ECEn 671: Mathematics of Signals and Systems Moon: Chapter 3.

Randal W. Beard

Brigham Young University

August 29, 2023

Table of Contents

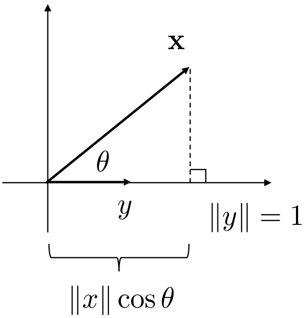
Section 1

Approximation Theory

Projection and Inner Product

- How does inner product represent a projection?
- Recall that

$$\langle x, y \rangle = \|x\| \, \|y\| \, cos\theta$$



- In 2-D $\langle x, y \rangle$ represents the length of the projection of x in the direction of y.
- In general, inner products represent (non-orthogonal) projection of one vector onto another.

Approximation Problem

- Let $\mathbb{S}$ be a Hilbert space, and let $x \in \mathbb{S}$.
- Let $\{p_1, \ldots, p_n\}$ be a set of vectors, all in $\mathbb{S}$.
- Find $\hat{x} \in \mathrm{span}\{p_1, \ldots p_n\}$ that minimizes $\|x - \hat{x}\|$.

Approximation Problem, cont

- Let $\hat{x} = c_1 p_1 + \ldots + c_n p_n \in \text{span}\{p_1, \ldots, p_n\}$.
- By the projection theorem, the error

$$\begin{aligned} e &= x - \hat{x} \\ &= x - c_1 p_1 - \ldots - c_n p_n \end{aligned}$$

is minimized if

$$e \perp \text{span}\{p_1, \ldots, p_n\}.$$

Approximation Problem, cont

$$e \perp \operatorname{span}\{p_1, \ldots, p_n\}.$$

iff

$$\begin{aligned} \langle e, p_1 \rangle &= 0 \\ \langle e, p_2 \rangle &= 0 \\ &\vdots \\ \langle e, p_n \rangle &= 0 \end{aligned}$$

iff

$$\begin{aligned} \langle x - c_1 p_1 - \ldots c_n p_n, p_1 \rangle &= 0 \\ &\vdots \\ \langle x - c_1 p_1 - \ldots c_n p_n, p_n \rangle &= 0 \end{aligned}$$

Approximation Problem, cont

By properties of the inner product we can write this as

$$\begin{aligned}\langle x, p_1\rangle - c_1\langle p_1, p_1\rangle - \cdots - c_n\langle p_n, p_1\rangle &= 0\\ &\vdots\\ \langle x, p_n\rangle - c_1\langle p_1, p_n\rangle - \cdots - c_n\langle p_n, p_n\rangle &= 0\end{aligned}$$

or in matrix notation,

$$\underbrace{\begin{pmatrix}\langle p_1, p_1\rangle & \cdots & \langle p_n, p_1\rangle\\ \vdots & & \vdots\\ \langle p_1, p_n\rangle & \cdots & \langle p_n, p_n\rangle\end{pmatrix}}_{R}\underbrace{\begin{pmatrix}c_1\\ \vdots\\ c_n\end{pmatrix}}_{\mathbf{c}} = \underbrace{\begin{pmatrix}\langle x, p_1\rangle\\ \vdots\\ \langle x, p_n\rangle\end{pmatrix}}_{\mathbf{p}}$$

R is called the Grammian of the set $\{p_1, \ldots, p_n\}$.

The Grammian of a set

Definition (Grammian)

Given a set $\{p_1, \ldots, p_n\}$ of vectors in $\mathbb{S}$, the Grammian of the set is the matrix

$$R = \begin{pmatrix} \langle p_1, p_1 \rangle & \cdots & \langle p_n, p_1 \rangle \\ \vdots & & \vdots \\ \langle p_1, p_n \rangle & \cdots & \langle p_n, p_n \rangle \end{pmatrix}$$

Note that $R^H = R$

We also have the following theorem:

Theorem (Moon, Theorem 3.1)

The Grammian R is positive definite iff the set of vectors $\{p_1, \ldots p_n\}$ are linearly independent.

Proof

Let $y \in \mathbb{S}$ then

$$y^H R y = (\bar{y}_1 \cdots \bar{y}_n) \begin{pmatrix} \langle p_1, p_1 \rangle & \dots & \langle p_n, p_1 \rangle \\ \vdots & & \vdots \\ \langle p_1, p_n \rangle & \dots & \langle p_n, p_n \rangle \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$= \left(\sum_{i=1}^{n} \bar{y}_i \langle p_1, p_i \rangle \dots \bar{y}_i \sum_{i=1}^{n} \langle p_n, p_i \rangle \right) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$= \sum_{j=1}^{n} \sum_{i=1}^{n} \bar{y}_i y_j \langle p_j, p_i \rangle$$

$$= \left\langle \sum y_j p_j, \sum y_i, p_i \right\rangle = \| \sum y_i p_i \|^2 \geq 0$$

Therefore R is always positive semi-definite.

Proof, cont.

($\Rightarrow$): Suppose that R is pd then

$$y^H R y = \left\| \sum y_i p_i \right\|^2 > 0$$
$$\Rightarrow \sum y_i p_i \neq 0 \text{ for all nonzero } y \in \mathbb{S}$$
$$\Rightarrow \{p_1, \cdots p_n\} \text{ is linearly independent}$$

($\Leftarrow$): Conversely suppose $\{p_1, \cdots p_n\}$ is linearly independent, but R is only psd. R is psd implies that $\exists y \neq 0$ such that

$$y^H R y = \left\| \sum y_i p_i \right\|^2 = 0$$
$$\Rightarrow \sum y_i p_i = 0$$
$$\Rightarrow \{p_1, \cdots p_n\} \text{is linearly dependent.}$$

Which contradicts the assumption that R is psd.

Orthogonality Theorem

Theorem (Moon, Theorem 3.2)

Let $p_1, p_2, \cdots p_n$ be data vectors (or basis vectors) in a Hilbert space $\mathbb{S}$. Let $x \in \mathbb{S}$. Let e be defined as

$$e \stackrel{\triangle}{=} x - \hat{x} = x - \sum_{j=1}^{n} c_j p_j,$$

then e is minimized when it is orthogonal to each of the data vectors, i.e.

$$\langle e, p_j \rangle = 0 \qquad j = 1, \ldots, n$$

Equivalently

$$R\mathbf{c} = \mathbf{p}.$$

Proof.

Follows directly from projection theorem. $\square$

Calculus-Based Approach (Alternative proof)

Rather than using the projection theorem, we can derive the same result using calculus.

Problem Statement: Let $\mathbf{e} = x - \sum_{i=1}^{n} c_i p_i$. Find $\mathbf{c} = (c_1, \ldots, c_n)^\top$ that minimizes $\|\mathbf{e}\|$.

Solution: First note that minimizing $\|\mathbf{e}\|^2$ is equivalent to minimizing $\|e\|$. Also note that

$$\begin{aligned}
\|e\|^2 &= \left\langle x - \sum c_j p_j, x - \sum c_j p_j \right\rangle \\
&= \|x\|^2 - 2Re\{\sum_{i=1}^{n} \bar{c}_i \langle x, p_i \rangle\} + \sum \sum c_j \bar{c}_i \langle p_j, p_i \rangle \\
&= \|x\|^2 - 2Re\{\mathbf{c}^H \mathbf{p}\} + \mathbf{c}^H R \mathbf{c}.
\end{aligned}$$

Calculus-Based Approach, cont.

To minimize

$$\|e\|^2 = \|x\|^2 - 2Re\{\mathbf{c}^H\mathbf{p}\} + \mathbf{c}^H R\mathbf{c}$$

differentiate with respect to **c** and set to zero. This will be a local minima if the second derivative is psd.

Calculus-Based Approach, cont.

From Moon Appendix we have

$$\frac{\partial}{\partial \bar{\mathbf{c}}} Re\{\mathbf{c}^H \mathbf{p}\} = \frac{1}{2}\mathbf{p}$$
$$\frac{\partial}{\partial \bar{\mathbf{c}}} \mathbf{c}^H R \mathbf{c} = R\mathbf{c}$$

Therefore

$$\frac{\partial \|e\|^2}{\partial \bar{\mathbf{c}}} = -\mathbf{p} + R\mathbf{c} = 0 \qquad \Rightarrow \qquad R\mathbf{c} = \mathbf{p}$$

In addition, we have that

$$\frac{\partial^2 \|e\|^2}{\partial \bar{\mathbf{c}}} = R \geq 0.$$

Therefore the solution of $R\mathbf{c} = \mathbf{p}$ minimize $\|e\|$.
$R\mathbf{c} = \mathbf{p}$ is the same equation we obtained using the projection theorem.

Matrix Representation

- Stack the vectors $\{p_1, \ldots p_n\}$ in a matrix

$$A = \begin{pmatrix} p_1 & p_2 & \ldots & p_n \end{pmatrix}$$
$$\mathbf{c} = \begin{pmatrix} c_1 & c_2 & \ldots & c_n \end{pmatrix}^\top$$

- Then $\hat{x} = \sum c_j p_j = A\mathbf{c}$.
- Therefore $\mathbf{e} = x - \hat{x} = x - A\mathbf{c}$.

Matrix Representation, cont.

- Project **e** onto $\{p_1 \dots p_n\}$:

$$\langle x - A\mathbf{c}, p_1 \rangle = p_1^H (x - A\mathbf{c}) = 0$$
$$\vdots$$
$$\langle x - A\mathbf{c}, p_n \rangle = p_n^H (x - A\mathbf{c}) = 0$$

- Note that $A^H = \begin{bmatrix} p_1^H \\ \vdots \\ p_n^H \end{bmatrix}$.
- Rewrite as

$$A^H (x - A\mathbf{c}) = 0$$
$$\Rightarrow \underbrace{A^H A}_{R} \mathbf{c} = \underbrace{A^H x}_{\mathbf{p}}$$

Matrix Representation, cont.

- If $\{p_1, \ldots p_n\}$ are linearly independent then $R > 0$ which implies that R^{-1} exists, so

$$\mathbf{c} = (A^H A)^{-1} A^H x$$

- Since $\hat{x} = A\mathbf{c}$ we have that

$$\hat{x} = A(A^H A)^{-1} A^H x$$

  is the best approximation to x in span$\{p_1, \ldots, p_n\}$.
- **Fact:** $P_A = A(A^H A)^{-1} A^H$ is a projection operator from S to $span\{p_1, \ldots, p_n\}$

Application:Polynomial Approximation

- Suppose you are given a real continuous function $f(t)$ and you would like to approximate it by an m^{th} order polynomial on the interval $[a, b]$.
- Let the basis vectors be $\{1, t, \dots, t^m\}$.
- Then $\hat{f}(t) = c_1 + c_2 t + \cdots + c_{m+1} t^m$
- Define the inner product as $\langle f, g \rangle = \int_a^b f(t)g(t)dt$

Application: Polynomial Approximation, cont.

Then the orthogonality theorem implies that the "best" approximation is given by

$$\left\langle f - \hat{f}, 1 \right\rangle = 0$$

$$\vdots$$

$$\left\langle f - \hat{f}, t^m \right\rangle = 0$$

or

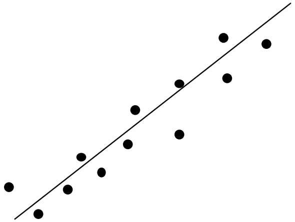
$$\underbrace{\begin{pmatrix} \langle 1, 1 \rangle & \cdots & \langle t^m, 1 \rangle \\ \vdots & & \vdots \\ \langle 1, t^m \rangle & \cdots & \langle t^m, t^m \rangle \end{pmatrix}}_{\text{Grammian Matrix}} \begin{pmatrix} c_1 \\ \vdots \\ c_{m+1} \end{pmatrix} = \begin{pmatrix} \langle f, 1 \rangle \\ \vdots \\ \langle f, t^m \rangle \end{pmatrix}$$

or

$$R\mathbf{c} = \mathbf{p}.$$

Application: Linear Regression

- Suppose you have a number of data points that you are trying to fit to a line.



- Given $(x_i, y_i) \qquad i = 1, \ldots N$
- The equation for a line is $y = ax + b$
- **Problem:** Find a and b that minimizes the mean squared error $\sum_{i=1}^{N} |y_i - ax_i - b|^2$

Application: Linear Regression, cont.

- For each data point we have

$$e_i = y_i - ax_i - b$$

where e_i is the error for the i^{th} data point.

- Let

$$\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix}, \quad \mathbf{e} = \begin{pmatrix} e_1 \\ \vdots \\ e_N \end{pmatrix}, \quad A = \begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_N & 1 \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} a \\ b \end{pmatrix}$$

- Then $\mathbf{e} = \mathbf{y} - A\mathbf{c}$.

Application: Linear Regression, cont.

- Project the error **e** on the data vector (columns of A) and set to zero:
$$A^H\mathbf{e} = A^H(\mathbf{y} - A\mathbf{c}) = 0$$
- Therefore
$$A^H A\mathbf{c} = A^H\mathbf{y}$$
- Giving the minimum least squares solution
$$\mathbf{c} = (A^H A)^{-1}A^H\mathbf{y}.$$

Section 2

Dual Approximation

Dual Approximation

This section develops an approach that allows approximation in infinite dimensional spaces with finite constraints.

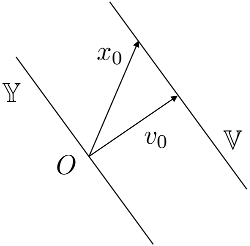
For matrices, we will solve the problem

$$\begin{aligned} \min \quad & \|x\| \\ \text{s.t.} \quad & Ax = b \end{aligned}$$

Dual Approximation, cont.

Definition (Affine Space)

Let $\mathbb{Y}$ be a subspace of $\mathbb{S}$ and let $x_o \in \mathbb{S}$. The set $\mathbb{V} = x_0 + \mathbb{Y}$ is called a <u>linear variety</u> or an <u>affine</u> space.



The projection theorem says that there exists a $v_0 \in \mathbb{V}$ such that $v_0 = \arg\min_{v \in \mathbb{V}} \|v\|$ such that $v_0 \perp \mathbb{Y}$.

Dual Approximation, cont.

Let $M = span\{y_1, \ldots, y_m\}$ then $\dim(M) < \infty$.

If $\dim(\mathbb{S}) = \infty$ then $\dim(M^\perp) = \infty$ where $M^\perp$ is the set of all $x \in \mathbb{S}$ such that

$$\langle x, y_1 \rangle = 0$$
$$\vdots$$
$$\langle x, y_m \rangle = 0$$

Dual Approximation, cont.

Now suppose that there are m inner product constraints:

$$\langle x, y_1 \rangle = a_1$$

$$\vdots$$

$$\langle x, y_m \rangle = a_n$$

If $\exists x_0$ that satisfies the constraints then so does $x_0 + v$ where $v \in M^{\perp}$ since

$$\begin{aligned}\langle x_0 + v, y_j \rangle &= \langle x_0, y_j \rangle + \langle v, y_j \rangle \\ &= \langle x_0, y_j \rangle \\ &= a_j\end{aligned}$$

Therefore all solutions are in the (infinite dimensional) affine space

$$v = x_0 + M^{\perp}$$

Dual Approximation, cont.

Theorem (Moon Theorem 3.4)

Let $\{y_1, \cdots, y_m\}$ be linearly independent in a Hilbert space $\mathbb{S}$, and let $M = span\{y_1, \cdots, y_m\}$. The solution of the problem

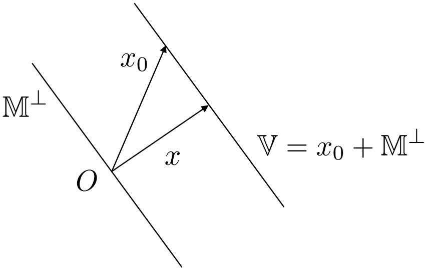
$$\begin{aligned} \min_{x \in \mathbb{S}} \quad & \|x\|^2 \\ \text{s.t.} \quad & \langle x, y_1 \rangle = \alpha_1, \\ & \vdots, \\ & \langle x, y_m \rangle = \alpha_m \end{aligned}$$

is an element of M, i.e., $\hat{x} = \arg\min_{x \in \mathbb{S}} \|x\|^2 = \sum_{i=1}^m c_i y_i$, where $\mathbf{c}$ satisfies $R\mathbf{c} = \boldsymbol{\alpha}$, where R is the Grammian and

$$\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_m)^\top.$$

Proof:

From the previous discussion, the solution lies in the affine space $\mathbb{V} = x_0 + M^\perp$ for some $x_0 \in \mathbb{S}$.



The minimum norm solution is orthogonal to $M^\perp$ i.e. $\hat{x} \perp M^\perp \Rightarrow \hat{x} \in M^{\perp\perp} = M$

So $\hat{x}$ is of the form $\hat{x} = \sum_{j=1}^{m} c_j y_j$

Proof, cont.

Now projecting x onto M gives

$$\begin{aligned}
\langle \hat{x}, y_1 \rangle = \left\langle \sum c_j y_j, y_1 \right\rangle &= \sum c_j \langle y_j, y_1 \rangle = \alpha_1 \\
\vdots \quad = \quad \vdots \quad &= \quad \vdots \quad = \vdots \\
\langle \hat{x}, y_m \rangle = \left\langle \sum c_j y_j, y_m \right\rangle &= \sum c_j \langle y_j, y_m \rangle = \alpha_m
\end{aligned}$$

rewriting in matrix notation gives

$$R\mathbf{c} = \boldsymbol{\alpha}$$

Dual Approximation, Example

Given the differential equation

$$\ddot{y} + 6\dot{y} + 8y = 4\dot{u} + 10u, \qquad y(0) = \dot{y}(0) = 0$$

Solve the following optimal control problem:

$$\begin{aligned} \min_{u \in L_2} \quad & \|u\|^2 \\ \text{s.t.} \quad & y(1) = 1, \\ & \int_0^1 y(t)dt = 0 \end{aligned}$$

Dual Approximation, Example, cont.

The corresponding transfer function is

$$H(s) = \frac{4s+10}{s^2+6s+8} = \frac{1}{s+2} + \frac{3}{s+4}$$
$$\Rightarrow h(t) = e^{-2t} + 3e^{-4t}$$
$$\Rightarrow y(t) = \int_0^t \left[e^{-2(t-\tau)} + 3e^{-4(t-\tau)} \right] u(\tau) d\tau$$

Define the following inner product

$$\langle f(t), g(t) \rangle = \int_0^1 f(\tau) g(\tau) d\tau$$

then $y(1) = 1$ can be written as

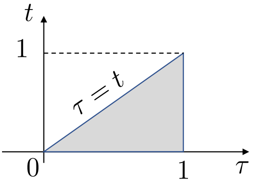
$$\int_0^1 \left[e^{-2(1-\tau)} + 3e^{-4(1-\tau)} \right] u(\tau) d\tau = \langle u, y_1 \rangle = 1$$

where $y_1(t) = e^{-1(1-t)} + 3e^{-4(1-t)}$

Dual Approximation, Example, cont.

The second constraint is of the form

$$\int_0^1 y(t)dt = \int_{t=0}^{t=1} \int_{\tau=0}^{\tau=t} h(t-\tau)u(\tau)d\tau dt = 0$$



Changing order of integration gives

$$= \int_{\tau=0}^{1} \left[\int_{t=\tau}^{1} h(t-\tau)dt \right] u(\tau)d\tau.$$

Dual Approximation, Example, cont.

Letting $\sigma = t - \tau \Rightarrow t = \sigma + \tau \Rightarrow dt = d\sigma$ gives

$$= \int_{\tau=0}^{1} \left[\int_{\sigma=0}^{\sigma=1-\tau} h(\sigma) d\sigma \right] u(\tau) d\tau$$

$$= \int_{\tau=0}^{1} \left(\frac{5}{4} - \frac{3}{4} e^{-4(1-\tau)} - \frac{1}{2} e^{-2(1-\tau)} \right) u(\tau) d\tau$$

$$= \langle u, y_2 \rangle = 0$$

where

$$y_2(t) = \frac{5}{4} - \frac{3}{4} e^{-4(1-\tau)} - \frac{1}{2} e^{-2(1-\tau)}$$

so we have that

$$\langle u, y_1 \rangle = 1$$
$$\langle u, y_2 \rangle = 0$$

and we want to minimize $\|u\|_{L_2[0,1]}^2$

Dual Approximation, Example, cont.

Let $M = span\{y_1, y_2\}$.
By Theorem 3.4

$$u \in M \Rightarrow u(t) = c_1 y_1(t) + c_2 y_2(t)$$

where

$$\begin{pmatrix} \langle y_1, y_1 \rangle & \langle y_2, y_1 \rangle \\ \langle y_1, y_2 \rangle & \langle y_2, y_2 \rangle \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Section 3

Underdetermined Problems

Section 3.15: Underdetermined Problems

Given $Ax = b$ where A is fat, i.e. fewer equations than unknowns, solve the following problem:

$$\begin{aligned} \min \quad & \|x\|_2 \\ \text{s.t.} \quad & Ax = b \end{aligned}$$

where $A = \begin{pmatrix} y_1^H \\ \vdots \\ y_m^H \end{pmatrix}$, $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$,
and $y_i \in \mathbb{C}^n$ and $b \in \mathbb{C}^m$.

Section 3.15: Underdetermined Problems, cont.

$Ax = b$ is a set of inner product constraints

$$y_1^H x = b_1$$

$$\vdots$$

$$y_m^H x = b_m$$

Let $M = span\{y_1, \cdots, y_m\}$.
Theorem 3.4 implies that $x_0 = \arg\min \|x\| \in M$

$$\Rightarrow x_0 = \sum c_j y_j = A^H c$$

and that c satisfies

$$R\mathbf{c} = \mathbf{b} \text{ where } R = AA^H$$

if $\{y_1, \cdots, y_m\}$ are linearly independent then

$$\mathbf{c} = (AA^H)^{-1}\mathbf{b} \qquad \Rightarrow \qquad x_0 = \underbrace{A^H(AA^H)^{-1}}_{\text{pseudo-inverse}}\mathbf{b}$$

Section 4

Generalized Fourier Series

Section 3.17: Generalized Fourier Series

Topic of interest: L_2 function approximation

Definition (Complete Basis)

An orthonormal set $\{p_i, i = 1, \ldots, \infty\}$ in a Hilbert space $\mathbb{S}$ is a complete basis or total basis if $\forall x \in \mathbb{S}$

$$x = \sum_{i=1}^{\infty} \langle x, p_i \rangle \, p_i$$

Note that if $x = \sum_{i=1}^{\infty} c_i p_i$ and $\langle p_i, p_j \rangle = \delta_{ij}$ then

$$\langle x, p_j \rangle = \sum_{i=1}^{\infty} c_i \langle p_i, p_j \rangle = c_j$$

$$\Rightarrow c_j = \langle x, p_j \rangle$$

Generalized Fourier Series, cont.

Therefore we can write

$$x = \sum_{i=1}^{\infty} \langle x, p_i \rangle \, p_i.$$

Most common example: standard Fourier basis

$$P_n(t) = \frac{1}{\sqrt{T}} e^{j\left(\frac{2\pi}{T}\right)nt}$$

Any function $f \in L_2[0, T]$ can be written as

$$f(t) = \sum_{n=-\infty}^{\infty} c_n \frac{1}{T} e^{j\left(\frac{2\pi}{T}\right)nt}$$

where the coefficients are given as

$$c_n = \left\langle f, \frac{1}{\sqrt{T}} e^{j\left(\frac{2\pi}{T}\right)nt} \right\rangle \triangleq \frac{1}{\sqrt{T}} \int_0^T f(t) e^{j\left(\frac{2\pi}{T}\right)nt} dt$$

Generalized Fourier Series, cont.

Actually it is common to place the $\frac{1}{\sqrt{T}}$'s together letting $f(t) = \sum_{n=-\infty}^{\infty} b_n e^{j\left(\frac{2\pi}{T}\right)nt}$ where

$$b_n = \left\langle f(t), \frac{1}{T} e^{j\left(\frac{2\pi}{T}\right)nt} \right\rangle = \frac{1}{T} \int_0^T f(t) e^{-j\left(\frac{2\pi}{T}\right)nt} dt$$

Generalized Fourier series hold for <u>any</u> complete basis, i.e.

$$x = \sum_{j=1}^{\infty} \langle x, p_j \rangle \, p_j$$

Generalized Fourier Series, cont.

There are two important relationship between a function and its Fourier transform.

Theorem (Bessel's Inequality)

Suppose $\{p_1, p_2, \ldots\}$ *is orthonormal but not necessarily complete and let*

$$c = \{\langle x, p_1\rangle, \langle x, p_2\rangle, \ldots\} = \{c_1, c_2, \ldots\}$$

then

$$\boxed{\|c\|_{\ell_2} \leq \|x\|_{L_2}}$$

Proof:

$$
\begin{aligned}
0 \le \left\| x - \sum c_j p_j \right\|_{L_2}^2 &= \left\langle x - \sum c_j p_j, x - \sum c_j p_j \right\rangle_{L_2} \\
&= \langle x, x \rangle_{L_2} - \sum \bar{c}_j \langle x, p_j \rangle_{L_2} \\
&- \sum c_j \overline{\langle x, p_j \rangle}_{L_2} + \sum \sum c_j \bar{c}_k \langle p_j, p_k \rangle_{L_2} \\
&= \|x\|_{L_2}^2 - \sum \bar{c}_j c_j - \sum c_j \bar{c}_j + \sum c_j \bar{c}_j \\
&= \|x\|_{L_2}^2 - \sum_{j=1}^{\infty} |c_j|^2 \\
&= \|x\|_{L_2}^2 - \|c\|_{\ell_2}^2 \\
&\Rightarrow \|c\|_{\ell_2}^2 \le \|x\|_{L_2}^2
\end{aligned}
$$

Generalized Fourier Series, cont.

Theorem (Parseval's Equality)

If $T = \{p_1, p_2, \ldots\}$ *is complete then*

$$\boxed{\|x\|_{L_2}^2 = \|c\|_{\ell_2}^2}$$

Proof.

If T is complete then

$$\left\| x - \sum c_j p_j \right\|^2 = 0$$

and the result follows from the proof of Bessel's inequality . $\square$

Significance of Parseval's Equality

$\|x\|_{L_2}^2 = \|c\|_{\ell_2}^2$ says that the energy in a signal (i.e. $\|x\|_{L_2}$) is equal to the energy in the Fourier coefficients (i.e. $\|c\|_{\ell_2}^2$).

This relationship between x and its transform c is written as

$$x \overset{\mathcal{F}}{\longleftrightarrow} c.$$

Significance of Parseval's Equality, cont.

Lemma (Moon Lemma 3.1)

If $x \overset{\mathcal{F}}{\longleftrightarrow} c$ *and* $y \overset{\mathcal{F}}{\longleftrightarrow} b$ *for the same complete basis* $\{p_1, p_2, \ldots\}$ *then*

$$\langle x, y \rangle_{L_2} = \langle c, b \rangle_{\ell_2} .$$

Proof.

Let $x = \sum_{i=1}^{\infty} c_i p_i$, and $y = \sum_{i=1}^{\infty} b_i p_i$ then

$$\begin{aligned}
\langle x, y \rangle_{L_2} &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} c_i \bar{b}_j \langle p_i, p_j \rangle \\
&= \sum_{i=1}^{\infty} c_i \bar{b}_i \\
&= \langle c, b \rangle_{\ell_2}
\end{aligned}$$

$\square$