |\alpha_k|.$$

Hence $(*)$ implies (C). ■

When we use (C) we need only to calculate maximum $\frac{(n-1)n}{2}$ pairs since we have symmetry.

If we have only one element $(n=1)$, $y=y_1$, then the Selberg inequality becomes the Cauchy-Schwarz inequality.
If we have an orthogonal system $\{y_j\}_{j=1}^{n}$ with several elements $(n \geq 2)$, $\langle y_j, y_k\rangle = 0$ if $j \neq k$, let $\left(e_j = \frac{y_j}{\|y_j\|}\right)_{j\geq 1}$, then the Selberg inequality becomes the Bessel inequality.

1.4.2 Examples of Selberg's inequality

(a) In $\mathbb{R}^3$, $x = \begin{bmatrix}1\\2\\1\end{bmatrix}$, $y_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}$, $y_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}$, $y_3 = \begin{bmatrix}0\\0\\1\end{bmatrix}$.

We have $x = \begin{bmatrix}1\\0\\0\end{bmatrix} + 2\begin{bmatrix}0\\1\\0\end{bmatrix} + \begin{bmatrix}0\\0\\1\end{bmatrix}$ and for each pair (j,k) in (1.5) we have (C1) since $y_1 \perp y_2$, $y_1 \perp y_3$, $y_2 \perp y_3$. By Selberg's inequality we have equality in (1.5) since (C) is satisfied. And it follows that we have Parseval's identity.

(b) In $\mathbb{R}^3$, $x = \begin{bmatrix}1\\2\\1\end{bmatrix}$, $y_1 = \begin{bmatrix}1\\2\\2\end{bmatrix}$, $y_2 = \begin{bmatrix}1\\1\\0\end{bmatrix}$, $y_3 = \begin{bmatrix}1\\1\\1\end{bmatrix}$.

We have $x = \begin{bmatrix}1\\2\\2\end{bmatrix} + \begin{bmatrix}1\\1\\0\end{bmatrix} - \begin{bmatrix}1\\1\\1\end{bmatrix}$ and for each pair (j,k) in (1.5) we have (C2).

By Selberg's inequality we have equality in (1.5) since (C) is satisfied.

(c) In $\mathbb{R}^3$, $x = \begin{bmatrix}2\\2\\1\end{bmatrix}$, $y_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}$, $y_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}$, $y_3 = \begin{bmatrix}1\\0\\1\end{bmatrix}$.

We have $x = \begin{bmatrix}1\\0\\0\end{bmatrix} + 2\begin{bmatrix}0\\1\\0\end{bmatrix} + \begin{bmatrix}1\\0\\1\end{bmatrix}$, and for the following pairs $(1,2)$, $(2,3)$ in (1.5) we have (C1) since $y_1 \perp y_2$, $y_2 \perp y_3$.
And for the following pair $(1,3)$ in (1.5), we have (C2).
By Selberg's inequality we have equality in (1.5) since (C) is satisfied.

(d) In $L^2[-1,1]$, $x = 1 + t + t^2$, $y_1 = 1$, $y_2 = t$, $y_3 = t^2$.
We have $x = 1 + t + t^2$ and for each pair (j,k) in (1.5) we have (C2).
By Selberg's inequality we have equality in (1.5) since (C) is satisfied.

(e) In $\mathbb{R}^3$, if $x = ay_1 + by_2 + cy_3$, $a \in \mathbb{R}$, $b \in \mathbb{R}$, $c \in \mathbb{R}$, then the Selberg inequality can be written as

$$\frac{(\langle ay_1,y_1\rangle+\langle by_2,y_1\rangle+\langle cy_3,y_1\rangle)^2}{|\langle y_1,y_1\rangle|+|\langle y_1,y_2\rangle|+|\langle y_1,y_3\rangle|} + \frac{(\langle ay_1,y_2\rangle+\langle by_2,y_2\rangle+\langle cy_3,y_2\rangle)^2}{|\langle y_2,y_1\rangle|+|\langle y_2,y_2\rangle|+|\langle y_2,y_3\rangle|} + \frac{(\langle ay_1,y_3\rangle+\langle by_2,y_3\rangle+\langle cy_3,y_3\rangle)^2}{|\langle y_3,y_1\rangle|+|\langle y_3,y_2\rangle|+|\langle y_3,y_3\rangle|}$$

$$\leq a^2\langle y_1, y_1\rangle + 2ab\langle y_1, y_2\rangle + 2ac\langle y_1, y_3\rangle + b^2\langle y_2, y_2\rangle + 2bc\langle y_2, y_3\rangle + c^2\langle y_3, y_3\rangle.$$

If $y_1 = \begin{bmatrix}1\\2\\2\end{bmatrix}$, $y_2 = \begin{bmatrix}1\\1\\1\end{bmatrix}$, $y_3 = \begin{bmatrix}2\\2\\3\end{bmatrix}$ and $x = a\begin{bmatrix}1\\2\\2\end{bmatrix} + b\begin{bmatrix}1\\1\\1\end{bmatrix} + c\begin{bmatrix}2\\2\\3\end{bmatrix}$,

then we have the following Selberg's inequality,

$$\frac{(9a+5b+12c)^2}{26} + \frac{(5a+3b+7c)^2}{15} + \frac{(12a+7b+17c)^2}{36} \leq 9a^2 + 10ab + 24ac + 3b^2 + 14bc + 17c^2.$$

If $a = b = c$, then we have equality $\left(77a^2 = 77a^2\right)$.

(f) In $L^2[-1,1]$, $x = a1 + bt + ct^2$, $y_1 = 1$, $y_2 = t$, $y_3 = t^2$, $a \in \mathbb{R}$, $b \in \mathbb{R}$, $c \in \mathbb{R}$.

We have $y_1 \perp y_2$, $y_2 \perp y_3$ and we have the following Selberg's inequality,

$$\frac{\left(2a+\frac{2}{3}c\right)^2}{\frac{8}{3}} + \frac{\left(\frac{2}{3}b\right)^2}{\frac{2}{3}} + \frac{\left(\frac{2}{3}a+\frac{2}{5}c\right)^2}{\frac{16}{15}} \le 2a^2 + \frac{4}{3}ac + \frac{2}{3}b^2 + \frac{2}{5}c^2.$$

$$\Updownarrow$$

$$\frac{\left(2a+\frac{2}{3}c\right)^2}{\frac{8}{3}} + \frac{\left(\frac{2}{3}a+\frac{2}{5}c\right)^2}{\frac{16}{15}} \le 2a^2 + \frac{4}{3}ac + \frac{2}{5}c^2.$$

If $a = c$, then we have equality $\left(\frac{56}{15}a^2 = \frac{56}{15}a^2\right)$.

Chapter 2

Positive semidefinite matrices and inequalities

In this chapter we give a new proof of Selberg's inequality. It is based on the theory of positive semidefinite matrices. This approach gives a new generalization of Selberg's inequality.

2.1 Positive semidefinite matrices

Positive semidefinite matrices are closely related to nonnegative real numbers.

2.1.1 Definition and basic properties of positive semidefinite matrices

A $n \times n$ matrix A is **normal**, if $A^*A = AA^*$.
A $n \times n$ matrix A is **Hermitian**, if $A^* = A$. Hermitian matrices are normal matrices.
A $n \times n$ matrix A is **positive semidefinite**, if A is Hermitian and $x^*Ax \geq 0$ for all nonzero $x \in \mathbb{C}^n$. We will then write $A \geq 0$.
A $n \times n$ matrix A is **positive definite**, if A is Hermitian and $x^*Ax > 0$ for all nonzero $x \in \mathbb{C}^n$. We will then write $A > 0$.
If $A - B \geq 0$, then we will write $A \geq B$.

The following is a list of some properties for positive semidefinite matrices that are needed in this chapter.
Let A, B, C, F, I, S, U denote $n \times n$ matrices and x, y denote $n \times 1$ vectors.

(i) If A is Hermitian, then $A = U^*$ diag$(\lambda_1, \ldots, \lambda_n)U$ where U is a unitary

matrix and λ_j are nonnegative real numbers on diagonal matrix.
Proof. See [Zhang,p65]. ■

(ii) If A is Hermitian, then S^*AS is Hermitian.
Proof. $(S^*AS)^* = S^*A^*S = S^*AS$. ■

(iii) $A \geq 0$ implies $S^*AS \geq 0$.
Proof. $x^*Ax \geq 0$ implies $y^*S^*ASy \geq 0$. ■

(iv) $A \geq 0$ implies $det(A) \geq 0$.
Proof. Let $Ax = \lambda x$ where λ is an eigenvalue and x is an eigenvector of A corresponding to λ. Then for each λ, we have $x^*Ax = x^*\lambda x \geq 0 \Rightarrow$
$\lambda = \frac{x^*Ax}{x^*x} \geq 0 \Rightarrow det(A) = \lambda_1\lambda_2 \ldots \lambda_n \geq 0$. ■

(v) $A \geq 0$ and $det(A) > 0 \Rightarrow A^{-1} \geq 0$.
Proof. For any $y \in \mathbb{C}^n$ there exists an $x \in \mathbb{C}^n$ such that
$Ax = y \Rightarrow x = A^{-1}y$ and $0 \leq x^*Ax = \left(A^{-1}y\right)^* A\,A^{-1}y = y^* \left(A^{-1}\right)^* y = y^*A^{-1}y$.
■

(vi) If $A \geq B$, then $S^*AS \geq S^*BS$.
Proof. $A - B \geq 0 \Rightarrow S^*(A - B)S \geq 0 \Rightarrow S^*AS \geq S^*BS$. ■

(vii) If $A \geq 0$, then there exists a matrix $B \geq 0$ such that $B^2 = A$.
B is denoted by $A^{\frac{1}{2}}$.
Proof. See [Zhang,p162]. ■

(viii) If $A \geq B$ and A^{-1} exists and B^{-1} exists, then $B^{-1} \geq A^{-1}$.
Proof. If $C \leq I$, then $I = C^{-\frac{1}{2}}CC^{-\frac{1}{2}} \leq C^{-\frac{1}{2}}IC^{-\frac{1}{2}} = C^{-1}$.
$A \geq B \Rightarrow I \geq A^{-\frac{1}{2}}BA^{-\frac{1}{2}} \Rightarrow A^{\frac{1}{2}}B^{-1}A^{\frac{1}{2}} \geq I \Rightarrow B^{-1} \geq A^{-1}$. ■

2.1.2 Further properties of positive semidefinite matrices

We also need properties of products of two and three positive semidefinite matrices. The product of two positive semidefinite matrices is not always positive semidefinite.

(ix) $(A \geq 0$ and $B \geq 0) \not\Rightarrow AB \geq 0$.

Proof. $A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$, $AB = \begin{bmatrix} 5 & 5 \\ 3 & 4 \end{bmatrix}$, then $A \geq 0$, $B \geq 0$, $AB \ngeq 0$. ■

(x) If $A \geq 0$ then $A^2 \geq 0$
Proof. $A \geq 0 \Rightarrow \left(y^* A^{\frac{1}{2}}\right) A \left(A^{\frac{1}{2}} y\right) = \left(A^{\frac{1}{2}} y\right)^* A \left(A^{\frac{1}{2}} y\right) = x^* A x \geq 0$. ■

(xi) $A \geq 0$ and $B \geq 0$, then AB is Hermitian $\Leftrightarrow$ $AB = BA$.
Proof. $(AB)^* = B^* A^* = BA$. ■

(xii) If $A \geq 0$ and $B \geq 0$ and $AB = BA$, then $A^{\frac{1}{2}} B^{\frac{1}{2}} = B^{\frac{1}{2}} A^{\frac{1}{2}}$.
Proof. We have $A \geq 0$, $B \geq 0$. AB is Hermitian.
Let $A = U^* C U$ and $B = U^* F U$ where C and F are diagonal matrices and U is a unitary matrix. U is the same for A and B since A and B commute, see [Zhang,p61]. Then we have
$A^{\frac{1}{2}} B^{\frac{1}{2}} = U^* C^{\frac{1}{2}} U U^* F^{\frac{1}{2}} U = U^* C^{\frac{1}{2}} F^{\frac{1}{2}} U = U^* F^{\frac{1}{2}} C^{\frac{1}{2}} U = U^* F^{\frac{1}{2}} U U^* C^{\frac{1}{2}} U = B^{\frac{1}{2}} A^{\frac{1}{2}}$. ■

(xiii) If $A \geq 0$ and $B \geq 0$ and AB is Hermitian, then $AB \geq 0$.
Proof. $y^* A B y = y^* A^{\frac{1}{2}} A^{\frac{1}{2}} B y = \left(A^{\frac{1}{2}} x\right)^* A^{\frac{1}{2}} B^{\frac{1}{2}} B^{\frac{1}{2}} y = \left(A^{\frac{1}{2}} y\right)^* B \left(A^{\frac{1}{2}} y\right) = x^* B x \geq 0$.
■

(xiv) $A \geq 0$, $B \geq 0$, B is invertible, $C \geq 0$ and ABC is Hermitian implies $ABC \geq 0$.
Proof. Let $F = \left(B^{\frac{1}{2}} A B^{\frac{1}{2}}\right)\left(B^{\frac{1}{2}} C B^{\frac{1}{2}}\right)$. F is Hermitian since $F = B^{\frac{1}{2}} (ABC) B^{\frac{1}{2}}$, see (ii). We have $\left(B^{\frac{1}{2}} A B^{\frac{1}{2}}\right) \geq 0$ and $\left(B^{\frac{1}{2}} C B^{\frac{1}{2}}\right) \geq 0$, see (iii). From (xiii) we have $F \geq 0$ and finally from (iii) we have $ABC = B^{-\frac{1}{2}} F B^{-\frac{1}{2}} \geq 0$. ■

(xv) If $A \geq 0$, then $\lambda_{\max}(A) I \geq A$.
$\lambda_{\max}(A)$ is the largest eigenvalue of matrix A.
Proof. Let $B = U^* A U$ where B is a diagonal matrix and U is a unitary matrix. Then we have $\lambda_{\max}(A) I - B \geq 0$. Hence $\lambda_{\max}(A) I \geq A$. ■

(xvi) If $A \geq 0$, then $\lambda_{\min>0}(A) A \leq A^2$.
$\lambda_{\min>0}(A)$ is the smallest positive eigenvalue of matrix A.
Proof. Let $B = U^* A U$ where B is a diagonal matrix and U is a unitary matrix. Then we have $B^2 - \lambda_{\min>0}(A) B = B^2 - \lambda_{\min>0}(B) B$
$= \text{diag}\left(\lambda_1^2 - \lambda_{\min>0}(B)\lambda_1, \ldots, \lambda_n^2 - \lambda_{\min>0}(B)\lambda_n\right) \geq 0$.
Hence $\lambda_{\min>0}(A) A \leq A^2$. ■

(xvii) We define the Gram matrix for $\{y_1, y_2, \ldots, y_n\}$ in a Hilbert space $\mathcal{H}$ in the following way:

$$G = \begin{bmatrix} \langle y_1, y_1 \rangle & \langle y_2, y_1 \rangle & \dots & \langle y_n, y_1 \rangle \\ \langle y_1, y_2 \rangle & \langle y_2, y_2 \rangle & \dots & \langle y_n, y_2 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle y_1, y_n \rangle & \langle y_2, y_n \rangle & \dots & \langle y_n, y_n \rangle \end{bmatrix} \tag{2.1}$$

We see that G is Hermitian.

$u^*Gu = \langle \beta_1 y_1 + \beta_2 y_2 + \ldots + \beta_n y_n \,,\, \beta_1 y_1 + \beta_2 y_2 + \ldots + \beta_n y_n \rangle \geq 0$

for all nonzero $u = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}$, $\beta_j \in \mathbb{C}$, that is $G \geq 0$.

If $y_1, y_2, \ldots, y_n$ are linearly independent, then $G > 0$.

(xviii) If we have a diagonal matrix $D = \begin{bmatrix} d_1 & & & \\ & d_2 & & \mathbf{0} \\ & \mathbf{0} & \ddots & \\ & & & d_n \end{bmatrix}$ where each d_j is a nonnegative real number, then we have $u^*D\,u = \sum_{j=1}^{n} |\beta_j|^2 d_j \geq 0$

for all nonzero $u = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}$, $\beta_j \in \mathbb{C}$, that is $D \geq 0$.

Remark 3. If $G = \begin{bmatrix} \langle y_1, y_1 \rangle & \langle y_2, y_1 \rangle \\ \langle y_1, y_2 \rangle & \langle y_2, y_2 \rangle \end{bmatrix}$, then $det(G) \geq 0$ is same the Cauchy-Schwarz inequality.

2.2 Inequalites

In this section we will use theory for positive semidefinite matrices to study inequalities. First we give a proof of a generalized Bessel's inequality following [Akhiezer,Glazman], then we use the same technique to give a new proof of Selberg's inequality. We conclude this section with a new generalization of Selberg's inequality with a proof.

2.2.1 Generalized Bessel's inequality

We will need the following generalized Bessel's inequality to prove Selberg's inequality in the next subsection.

Theorem 6 (Generalized Bessel's inequality).
Let $\{y_j\}_{j=1}^n$ be a linearly independent system in a Hilbert space $\mathcal{H}$ and let G be the corresponding Gram matrix. Then

$$v^* G^{-1} v \leq ||x||^2 \quad \forall x \in \mathcal{H} \tag{2.2}$$

where $v = \begin{bmatrix} \langle x, y_1 \rangle \\ \langle x, y_2 \rangle \\ \vdots \\ \langle x, y_n \rangle \end{bmatrix}$.

See [Akhiezer,Glazman,p24].

Proof. Let $x = \sum_{j=1}^{n} \alpha_j y_j + h$ where $\alpha_j \in \mathbb{C}$, $h \in \mathcal{H}$, $h \perp y_j$, for any $j \in \{1, 2, \ldots, n\}$.

From $\sum_{k=1}^{n} \langle x, y_k \rangle = \sum_{k=1}^{n} \sum_{j=1}^{n} \alpha_j \langle y_j, y_k \rangle + \sum_{k=1}^{n} \langle h, y_k \rangle$, we have $v = G w$ where $w = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$.

$$\begin{aligned} ||x||^2 &= \left\langle \sum_{j=1}^{n} \alpha_j y_j + h, \sum_{j=1}^{n} \alpha_j y_j + h \right\rangle \\ &= \left\langle \sum_{j=1}^{n} \alpha_j y_j, \sum_{k=1}^{n} \alpha_k y_k \right\rangle + \left\langle h, \sum_{k=1}^{n} \alpha_k y_k \right\rangle + \left\langle \sum_{j=1}^{n} \alpha_j y_j, h \right\rangle + \langle h, h \rangle \\ &= \sum_{j=1}^{n} \sum_{k=1}^{n} \alpha_j \overline{\alpha_k} \, \overline{\langle y_k, y_j \rangle} + ||h||^2 \end{aligned}$$

$= w^* G^* w + ||h||^2$

$= w^* G w + ||h||^2.$

For a linearly independent system $\{y_j\}_{j=1}^n$, the matrix G^{-1} exists, and we have $v^* G^{-1} v = (Gw)^* G^{-1} G w = w^* G^* w = w^* G w$.

We have equality when $h = 0$. ■

If $y_j \perp y_k$, $j \neq k$ and $||y_j|| = 1$, then $G = I$ and the generalized Bessel inequality can be written as $\sum_{j=1}^n |\langle x, y_j\rangle|^2 \leq ||x||^2$.

2.2.2 Selberg's inequality

Here we give a new proof of Selberg's inequality based on the theory of positive semidefinite matrices.

Theorem 7 (Selberg's inequality).
In a Hilbert space $\mathcal{H}$,

$$\sum_{j=1}^n \frac{|\langle x, y_j\rangle|^2}{\sum_{k=1}^n |\langle y_j, y_k\rangle|} \leq ||x||^2 \quad \forall x \in \mathcal{H} \ y_j \neq 0 \ y_j \in \mathcal{H} \tag{2.3}$$

Proof. We have $\displaystyle\sum_{j=1}^n \frac{|\langle x, y_j\rangle|^2}{\sum_{k=1}^n |\langle y_j, y_k\rangle|} \leq ||x||^2 \Leftrightarrow v^* D^{-1} v \leq ||x||^2$ where

$$v = \begin{bmatrix} \langle x, y_1\rangle \\ \langle x, y_2\rangle \\ \vdots \\ \langle x, y_n\rangle \end{bmatrix}, D = \begin{bmatrix} d_1 & & & \\ & d_2 & \mathbf{0} & \\ & \mathbf{0} & \ddots & \\ & & & d_n \end{bmatrix}, d_j = \sum_{k=1}^n |\langle y_j, y_k\rangle|.$$ See Appendix A.

We will prove that $v^* D^{-1} v \leq ||x||^2$.

Let $x = \displaystyle\sum_{j=1}^n \alpha_j y_j + h$ where $\alpha_j \in \mathbb{C}$, $h \in \mathcal{H}$, $h \perp y_j$, for any $j \in \{1, 2, \ldots, n\}$.

From the proof of the generalized Bessel's inequality we have
$||x||^2 = w^* G w + ||h||^2$ and $v = Gw$.

$$v^* D^{-1} v \leq ||x||^2$$

$$\Updownarrow$$

$$(Gw)^* D^{-1} G w \leq w^* G w + ||h||^2$$

$\Updownarrow$

$w^* G^* D^{-1} G w \leq w^* G w + ||h||^2$

$\Updownarrow$

$w^* G D^{-1} G w \leq w^* G w + ||h||^2$.

Hence it is sufficient to show that $G - GD^{-1}G \geq 0$.
First we will show that $G \leq D$.

$$u^* G u = \langle \beta_1 y_1 + \beta_2 y_2 + \ldots + \beta_n y_n\,,\, \beta_1 y_1 + \beta_2 y_2 + \ldots + \beta_n y_n \rangle = \sum_{j=1}^{n} \sum_{k=1}^{n} \beta_j \overline{\beta_k} \langle y_j, y_k \rangle$$

$$\leq \sum_{j=1}^{n} \sum_{k=1}^{n} |\beta_j| \, |\overline{\beta_k}| \, |\langle y_j, y_k \rangle| \leq \sum_{j=1}^{n} \sum_{k=1}^{n} \left(\tfrac{1}{2} |\beta_j|^2 + \tfrac{1}{2} |\overline{\beta_k}|^2 \right) |\langle y_j, y_k \rangle|$$

$$= \sum_{j=1}^{n} \sum_{k=1}^{n} |\beta_j|^2 |\langle y_j, y_k \rangle| \text{ for all nonzero } u = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}, \beta_j \in \mathbb{C}.$$

$$u^* D u = \sum_{j=1}^{n} |\beta_j|^2 d_j = \sum_{j=1}^{n} \sum_{k=1}^{n} |\beta_j|^2 |\langle y_j, y_k \rangle|. \text{ Hence } G \leq D.$$

Further $0 \leq G - GD^{-1}G \Leftrightarrow 0 \leq G\left(I - D^{-1}G\right) \Leftrightarrow 0 \leq G\left(D^{-1}(D - G)\right)$.

We have $G \geq 0$, $D^{-1} \geq 0$, D^{-1} is invertible, $D - G \geq 0$, and $G - GD^{-1}G$ is Hermitian, so (xiv) is satisfied. Hence $G - GD^{-1}G \geq 0$. ■

Remark 4. The proof for Selberg's inequality is valid for all systems $\{y_j\}$ including linearly dependent systems that has singular Gram matrices.

2.2.3 Generalized Selberg's inequality

Here we introduce a new inequality based on the results from the proof of Selberg's inequality in Chapter 2.2.2.

Theorem 8 (Generalized Selberg's inequality).
In a Hilbert space $\mathcal{H}$. If $y_1, \ldots, y_n \in \mathcal{H}$, G is the Gram matrix for $y_1, \ldots, y_n$, $E \geq G$, and E is invertible, then

$$v^* E^{-1} v \leq ||x||^2 \quad \forall x \in \mathcal{H} \tag{2.4}$$

where $v = \begin{bmatrix} \langle x, y_1 \rangle \\ \langle x, y_2 \rangle \\ \vdots \\ \langle x, y_n \rangle \end{bmatrix}$.

Proof. Let $x = \sum_{j=1}^{n} \alpha_j y_j + h$ where $\alpha_j \in \mathbb{C}$, $h \in \mathcal{H}$, $h \perp y_j$, for any $j \in \{1, 2, \dots, n\}$.

From the proof of the generalized Bessel's inequality we have

$||x||^2 = w^* G w + ||h||^2$, and $v = G w$.

$v^* E^{-1} v \leq ||x||^2$

$$\Updownarrow$$

$(G w)^* E^{-1} G w \leq w^* G w + ||h||^2$

$$\Updownarrow$$

$w^* G^* E^{-1} G w \leq w^* G w + ||h||^2$

$$\Updownarrow$$

$w^* G E^{-1} G w \leq w^* G w + ||h||^2$.

Hence it is sufficient to show that $G - G E^{-1} G \geq 0$.

Further $0 \leq G - G E^{-1} G \Leftrightarrow 0 \leq G\left(I - E^{-1} G\right) \Leftrightarrow 0 \leq G\left(E^{-1}(E - G)\right)$.

We have $G \geq 0$, $E^{-1} \geq 0$, E^{-1} is invertible, $E - G \geq 0$, and $G - G E^{-1} G$ is Hermitian, so (xiv) is satisfied. Hence $G - G E^{-1} G \geq 0$. ■

Remark 5. The Inequality in Theorem 8 is called Generalized Selberg's inequality because by choosing

$$E = \begin{bmatrix} d_1 & & & \\ & d_2 & & \mathbf{0} \\ & \mathbf{0} & \ddots & \\ & & & d_n \end{bmatrix}$$ where $d_j = \sum_{k=1}^{n} |\langle y_j, y_k \rangle|$ we have

Selberg's inequality.

2.3 Frames

In this section we show how the matrix approach developed in Chapter 2.1 and Chapter 2.2 can be used to obtain optimal frame bounds.
We introduce a new notation for frame bounds, see page vii.

2.3.1 Definition and Cauchy-Schwarz upper bound

We will follow [Christensen] in this subsection.

Definition 2.
A countable system of elements $\{y_j\}_{j\geq 1}$ in a Hilbert space $\mathcal{H}$
is a **frame** for $\mathcal{H}$ if there exist constants $0 < a \leq b < \infty$, such that

$$a\,||x||^2 \leq \sum_{j\geq 1} |\langle x, y_j\rangle|^2 \leq b\,||x||^2 \quad \forall x \in \mathcal{H}. \tag{2.5}$$

The constants a, b are called frame bounds.

a is a lower frame bound, and b is an upper frame bound.
Frame bounds are not unique. The optimal lower frame bound is the supremum over all lower frame bounds, and the optimal upper frame bound is the infimum over all upper frame bounds. We can have infinitely many elements $\{y_j\}$ in a frame. We will here assume that we have finitely many elements $\{y_j\}$ in a frame. An important task is to estimate the frame bounds. We start with the upper bound. The first estimate is quite obvious.

From the Cauchy-Schwarz inequality we have

$\sum_{j=1}^{m} |\langle x, y_j\rangle|^2 \leq \sum_{j=1}^{m} ||y_j||^2\,||x||^2$, which gives $b = \sum_{j=1}^{m} ||y_j||^2$.

From the Cauchy-Schwarz inequality it follows that we always have an upper bound.

2.3.2 Upper bound

We will in this subsection use the generalized Selberg inequality to find the optimal upper frame bound.

Lemma 1. Let $\{y_j\}_{j=1}^{m}$ be a system of elements in a Hilbert space $\mathcal{H}$ and let

G be the corresponding Gram matrix. Then for all $x \in \mathcal{H}$ we have

$$\sum_{j=1}^{m} |\langle x, y_j \rangle|^2 \leq \lambda_{\max}(G)\, \|x\|^2 \tag{2.6}$$

Moreover, we have

$$\sum_{j=1}^{m} |\langle x, y_j \rangle|^2 \leq b\, \|x\|^2 \Rightarrow \lambda_{\max}(G) \leq b \tag{2.7}$$

Proof. First we will prove (2.6).
From (xv) in Chapter 2.1.2 we have $\lambda_{\max}(G)\, I \geq G$, and then by applying generalized Selberg inequality we have (2.6).

To prove (2.7) assume that $\sum_{j=1}^{m} |\langle x, y_j \rangle|^2 \leq b\, \|x\|^2$.

Choose $x = \sum_{j=1}^{m} \alpha_j y_j$ such that $G w = \lambda_{\max}(G)\, w$ where $w = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix}$, then

$\sum_{j=1}^{m} |\langle x, y_j \rangle|^2 = (G w)^* G w$, and $\|x\|^2 = w^* G w$. We have

$$\sum_{j=1}^{m} |\langle x, y_j \rangle|^2 \leq b\, \|x\|^2$$

$$\Downarrow$$

$$(\lambda_{\max}(G))^2\, w^* w \leq b\, \lambda_{\max}(G)\, w^* w$$

$$\Downarrow$$

$\lambda_{\max}(G) \leq b$. ■

(2.6) means that $\lambda_{\max}(G)$ is an upper bound.
(2.6) and (2.7) means that $\lambda_{\max}(G)$ is an optimal upper bound.

Remark 6. We have not used any properties of frame for describing the optimal upper bound. (2.6) and (2.7) holds for any finite system $\{y_j\}_{j=1}^{m}$.

From the Selberg inequality we have

$$\sum_{j=1}^{m} |\langle x, y_j \rangle|^2 \leq \left(\max_{j=1,\ldots,m} \sum_{k=1}^{m} |\langle y_j, y_k \rangle| \right) \|x\|^2 \tag{2.8}$$

This implies

$$\lambda_{\max}(G) \leq \max_{j=1,\ldots,m} \sum_{k=1}^{m} |\langle y_j, y_k \rangle| \tag{2.9}$$

If $G = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$, then $\lambda_{\max}(G) < \max_{j=1,\ldots,m} \sum_{k=1}^{m} |\langle y_j, y_k \rangle|$.

Remark 7. $b = \max_{j=1,\ldots,m} \sum_{k=1}^{m} |\langle y_j, y_k \rangle|$ is an upper frame bound for $\{y_j\}_{j=1}^{m}$ that always exists. It may not be optimal but it is easier to compute than $\lambda_{\max}(G)$.

2.3.3 Lower bound

To find the optimal lower bound we use the condition that $\{y_j\}_{j=1}^{m}$ is a frame.

Proposition 1. Let $\{y_j\}_{j=1}^{m}$ be a sequence in a Hilbert space $\mathcal{H}$.
Then $\{y_j\}_{j=1}^{m}$ is a frame for *span*$\{y_j\}_{j=1}^{m}$.

Proof. See [Christensen,p4]. ■

Corollary 1. A system of elements $\{y_j\}_{j=1}^{m}$ in a Hilbert space $\mathcal{H}$ is a frame for $\mathcal{H}$ if and only if *span*$\{y_j\}_{j=1}^{m} = \mathcal{H}$.

Proof. See [Christensen,p4]. ■

Lemma 2. Let $\{y_j\}_{j=1}^{m}$ be a frame for a Hilbert space $\mathcal{H}$ and let G be the corresponding Gram matrix. Then for all $x \in \mathcal{H}$ we have

$$\lambda_{\min>0}(G) \|x\|^2 \leq \sum_{j=1}^{m} |\langle x, y_j \rangle|^2 \tag{2.10}$$

where $\lambda_{\min>0}(G)$ is the smallest positive eigenvalue of G.

Moreover, we have

$$a\,\|x\|^2 \le \sum_{j=1}^{m} |\langle x, y_j\rangle|^2 \Rightarrow a \le \lambda_{\min>0}(G) \tag{2.11}$$

Proof. Let $x = \sum_{j=1}^{m} \alpha_j y_j$, since from Corollary 1, we have $x \in span\{y_1, \dots, y_m\}$.

Then we have $\sum_{j=1}^{m} |\langle x, y_j\rangle|^2 = (G\,w)^* G\,w$, and $\|x\|^2 = w^* G\,w$ where $w = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix}$.

$\lambda_{\min>0}(G)\,\|x\|^2 \le \sum_{j=1}^{m} |\langle x, y_j\rangle|^2$

$\Updownarrow$

$\lambda_{\min>0}(G)\,w^* G\,w \le (G\,w)^* G\,w$

$\Updownarrow$

$w^* \lambda_{\min>0}(G)\,G\,w \le w^* G^2\,w.$

We have $\lambda_{\min>0}(G)\,G \le G^2$, see (xvi) in Chapter 2.1.2.

Next, assume $a\|x\|^2 \le \sum_{j=1}^{m} |\langle x, y_j\rangle|^2$. Choose $x = \sum_{j=1}^{m} \alpha_j y_j$ such that

$G\,w = \lambda_{\min>0}(G)\,w$ where $w = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix}$, then

$\sum_{j=1}^{m} |\langle x, y_j\rangle|^2 = (G\,w)^* G\,w$, and $\|x\|^2 = w^* G\,w$. We have

$a\|x\|^2 \le \sum_{j=1}^{m} |\langle x, y_j\rangle|^2$

$\Downarrow$

$a\lambda_{\min>0}(G)\,w^* w \le (\lambda_{\min>0}(G))^2\; w^* w$

$\Downarrow$

$a \leq \lambda_{\min>0}(G)$. ■

(2.10) means that $\lambda_{\min>0}(G)$ is a lower bound.
(2.10) and (2.11) means that $\lambda_{\min>0}(G)$ is an optimal lower bound.

2.3.4 Tight frames

Definition 3. A frame is a **tight frame** if (2.5) is satisfied with $a = b$, that is if the optimal upper frame bound and the optimal lower frame bound are equal.
a is then called the frame bound.

When we have a tight frame $\{y_j\}_{j=1}^m$ in a Hilbert space $\mathcal{H}$, (2.5) becomes

$$\sum_{j=1}^{m} |\langle x, y_j \rangle|^2 = a \, \|x\|^2 \quad \forall x \in \mathcal{H}. \tag{2.12}$$

Proposition 2. Assume $\{y_j\}_{j=1}^m$ is a tight frame for a Hilbert space $\mathcal{H}$ with frame bound a. Then

$$x = \frac{1}{a} \sum_{j=1}^{m} \langle x, y_j \rangle \, y_j \quad \forall x \in \mathcal{H}. \tag{2.13}$$

Proof. See [Christensen,p5] ■

Theorem 9 (Casazza,Fickus,Kovačević,Leon,Tremain).
Given an n-dimensional Hilbert space $\mathcal{H}$ and a sequence of positive scalars $\{a_j\}_{j=1}^m$, there exists a tight frame $\{y_j\}_{j=1}^m$ for $\mathcal{H}$ of lengths $\|y_m\| = a_m$ for all
$j = 1, \ldots, m$ if and only if,

$$\max_{j=1,\ldots,m} a_j^2 \leq \frac{1}{n} \sum_{j=1}^{m} a_j^2 \tag{2.14}$$

Proof. See [Casazza,Fickus,Kovačević,Leon,Tremain,p33]. ■

Theorem 10.
$\{y_j\}_{j=1}^m$ is a tight frame for a Hilbert space $\mathcal{H}$ if and only if *span*$\{y_j\}_{j=1}^m = \mathcal{H}$ and
$\lambda_{\min>0}(G) = \lambda_{\max}(G)$. G is the corresponding Gram matrix.

Proof. Follows from Corollary 1, Lemma 1 and Lemma 2. ■

2.3.5 Examples of tight frames

The following are examples of tight frames for $\mathbb{R}^3$.

(a) $y_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}$, $y_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}$, $y_3 = \begin{bmatrix}0\\0\\1\end{bmatrix}$, $G = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}$.

(b) $y_1 = \begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}$, $y_2 = \begin{bmatrix}-\frac{1}{2}\\-\frac{1}{2}\\\frac{1}{2}\end{bmatrix}$, $y_3 = \begin{bmatrix}-\frac{1}{2}\\\frac{1}{2}\\-\frac{1}{2}\end{bmatrix}$, $y_4 = \begin{bmatrix}\frac{1}{2}\\-\frac{1}{2}\\-\frac{1}{2}\end{bmatrix}$,

$$G = \begin{bmatrix}\frac{3}{4} & -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4}\\-\frac{1}{4} & \frac{3}{4} & -\frac{1}{4} & -\frac{1}{4}\\-\frac{1}{4} & -\frac{1}{4} & \frac{3}{4} & -\frac{1}{4}\\-\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{3}{4}\end{bmatrix} \sim \begin{bmatrix}1 & 0 & 0 & 0\\0 & 1 & 0 & 0\\0 & 0 & 1 & 0\\0 & 0 & 0 & 0\end{bmatrix}.$$

(c) $y_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}$, $y_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}$, $y_3 = \begin{bmatrix}0\\0\\1\end{bmatrix}$, $y_4 = \begin{bmatrix}1\\0\\0\end{bmatrix}$, $y_5 = \begin{bmatrix}0\\1\\0\end{bmatrix}$, $y_6 = \begin{bmatrix}0\\0\\1\end{bmatrix}$,

$$G = \begin{bmatrix}1 & 0 & 0 & 1 & 0 & 0\\0 & 1 & 0 & 0 & 1 & 0\\0 & 0 & 1 & 0 & 0 & 1\\1 & 0 & 0 & 1 & 0 & 0\\0 & 1 & 0 & 0 & 1 & 0\\0 & 0 & 1 & 0 & 0 & 1\end{bmatrix} \sim \begin{bmatrix}2 & 0 & 0 & 0 & 0 & 0\\0 & 2 & 0 & 0 & 0 & 0\\0 & 0 & 2 & 0 & 0 & 0\\0 & 0 & 0 & 0 & 0 & 0\\0 & 0 & 0 & 0 & 0 & 0\\0 & 0 & 0 & 0 & 0 & 0\end{bmatrix}.$$

(d) $y_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}$, $y_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}$, $y_3 = \begin{bmatrix}0\\0\\1\end{bmatrix}$, $y_4 = \begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}$, $y_5 = \begin{bmatrix}-\frac{1}{2}\\-\frac{1}{2}\\\frac{1}{2}\end{bmatrix}$, $y_6 = \begin{bmatrix}-\frac{1}{2}\\\frac{1}{2}\\-\frac{1}{2}\end{bmatrix}$, $y_7 = \begin{bmatrix}\frac{1}{2}\\-\frac{1}{2}\\-\frac{1}{2}\end{bmatrix}$,

$$G = \begin{bmatrix} 1 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{3}{4} & -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{4} & \frac{3}{4} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{4} & -\frac{1}{4} & \frac{3}{4} & -\frac{1}{4} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{3}{4} \end{bmatrix} \sim \begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

We see that *span*$\{y_j\}_{j=1}^m = \mathbb{R}^3$ and $\lambda_{\min>0}(G) = \lambda_{\max}(G)$ in all the four examples. By Theorem 10 we have tight frames in all the four examples.
In example (a) $a = 1$, in example (b) $a = 1$, in example (c) $a = 2$ and in example (d) $a = 2$.

Conclusion

We have shown that by using a generalized form of nonnegative real numbers called positive semidefinite matrices we get a nontrivial generalization of the Selberg inequality.

Appendices

Appendix A

Here we will show the first equivalence in the proof of Theorem 7 in Chapter 2.2.2.

$$\overline{v}^T D^{-1} v \leq ||x||^2$$

$$\Updownarrow$$

$$\begin{bmatrix} \overline{\langle x, y_1 \rangle} & \overline{\langle x, y_2 \rangle} & \dots & \overline{\langle x, y_n \rangle} \end{bmatrix} \begin{bmatrix} \frac{1}{d_1} & & & \\ & \frac{1}{d_2} & & \mathbf{0} \\ \mathbf{0} & & \ddots & \\ & & & \frac{1}{d_n} \end{bmatrix} \begin{bmatrix} \langle x, y_1 \rangle \\ \langle x, y_2 \rangle \\ \vdots \\ \langle x, y_n \rangle \end{bmatrix} \leq ||x||^2$$

$$\Updownarrow$$

$$\begin{bmatrix} \frac{1}{d_1} \overline{\langle x, y_1 \rangle} & \frac{1}{d_2} \overline{\langle x, y_2 \rangle} & \dots & \frac{1}{d_n} \overline{\langle x, y_n \rangle} \end{bmatrix} \begin{bmatrix} \langle x, y_1 \rangle \\ \langle x, y_2 \rangle \\ \vdots \\ \langle x, y_n \rangle \end{bmatrix} \leq ||x||^2$$

$$\Updownarrow$$

$$\frac{1}{d_1} \overline{\langle x, y_1 \rangle} \langle x, y_1 \rangle + \frac{1}{d_2} \overline{\langle x, y_2 \rangle} \langle x, y_2 \rangle + \dots + \frac{1}{d_n} \overline{\langle x, y_n \rangle} \langle x, y_n \rangle \leq ||x||^2$$

$$\Updownarrow$$

$$\frac{|\langle x, y_1 \rangle|^2}{\sum_{k=1}^{n} |\langle y_1, y_k \rangle|} + \frac{|\langle x, y_2 \rangle|^2}{\sum_{k=1}^{n} |\langle y_2, y_k \rangle|} + \dots + \frac{|\langle x, y_n \rangle|^2}{\sum_{k=1}^{n} |\langle y_n, y_k \rangle|} \leq ||x||^2$$

$$\Updownarrow$$

$$\sum_{j=1}^{n} \frac{|\langle x, y_j \rangle|^2}{\sum_{k=1}^{n} |\langle y_j, y_k \rangle|} \leq ||x||^2.$$

References

Akhiezer N. I., Glazman I. M., Theory of Linear Operators in Hilbert Space Volume I, Pitman, London, 1981.

Casazza P. G., Fickus M., Kovačević J., Leon M. T., and Tremain J. C., A Physical Interpretation for Finite Tight Frames, Preprint submitted to Elsevier Science, 15 October 2003.
http://www.math.missouri.edu/ pete/pdf/85.fiftf.pdf

Christensen Ole, An Introduction to Frames and Riesz Bases, Birkhäuser, Boston, 2003.

Folland Gerald B., Real Analysis: Modern Techniques and Their Applications, 2nd ed., Wiley-Interscience, 1999.

Furuta Takayuki, Invitiation to Linear Operators: From Matrices to Bounded Linear Operators on a Hilbert Space, Taylor & Francis, London, 2002.

Gelbaum Bernard R., Olmsted John M.H. Counterexamples in Analysis, Dover Publications, 2003.

Griffel D. H., Applied Functional Analysis, Ellis Horwood, 1985.

Lang Serge, Real and Functional Analysis, 3rd ed., Springer-Verlag, New York, 1993.

Schechter Martin, Principles of Functional Analysis, 2nd ed., American Mathematical Society, 2002.

Weidmann Joachim, Linear Operators in Hilbert Spaces, Springer-Verlag, New York, 1980.

Zhang Fuzhen, Matrix Theory: Basic Results and Techniques, Springer-Verlag, New York, 1999.