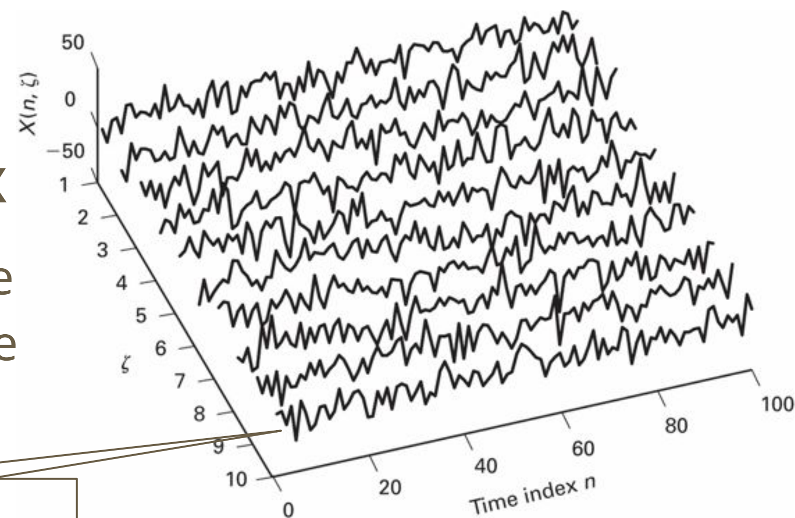
Chapter - 8: Random Sequences

Aveek Dutta
Assistant Professor
Department of Electrical and Computer Engineering
University at Albany
Fall 2019

Basic Concepts

Definition 6.1-1 Let $(\Omega, \mathscr{F}, P)$ be a probability space. Let $\zeta \in \Omega$. Let $X[n, \zeta]$ be a mapping of the sample space Ω into a space of complex-valued sequences on some index set Z. If, for each fixed integer $n \in Z$, $X[n, \zeta]$ is a random variable, then $X[n, \zeta]$ is a random (stochastic) sequence. The index set Z is usually all the integers, $-\infty < n < +\infty$, but can be just a subset of the integers.

- For each outcome (**ζ**) in the sample space (**Ω** = {1…10}), X[n,**ζ**] is a deterministic (nonrandom) function of the discrete parameter n.'
  - Each sequence is referred to as a *realization* or *sample sequence*. $X[n, \zeta] \triangleq A(\zeta)\sin(\pi n/10 + \Theta(\zeta))$
  - But, for a fixed ***n***, X[n,**ζ**] is a random variable
- If X[n] has finite support then we can model this sequence as a random vector **X**
- **Def:** Independent random sequence is one whose random variables at any time n_N are independent for all positive integer N.



Each sequence is called a sample sequence or realization (x[n]) of the random sequence for the outcome **ζ** = 10

Axiom 4: Continuity of Prob. Measure

- Axiom 3 of probability states that P[A U B] = P[A] + P[B] if AB = ϕ
  - But it does not allow us to define probabilities for events like $\cap_{n=1}^{+\infty}\{X[n] < 5\}$
- Axiom 4 extends it to define *"Countable Additivity"*. For infinite collection of events satisfying $A_iA_j = \phi$ for $i \neq j$

$$P\left[\bigcup_{n=1}^{\infty} A_n\right] = \sum_{n=1}^{+\infty} P[A_n]$$

- Theorem: Consider an increasing sequence of events B_n, that is, $B_n \subset B_{n+1}$ for all $n \geq 1$ as shown in Figure below. Define $B_\infty \triangleq \cup_{n=1}^{\infty} B_n$, then $\lim_{n\to\infty} P[B_n] = P[B_\infty]$.

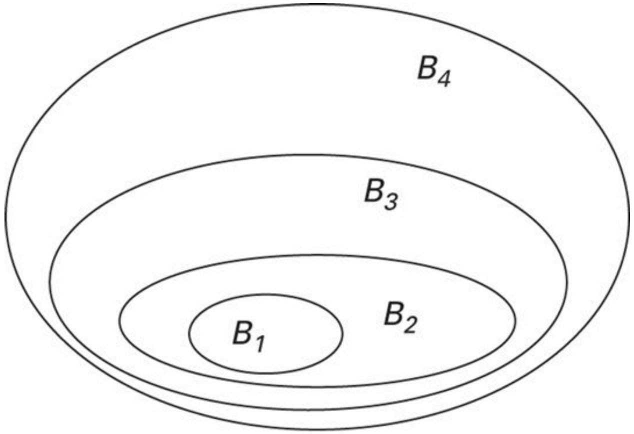
- Proof: $P[B_N] = P\left[\bigcup_{n=1}^{N} B_n\right] = P\left[\bigcup_{n=1}^{N} A_n\right] = \sum_{n=1}^{N} P[A_n]$

$$\lim_{n\to\infty} P[B_N] = \lim_{n\to\infty} \sum_{n=1}^{N} P[A_n]$$

$$= \sum_{n=1}^{+\infty} P[A_n] \quad \text{by definition of the limit of a sum,}$$

$$= P\left[\bigcup_{n=1}^{\infty} A_n\right] \quad \text{by Axiom 4}$$

$$= P[B_\infty] \quad \text{by definition of the } A_n$$



The theorem provides a way to calculate events involving infinite random variables by just taking the limit of the probability involving finite number of random variable.

Statistical Representation of Random Sequences

- **Nth order distribution** and density, for all times, n, n+1..., N+n-1.
  - Infinite set of pdfs for each order N because we must know pdf at *all* times

$$F_X(x_n, x_{n+1}, x_{n+2}, \ldots, x_{n+N-1}; n, n+1, \ldots, n+N-1) \stackrel{\Delta}{=} P\{X[n] \le x_n, X[n+1] \le x_{n+1}, \ldots, X[n+N-1] \le x_{n+N-1}\}$$

  - Low-order distributions must agree with higher orders. E.g., N = 2 and 3

$$F_X(x_n, x_{n+2}; n, n+2) = F_X(x_n, \infty, x_{n+2}; n, n+1, n+2)$$

- **Nth order density** is given by

$$f_X(x_n, x_{n+1}, \ldots, x_{n+N-1}; n, n+1, \ldots, n+N-1) = \frac{\partial^N F_X(x_n, x_{n+1}, \ldots, x_{n+N-1}; n, n+1, \ldots, n+N-1)}{\partial x_n \partial x_{n+1} \ldots \partial x_{n+N-1}}$$

- **Mean function** of a random sequence

$$\mu_X[n] \triangleq E\{X[n]\} = \int_{-\infty}^{+\infty} x f_X(x; n) dx = \int_{-\infty}^{+\infty} x_n f_X(x_n) dx_n \qquad \mu_X[n] = E\{X[n]\} = \sum_{k=-\infty}^{+\infty} x_k P[X[n] = x_k]$$

- **Autocorrelation -** mean of the product of the random seq. at two times

$$R_{XX}[k, l] \triangleq E\{X[k] X^*[l]\} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x_k x_l^* f_X(x_k, x_l; k, l) dx_k dx_l$$

- **Auto-covariance**

$$K_{XX}[k, l] \triangleq E\{(X[k] - \mu_X[k])(X[l] - \mu_X[l])^*\}$$

$$K_{XX}[k, l] = R_{XX}[k, l] - \mu_X[k] \mu_X^*[l]$$

$$\sigma_X^2[n] \triangleq K_{XX}[n, n]$$

Variance function or average power in the sequence

$$R_{XX}[k, l] = R_{XX}^*[l, k]$$
$$K_{XX}[k, l] = K_{XX}^*[l, k]$$

Hermitian symmetry

Examples of Random Sequences

- **Ex 8.1-11:** Convolving multiple exponential rv. is a Erlang (Gamma) PDF

$$f_\tau(t;n) = f_\tau(t) = \lambda \exp(-\lambda t) u(t) \longrightarrow \boxed{T[n] \triangleq \sum_{k=1}^{n} \tau[k]} \longrightarrow f_T(t;2) = f_\tau(t) * f_\tau(t) = \lambda^2 t \exp(-\lambda t) u(t)$$

- **Ex 8.1-12:** Autocorrelation of sum of iid sequences

$$R_W[k,l] = \sigma^2 \delta[k-l], \sigma > 0 \longrightarrow X[n] \triangleq W[n] + W[n-1] \longrightarrow R_{XX}[k,l] = \sigma^2(\delta[k-l] + \delta[k-l+1] + \delta[k-l-1] + \delta[k-l])$$

- **Ex 8.1-13: Random Walk - running sum of no. of heads minus tails**

$$X[n] = \sum_{k=1}^{n} W[k] \quad \text{with} \quad X[0] = 0$$

where we redefine $W[k] = +s$ for $\zeta = H$ and $W[k] = -s$ for $\zeta = T$

  - The sequence models a random walk with a unit step size ***s*** taken either to the right or left
  - After n steps the position is **rs** for some integer **r.** For **k successes** and **(n-k) failures**

$$rs = ks - (n-k)s = (2k-n)s \longrightarrow P\{X[n] = r \cdot s\} = P[(n+r)/2 \text{ successes }] = \begin{cases} \binom{n}{(n+r)/2} 2^{-n}, & (n+r)/2 \text{ an integer}, r \le n \\ 0, \text{ else.} \end{cases}$$

$$E\{X[n]\} = \sum_{k=1}^{n} E\{W[k]\} = \sum_{k=1}^{n} 0 = 0$$

$$E\{X^2[n]\} = \sum_{k=1}^{n} E\{W^2[k]\} = \sum_{k=1}^{n} 0.5\left[(+s)^2 + (-s)^2\right] = ns^2$$

$$\longrightarrow \tilde{X}[n] \triangleq \frac{1}{\sqrt{n}} X[n]$$

Normalizing the sequence

From CLT $\longrightarrow$

$$\tilde{X}[n] \sim N(0, s^2)$$

$$P[a < \tilde{X}[n] \le b] = P[a\sqrt{n} < X[n] \le b\sqrt{n}] \simeq \text{erf}(b/s) - \text{erf}(a/s)$$

$$P[(r-2)s < X[n] \le rs] = P\left[\frac{(r-2)s}{\sqrt{n}} < \tilde{X}[n] \le \frac{rs}{\sqrt{n}}\right] = \frac{1}{\sqrt{2\pi}} \int_{(r-2)/\sqrt{n}}^{r/\sqrt{n}} \exp(-0.5v^2)dv \approx 1/\sqrt{\pi(n/2)} \exp(-r^2/2n)$$

Stationarity and WSS

Definition: A random sequence is said to have independent increments if for all integer parameters $n_1 < n_2 < \ldots < n_N$, the increments $X[n_1], X[n_2] - X[n_1], X[n_3] - X[n_2], \ldots, X[n_N] - X[n_{N-1}]$ are jointly independent for all integers $N > 1$

Definition: If for all orders N and for all shift parameters k, the joint PDFs of $(X[n], X[n+1], \ldots, X[n+N-1])$ and $(X[n+k], X[n+k+1], \ldots, X[n+k+N-1])$ are the same functions, then the random sequence is said to be stationary, i.e., for all $N \geq 1$

$$\begin{aligned} &F_X(x_n, x_{n+1}, \ldots, x_{n+N-1}; n, n+1, \ldots, n+N-1) \\ &\quad = F_X(x_n, x_{n+1}, \ldots, x_{n+N-1}; n+k, n+1+k, \ldots, n+N-1+k) \end{aligned}$$

for all $-\infty < k < +\infty$ and for all x_n through x_{n+N-1}. This definition also holds for pdf's when they exist and PMFs in the discrete amplitude case.

Definition: A random sequence $X[n]$ defined for $-\infty < n < +\infty$, is called wide-sense stationary (WSS) if

(1) The mean function of $X[n]$ is constant for all integers $n, -\infty < n < +\infty$

$\mu_X[n] = \mu_X[0]$ and

(2) For all times $k, l, -\infty < k, l < +\infty$, and integers $n, -\infty < n < +\infty$ (correlation) function is independent of the shift n,

$K_{XX}[k, l] = K_{XX}[k+n, l+n]$

Shift Invariant

WSS

- **All stationary random sequences are wide-sense stationary**
- Proof: First show that mean is constant for a stationary random sequence. Since $\boldsymbol{f_X}$ does not depend on *n.*

$$\mu_X[n] = E\{X[n]\} = \int_{-\infty}^{+\infty} x f_X(x;n)dx = \int_{-\infty}^{+\infty} x f_X(x;0)dx = \mu_X[0]$$

- Next show covariance function is shift invariant using R_{XX}

$$\begin{aligned} R_{XX}[k,l] &= E\{X[k]X^*[l]\} \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x_k x_l^* f_X(x_k,x_l)dx_k dx_l \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x_{n+k} x_{n+l}^* f_X(x_{n+k},x_{n+l})dx_{n+k} dx_{n+l}^{\dagger} \\ &= R_{XX}[n+k,n+l] \end{aligned}$$

- Covariance is also shift invariant, hence use one *shift parameter*

$$\begin{aligned} K_{XX}[m] &\triangleq E\{X_c[k+m]X_c^*[k]\} = K_{XX}[k+m,k] \\ &= K_{XX}[m,0] \end{aligned}$$

$$R_{XX}[m] = R_{XX}[k+m,k] = R_{XX}[m,0]$$

Random sequences and Linear Systems

- **Self Study: Section 8.2 - Discrete-time linear systems, Linearity, Impulse response, Discrete time Fourier transform, Convolution theorem, z-transform**

- Definition: We say a system with operator L is linear if for all permissible input sequences $x_1[n]$ and $x_2[n]$, and for all permissible pairs of scalar gains a_1 and a_2 we have
  $$L\{a_1x_1[n] + a_2x_2[n]\} = a_1L\{x_1[n]\} + a_2L\{x_2[n]\}$$

- Definition: When we write $Y[n] = L\{X[n]\}$ for a random sequence $X[n]$ and a linear system L, we mean that for each $\zeta \in \Omega$ we have
  $$Y[n,\zeta] = L\{X[n,\zeta]\}$$
  Equivalently, for each sample function $x[n]$ taken on by the input random sequence $X[n]$, we set $y[n]$ as the corresponding sample sequence of the output random sequence $Y[n]$, i.e., $y[n] = L\{x[n]\}$.

- Impulse response

Time-variant impulse response: The response at time *n* to an impulse at time *k*

$$h[n,k] \triangleq L\{\delta[n-k]\}$$

$$y[n] = L\left\{\sum_{k=-\infty}^{+\infty} x[k]\delta[n-k]\right\}$$
$$= \sum_{k=-\infty}^{+\infty} x[k]L\{\delta[n-k]\}$$
$$= \sum_{k=-\infty}^{+\infty} x[k]h[n,k]$$

In LTI systems, convolution is an important property

$$h[n] \triangleq L\{\delta[n]\}$$

$$y[n] = h[n] * x[n] = x[n] * h[n]$$
$$h[n] * x[n] \triangleq \sum_{k=-\infty}^{+\infty} h[k]x[n-k]$$

Statistics of Linear Systems

Theorem: For a linear system L and a random sequence $X[n]$, the mean of the output random sequence $Y[n]$ is $E\{Y[n]\} = L\{E\{X[n]\}\}$ as long as both sides are well defined.

- Since L is a linear operator we can write and taking Expectation

$$y[n] = \sum_{k=-\infty}^{+\infty} h[n,k]x[k] \longrightarrow E\{Y[n]\} = E\left\{\sum_{k=-\infty}^{+\infty} h[n,k]X[k]\right\}$$

- Assuming we can bring the E[ ] inside the sum (not always possible)

$$E\{Y[n]\} = \sum_{k=-\infty}^{+\infty} h[n,k]E\{X[k]\} = L\{E\{X[n]\}\} \longrightarrow \mu_Y[n] = \sum_{k=-\infty}^{+\infty} h[n,k]\mu_X[k]$$

  - A necessary and sufficient condition is h(k) has to be *absolutely summable*

- If the input is WSS and if, $\sum_{k=-\infty}^{+\infty} |h[k]|$ exists, we can write $E\{Y[n]\} = \left[\sum_{k=-\infty}^{+\infty} h[k]\right]\mu_X = \mathrm{H}(z)|_{z=1}\mu_X$

- Also, cross-correlation between input and output

$$R_{XY}[m,n] \triangleq E\{X[m]Y^*[n]\} = E\{X[m](L\{X[n]\})^*\} = E\{X[m]L_n^*\{X^*[n]\}\} = L_n^*\{E\{X[m]X^*[n]\}\} = L_n^*\{R_{XX}[m,n]\}$$

$$R_{YY}[m,n] = E\{Y[m]Y^*[n]\} = E\{L_m\{X[m]Y^*[n]\}\} = L_m\{E\{X[m]Y^*[n]\}\} = L_m\{R_{XY}[m,n]\} = L_m\{L_n^*\{R_{XX}[m,n]\}\}$$

- The covariance functions for zero mean sequences is given by

$$\begin{cases} K_{XY}[m,n] = L_n^*\{K_{XX}[m,n]\} \\ K_{YY}[m,n] = L_m\{K_{XY}[m,n]\} \\ K_{YY}[m,n] = L_m\{L_n^*\{K_{XX}[m,n]\}\} \end{cases}$$

Operator L^*_n has impulse response h*(n,k) and L_m is the linear operator with time index m that treat n as a constant

Example 8.3-2

- A linear system with $Y[n] \triangleq X[n] - X[n-1] = L\{X[n]\}$
- The output correlation function is

$$\begin{aligned} R_{YY}[m,n] &= L_m\{R_{XY}[m,n]\} \\ &= R_{XY}[m,n] - R_{XY}[m-1,n] \\ &= R_{XX}[m,n] - R_{XX}[m-1,n] - R_{XX}[m,n-1] + R_{XX}[m-1,n-1] \end{aligned}$$

- If the input sequence were WSS with autocorrelation

$$R_{XX}[m,n] = a^{|m-n|}, \quad 0 < a < 1$$

- Then

$$R_{XY}[m,n] = a^{|m-n|} - a^{|m-n+1|}$$
$$R_{YY}[m,n] = 2a^{|m-n|} - a^{|m-1-n|} - a^{|m-n+1|}$$

- Note: R_{YY} only depends on the shifts *(m-n)* and hence is WSS

WSS random sequences

- For WSS sequences, we have
$$\begin{aligned}(1)\quad E\{X[n]\} &= \mu_X,\ \text{a constant},\\ (2)\ R_{XX}[k+m,k] &= E\{X[k+m]X^*[k]\}\\ &= R_{XX}[m] = E[X(m+0)X^*(m)]\end{aligned}$$
- Properties of WSS
  - $|R_{XX}[m]| \leq R_{XX}[0] \geq 0$
  - $|R_{XY}[m]| \leq \sqrt{R_{XX}[0]R_{YY}[0]}$

    Use the Cauchy-Schwarz inequality to prove $|E[h(X)g(X)]| \leq \left(E[h^2(X)]\right)^{1/2}\left(E[g^2(X)]\right)^{1/2}$
  - $R_{XX}[m] = R^*_{XX}[-m]$
  - For all n and complex a_i we must have $\sum_{n=1}^{N}\sum_{k=1}^{N} a_n a_k^* R_{XX}[n-k] \geq 0$
    - This property is the *positive semidefinite property* for autocorrelation function
- Few derivations for LTI systems

$$Y[n] = \sum_{k=-\infty}^{+\infty} h[n-k]X[k]$$

$$\begin{aligned}R_{XY}[m,n] &= E\{X[m]Y^*[n]\}\\ &= \sum_{k=-\infty}^{+\infty} h^*[n-k]E\{X[m]X^*[k]\}\\ &= \sum_{k=-\infty}^{+\infty} h^*[n-k]R_{XX}[m-k]\\ &= \sum_{k=-\infty}^{+\infty} h^*[-l]R_{XX}[(m-n)-l], \quad \text{with } l \triangleq k-n\end{aligned}$$

Similarly,

$$\begin{aligned}R_{YY}[n+m,n] &\triangleq E\{Y[n+m]Y^*[n]\}\\ &= \sum_{k=-\infty}^{+\infty} h[k]E\{X[n+m-k]Y^*[n]\}\\ &= \sum_{k=-\infty}^{+\infty} h[k]R_{XY}[m-k]\\ &= h[m]*R_{XY}[m]\end{aligned}$$

$$R_{YY}[m] = h[m]*R_{XY}[m]$$

Using the single shift parameter for WSS (if input is WSS then R_{XX} is shift invariant)

$$\begin{aligned}R_{XY}[m] &= \sum_{l=-\infty}^{+\infty} h^*[-l]R_{XX}[m-l]\\ &= h^*[-m]*R_{XX}[m]\end{aligned}$$

$$\begin{aligned}R_{YY}[m] &= h[m]*h^*[-m]*R_{XX}[m]\\ &= (h[m]*h^*[-m])*R_{XX}[m]\\ &= g[m]*R_{XX}[m], \quad \text{with } g[m] \triangleq h[m]*h^*[-m]\end{aligned}$$

Power Spectral Density

- PSD is the Fourier transform of $R_{XX}(m)$ of a WSS random sequence *X[n]*

$$S_{XX}(\omega) \triangleq \sum_{m=-\infty}^{+\infty} R_{XX}[m]\exp(-j\omega m), \quad \text{for } -\pi \le \omega \le +\pi \quad \longleftrightarrow \quad R_{XX}[m] = \text{IFT}\{S_{XX}(\omega)\} = \frac{1}{2\pi}\int_{-\pi}^{+\pi} S_{XX}(\omega)e^{j\omega m}d\omega$$

$$S_{YY}(\omega) = |H(\omega)|^2 S_{XX}(\omega) = G(\omega)S_{XX}(\omega)$$ (From previous slide)

$$E\left\{|X[n]|^2\right\} = R_{XX}[0] = \frac{1}{2\pi}\int_{-\pi}^{+\pi} S_{XX}(\omega)d\omega$$

$$S_{XY}(\omega) \triangleq \sum_{m=-\infty}^{+\infty} R_{XY}[m]\exp(-j\omega m), \quad \text{for } -\pi \le \omega \le +\pi$$

Integral of PSD over [-π,π] is the ensemble average power

- Interpretation of psd
  - Since $R_{XX}(0)$, which is the average power, is constant, Fourier transform may not be computed for some sequences, so a truncated sequence is used and PSD is defined under limit.

$$X_N(\omega) \triangleq FT\{w_N[n]X[n]\} \qquad w_N[n] \triangleq \begin{cases} 1 & |n| \le N \\ 0 & \text{else} \end{cases}$$

Window function that limits the sequence to ± N timesteps

- Under such an assumption, PSD represents the ensemble average power at frequency ω

$$S_{XX}(\omega) = \lim_{N\to\infty} \frac{1}{2N+1} E\left\{|X_N(\omega)|^2\right\}$$

Markov Random Sequences

- A continuous valued Markov random sequence X[n] satisfies the conditional pdf expression (for all k>0, but sufficient for K=1)

$$f_X(x_{n+k}|x_n, x_{n-1}, \dots, x_0) = f_X(x_{n+k}|x_n)$$

- The Nth order pdf can be written using using the chain rule

$$f_X(x_0, x_1, \dots, x_N) = f_X(x_0) f_X(x_1|x_0) f_X(x_2|x_1, x_0) \dots f_X(x_N|x_{N-1}, \dots, x_0)$$

  - Substituting the basic one-step (k=1) version of Markov definition

$$\begin{aligned} f_X(x_0, x_1, \dots, x_N) &= f_X(x_0) f_X(x_1|x_0) f_X(x_2|x_1) \dots f_X(x_N|x_{N-1}) \\ &= f_X(x_0) \prod_{k=1}^{N} f_X(x_k|x_{k-1}) \end{aligned}$$

- Markov-p random sequence satisfies the conditional pdf expressions as

$$f_X(x_{n+k}|x_n, x_{n-1}, \dots, x_0) = f_X(x_{n+k}|x_n, x_{n-1}, \dots, x_{n-p+1})$$

- Therefore, the unconditional pdf can be approximated as

$$\begin{aligned} f_X(x_0, x_1, \dots, x_N) =& f_X(x_0) f_X(x_1|x_0) f_X(x_2|x_1, x_0) \dots f_X(x_N|x_{N-1}, \dots, x_0) \\ \approx& f_X(x_0) f_X(x_1|x_0) f_X(x_2|x_1, x_0) \dots f_X(x_p|x_{p-1}, \dots, x_0) \\ &\times \prod_{k=p+1}^{N} f_X(x_k|x_{k-1}, \dots, x_{k-p+1}) \end{aligned}$$

Markov Chains

- Discrete-time Markov sequences are called Markov chains with PMF

$$P_X(x[n]|x[n-1],\ldots,x[n-N]) = P_X(x[n]|x[n-1])$$

- The value of X[n] at time n is called *"the state"*, because current value determines future value taken on by X[n]
    - If X[n] takes finite set of values {0, M-1}, it is a *finite state* Markov chain (*finite state-space*)
    - The state transition information is represented as a state transition matrix **P**

    $$p_{ij} = P_{X[n]|X[n-1]}(j|i)$$

    - Each row adds up to 1 and initial probabilities at n=0 is vector **p[0]** with elements

    $$(\mathbf{p}[0])_i = P_X(i;0), 1 \le i \le M$$

- 2-state markov chain (ex - 8.5-5)

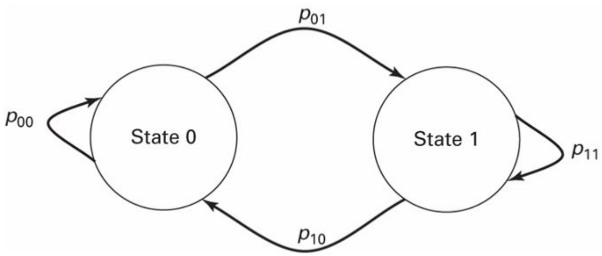
two-element probability row vector $\mathbf{p}[n] = (p_0[n], p_1[n])$

$$\mathbf{p}[1] = \mathbf{p}[0]\mathbf{P}$$
$$\mathbf{p}[2] = \mathbf{p}[1]\mathbf{P} = \mathbf{p}[0]\mathbf{P}^2$$
$$\mathbf{p}[n] = \mathbf{p}[0]\mathbf{P}^n$$

$$\mathbf{p}[\infty] = \mathbf{p}[\infty]\mathbf{P}, \text{ where } \mathbf{p}[\infty] = \lim_{n\to\infty} \mathbf{p}[n]$$

$$\mathbf{p} \triangleq \mathbf{p}[\infty], \text{ we have } \mathbf{p}(\mathbf{I}-\mathbf{P}) = \mathbf{0}$$
$$\mathbf{p1} = 1$$

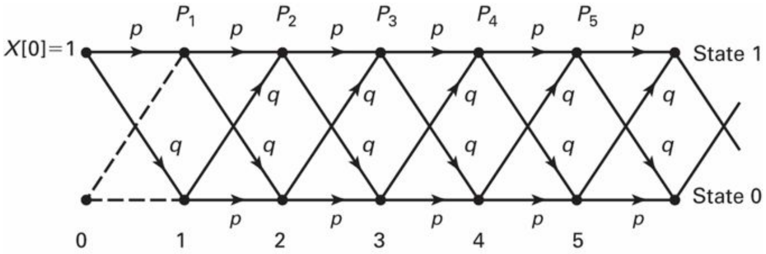
Solving these two equations provide the steady state probabilities **p**



$$\mathbf{P} = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix} = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}$$

Ex 8.5-6 Trellis for Markov Chain

- Each node denote the state at time [n]
- The links are possible transitions with transition probabilities (it is symmetric in this case)
- The probability of a path through the trellis is the product of the corresponding transition probabilities.

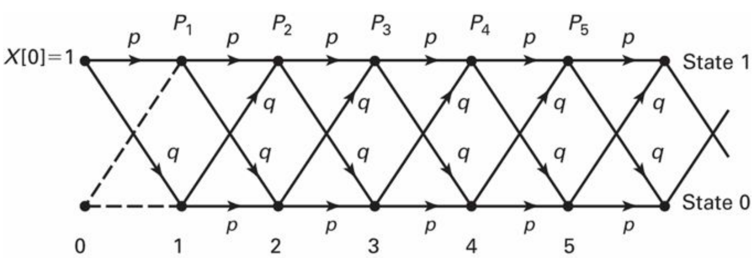


- If we know that the chain is in state "1" at n=0, then the trellis will be conditioned on this initial state
- Then we can find $P_n \triangleq P\{X[n]=1|X[0]=1\}$

$$P_1 = p, P_2 = p^2 + q^2, P_3 = p^3 + 3pq^2, \text{ etc.}$$

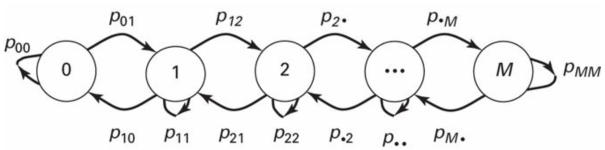
- The steady state autocorrelation function (*Asymptotically Stationary Autocorrelation (ASA)*) is

$$\begin{aligned} R_{XX}[m] &\approx P\{X[k]=1, X[m+k]=1\} \quad \text{for sufficiently large } k \\ &= P\{X[k]=1\}P\{X[m+k]=1|X[k]=1\} \\ &= p_1[\infty]P\{X[m]=1|X[0]=1\} \end{aligned}$$



Ex 8.5-7 Buffer Fullness Problem

- M+1 States of a buffer. Transitions occur only between neighboring states



$$[p_0[n+1], \quad p_1[n+1]] = [p_0[n], \quad p_1[n]] \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix}$$

The transition matrix is related by a simultaneous difference equation

- The solution has the form $p_0[n] = C_0 z^n, \quad p_1[n] = C_1 z^n$, leads to the following

$$\begin{aligned} C_0 z &= C_0 p_{00} + C_1 p_{10} \\ C_1 z &= C_0 p_{01} + C_1 p_{11} \end{aligned} \longrightarrow C_1 = C_0\left(\frac{z - p_{00}}{p_{10}}\right) = C_0\left(\frac{p_{01}}{z - p_{11}}\right) \longrightarrow \begin{aligned} (z - p_{00})(z - p_{11}) - p_{10}p_{01} = 0 \\ \det(z\mathbf{I} - \mathbf{P}) = 0 \end{aligned}$$

- Since, $(1 - p_{00}) = p_{01}$ at least one solution is z=1. The combined solution is

$$\mathbf{p}[n] = A_1\left[1, \frac{z_1 - p_{00}}{p_{10}}\right] z_1^n + A_2\left[1, \frac{z_2 - p_{00}}{p_{10}}\right] z_2^n$$

(See example 8.2-1 on difference equations)

- Example: $\mathbf{P} = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}$, with $\mathrm{p}[0] = [1/2, 1/2]$

$$\det(z\mathbf{I} - \mathbf{P}) = \det\begin{pmatrix} z - 0.9 & -0.1 \\ -0.2 & z - 0.8 \end{pmatrix} = z^2 - 1.7z + 0.7 = 0$$

$z_1 = 0.7$ and $z_2 = 1.0$

$$\mathbf{p}[n] = C_1[1, -1]0.7^n + C_2[1, 0.5]1^n$$

Invoking initial condition p[0]

$$C_2 = \frac{2}{3} \qquad C_1 = -\frac{1}{6}$$

$$\mathbf{p}[n] = \left[-\frac{1}{6}, \frac{1}{6}\right]0.7^n + \left[\frac{2}{3}, \frac{1}{3}\right], \text{ or in scalar form}$$

$$p_0[n] = -\frac{1}{6}0.7^n + \frac{2}{3}$$

$$p_1[n] = \frac{1}{6}0.7^n + \frac{1}{3}$$

Steady-state probabilities

$$p_0(\infty) = \frac{2}{3} \text{ and } p_1[\infty] = \frac{1}{3}$$

Convergence

Definition 6.7 – 1 A sequence of complex (or real) numbers x_n converges to the complex (or real) number x if given any $\varepsilon > 0$, there exists an integer n_0 such that whenever $n > n_0$, we have $|x_n - x| < \varepsilon$

$\longrightarrow$ $\lim_{n\to\infty} x_n = x$ or as $x_n \to x$ as $n \to \infty$

- If the limit *x* does not exist or is difficult to ascertain, use Cauchy criterion

Theorem (Cauchy criterion) – A sequence of complex (or real) numbers x_n converges to a limit if and only if (iff)

$$|x_n - x_m| \to 0 \text{ as both } n \text{ and } m \to \infty$$

- Convergence of functions
  - The Cauchy criterion applies for pointwise convergence of functions if the set of functions is considered ***complete***

Definition 6.7 – 2 The sequence of functions $f_n(x)$ converges (pointwise) to the function $f(x)$ if for each x_0 the sequence of complex numbers $f_n(x_0)$ converges to $f(x_0)$.

- Example convergence of sequence and functions

(a) $x_n = (1 - 1/n)a + (1/n)b \to a$ as $n \to \infty$

(b) $x_n = \sin(\omega + e^{-n}) \to \sin\omega$ as $n \to \infty$.

(c) $f_n(x) = \sin[(\omega + 1/n)x] \to \sin(\omega x)$, as $n \to \infty$ for any (fixed) x

(d) $f_n(x) = \begin{Bmatrix} e^{-n^2 x}, & \text{for } x > 0 \\ 1, & \text{for } x \le 0 \end{Bmatrix} \to u(-x)$, as $n \to \infty$ for any (fixed) x

Sure Convergence

Definition (Sure convergence.) The random sequence $X[n]$ converges surely to the random variable X if the sequence of functions $X[n,\zeta]$ converges to the function $X(\zeta)$ as $n \to \infty$ for all $\zeta \in \Omega$.

- Most of the time we may not be interested in defining random variables for sets in Ω of probability zero. So we use almost-sure convergence
  - Also called ***probability-1*** convergence and sometimes written as $P\left\{\lim_{n\to\infty} X[n,\zeta] = X(\zeta)\right\} = 1$

Definition (Almost-sure convergence.) The random sequence $X[n]$ converges almost surely to the random variable X if the sequence of functions $X[n,\zeta]$ converges for all $\zeta \in \Omega$ except possibly on a set of probability zero.

- There is a set A, with P[A]=1 and X[n] converges to X for all **ζ** ϵ A or $A \triangleq \left\{\zeta : \lim_{n\to\infty} X[n,\zeta] = X(\zeta)\right\}$
- Notation as $X[n] \to X$ a.s. and $X[n] \to X$ pr.1

Definition (Mean-square convergence.) A random sequence $X[n]$ converges in the mean-square sense to the random variable X if $E\left\{|X[n] - X|^2\right\} \to 0$ as $n \to \infty$.

- Depends only on the second order properties of X[n]

... contd

Definition (Convergence in probability.) Given the random sequence $X[n]$ and the limiting random variable X, we say that $X[n]$ converges in probability to X if for every $\varepsilon > 0$, $\lim_{n\to\infty} P[|X[n] - X| > \varepsilon] = 0$.

- Also called p-convergence
  - Convergence in mean-square implies convergence in probability
    - Use Chebyshev inequality $P[|Y| > \varepsilon] \le E\left[|Y|^2\right]/\varepsilon^2$ for $\varepsilon > 0$

$$P[|X[n] - X| > \varepsilon] \le E\left[|X[n] - X|^2\right]/\varepsilon^2$$

  - Convergence in a.s (probability-1) implies convergence in probability
  - So, conv. in probability is weaker than mean square and even weaker than probability 1
  - **Key difference - Limit of probability vs probability of limit**

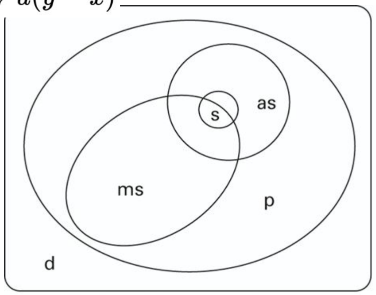
Definition: A random sequence $X[n]$ with probability distribution function $F_n(x)$ converges in distribution to the random variable X with probability distribution function $F(x)$ if $\lim_{n\to\infty} F_n(x) = F(x)$ at all x for which F is continuous.

Consider the conditional distribution → $F_{X[n]|X}(y|x) = P\{X[n] \le y | X = x\}$

P-convergence means seq. X[n] converges to rv X, as n -> ∞ therefore, → $F_{X[n]|X}(y|x) \to \begin{cases} 1, y > x \\ 0, y < x \end{cases}$ as $n \to \infty$ → $F_{X[n]|X}(y|x) \to u(y-x)$

Using definition of conditional distribution for continuous r.v **(see (2.6-4) in text)** →

$$F_{X[n]}(y) = P\{X[n] \le y\} = \int_{-\infty}^{+\infty} F_{X[n]|X}(y|x) f_X(x)dx$$
$$\to \int_{-\infty}^{+\infty} u(y-x) f_X(x)dx$$
$$= \int_{-\infty}^{y} f_X(x)dx = F_X(y)$$



s- surely
as- almost surely
ms- mean square
p- probability
d- distribution

Law of Large Numbers

- LLN deals with the convergence of a ***sequence of estimates of the mean*** of a random variable to a constant value
  - Weak law obtain convergence in probability
  - Strong law yield convergence with probability -1

Theorem (Weak Law of Large Numbers). Let $X[n]$ be an independent random sequence with mean μ_X and variance σ_X^2 defined for $n \geq 1$. Define another random sequence as $\hat{\mu}_X[n] \triangleq (1/n)\sum_{k=1}^{n} X[k]$ for $n \geq 1$Then $\hat{\mu}_X[n] \rightarrow \mu_x$ (p) as $n \rightarrow \infty$.

Theorem (Strong Law of Large Numbers.) Let $X[n]$ be a WSS independent random sequence with mean μ_X and variance σ_X^2 defined for $n \geq 1$. Then as $n \rightarrow \infty$

$$\hat{\mu}_X[n] = \frac{1}{n}\sum_{k=1}^{n} X[k] \rightarrow \mu_X \quad \text{(a.s.)}$$