Review of Correlation Theory

A random process can be thought of as a function of two variables

\[ X(t, \omega) \]

\( t = \text{time, with } -\infty < t < \infty \) \( \omega \in \Omega \) = the sample space

we may define a distribution function for the process as

\[ P_t \left[ \{ \omega : X(t, \omega) \leq x \} \right] = P_x(x; t) \]

and the density as

\[ P_x(x; t) = \frac{d}{dx} \left[ P_x(x; t) \right] \]

These are the first order statistics of the process. The \( n \)th order statistics are defined by.

\[ P_x(x_1, \ldots, x_n) = P_t \left[ \{ \omega : X(t, \omega) \leq x_1, \ldots, X(t, \omega) \leq x_n \} \right] \]

and

\[ P_x(x_1, \ldots, x_n) = \frac{d^n}{dx_1 \cdots dx_n} \left[ P_x(x_1, \ldots, x_n) \right] \]

A process is strict sense stationary (SSS)

\[ P_x(x; t) = P_x(x; t + \delta), \forall t, x, n \text{ and } \delta \in \mathbb{C}, (\mathbb{C} = \{\epsilon, e, \ldots, \epsilon\}) \]
For \( n=1 \) this means
\[ P_x(x_j \pm t) = P_x(x_j \pm t + \epsilon) \]

Thus
\[ P_x(x_j \pm t) = P_x(x) \quad \text{(independent of time)} \]

\( n=2 \)
\[ P_x(x_j, x_{2j} \pm t_1, t_2) = P_x(x_1, x_2, j \pm t_1, t_2 + \epsilon) \]

Letting \( \epsilon = -t_2 \) yields
\[ P_x(x_j, x_{2j} \pm t_1, t_2) = P_x(x_j, x_{2j} \pm t_1, -t_2) \]

**Moments:**

a) **First order moments**: The first order moments require only the first order statistics.

**Mean:**
\[ m_x(\pm t) = E[x(\pm t)] = \int_{-\infty}^{\infty} x P_x(x \pm t) \, dx \]

**Variance:**
\[ \sigma_x^2(\pm t) = E[(x(\pm t) - m_x(\pm t))^2] = E[x^2(\pm t)] - E^2[x(\pm t)] \]
\[ = \int_{-\infty}^{\infty} x^2 P_x(x \pm t) \, dx - m_x^2(\pm t) \]

All other moments may be defined similarly (with, of course, requirements that they exist).
6) **Second order moments**: Now calculate the moments of the two random variables $X(t_1)$ and $X(t_2)$.

**Correlation function** (ACF):

$$R_X(t_1,t_2) = E[(X(t_1) - m_X(t_1))(X(t_2) - m_X(t_2))]$$

(complex conj. required for complex processes)

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 P_X(x_1, x_2; t_1, t_2) \, dx_1 \, dx_2$$

**Covariance function**:

$$k_X(t_1, t_2) = E[(X(t_1) - m_X(t_1))(X^*(t_2) - m_X^*(t_2))]$$

$$= R_X(t_1, t_2) - m_X(t_1) m_X^*(t_2)$$

**Properties of correlation functions**

a) We define a process for which

a) $m_X(t) = m_X$

b) $R_X(t_1, t_2) = R_X(t_2, t_1)$

as **wide sense stationary** or **correlation stationary**.

Clearly, $X(t)$ is WSS if $X(t,t)$ is SSS.
0) Symmetry:
\[ R_x(z_1, z_2) = R_x^*(z_2, z_1) \]

and
\[ K_x(z_1, z_2) = K_x^*(z_2, z_1) \]

If \( x(t) \) is real then the complex conjugate sign is unnecessary. If \( x(t) \) is stationary then
\[ R_x(z) = R_x^*(-z), \quad K_x(z) = K_x^*(-z). \]

e) Positive Semi-definite (non-negative definite)
Let \( f(t) \) be any function with bounded energy \( (\int_0^T |f(t)|^2 dt < \infty) \) then we define the random variable \( x_f \) as
\[ x_f = \int_0^T x(t) f(t) \, dt \]

\[ m_{x_f} = E[x_f] = E \int_0^T x(t) f(t) \, dt = \int_0^T m_x(t) f(t) \, dt \]

\[ \Sigma_{x_f}^2 = E\left[ (x_f - E[x_f])^2 \right] \]

\[ = E \left[ \int_0^T \left[ x(t) - m_x(t) \right] f(t) \, dt \right] \]

\[ = \int_0^T \int_0^T K_x(\tau, \tau') f(\tau) f^*(\tau') \, d\tau \, d\tau' \]

Since \( \Sigma_{x_f}^2 \geq 0 \) we obtain that
\[ \sum_{a} \sum_{b} \int \mathbb{R} \, k_x(a, b) f(u) \, \delta(u) \, du \geq 0 \quad \text{if } f \in L^2(\mathbb{R}) \]

Functions \( k_x(a, b) \) which satisfy such a property are called positive semidefinite. If the inequality is strict then it is positive definite.

Consider the function
\[
E[|x(u) + x(v)|^2] = E[(x(u) + x(v))(x^*(u) + x^*(v))] \geq 0
\]

\[
= R_x(u, v) + R_x(u, u) + R_x(v, v) + R_x(u, v)
\]

\[
= R_x(u, v) + R_x(u, u) + 2 \Re \sum R_x(u, v) \geq 0.
\]

If \( x \) is real and stationary then
\[
R_x(0) = |R_x(\pm 2)|, \quad \text{all } \pm 2
\]

and
\[
k_x(0) = |k_x(\pm 2)|, \quad \text{all } \pm 2
\]

Finally consider
\[
E\left[\left|\sum_{i=1}^{N} a_i x(t_i + u)\right|^2\right] \geq 0
\]

Thus we have
\[
\sum_{i=1}^{N} \sum_{j=1}^{N} a_i a_j^* R_x(t_i + t_j) \geq 0, \quad \text{all } t_i, t_j
\]

In the stationary case \( R_x(t_i + t_j) = R_x(t_i - t_j) \). This equation can be shown to be equivalent to the equation at the top of this page. \( \Delta^T R \Delta > 0 \).
1) Cross-correlation (covariance) functions:

Sometimes we must consider the joint statistic of two random processes $x(t)$ and $y(t)$. For example, the input and output of a system.

1) Cross-correlation function (CCF)

$$
\begin{align*}
R_{xy}(t, u) &= E[x(t) y^*(u)] \\
R_{yx}(t, u) &= E[y(t) x^*(u)]
\end{align*}
$$

Thus we may show that

$$
R_{xy}(t, u) = R_{yx}^*(u-t)
$$

If $x(t)$ and $y(t)$ are jointly stationary, then

$$
R_{xy}(t) = R_{yx}^*(-t)
$$

2) Cross-covariance function

$$
K_{xy}(t, u) = E[(x(t) - m_x(t))(y^*(u) - m_y(u))]
$$

$$
= K_{yx}^*(u-t)
$$

For the stationary case

$$
K_{uv}(t) = R_{uv}(t) - m_x m_y
$$
Considering
\[ E \left[ \left| x(t) + y(u) \right|^2 \right] \geq 0 \]
yield that
\[ R_x(t, t) + R_y(u, u) \geq 2 \Re \{ R_{xy}(\tau, \kappa) \} \]
If \( x(t) \) and \( y(u) \) is jointly stationary, then
\[ R_x(0) R_y(0) \geq |R_{xy}(\tau)|^2 \]

**Time-averages; Ergodicity**

Consider the random variable for a stationary \( x(t) \)
\[ X_T = \frac{1}{2T} \int_{-T}^{T} x(t) \, dt \]
we obtain
\[ E [X_T] = E \left[ \frac{1}{2T} \int_{-T}^{T} x(t) \, dt \right] = \frac{1}{2T} \int_{-T}^{T} M_x \, dt = M_x \]
The process is **Ergodic in the mean** if
\[ \lim_{T \to \infty} P \left[ \left| X_T - M_x \right| > \epsilon \right] = 0, \quad \text{all } \epsilon > 0 \]

Thus the law of large numbers holds for the mean.
We indicate this by
\[ X_T \xrightarrow{P} M_x \]
In order to prove this result it is sufficient to show that \( \frac{T^{2}}{\varepsilon^2} \to 0 \) and use the Chebyshev inequality. Thus we use the second moments to prove first moment ergodicity. (This is sufficient but not necessary.)

In a similar manner we may define

\[
\hat{R}_x(t) = \frac{1}{2T} \int_{-T}^{T} x(t) x(t + 2) dt
\]

and show that

\[
E[\hat{R}_x(t)] = R_x(t)
\]

and \( \hat{R}_x(t) \to R_x(t) \) as \( T \to \infty \), which is ergodicity of the correlation function.

* Since ergodicity is generally difficult to prove we usually assume general ergodicity unless we reason not to. It should be clear that the process must be at least stationary to be ergodic.

* The basic assumption that usually allows us to do this is the random phase argument. This means we do not know when the process starts and all possible shifts in time are ensemble members.

**Example:**

\[ x(t) = \cos(\omega_0 t + \Theta) \quad \Theta \text{ uniform on } (0, \pi) \]

\[ y(t) = \cos(\omega_0 t + \Theta) \quad \Theta \text{ uniform on } (0, \pi) \quad \text{ erg.} \]
Linear Systems:

A linear system may be characterized by its impulse response

\[ h(t,u) \]

If the system is time-invariant (stationary) then

\[ h(t,u) = h(t-u) = h(u), \quad u = t-u \]

Suppose we have an input process \( x(t) \) and the output is \( y(t) \) another process.

\[ y(t) = \int_{-\infty}^{\infty} h(t,u) x(u) du \]

Using this formula it is not hard to show

\[ R_y(t,u) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_x(t - z_1, u - z_2) h(z_1) h(z_2) d z_1 d z_2 \]

if the system is time-invariant, with equivalent results for \( R_y(t, u) \).

Should \( x(t) \) be stationary then so is \( y(t) \) and has ACF

\[ R_y(z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(z_1, z_2) R_x(z - z_1 + z_2) d z_1 d z_2 \]
In the general case when nothing is either stationary or real, then

\[ R_y(z, u) = \frac{1}{2\pi} \int_{-\infty}^{\infty} R_x(z_1, z_2) h^*(u, z_2) \, dz_1 \, dz_2 \]

**Power Spectral Density (PSD):**

We define the PSD of a stationary process \( x(t) \) as

\[
\begin{align*}
S_x(\omega) &= \int_{-\infty}^{\infty} R_x(z) e^{-j\omega z} \, dz \\
R_x(z) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} S_x(\omega) e^{j\omega z} \, d\omega
\end{align*}
\]

Equivalent definitions exist for \( S_{xy}(\omega) = \mathbb{F} \left[ R_{xy}(z) \right](\omega) \).

**Properties of PSD:**

a) \( S_x(\omega) \geq 0 \)

b) Given a stable, linear time invariant system, \( h(t) \), we have

\[ S_y(\omega) = |H(j\omega)|^2 S_x(\omega) \]

c) \( S_{yx}(\omega) = S_{xy}^*(\omega) \)

d) \( E\left[ x^2(z) \right] = R_x(0) = \int_{-\infty}^{\infty} S_x(\omega) \, d\omega \)
Expansions and Representations of Signals:

Here we shall discuss the general expansion of waveforms in terms of orthonormal basis functions. For the purpose of this course we shall discuss signals defined on a bounded interval, say \([0, T]\), but all results are true for the infinite interval.

1) A set of functions, \(\{ f_i(t) : i = 1, \ldots, N \}\) is linearly dependent if there exists a set of numbers \(\{ c_i : i = 1, \ldots, N \}\) such that

\[
\sum_{i=1}^{N} c_i f_i(t) = 0, \quad 0 \leq t \leq T
\]

and not all \(c_i\)'s zero. If no such constants exist then the set is called linearly independent. (L.I.).

Thus if \(\{ f_i(t) \}\) is L.I. then none of the signals can be written as a linear combination of the other.

**Examples**

\(\{ \sqrt{t}, \sqrt{t}, x^2 \}\) is L.I. but \(\{ \sqrt{t}, \sqrt{t}, x^2, 3 + 4t + 2x^2 \}\) is L.D. since

\[
\sqrt{t}(t) = 3 + 4t + 2t^2 = \sum_{i=1}^{3} c_i f_i(t)
\]

where \(c_1 = 3, c_2 = 4, c_3 = 2\).
2) The inner product (scalar product) of two functions is defined as

\[(f, g) = \int_{0}^{T} f(t)g^*(t) dt\]

3) The \((L^2)\) norm of a function, \(f(t)\), is

\[\|f\| = (f, f)^{1/2} = \left[\int_{0}^{T} |f(t)|^2 dt\right]^{1/2}\]

*we shall discuss only those functions with bounded norm (energy).* Thus

\[E_f = \|f\| < \infty\]

Such functions are said to be in \(L^2(0,T)\). Symbolically,

\[f \in L^2(0,T)\]

4) Two functions in \(L^2(0,T)\) are orthogonal if

\[(f, g) = 0\]

5) A function is normalized if

\[\|f\| = 1\]

6) Any two functions satisfying 4) and 5) are orthonormal
2) Any \( L^2 \) set of functions in \( L^2(\mathbb{R}^n) \) may be orthonormalized. One way of accomplishing this is the Gram-Schmidt procedure. Given a \( L^2 \) set \( \{ f_2(t) \}^\infty \), we shall form an \( L^2 \) set \( \{ \phi_n(t) \}^\infty \), by

\[
\begin{align*}
\phi_1(t) &= \frac{f_1(t)}{\| f_1(t) \|}, \\
\phi_2(t) &= \frac{f_2(t) - (f_2, \phi_1) \phi_1(t)}{\| f_2 - (f_2, \phi_1) \phi_1 \|}, \\
&\quad \vdots \\
\phi_n(t) &= \frac{f_n(t) - \sum_{i=1}^{n-1} (f_n, \phi_i) \phi_i(t)}{\| f_n - \sum_{i=1}^{n-1} (f_n, \phi_i) \phi_i \|}.
\end{align*}
\]

**Example:** \( \{ 1, \phi_2 \} \), \( f_1(t) = 1 \), \( f_2(t) = \chi \)

\[
\phi_1(t) = \frac{1}{\| f_1 \|} = \frac{1}{\left[ \int_{-\infty}^{\infty} f_1^2(t) \, dt \right]^{1/2}} = \frac{1}{\sqrt{\pi}} = \phi_1(t)
\]

\[
\begin{align*}
(f_2, \phi_1) \phi_1(t) &= \chi - \left( \frac{1}{\sqrt{\pi}} \right) \frac{1}{\sqrt{\pi}} \\
&= \chi - \frac{1}{\pi} \left( \frac{1}{\sqrt{\pi}} \right) \frac{1}{\sqrt{\pi}} \\
&= \chi - \frac{1}{\pi} \left( \frac{1}{\sqrt{\pi}} \right) \frac{1}{\sqrt{\pi}} \\
&= \chi - \frac{1}{\pi} \left( \frac{1}{\sqrt{\pi}} \right) \frac{1}{\sqrt{\pi}} \\
\| f_2 - (f_2, \phi_1) \phi_1 \| &= \left[ \int_{-\infty}^{\infty} \left( f_2 - \frac{1}{\| f_1 \|} \sum_{i=1}^{n-1} (f_n, \phi_i) \phi_i(t) \right)^2 \, dt \right]^{1/2} = \frac{\sqrt{\pi}}{\sqrt{\pi}} \\
\phi_2(t) &= \frac{f_2(t) - \sum_{i=1}^{n-1} (f_n, \phi_i) \phi_i(t)}{\| f_2 - \sum_{i=1}^{n-1} (f_n, \phi_i) \phi_i \|}.
\end{align*}
\]

\*
show \( \pm \quad (\phi_n^*, \phi_n) = \frac{1}{\pi} (\pm - \frac{1}{\pi}) \)

\*
8. A set of functions in $L^2([0,\pi])$ is LI if and only if its Grammian determinant is non-zero

$$G = \left[ (f_i, f_j) = \int_{\pi} f_i f_j \right], \quad \det G \neq 0$$

9. A set of functions (not nec in $L^2$), each of which has $N-1$ derivatives, is LI if and only if its Wronskian is not identically zero

$$W(t) = \begin{vmatrix} f_1 & f_2 & \cdots & f_N \\ \frac{d}{dt} f_1 & \frac{d}{dt} f_2 & \cdots & \frac{d}{dt} f_N \\ \vdots & \vdots & \ddots & \vdots \\ \frac{d^{N-1}}{dt^{N-1}} f_1 & \frac{d^{N-1}}{dt^{N-1}} f_2 & \cdots & \frac{d^{N-1}}{dt^{N-1}} f_N \end{vmatrix} \neq 0$$

**Example:** \(\{ f_1(t), f_2(t) \} \equiv \{ e^{-at}, e^{-bt} \} \)

$$W(t) = \begin{vmatrix} e^{-at} & e^{-bt} \\ -ae^{-at} & -be^{-bt} \end{vmatrix} = e^{-(a+b)t} [a-b] \neq 0, a \neq b$$

but if $a = b$ then $W(t) = 0$.

* This can be generalized to \(\{ e^{s_i t}, s_i \neq s_j, i \neq j \} \)

10. Given a function $f(t)$ and an ON set in $L^2$, $\{ \phi_i \}$, the minimum value of $E_N(\xi)$, defined as

$$E_N(\xi) = \| \xi - \sum_{i=1}^{N} \langle \xi, \phi_i \rangle \phi_i \|^2,$$
is attained when

\[ \langle \alpha_i \rangle = \int_0^T \phi_i(t) \phi_i^*(t) \, dt = (\alpha_i, \phi_i) \]

**Proof:**

\[ e_N(\alpha) = \sum_{i=1}^N \left( \int_0^T \phi_i(t) \phi_i^*(t) \, dt - \sum_{j=1}^N \langle \psi_j, \phi_i \rangle \langle \psi_j, \phi_i \rangle^* \right) \]

\[ = \sum_{i=1}^N \left| \phi_i \right|^2 + \sum_{j=1}^N \langle \psi_j, \phi_i \rangle \langle \psi_j, \phi_i \rangle^* - \sum_{i=1}^N \left( \langle \psi_i, \phi_i \rangle \langle \psi_i, \phi_i \rangle^* - \langle \psi_i, \phi_i \rangle \langle \psi_i, \phi_i \rangle^* + \left| \phi_i - (\psi, \phi_i) \right|^2 \right) \]

Now add and subtract \( \sum_{i=1}^N \left| (\psi, \phi_i) \right|^2\) so that

\[ e_N(\alpha) = \sum_{i=1}^N \left| (\psi, \phi_i) \right|^2 - \sum_{i=1}^N \left| \phi_i \right|^2 \]

\[ + \sum_{i=1}^N \left[ \left| (\psi, \phi_i) \right|^2 - \langle \psi_i, \phi_i \rangle \langle \psi_i, \phi_i \rangle^* - \langle \psi_i, \phi_i \rangle \langle \psi_i, \phi_i \rangle^* + \left| \phi_i - (\psi, \phi_i) \right|^2 \right] \]

\[ = \sum_{i=1}^N \left| \phi_i \right|^2 - \sum_{i=1}^N \left| (\psi, \phi_i) \right|^2 + \sum_{i=1}^N \left| \phi_i - (\psi, \phi_i) \right|^2 \]

\( \square \) independent of \( \psi \), thus choose \( \phi_i = (\psi, \phi_i) \).
Thus

\[ \min_{\Omega} \varepsilon_n(\Omega) = \| f \|^2 - \frac{1}{N} \sum_{i=1}^{N} |c_i|^2 \]

but \( \varepsilon_n(\Omega) \) is always greater than or equal to zero so that we have

11) \[ \| f \|^2 \geq \sum_{i=1}^{N} |c_i|^2 \]

This is Bessel's Inequality.

12) An orthonormal set \( \{ \phi_i \}_{i=1}^{\infty} \) is said to be complete (and on; complete) if for all functions \( f(t) \in L^2(\mathbb{R}) \), and any \( \varepsilon > 0 \) there exists an \( N \) such that

\[ \| f - \sum_{i=1}^{N} c_i \phi_i \| < \varepsilon \]

* You choose \( \varepsilon \) and I choose \( N(\varepsilon) \); if you make \( \varepsilon \) small I'll make \( N \) big.

13) If the set \( \{ \phi_i \}_{i=1}^{\infty} \) is complete on \( L^2(\mathbb{R}) \) then:

a) \[ \lim_{N \to \infty} \| f(t) - \sum_{i=1}^{N} c_i \phi_i(t) \| = 0 \], all \( f(t) \in L^2(\mathbb{R}) \).

b) \[ \| f \|^2 = \sum_{i=1}^{\infty} |c_i|^2 \] (Parseval's formula).

* recall Fourier Series.
e) Let \( f \in L^2(0, T) \), \( g \in L^2(0, T) \) and

\[
\alpha_i = (\phi_i, \phi_i), \quad \beta_i = (\phi_i, \phi_i)
\]

then:

\[
(\alpha_i, \beta_i) = \sum_{i=1}^{\infty} \alpha_i \beta_i^* \]

*This is known as the completeness relation or Parseval's formula.* The result has analogous forms in Fourier series and transforms, e.g.,

\[
\int_{-\infty}^{\infty} |\hat{f}(t)|^2 \, dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\hat{f}(\omega)|^2 \, d\omega
\]

\[
\int_{-\infty}^{\infty} \hat{g}(t) \, dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega) \, g(-\omega) \, d\omega
\]

These results constitute a review of the expansion of bounded energy signals. The signals were assumed to be deterministic (non-random). We now present an equivalent theory for the representation of random signals.

* Before leaving deterministic signals, we give two methods of generating the coefficients \( \xi_i \); from linear time-invariant systems:

a) \( f(t) \rightarrow \mathbb{X} \rightarrow \int_0^T \, dt \rightarrow c_i \)

b) \( s(t) \rightarrow h(t) = \phi_i(t-t) \rightarrow y(t) \)

\( \xi_i = y(t) \)
Representation of Random Processes

Now let us assume we have a random process \( x(t) \) and, for the moment, let us recall the sample space variable \( \xi \). We also assume we have a count set of functions for \( L^2(0,T) \). For any sample point \( \xi \), we can expand \( x(t|\xi) \) as

\[
x(t|\xi) = \sum_{i=1}^{\infty} \xi_i(\xi) \phi_i(t)
\]

Observe that \( \{\xi_i(\xi)\} \) is a set of random variables. This technique will work if \( x(t|\xi) \) is in \( L^2(0,T) \) for every point \( \xi \in \Xi \). It is more practical to require that

\[
\lim_{N \to \infty} E\left[ \left( x(t|\xi) - \sum_{i=1}^{N} \xi_i(\xi) \phi_i(t) \right)^2 \right] = 0, \quad 0 \leq t \leq T
\]

This is called the limit-in-the-mean (l.i.m.).

Now from the discussion for the general Gaussian problem it would be convenient to ask for the expansion coefficients (random variables) \( \xi_i(\xi) \) to be uncorrelated. Thus we want

\[
E\left[ (\xi_i - m_i)(\xi_j - m_j) \right] = \lambda_i \delta_{ij}
\]

where

\[
m_i = E[\xi_i], \quad \lambda_i \geq 0 \quad \text{(since } \lambda_i = \sigma^2_{\xi_i})
\]
Let us assume that $E[\tau_i] = 0$, which implies that $E[\tau_i] = m_i = 0, i = 1, 2, \ldots$. From this we also observe that

$$R_x(t, u) = k_x(t, u).$$

Now from our previous discussion

$$\tau_i = \int_0^T R_x(t, \tau_i \theta_i) \phi_i(t) dt$$

Thus

$$E[\tau_i \tau_j] = E \left[ \int_0^T R_x(t, \tau_i \theta_i) \phi_i(t) \int_0^T R_x(u, \tau_j \theta_j) \phi_j(u) du \right]$$

$$= \int_0^T \int_0^T R_x(t, u) \phi_i(t) \phi_j(u) dt du$$

$$= \int_0^T \phi_i(t) \left[ \int_0^T R_x(t, u) \phi_j(u) du \right] dt$$

$$= \lambda_i \phi_i(t).$$

Now the bracket must be orthogonal to $\phi_i(t)$ for $i \neq j$.

Thus

$$\int_0^T R_x(t, u) \phi_j(u) du = \lambda_i \phi_i(t), \quad 0 \leq t \leq T$$

Thus a necessary and sufficient condition for the cost to be orthogonal is that $\phi_i(t)$ be chosen as the solution to

$$\lambda_i \phi_i(t) \text{ are eigenvalues of the integral equation,}$$

$$\phi_i(t) \text{ are eigenfunctions of the integral equation.}$$
Because the orthonormality of the expansion coefficients was equivalent to determining the eigenvalues and eigenfunctions of an integral equation, let us now review some results from the theory of (homogeneous) integral equations. This problem is similar to the eigenvalue problem for matrices ($k\phi = \lambda \phi$).

We shall deal with a specific class of kernels (covariance functions). Those which are square-integrable:

$$\int_0^T \int_0^T k_x(t,u) dt du < \infty$$

**Properties of Homogeneous Integral Equations**

We shall consider the general integral equation:

$$\int_0^T k_x(t,u) \phi(u) du = \lambda \phi(t), \quad 0 \leq t \leq T \quad \text{and} \quad k_x(t,u) = k_x(u,t)$$

* Some of the following results have their parallel in matrix theory.

1. There exist at least one $\lambda \neq 0$ and one $\phi \in L^2(0,T)$ satisfying (5).

   There may be no more than one for degenerate kernels, e.g. $k_x(t,u) = t^2 \delta(t) \delta(u), \quad 0 \leq t \leq T$

2. If $\phi(t)$ is a solution then so is $t \phi(t)$. Thus we can make $\|\phi\|_1 = 1$
3) let \( \phi_1(t) \) and \( \phi_2(t) \) correspond to the same \( \lambda \) then
\( \phi_1(t) + \gamma \phi_2(t) \) also corresponds to \( \lambda \).

4) let \( \lambda_1 \) correspond to \( \phi_1(t) \) and \( \lambda_2 \) correspond to \( \phi_2(t) \) then \( \langle \phi_1, \phi_2 \rangle = 0 \) if \( \lambda_1 \neq \lambda_2 \). They are orthogonal.

5) there are at most a countable number of possible eigenvalues, \( \lambda \), and all are bounded.

6) for any single \( \lambda \), there are only a finite number of \( L2 \) eigenfunctions \( \phi(t) \). Thus they can be made on.

7) \textbf{(Mercer's Theorem)}. The covariance function (kernel of \( \mathcal{E} \)) can be expanded as follows

\[
k(t,u) = \sum_{i=1}^{\infty} \lambda_i \phi_i(t) \phi_i(u), \quad 0 \leq t, u \leq T
\]

the expansion is uniformly convergent.

8) If \( k(t,u) \) is \underline{positive definite} then the set of eigenfunctions is \underline{con} for \( L2(0, T) \). i.e. \( \phi_i(t) \) is \underline{con}.

9) If \( k(t,u) \) is not \underline{positive definite} (sign holds for some \( \phi(t) \neq 0, \phi(t) \)) then \( \phi_i(t) \) is \underline{not con} but we can complete \( \phi_i(t) \) by adding functions corresponding to \( \lambda = 0 \).
10. We observe
\[ \int_0^T k_x(t, t) \, dt = \int_0^T E[ \xi^2(\xi)] \, dt = E \left[ \int_0^T \xi^2(\xi) \, dt \right] \]

but
\[ \int_0^T k_x(t, t) \, dt = \int_0^T \sum_{i=1}^\varnothing \lambda_i \phi_i(t) \phi_i(t) \, dt = \sum_{i=1}^\varnothing \lambda_i \]

Thus,
\[ E \left[ \int_0^T \xi^2(\xi) \, dt \right] = \sum_{i=1}^\varnothing \lambda_i = \text{expected energy} \]

This completes the necessary review of integral equations. Let us now return to the expansion of random processes.

Let us now prove that if the expansion functions are chosen to be the eigen functions of the covariance function \( k_x(t, u) \), then \( \xi(t) \) converges as a limit in the mean to its expansion, hence

\[ \lim_{N \to \infty} E \left[ \xi(t) - \sum_{i=1}^N \zeta_i \phi_i(t) \right]^2 = 0, \quad 0 \leq t \leq T \]

where \( \zeta_i = (\zeta, \phi_i) \) and \( \{\phi_i\} \) are the eigen functions of \( R_x(t, u) \). Let

\[ f_n(t) = E \left[ \xi(t) - \sum_{i=1}^N \zeta_i \phi_i(t) \right]^2 \]

Expanding, we obtain
\[ \xi_N(t) = E \left[ X^2(t) - X(t) \sum_{i=1}^{N} \phi_i(t) \phi_i(t) + \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{N} \sum_{l=1}^{N} \phi_i(t) \phi_j(t) \phi_k(t) \phi_l(t) \right] \]

\[ = E[X^2(t)] - \sum_{i=1}^{N} E[\phi_i(t) \phi_i(t)] \phi_i(t) \]

Now we calculate

\[ \circ = R_X(t, t) \]

\[ \circ = -\sum_{i=1}^{N} E[\phi_i(t) \phi_i(t)] \sum_{j=1}^{N} \phi_j(t) \phi_j(t) \]

\[ = \sum_{i=1}^{N} \sum_{j=1}^{N} \phi_i(t) \phi_j(t) \phi_i(t) \phi_j(t) = \sum_{i=1}^{N} \lambda_i \phi_i(t) \phi_i(t) \]

\[ = \sum_{i=1}^{N} \lambda_i \phi_i(t) \phi_i(t) \]

Thus using these results

\[ \xi_N(t) = R_X(t, t) - \sum_{i=1}^{N} \lambda_i \phi_i(t) \phi_i(t) \]

Now by Mercer's Theorem the sum converges uniformly for \( R_X(t, t) \). Thus

\[ \lim_{N \to \infty} \xi_N(t) = 0 \]

This result is the Karhunen–Loève expansion:
There are several examples of the Karhunen-Loé expansion in Van Trees pp 105-106. We shall do just one but read the others.

**Example: The Wiener Process:** This is usually obtained from the one-dimensional random walk in the limit. It can also be thought of as integrated white noise. (See Papoulis pp 290-293, also ch. 14)

From Papoulis

\[
K_x(t,u) = \begin{cases} 
\sigma^2 u, & u \leq t \\
\sigma^2 \frac{t}{u}, & u \leq t
\end{cases}
\]

Thus \(\frac{\sigma}{\tau}\) is

\[
\int_0^T K_x(t,u) \phi(u) du = \lambda \phi(t), \quad 0 \leq t \leq T.
\]

Now we have two regions of \(u\):

- \(u \leq t, u \geq t\)

\[
\lambda \phi(t) = \int_0^t \sigma^2 u \phi(u) du + \int_t^T \sigma^2 \frac{t}{u} \phi(u) du
\]

Now we differentiate to obtain a d.e.

\[
\lambda \phi(t) = \sigma^2 \frac{t}{\lambda} \phi(t) - \sigma^2 \int_0^t \phi(u) du + \int_t^T \sigma^2 \phi(u) du
\]

\[
= \sigma^2 \int_t^T \phi(u) du.
\]

And again

\[
\lambda \phi = -\sigma^2 \phi(t)
\]
Finally we get
\[ \dot{\phi} + \frac{\alpha^2}{\lambda} \phi = 0 \]

We can show that there are no solutions for \( \lambda \leq 0 \) thus we let
\[ b^2 = \frac{\alpha^2}{\lambda}, \quad \lambda > 0 \]
and the solutions are
\[ \phi(t) = c_1 \sin(bt) + c_2 \cos(bt) \]

Now we substitute \( \phi(t) \) into (A) and obtain
\[ \lambda \left[ c_1 \sin(bt) + c_2 \cos(bt) \right] = \int_0^T \sigma^2 u \left[ c_1 \sin(\theta) + c_2 \cos(\theta) \right] d\theta \]
\[ + \int_0^T \tau^2 \left[ c_1 \sin(\theta) + c_2 \cos(\theta) \right] d\theta \]

Integrating
\[ (A) = \int_0^T \sigma^2 c_1 \left[ \frac{\sin(bt)}{b^2} - \frac{\cos(bt)}{b^3} \right] + \int_0^T \tau^2 c_2 \left[ \frac{\cos(bt)}{b} + \frac{\sin(bt)}{b^2} - \frac{1}{b^3} \right] \]
\[ (B) = \int_0^T \tau^2 c_1 \left[ \frac{\cos(bt)}{b} - \frac{\sin(bt)}{b^3} \right] + \int_0^T \tau^2 c_2 \left[ \frac{\sin(bt)}{b} - \frac{\cos(bt)}{b^3} \right] \]

Now equating terms in \( \sin(bt) \) and \( \cos(bt) = 0 \) yields
\[ b^2 = \frac{\alpha^2}{\lambda} \]

and equating terms in \( t = 0 \) yields
\[ \frac{T^2 c_2}{b^2} = 0 \Rightarrow c_2 = 0 \]

Finally from the constant terms \( \frac{c_1 \cos(bT)}{b} = 0 \)
Thus \( b^n = \left( n - \frac{1}{2} \right) \pi \), \( n = 1, 2, 3, \ldots \) \( (\cos b t = 0) \).

\[
\therefore \quad b^2 = \frac{c^2}{\lambda} = \left( n - \frac{1}{2} \right)^2 \pi^2 \frac{\pi^2}{\lambda} \quad \Rightarrow \quad \lambda n = \frac{c^2 \pi^2}{(n - \frac{1}{2})^2 \pi^2} \quad n = 1, 2, \ldots
\]

The eigenfunctions are

\[
\phi_n(t) = c_1 \sin \left( \frac{(n - \frac{1}{2}) \pi t}{\lambda} \right) \quad 0 \leq t \leq T
\]

and requiring \( \| \phi_n(t) \| = 1 \)

\[
\int_0^T c_1^2 \sin^2 \left( \frac{(n - \frac{1}{2}) \pi t}{\lambda} \right) dt = c_1^2 \int_0^T \left[ \frac{1}{2} - \sin \left( \frac{2(n - 1) \pi t}{T} \right) \right] dt
\]

\[
= c_1^2 \frac{T}{\pi} = 1 \quad \Rightarrow \quad c_1 = \frac{\sqrt{\pi}}{T}
\]

and thus

\[
\phi_n(t) = \frac{\sqrt{\pi}}{T} \sin \left[ \frac{(n - \frac{1}{2}) \pi t}{T} \right], \quad 0 \leq t \leq T
\]

Now let us derive important results for the derivative of the Wiener process.
white noise

From the previous discussion, the Karhunen-Loève expansion for the Wiener process would be

\[ x(t) = \sum_{k=-\infty}^{\infty} \xi_k \phi_k(t) \]

where \[ E[\xi_n^2] = \lambda_n = \frac{\sigma^2}{(n-\frac{1}{2})^2 \pi^2} \]

Suppose we differentiate the first \( k \) terms of the expansion

\[ x_k(t) = \sum_{n=1}^{k} \xi_n \sqrt{\frac{2}{\pi}} \left(\frac{n-\frac{1}{2}}{\pi}\right) \cos \left(\frac{(n-\frac{1}{2})\pi t}{\pi}\right) \]

now \( \phi_{n,k}(t) = \sqrt{\frac{2}{\pi}} \cos \left(\frac{(n-\frac{1}{2})\pi t}{\pi}\right) \) is still normalized

so that \( d_n = \left(\frac{n-\frac{1}{2}}{\pi}\right) \xi_n \) is the expansion coeffic

for \( x(t) \)

Observe that

\[ E[d_n^2] = \frac{(n-\frac{1}{2})^2 \pi^2}{\pi^2} \quad E[\xi_n^2] = \frac{(n-\frac{1}{2})^2 \pi^2}{(n-\frac{1}{2})^2 \pi^2} \cdot \frac{\sigma^2}{\pi} \]

Thus every eigenvalue of the new covariance is the same.
Formally we can obtain the covariance function for \( x(t) \) by forming

\[
k(x, t) = \frac{\partial}{\partial s} \left\{ \frac{\sigma^2}{\sigma^2} \delta(t - u) \right\} = \sigma^2 \delta(t - u)
\]

Consider the k-th expansion of such a covariance

\[
\lambda \phi(t) = \sigma^2 \int_0^T \delta(t - u) \phi(u) \, du = \sigma^2 \phi(t)
\]

\[
\lambda = \sigma^2 \quad \text{for any } \phi(t)
\]

Thus

\[
k^2 \delta(t - u) = \sum_{i=1}^\infty k^2 \phi_i(t) \phi_i(u)
\]

and

\[
\delta(t - u) = \sum_{i=1}^\infty \phi_i(t) \phi_i(u) \quad 0 \leq t, u \leq T
\]

* Thus, at least formally, any white noise process can be expanded using any con set \( \Phi(t) \) over the interval \( (0, T) \).
The optimum linear filter (Calculus of Variations)

Although the results obtained here can be derived using the "orthogonality principle" (see Papoulis), it is important to see the calculus of variations approach since the minimization techniques can be useful for other problems. Also, the solutions for the optimum linear filter can be expressed in terms of the eigenfunction expansions discussed previously.

We have a zero-mean message process \( c(t) \) to which is added zero-mean noise \( n(t) \). The received waveform is

\[
r(t) = c(t) + n(t) \quad 0 \leq t \leq T
\]

\( n(t) \) is passed through a linear time-variant filter \( h(t,u) \) to obtain an estimate, \( \hat{c}(t) \). Thus

We assume \( h(t,u) = 0 \), \( t < u \). Thus the linear system is causal. We may write

\[
\hat{c}(t) = \int_{0}^{T} h(t,u) c(u) \, du
\]
We want to choose $h(\tau, u)$ such that

$$\mathbb{E}\left[ \frac{1}{T} \int_0^T \left( a(t) - \hat{a}(t) \right)^2 dt \right]$$

This is the integral estimation error. Clearly if we minimize the point estimation error over $[0, T]$

$$\mathbb{E}\left[ a(t) - \hat{a}(t) \right]^2$$

we shall also minimize its time-average $\mathbb{E}$. We shall assume that the optimum processor, $h(\tau, u)$, is continuous in both variables on $0 \leq \tau, u \leq T$.

Now consider an arbitrary filter in the allowable class

$$h(\tau, u) = h_0(\tau, u) + \varepsilon h_\varepsilon(\tau, u), \quad 0 \leq \tau, u \leq T$$

where $\varepsilon$ is real and $h_\varepsilon(\tau, u)$ is allowable.

Now consider $\mathbb{E}_p(\tau; \varepsilon)$

$$\mathbb{E}_p(\tau; \varepsilon) = \mathbb{E}\left[ \int_0^T h(\tau, u) \varepsilon(u) du \right]^2$$

$$= \mathbb{E}\left[ a^2(t) - 2 a(t) \int_0^T h(\tau, u) \varepsilon(u) du \right.$$}

$$\left. + \int_0^T \int_0^T h(\tau, u) h(\tau, v) \varepsilon(u) \varepsilon(v) du dv \right]$$

$$= ka(\tau, t) - 2 \int_0^T h(\tau, u) k a(\tau, u) du$$

$$+ \int_0^T \int_0^T h(\tau, u) h(\tau, v) k(\tau, \nu) du dv$$
but \[ k_{\alpha r}(t; u) = \Xi [ a(t, u) (a(t, u) + n(u))] \]
\[ = k_{\alpha}(t; u) + k_{\alpha n}(t; u) \]

Thus
\[ \bar{S}_p(t; \varepsilon) = k_{\alpha}(t; t) - 2 \int_0^T h(t, u) k_{\alpha}(t; u) \, du + \int_0^T \int_0^T h(t, u) h(t, v) k_r(u, v) \, du \, dv \]

Now let \( h = h_0 + \varepsilon h_1 \): \[ \bar{S}_p(t; \varepsilon) = k_{\alpha}(t; t) - 2 \int_0^T h_0(t, u) k_{\alpha}(t; u) \, du + 2 \varepsilon \int_0^T \int_0^T h(t, u) k_{\alpha}(t; u) \, du \, dv \]
\[ + 2 \varepsilon \int_0^T \int_0^T h(t, u) h(t, v) k_r(u, v) \, du \, dv \]
\[ + 2 \varepsilon^2 \int_0^T \int_0^T h(t, u) h(t, v) k_r(u, v) \, du \, dv \]

Now group together terms in powers of \( \varepsilon \):
\[ \bar{S}_p(t; \varepsilon) = \left[ k_{\alpha}(t; t) - 2 \int_0^T h(t, u) k_{\alpha}(t; u) + 2 \int_0^T \int_0^T h(t, u) k_r(u, v) \, du \, dv \right] \]
\[ - 2 \varepsilon \left[ \int_0^T h(t, u) \left\{ k_{\alpha}(t, u) - 2 \int_0^T h(t, v) k_r(u, v) \, dv \right\} \, du \right] \]
\[ + \varepsilon^2 \left[ \int_0^T \int_0^T h(t, u) h(t, v) k_r(u, v) \, du \, dv \right] \]
This is of the form

$$p_{o}(t; \varepsilon) = p_{o}(t; +) + \Delta p(t; \varepsilon)$$

Now if \( h_{o}(t; u) \) is optimum \( \Delta g(t; \varepsilon) \geq 0 \) all \( h_{o}(t; u) \)
and \( \varepsilon \neq 0 \), \( \Delta g(t; \varepsilon) \) has two terms.

$$\Theta = e^{2} \int_{0}^{\infty} \int_{0}^{\infty} h_{o}(t; u) h_{e}(t; v) k_{r}(v; v) \, du \, dv \geq 0$$

since \( k_{r}(v; v) \) is positive semidefinite

$$\Theta = -2e \int_{0}^{\infty} h_{e}(t; u) \left[ k_{a}(t; u) - \int_{0}^{\infty} h_{o}(t; v) k_{r}(v; v) \, dv \right] \, du$$

since \( \Delta g \) must be nonnegative for all \( \varepsilon + h_{e}(t; u) \)

it is necessary and sufficient that

$$k_{a}(t; u) - \int_{0}^{\infty} h_{o}(t; v) k_{r}(v; v) \, dv = 0 \quad 0 \leq t \leq T \quad 0 < u < T$$

This is the integral equation for the optimum

linear filter. The inequality for \( u \) is strict

since the second term of \( \Theta \) can be discontinuous

at \( u = 0 \) if \( k_{r}(v; v) \) has delta functions (white

noise components)
**Additive white noise:**

Let us now assume that \( n(t) \) is white noise so that

\[
kr(t,u) = \frac{N_0}{2} \delta(t-u) + ka(t,u)
\]

Thus (ii) becomes

\[
\text{iii'}: \quad \frac{N_0}{2} h_0(t,u) + \int_0^T h(t,v) \, ka(u,v) \, dv = ka(t,u), \quad 0 \leq t \leq T
\]

Even though the inequality is strict, \( h_0(t,u) \) is assumed continuous, thus

\[
h_0(t,0) = \lim_{u \to 0} h_0(t,u), \quad h_0(t,T) = \lim_{u \to T} h_0(t,u)
\]

Actually since \( E[a(t)]^2 \), these results imply that (ii') is valid for \( 0 \leq u \leq T \)

**Errors:** Since \( \delta_p(t;\xi) \) is smallest for \( \xi = 0 \) we have

\[
\delta_p(t) = ka(t,u) - 2 \int_0^T h_0(t,u) \, ka(t,u) \, du
\]

\[
+ \int_0^T \int_0^T h_0(t,u) \, h_0(t,v) \, kr(u,v) \, du \, dv
\]
which is
\[ \xi_0(t) = k(a, t) - \int_0^T h(t, u) k(a, t, u) \, du \]
\[ - \int_0^T h(t, u) [k(a, t, u) - \int_0^T h(t, v) k(a, v, u) \, dv] \, du \]

\[ \xi_0(t) = k(a, t) - \int_0^T h(t, u) k(a, t, u) \, du \]

white noise error:

Since we have \( H^* \), we obtain

\[ \xi_0(t) = \frac{N_0}{2} h(t, t) \]

Solutions for optimum linear filter using Eigenfunctions:

We now solve \( H^* \), assuming we have solved for the eigenfunctions of \( k(a, t, u) \). We assume

\[ k(a, t, u) = \sum_{i=1}^b \lambda_i \phi_i(t) \phi_i(u) \]

If \( \{ \phi_i(t) \} \) is not con we shall have to complete the set by property 2) of our integral equation review. Also

\[ \frac{N_0}{2} \delta(t-u) = \sum_{i=1}^b \frac{N_0}{2} \phi_i(t) \phi_i(u) \]

from our discussion of white noise.
Thus \( k_r(t,u) \) has expansion

\[
k_r(t,u) = \sum_{i,j=1}^{\infty} \left( \frac{N_o}{2} + \lambda a \right) \phi_i(t) \phi_j(u)
\]

Now since \( \{\phi_i(t)\} \) is con then the set \( \{\phi_i(t), \phi_j(u)\} \) is con on the two dimensional interval \( [0,T] \times [0,T] \).

(See Courant and Hilbert about two-dim expansions)

Thus

\[
h_o(t,u) = \sum_{i,j=1}^{\infty} h_{ij} \phi_i(t) \phi_j(u)
\]

where

\[
h_{ij} = \int_0^T \int_0^T h_o(t,u) \phi_i(t) \phi_j(u) \, dt \, du = (h \phi_i \phi_j)
\]

provided \( \int_0^T \int_0^T h_o(t,u) \, dt \, du < \infty \).

Now plugging for \( k_o, k_r \) and \( h_o \) into \( \Pi \)', we obtain

\[
\left( \frac{N_o}{2} \right) \sum_{i,j=1}^{\infty} h_{ij} \phi_i(t) \phi_j(u) + \int_0^T \sum_{i,j=1}^{\infty} h_{ij} \phi_i(t) \phi_j(u) \, \sum_{k=1}^{\infty} \phi_k(a) \phi_k \, \nu \, \left( \int_0^T k_o(t,v) \, dv \right) = \sum_{i=1}^{\infty} \tau_i \phi_i(t) \phi_i(u)
\]

\[
- k_o(t,u)
\]
Now we integrate the second term to yield
\[ \frac{N_0}{2} \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} h_{ij} \phi_i(t) \phi_j(t) + \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \lambda_i \delta_{ij} \phi_i(t) \phi_j(t) \]
\[ = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \lambda_i \delta_{ij} \phi_i(t) \phi_j(t) \quad \text{for } i \neq j \] (we added nothing since this must be true for \( 0 \leq t \leq T \). We equate coefficients of \( \phi_i(t) \phi_j(t) \) and obtain
\[ \frac{N_0}{2} h_{ij} + h_{ij} \lambda_j = \delta_{ij} \lambda_i \]

or
\[ h_{ij} = \frac{\delta_{ij} \lambda_i}{\frac{N_0}{2} + \lambda_i} \]

so that
\[ h_{0l}(t, u) = \sum_{i=1}^{\infty} \frac{\lambda_i}{\frac{N_0}{2} + \lambda_i} \phi_i(t) \phi_i(u) \]

The errors are now easily obtained to be
\[ S_{P_0}(t) = \frac{N_0}{2} \sum_{i=1}^{\infty} h_{ii} \phi_i^2(t) \]
\[ S_{P_0}(t) = \frac{N_0}{2} \sum_{i=1}^{\infty} \frac{\lambda_i^2}{\frac{N_0}{2} + \lambda_i} \phi_i^2(t), \quad 0 \leq t \leq T \]

and
\[ S_x = \frac{1}{T} \int_0^T S_{P_0}(t) \, dt = \frac{N_0}{2T} \sum_{i=1}^{\infty} \frac{\lambda_i^2}{\frac{N_0}{2} + \lambda_i} \]