32

MEAN SQUARE APPROXIMATION

Let \( C(\mathbb{T}) \) be the set of continuous functions \( f: \mathbb{T} \rightarrow \mathbb{C} \). Write \( (f, g) = (2\pi)^{-1} \int_{\mathbb{T}} f(t)g(t)^* \, dt \) where \( z^* \) denotes the complex conjugate of \( z \).

**Lemma 32.1.** If \( f, g, h \in C(\mathbb{T}) \) and \( \lambda, \mu \in \mathbb{C} \) then

(i) \( (f, g) = (g, f)^* \),

(ii) \( (\lambda f + \mu g, h) = \lambda (f, h) + \mu (g, h) \),

(iii) \( (f, f) \) is real and \( (f, f) \geq 0 \),

(iv) If \( (f, f) = 0 \) then \( f = 0 \).

**Proof.**

(i) \( (g, f)^* = \frac{1}{2\pi} \left( \int_{\mathbb{T}} g(t)f(t)^* \, dt \right)^* = \frac{1}{2\pi} \int_{\mathbb{T}} (g(t)f(t)^*)^* \, dt \)

\[ = \frac{1}{2\pi} \int_{\mathbb{T}} f(t)^* g(t)^* \, dt = \frac{1}{2\pi} \int_{\mathbb{T}} f(t)g(t)^* \, dt = (f, g). \]

(ii) \( (\lambda f + \mu g, h) = \frac{1}{2\pi} \int_{\mathbb{T}} (\lambda f(t) + \mu g(t))h(t)^* \, dt \)

\[ = \frac{1}{2\pi} \int_{\mathbb{T}} \lambda f(t)h^*(t) + \mu g(t)h^*(t) \, dt = \lambda (f, h) + \mu (g, h). \]

(iii) \( (f, f) = \frac{1}{2\pi} \int_{\mathbb{T}} f(t)f(t)^* \, dt = \frac{1}{2\pi} \int_{\mathbb{T}} |f(t)|^2 \, dt \geq 0. \)

(iv) Recall that, if \( g \) is a continuous positive function on an interval \( [a, b] \) with \( a < b \), then \( \int_{a}^{b} g(x) \, dx = 0 \) implies \( g(x) = 0 \) for all \( x \in [a, b] \). Setting \( g(x) = |f(x)|^2 \) we see that if \( (f, f) = 0 \) then \( |f(x)|^2 = 0 \) for all \( x \in \mathbb{T} \), so \( f(x) = 0 \) for all \( x \in \mathbb{T} \) and we are done.

In the language of abstract algebra \( C(\mathbb{T}) \) is an (infinite dimensional) vector space.
with ( , ) as an inner product. All inner products satisfy a very important inequality, special cases of which were discovered by Cauchy, Schwarz, Buniakowski and, probably, many others.

**Lemma 32.2 (The Cauchy, Schwarz, Buniakowski inequality).** If \( f, g \in C(\mathbb{T}) \) then 
\[ |(f,g)|^2 \leq (f,f)(g,g) \text{ with equality if and only if } \lambda f + \mu g = 0 \text{ for some } \lambda, \mu \in \mathbb{C} \text{ not both zero.} \]

**Proof.** If \( f - g = 0 \), then there is nothing to prove, so we may suppose without loss of generality that \( f \neq 0 \) and so, by Lemma 32.1 (iv), \( (f,f) \neq 0 \). Thus using the various results of Lemma 32.1, we have

\[
0 \leq (\lambda f + \mu g, \lambda f + \mu g) = \lambda(f,f) + \mu(g,g) + \lambda \mu (f,g) + \mu \lambda (g,f)
\]

\[
= \lambda \lambda^* (f,f) + \mu \mu^* (g,g) + \lambda \mu (f,g) + \mu \lambda (g,f)
\]

\[
= \left| \lambda (f,f) + \mu \left( \frac{g}{(f,f)^{1/2}} \right)^* \right|^2 + \mu^* \left( \frac{(g,g) - \frac{|(g,f)|^2}{(f,f)}}{(f,f)} \right) \quad \text{for all } \lambda, \mu \in \mathbb{C}.
\]

In particular, taking \( \mu = 1, \lambda = -\frac{(g,f)}{(f,f)} \), we see that

\[
0 \leq \left( \frac{(g,g) - \frac{|(g,f)|^2}{(f,f)}}{(f,f)} \right),
\]

and so \( |(g,f)|^2 \leq (f,f)(g,g) \), i.e. \( |(f,g)|^2 \leq (f,f)(g,g) \) (since \( (f,g) = (g,f)^* \)) with equality only if

\[
0 = (\lambda f + \mu g, \lambda f + \mu g),
\]

i.e. (using Lemma 32.1 (iv)) only if

\[
\lambda f + \mu g = 0,
\]

where \( \mu \) and \( \lambda \) have the values chosen at the beginning of the paragraph.

Conversely, if \( \lambda f + \mu g = 0 \) for some \( \lambda \) and \( \mu \) with, say, \( \lambda \neq 0 \) then \( f = -\frac{1}{\lambda} \mu g \) and

\[
|(f,g)|^2 = |\frac{1}{\lambda} \mu|^2 (g,g)^2 = |\frac{1}{\lambda} \mu|^2 ((\lambda \mu^*)(g,g))(g,g) = (f,f)(g,g),
\]

so we are done.

Every inner product has an associated norm \( ||| \cdot ||_2 \) given by \( ||f||_2 = (f,f)^{1/2} \) (where the positive square root is taken).

**Lemma 33.1.** If \( f, g \in C(\mathbb{T}) \) and \( \lambda \in \mathbb{C} \) then

(i) \( ||\lambda f||_2 = |\lambda| ||f||_2 \).
Mean square approximation

(i) \( \| f \|_2 = 0 \) with equality if and only if \( f = 0 \).

(ii) (triangle inequality) \( \| f + g \|_2 \geq \| f \|_2 + \| g \|_2 \).

Proof. (i) \( \| \lambda f \|_2 = |\lambda| \| f \|_2 = \lambda \| f \|_2 = \lambda (\lambda f, f)^* = \lambda^* (f, f)^* = \| \lambda^* f \|_2 = \| f \|_2 \).

(ii) Immediate from the definition and Lemma 32.1 (iv).

(iii) Using Lemma 32.1 and Lemma 32.2 (the Cauchy, Schwartz, Buniaowski inequality) we have

\[
\| f + g \|_2^2 = (f, f + g) = (f, f) + (g, f) + (f, g) + (g, g) \\
= (f, f) + (f, g) + (g, f) + (g, g) = \| f \|_2^2 + 2 \text{Re}(f, g) + \| g \|_2^2 \\
\leq \| f \|_2^2 + 2 \| f \|_2 \| g \|_2 + \| g \|_2^2 \\
= (\| f \|_2 + \| g \|_2)^2.
\]

and the result follows.

In the context of this part of the book the exponentials \( e_n(t) = \exp(\text{int}t \in \mathbb{T}) \) have one outstanding characteristic: they are orthonormal.

**Lemma 32.4.** (i) \( (e_n, e_n) = 1 \),
(ii) \( (e_n, e_m) = 0 \) for \( n \neq m \).

**Proof.** Trivial.

The great advantage of inner product spaces is that they allow the use of geometric analogy. For example, consider

**Question A.** What values of \( \lambda_{-n} \ldots \lambda_n \in \mathbb{C} \) (if any) minimise \( \| f - \sum_{-n}^n \lambda_j e_j \|_2 \)?

In geometrical language this may be restated as

**Question B.** Which points \( g \) (if any) in the subspace \( E = \{ \sum_{-n}^n \lambda_j e_j : \lambda_j \in \mathbb{C} \} \) minimise \( \| f - g \|_2 \)?

Our geometrical intuition suggests the following answer.

**Hypothesis C.** There is a unique point \( g_0 \in E \) with \( \| f - g \|_2 > \| f - g_0 \|_2 \) for all \( g \in E, \ g \neq g_0 \). This \( g_0 \) is given by the condition that \( f - g_0 \) be perpendicular to \( E \).

What does it mean to say that \( f - g \) is perpendicular to \( E \)? Geometrically it means that \( f - g_0 \) is perpendicular to each \( h \in E \), or, following the analogy with the physicist's inner product on \( \mathbb{R}^3 \), that \( (f - g_0, h) = 0 \) for each \( h \in E \). In particular since \( e_j \in E \) it follows that \( (f - g_0, e_j) = 0 \) and so \( \langle f, e_j \rangle = (g_0, e_j) \) for each \( -n \leq j \leq n \).

But if \( g_0 \in E \) then \( g_0 = \sum_{-n}^n \mu_k e_k \) for some \( \mu_k \in \mathbb{C} \) \( -n \leq k \leq n \), where we have \( (g_0, e_j) = \sum_{-n}^n \mu_k (e_k, e_j) = \mu_j \) so that \( \mu_j = (f, e_j) \) \( -n \leq j \leq n \) and we have
Orthogonal series

\[ g_0 = \sum_{j=-n}^{n} (f, e_j) e_j \]

Hypothesis C can thus be rewritten algebraically to give the following theorem.

**Theorem 32.5.** If \( g_0 = \sum_{j=-n}^{n} (f, e_j) e_j \) and \( g = \sum_{j=-n}^{n} \lambda_j e_j \)

\[ \| f \|_2 \geq \sqrt{\sum_{j=-n}^{n} |(f, e_j)|^2} \]

and

\[ \| f - g \|_2 = \sqrt{\left( \| f \|_2^2 - \sum_{j=-n}^{n} |(f, e_j)|^2 \right)} \]

with equality if and only if \( \lambda_j = (f, e_j)[-n \leq j \leq n] \).

**Proof.**

\[ \| f - g \|_2^2 = \left( f - \sum_{j=-n}^{n} \lambda_j e_j, f - \sum_{j=-n}^{n} \lambda_j e_j \right) \]

\[ = (f, f) - \sum_{j=-n}^{n} \lambda_j (e_j, f) - \sum_{j=-n}^{n} \lambda_j^* (f, e_j) + \sum_{j=-n}^{n} \sum_{k=-n}^{n} \lambda_j \lambda_k^* (e_j, e_k) \]

\[ = (f, f) - \sum_{j=-n}^{n} \lambda_j (f, e_j) - \sum_{j=-n}^{n} \lambda_j^* (f, e_j) + \sum_{j=-n}^{n} \lambda_j \lambda_j^* \]

\[ = (f, f) - \sum_{j=-n}^{n} (f, e_j)^2 + \sum_{j=-n}^{n} |(f, e_j)|^2 \]

\[ \geq (f, f) - \sum_{j=-n}^{n} |(f, e_j)|^2 = \| f \|_2^2 - \sum_{j=-n}^{n} |(f, e_j)|^2 \]

with equality if and only if \( \lambda_j = (f, e_j)[-n \leq j \leq n] \). Setting \( \lambda_j = (f, e_j) \), we thus obtain

\[ \| f - g_0 \|_2 = \| f \|_2^2 - \sum_{j=-n}^{n} |(f, e_j)|^2 \]

so

\[ \| f \|_2^2 - \sum_{j=-n}^{n} |(f, e_j)|^2 \geq 0 \]

and the full result now follows.

Let us call

\[ \| f - g \|_2 = \sqrt{\left( \frac{1}{2\pi} \int \left| f(t) - g(t) \right|^2 dt \right)} \]

the mean square distance between \( f \) and \( g \). Theorem 32.5 can now be rewritten as follows.

**Theorem 32.6.** Let \( f : Y \to \mathbb{C} \) be continuous. Then, among all functions of the form
Mean square approximation

\[ \sum_{j=-N}^{N} \lambda_j e_j, \text{ the best mean square approximation to } f \text{ is given by } S_N(f). \text{ Further,} \]

\[ \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(t) - \sum_{j=-N}^{N} \lambda_j \exp(\text{i}jt)|^2 dt \geq \frac{1}{2\pi} \int_{-\pi}^{\pi} |\tilde{f}(t)|^2 dt - \sum_{j=-N}^{N} |\tilde{f}(j)|^2, \text{ for all } \lambda_j \in \mathbb{C}. \]

**Proof.** Observe that \( \tilde{f}(j) = (f, e_j) \) and apply Theorem 32.5.

Thus, although the Fourier sum \( S_N(f) \) may not be a good uniform approximation (see Chapter 11 and elsewhere) it is always the best mean square approximation to \( f \).