Stationary Processes and Models

Partial Characterization of a DT Stochastic Process

Deterministic Signals

A deterministic signal is characterized by a well-defined mathematical function of time.

Example: $x(n) = \cos(\omega_0 n),\ x(t) = \cos(\Omega_0 t)$

We can define exactly the way it evolves in time, i.e. its waveform.

Short Review on Random Variables

In probability theory, we employ the concept of a "random variable." Unlike its deterministic counterpart, a random variable assumes certain values with a certain probability. Random variables may be continuous or discrete. The probability of a particular random variable assuming certain values is given by its probability density function (pdf). For discrete random variables the pdf is sometimes also called a probability mass function.

Example. Consider the discrete random variable, *U* with the following pdf

n	$f_U(n)$
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6

Figure: Probability density function for a discrete random variable.

Example. Compute the probability that $U \le 3$.

$$P[U \le 3] = \sum_{\nu=1}^{3} \phi_V(\nu)$$
$$= \frac{1}{6} + \frac{1}{6} + \frac{1}{6}$$
$$= \frac{1}{2}$$

Note that for all discrete pdfs

$$\sum_{\nu} \{P[U = \nu]\} = 1 .$$

Example. Consider the continuous random variable, *U* with the following "uniform" pdf

$$f_U(u) = \begin{cases} 1/2, & -1 \le \upsilon < 1 \\ 0, & o.\omega. \end{cases}$$

Figure: Probability density function for a continuous random variable.

Example. Compute the probability that $0.25 \le U \le 0.75$.

$$P[0.25 \le Y \le 0.75] = \int_{0.25}^{0.75} \phi_Y(\upsilon)\delta\upsilon = \frac{1}{4}$$

Note that for all continuous pdfs

$$P[-\infty \le Y \le \infty] = \int_{-\infty}^{\infty} \phi_Y(\upsilon)\delta\upsilon = 1 .$$

Short Review on Random Signals

A stochastic or random process is used to describe the time evolution of a statistical phenomenon according to probabilistic laws. The time evolution of the phenomenon means that the stochastic process is a function of time. The statistical nature of the phenomenon means that before conducting an experiment, it is not possible to define exactly the way it evolves in time, i.e. its waveform.

Examples. Speech signals, television signals, radar returns, digital communications signals

A stochastic process also represents an infinite number of different realizations of the process. One particular realization is the *time series* $u(n),\ldots,u(n-M+1)$ where $u(n)$ is the current observation and $u(n-k)$ is a past observation at time $n-k$ ($k > 0$). We can only characterize these random signals with statistics or probability models.

In order for a stochastic process to be strictly stationary (SS) i.e. its statistical properties are invariant to a shift of time, the joint probability density functions (pdfs) of the observations must remain the same for all time.

Definition: The mean-value function of the process is defined as

$$\mu(n) \equiv E[u(n)]$$

where E is the expectation operator. For discrete random variables,

$$E[U] = \sum_{\upsilon:\pi(\upsilon)>0} \upsilon\pi(\upsilon) .$$

Example.

For continuous random variables,

$$E[U] = \int \upsilon\phi_Y(\upsilon)\,\delta\upsilon .$$

Definition: The autocorrelation function of the process is defined as

$$r(n, n-k) \equiv E[u(n)u^*(n-k)],\ k = 0,\ \pm 1,\ \pm 2, \ldots$$

For a partial characterization of the process, we specify the mean-value function (1st order statistic) and the autocorrelation function (2nd order statistic). This partial characterization offers important advantages.

1) easy to measure
2) well suited to linear operations on stochastic processes.

Definition: A process is wide sense stationary (WSS) if

$$\mu(n) = \mu$$
$$r(n, n-k) = r(k)$$

for all *n*. Note that these conditions are necessary but not sufficient for SS.

Mean Ergodic Theorem

Statistical parameters of real data are obtained through ensemble averages (or expected values). The estimation of any parameter of the stochastic process can be obtained by averaging a large number of realizations, *N* of the given process at a specific instant of time, i.e. "across the process".

Definition: The ensemble average of a process at time *n* is defined as

$$\hat{\mu}(n) = \frac{1}{N}\sum_{i=0}^{N-1} u(n, \zeta_i)$$

where ζ is the outcome variable. We have

$$\begin{aligned}\lim_{N\to\infty}\hat{\mu}(n) &= \mu(n) \\ &= E[u(n)]\end{aligned}.$$

Often though we only have a few samples of the time-series available, $u(n), \ldots, u(n-N+1)$. In this case we need to find out which statistical parameters of the process can be estimated by time-averaging of a single realization of the process. The equivalence between ensemble average and time average is called ergodicity.

Definition: Let $u(n)$ be a WSS process. The time average of a process is defined as

$$\hat{\mu}(N) = \frac{1}{N}\sum_{n=0}^{N-1} u(n),$$

i.e. "along the process."

Definition: A WSS process is *mean ergodic in the mean-square sense* if

$$\lim_{N\to\infty}[(\mu - \hat{\mu}(N))^2] = 0$$

i.e., mean-square of the difference between ensemble average, μ and time average $\hat{\mu}(N)$ approaches zero as *N* approaches ∞.

Correlation Matrix

Let $\mathbf{u}(n)$ be the observation vector

$$\mathbf{u}(n) = [u(n) \quad \ldots \quad u(n-M+1)]^T.$$

The correlation matrix is defined as the expectation of the outer product of **u**(n) with itself

$$
\begin{aligned}
\mathbf{R} &= E[\mathbf{u}(n)\mathbf{u}^H(n)] \\
&= E\begin{bmatrix} u(n)u^*(n) & \cdots & u(n)u^*(n-M+1) \\ u(n-1)u^*(n) & \cdots & u(n-1)u^*(n-M+1) \\ \vdots & \ddots & \vdots \\ u(n-M+1)u^*(n) & \cdots & u(n-M+1)u^*(n-M+1) \end{bmatrix}
\end{aligned}
$$

where ^{H} is the Hermitian or complex-conjugate transpose or ' in MATLAB. Using the condition of WSS (correlations depend only on time differences) we have the form

$$
\mathbf{R} = \begin{bmatrix} r(0) & r(1) & \cdots & r(M-1) \\ r(-1) & r(0) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ r(-M+1) & \cdots & \cdots & r(0) \end{bmatrix}.
$$

Note: in the following, take "stationary" to mean either SS or WSS.

Property 1: The correlation matrix of a stationary DT process is Hermitian

$$\mathbf{R}^H = \mathbf{R}$$

since

$$r(n, n-k) = r(k) \equiv E[u(n)u^*(n-k)]$$

and

$$
\begin{aligned}
r(n-k, n) &= r(-k) \\
&= E[u(n-k)u^*(n)] \\
&= E\left[\{u^*(n-k)u(n)\}^*\right] \\
&= \{E[u(n)u^*(n-k)]\}^* \\
&= r^*(k)
\end{aligned}
$$

which implies $r(-k) = r^*(k)$. Thus

$$
\mathbf{R} = \begin{bmatrix} r(0) & r(1) & \cdots & r(M-1) \\ r^*(1) & r(0) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \\ r^*(M-1) & r^*(M-2) & & r(0) \end{bmatrix}.
$$

Note we need only M correlations, $r(0), \ldots, r(M-1)$ to determine **R**. We note the Hermitian property is a direct result of SS or WSS.

Property 2: The correlation matrix of a stationary DT process is Toeplitz, i.e. elements along any particular diagonal are the equal.

Property 3: The correlation matrix of a DT process is always positive semidefinite (**R** ≥ **0**) and almost always positive definite (**R** > **0**).

Consider the inner product of **x** (deterministic) and the observation vector **u**(n) (statistical)

$$y = \boldsymbol{\xi}^H \boldsymbol{\upsilon}(\nu) \ .$$

We also have

$$y^* = \boldsymbol{\upsilon}^H(\nu)\boldsymbol{\xi} \ .$$

Therefore

$$\begin{aligned} 0 &\le E[|\psi|^2] \\ &= E[\psi\psi^*] \\ &= E[\boldsymbol{\xi}^H \boldsymbol{\upsilon}(\nu)\boldsymbol{\upsilon}^H(\nu)\boldsymbol{\xi}] \\ &= \boldsymbol{\xi}^H E[\boldsymbol{\upsilon}(\nu)\boldsymbol{\upsilon}^H(\nu)]\boldsymbol{\xi} \\ &= \boldsymbol{\xi}^H \mathbf{P}\boldsymbol{\xi} \end{aligned}$$

A Hermitian form that satisfies this condition for every nonzero **x** is said to be positive semidefinite. Unless the signals that compose the observation vector are linearly dependent (rare), then **R** is positive definite or $\mathbf{x}^H\mathbf{R}\mathbf{x} > 0$.

R > **0** implies nonsingularity or that $\mathbf{R}^{-1}$ exists.