Review of Random Process Concepts

K Sri Rama Murty

IIT Hyderabad

ksrm@ee.iith.ac.in

September 17, 2020

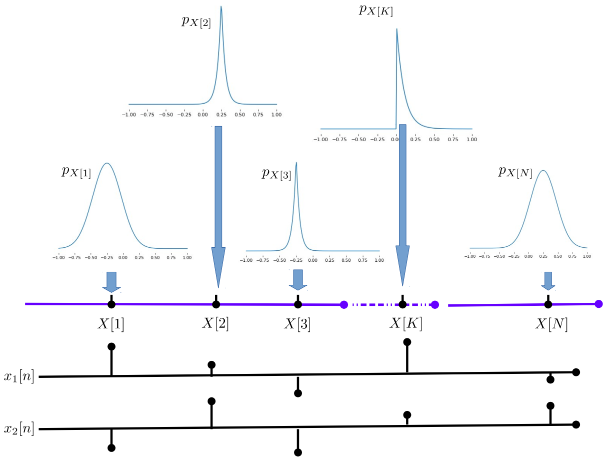
Random Process

- A random process is an infinite sequence of random variables

$$(\cdots,\ X[-1], X[0],\ X[1], \cdots\ X[n]\ \cdots)\ n \in \mathcal{T}$$

- The index parameter n is typically time, but also can be spatial dimension.
- RPs are used to model random experiments that evolve in time
  - Received signal at the output of a communication channel
  - Packet arrival times at a node in the computer network
  - Thermal noise in a resistor

Random Process - Illustration



- A random variable $X[k]$ is associated with every time-instant k
- A sample function is obtained by sampling the RVs in that sequence.

Bernoulli Process

- Sequence of Bernoulli RVs

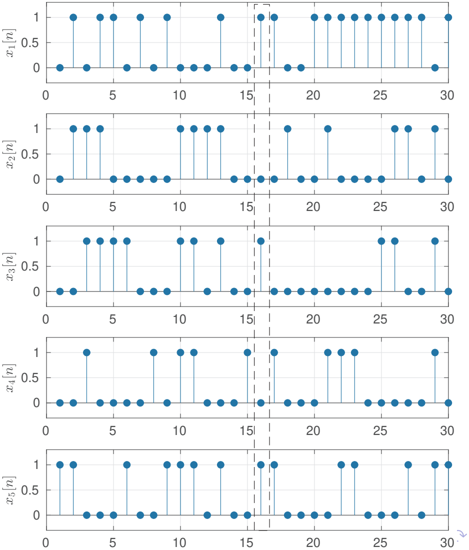
$$P_{X[k]}[1] = P[X[k] = 1] = p$$

- Dashed box shows RV X[16]
- The component RVs are i.i.d

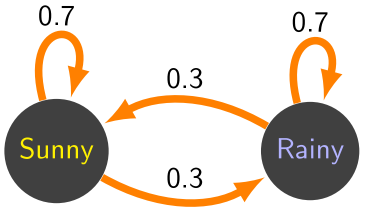
$$P_{X[1],\cdots,X[N]} = \prod_{k=1}^{N} P_{X[k]}$$

- No information in the sequence
- Used to model binary noise

Realizations of BP with $p = 0.5$

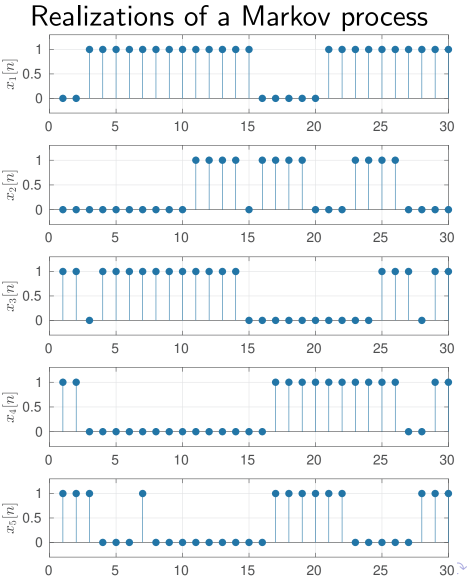


Markov Process



$P[S/S] = 0.7 \qquad P[R/R] = 0.7$

- RVs are not independent
- Stationary distribution is same as BP, but it is more predictable
- Sequence information helps in the prediction.



Stationary Random Process

- The relationships among the RVs is captured in their joint pdf

$$p_{X[n_1],X[n_2],\cdots X[n_N]}(x[n_1], x[n_2]\cdots x[n_N])$$

- A random process is stationary if

$$p_{X[n_1+K],X[n_2+K],\cdots X[n_N+K]} = p_{X[n_1],X[n_2],\cdots X[n_N]} \qquad \forall N, K$$

Stationary Random Process

- The relationships among the RVs is captured in their joint pdf

$$p_{X[n_1],X[n_2],\cdots X[n_N]}(x[n_1], x[n_2]\cdots x[n_N])$$

- A random process is stationary if

$$p_{X[n_1+K],X[n_2+K],\cdots X[n_N+K]} = p_{X[n_1],X[n_2],\cdots X[n_N]} \qquad \forall N, K$$

  - For a stationay RP, marginal densities are same for all the RVs.

$$p_{x[k]} = p_{X[0]}, \qquad \forall K$$

Stationary Random Process

- The relationships among the RVs is captured in their joint pdf

$$p_{X[n_1],X[n_2],\cdots X[n_N]}(x[n_1],x[n_2]\cdots x[n_N])$$

- A random process is stationary if

$$p_{X[n_1+K],X[n_2+K],\cdots X[n_N+K]} = p_{X[n_1],X[n_2],\cdots X[n_N]} \qquad \forall N, K$$

  - For a stationay RP, marginal densities are same for all the RVs.

$$p_{x[k]} = p_{X[0]}, \qquad \forall K$$

  - Joint moments and expected values of functions of RP are stationary

$$E_{X[n_1+K],X[n_2+K],\cdots X[n_N+K]}(.) = E_{X[n_1],X[n_2],\cdots X[n_N]}(.)$$

Stationary Random Process

- The relationships among the RVs is captured in their joint pdf

$$p_{X[n_1],X[n_2],\cdots X[n_N]}(x[n_1],x[n_2]\cdots x[n_N])$$

- A random process is stationary if

$$p_{X[n_1+K],X[n_2+K],\cdots X[n_N+K]} = p_{X[n_1],X[n_2],\cdots X[n_N]} \qquad \forall N,K$$

  - For a stationay RP, marginal densities are same for all the RVs.

$$p_{x[k]} = p_{X[0]}, \qquad \forall K$$

  - Joint moments and expected values of functions of RP are stationary

$$E_{X[n_1+K],X[n_2+K],\cdots X[n_N+K]}(.) = E_{X[n_1],X[n_2],\cdots X[n_N]}(.)$$

  - Joint and marginal densities should not depend on time, n.

Moments of a Random Process

- Mean sequence:

$$\mu_X[n] = \mathbb{E}[X[n]] \qquad -\infty < n < \infty$$

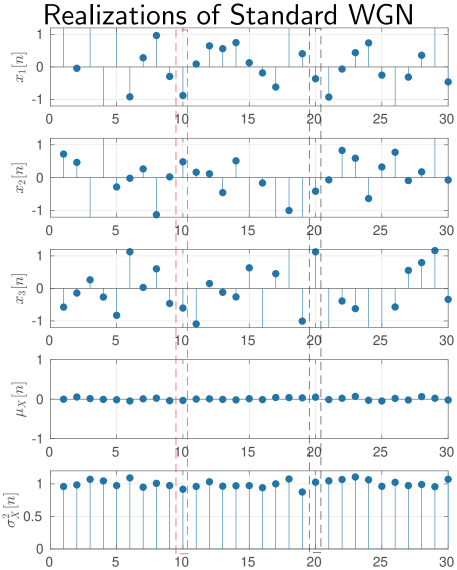
- Variance sequence:

$$\sigma_X^2[n] = \mathbb{E}[(X[n] - \mu_X[n])^2] \qquad -\infty < n < \infty$$

- Covariance sequence between the RVs at time instants n_1 and n_2

$$\begin{aligned} c_X[n_1, n_2] &= cov[X[n_1], X[n_2]] \\ &= \mathbb{E}[(X[n_1] - \mu_X[n_1])\,(X[n_2] - \mu_X[n_2])] \\ &\qquad -\infty < n_1, n_2 < \infty \end{aligned}$$

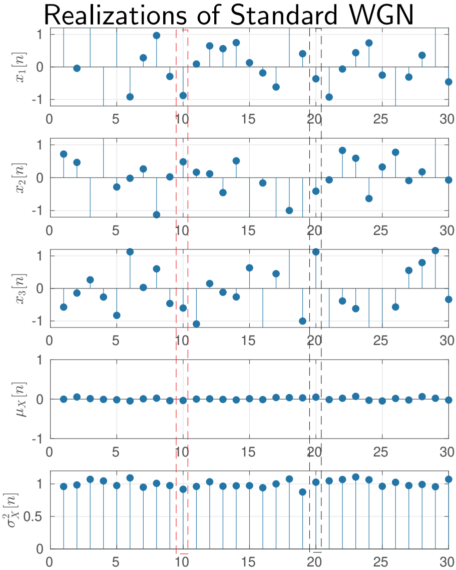
White Gaussian Noise



- RVs are Gaussian and i.i.d

$$P_{U[n]} \sim \mathcal{N}(0, \sigma^2)$$

White Gaussian Noise



- RVs are Gaussian and i.i.d

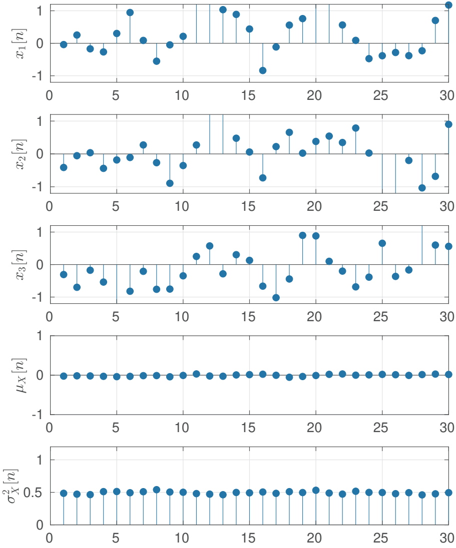
$$P_{U[n]} \sim \mathcal{N}(0, \sigma^2)$$

- Moments are computed across the ensemble

$$\mu_U[n] = 0$$
$$\sigma_U^2[n] = 1$$
$$c_U[n_1, n_2] = \delta[n_1 - n_2]$$

Moving Average Process

Realizations of MA Process on WGN

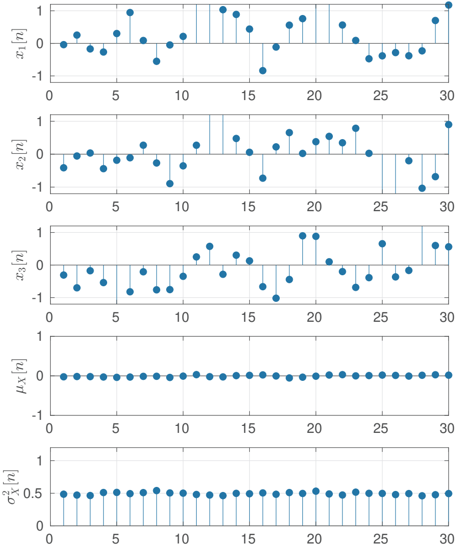


- Moving avg. of successive RVs

$$X[n] = \frac{1}{2}(U[n] + U[n-1])$$

Moving Average Process

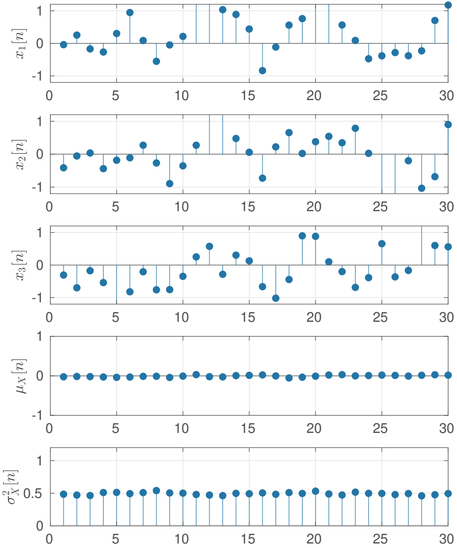
Realizations of MA Process on WGN



- Moving avg. of successive RVs
  $X[n] = \frac{1}{2}(U[n] + U[n-1])$
- Moments are computed across the ensemble
  $\mu_X[n] = 0 \quad \sigma_X^2[n] = 0.5$

Moving Average Process

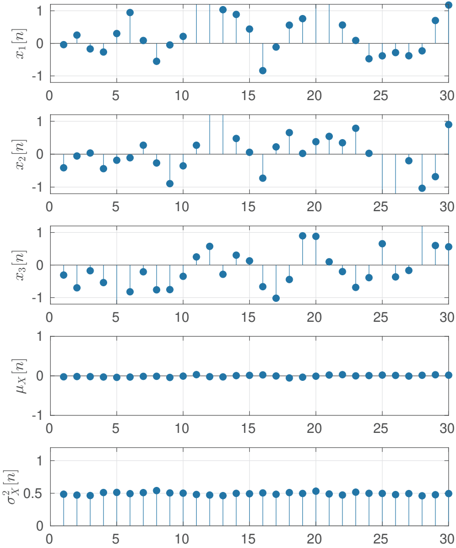
Realizations of MA Process on WGN



- Moving avg. of successive RVs
  $X[n] = \frac{1}{2}(U[n] + U[n-1])$
- Moments are computed across the ensemble
  $\mu_X[n] = 0 \quad \sigma_X^2[n] = 0.5$
- Covariance

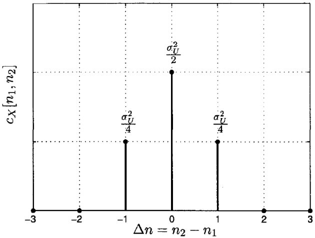
Moving Average Process

Realizations of MA Process on WGN



- Moving avg. of successive RVs $X[n] = \frac{1}{2}(U[n] + U[n-1])$
- Moments are computed across the ensemble

  $\mu_X[n] = 0 \quad \sigma_X^2[n] = 0.5$
- Covariance



Predict $X[n_2]$ from $X[n_1]$

- Can we predict $X[n_2]$ based on observation $X[n_1] = x[n_1]$?
- Assuming linear relation: $\hat{X}[n_2] = aX[n_1] + b$
- Estimate a and b such that overall error is minimized.

Predict $X[n_2]$ from $X[n_1]$

- Can we predict $X[n_2]$ based on observation $X[n_1] = x[n_1]$?
- Assuming linear relation: $\hat{X}[n_2] = aX[n_1] + b$
- Estimate a and b such that overall error is minimized.
- One of the objective criteria can be mean squared error

$$J(a, b) = \mathbb{E}[(X[n_2] - \hat{X}[n_2])^2]$$

Predict $X[n_2]$ from $X[n_1]$

- Can we predict $X[n_2]$ based on observation $X[n_1] = x[n_1]$?
- Assuming linear relation: $\hat{X}[n_2] = aX[n_1] + b$
- Estimate a and b such that overall error is minimized.
- One of the objective criteria can be mean squared error

$$J(a, b) = \mathbb{E}[(X[n_2] - \hat{X}[n_2])^2]$$

- Equating partial derivatives w.r.t a and b to zero, we get

$$\hat{X}[n_2] = \frac{c_x[n_1, n_2]}{c_x[n_1, n_1]} (X[n_1] - \mu_X[n_1]) + \mu_X[n_2]$$

Predict $X[n_2]$ from $X[n_1]$

- Can we predict $X[n_2]$ based on observation $X[n_1] = x[n_1]$?
- Assuming linear relation: $\hat{X}[n_2] = aX[n_1] + b$
- Estimate a and b such that overall error is minimized.
- One of the objective criteria can be mean squared error

$$J(a, b) = \mathbb{E}[(X[n_2] - \hat{X}[n_2])^2]$$

- Equating partial derivatives w.r.t a and b to zero, we get

$$\hat{X}[n_2] = \frac{c_x[n_1, n_2]}{c_x[n_1, n_1]} (X[n_1] - \mu_X[n_1]) + \mu_X[n_2]$$

- How to get $\mu_X[n_2]$ and $c_X[n_1, n_2]$?

Feasibility of Predicting $X[n_2]$ from $X[n_1]$

- For nonstationary process, it is exceedingly difficult to estimate $\mu_X[n_2]$ and $c_X[n_1, n_2]$.

Feasibility of Predicting $X[n_2]$ from $X[n_1]$

- For nonstationary process, it is exceedingly difficult to estimate $\mu_X[n_2]$ and $c_X[n_1, n_2]$.
- For stationary process, they can be replaced with $\mu_X[n_2 - K]$ and $c_X[n_1 - K, n_2 - K]$

Feasibility of Predicting $X[n_2]$ from $X[n_1]$

- For nonstationary process, it is exceedingly difficult to estimate $\mu_X[n_2]$ and $c_X[n_1, n_2]$.
- For stationary process, they can be replaced with $\mu_X[n_2 - K]$ and $c_X[n_1 - K, n_2 - K]$
- Stationarity is too strict a condition to be of any practical use.
- Moreover, we just need to estimate 1^{st} and 2^{nd} order statistics.

Feasibility of Predicting $X[n_2]$ from $X[n_1]$

- For nonstationary process, it is exceedingly difficult to estimate $\mu_X[n_2]$ and $c_X[n_1, n_2]$.
- For stationary process, they can be replaced with $\mu_X[n_2 - K]$ and $c_X[n_1 - K, n_2 - K]$
- Stationarity is too strict a condition to be of any practical use.
- Moreover, we just need to estimate 1st and 2nd order statistics.
- To extend the practical utility, it is enough if RP $X[n]$ satisfies

$$\mu_X[n] = \mu \qquad -\infty < n < \infty$$

$$c_X[n_1, n_2] = g(|n_2 - n_1|) \qquad -\infty < n_1, n_2 < \infty$$

Feasibility of Predicting $X[n_2]$ from $X[n_1]$

- For nonstationary process, it is exceedingly difficult to estimate $\mu_X[n_2]$ and $c_X[n_1, n_2]$.
- For stationary process, they can be replaced with $\mu_X[n_2 - K]$ and $c_X[n_1 - K, n_2 - K]$
- Stationarity is too strict a condition to be of any practical use.
- Moreover, we just need to estimate 1$^{\text{st}}$ and 2$^{\text{nd}}$ order statistics.
- To extend the practical utility, it is enough if RP $X[n]$ satisfies

$$\mu_X[n] = \mu \qquad -\infty < n < \infty$$
$$c_X[n_1, n_2] = g(|n_2 - n_1|) \qquad -\infty < n_1, n_2 < \infty$$

- Mean should be constant, and covariance should depend only on lag.

Widesense Stationary Process

- A random process is Widesense stationary if it satisfies

$$\begin{aligned} \mu_X[n] &= \mu & -\infty < n < \infty \\ \mathbb{E}[X[n_1]X[n_2]] &= h(|n_2 - n_1|) & -\infty < n_1, n_2 < \infty \end{aligned}$$

Widesense Stationary Process

- A random process is Widesense stationary if it satisfies

$$\mu_X[n] = \mu \qquad -\infty < n < \infty$$
$$\mathbb{E}[X[n_1]X[n_2]] = h(|n_2 - n_1|) \qquad -\infty < n_1, n_2 < \infty$$

- A stationary process is WSS, but converse is generally not true.
- WSS comes handy only when
  - The estimator is linear, and
  - The objective function is mean squared error, or equivalently

Widesense Stationary Process

- A random process is Widesense stationary if it satisfies

$$\mu_X[n] = \mu \qquad -\infty < n < \infty$$
$$\mathbb{E}[X[n_1]X[n_2]] = h(|n_2 - n_1|) \qquad -\infty < n_1, n_2 < \infty$$

- A stationary process is WSS, but converse is generally not true.
- WSS comes handy only when
  - The estimator is linear, and
  - The objective function is mean squared error, or equivalently
  - The error is assumed to be Gaussian. Minimizing $\sum_k v_k^2$ is same as maximizing $\exp(-\sum_k v_k^2) = \prod_k \exp(-v_k^2)$

Widesense Stationary Process

- A random process is Widesense stationary if it satisfies

$$\mu_X[n] = \mu \qquad -\infty < n < \infty$$
$$\mathbb{E}[X[n_1]X[n_2]] = h(|n_2 - n_1|) \qquad -\infty < n_1, n_2 < \infty$$

- A stationary process is WSS, but converse is generally not true.
- WSS comes handy only when
  - The estimator is linear, and
  - The objective function is mean squared error, or equivalently
  - The error is assumed to be Gaussian. Minimizing $\sum_k v_k^2$ is same as maximizing $\exp(-\sum_k v_k^2) = \prod_k \exp(-v_k^2)$
- WSS may not be enough for nonlinear systems and nongaussian error distributions

Autocorrelation Sequence ($n_1 = n$ & $n_2 = n + k$)

- The ACS of a WSS process is defined as

$$r_{XX}[k] = \mathbb{E}\left[X[n]\ X[n+k]\right] \qquad -\infty < k < \infty$$

- Properties of the ACS:
    - ACS is positive for the zero lag: $r_{XX}[0] = \mathbb{E}[X^2[n]] > 0$

Autocorrelation Sequence ($n_1 = n$ & $n_2 = n + k$)

- The ACS of a WSS process is defined as

$$r_{XX}[k] = \mathbb{E}\left[X[n]\ X[n+k]\right] \qquad -\infty < k < \infty$$

- Properties of the ACS:
  - ACS is positive for the zero lag: $r_{XX}[0] = \mathbb{E}[X^2[n]] > 0$
  - ACS is an even sequence: $r_{XX}[-k] = r_{XX}[k]$

Autocorrelation Sequence ($n_1 = n$ & $n_2 = n + k$)

- The ACS of a WSS process is defined as

$$r_{XX}[k] = \mathbb{E}\left[X[n]\ X[n+k]\right] \qquad -\infty < k < \infty$$

- Properties of the ACS:
  - ACS is positive for the zero lag: $r_{XX}[0] = \mathbb{E}[X^2[n]] > 0$
  - ACS is an even sequence: $r_{XX}[-k] = r_{XX}[k]$
  - ACS is maximum at 0^{th} lag: $|r_{XX}[k]| \leq r_{XX}[0]$

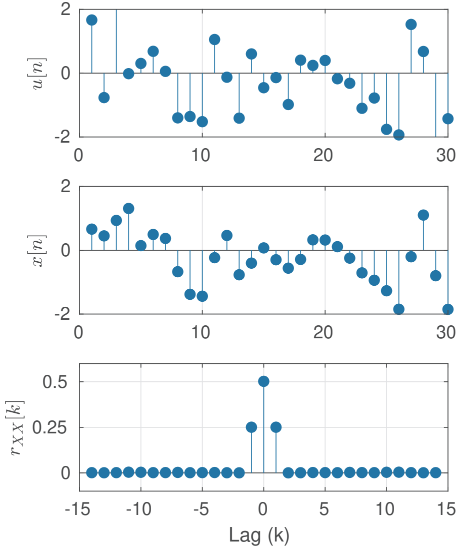
Autocorrelation Sequence ($n_1 = n$ & $n_2 = n + k$)

- The ACS of a WSS process is defined as

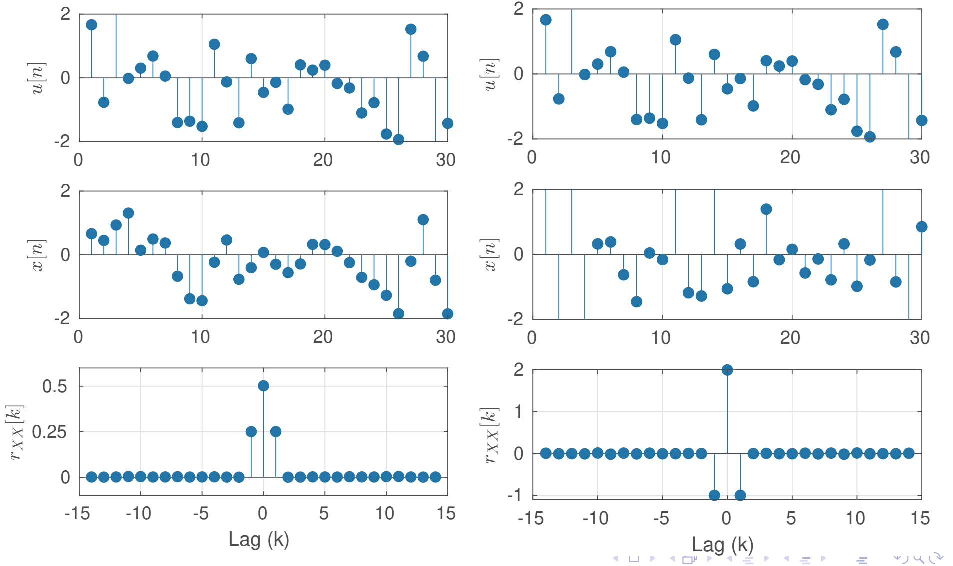
$$r_{XX}[k] = \mathbb{E}\left[X[n]\ X[n+k]\right] \qquad -\infty < k < \infty$$

- Properties of the ACS:
  - ACS is positive for the zero lag: $r_{XX}[0] = \mathbb{E}[X^2[n]] > 0$
  - ACS is an even sequence: $r_{XX}[-k] = r_{XX}[k]$
  - ACS is maximum at 0^{th} lag: $|r_{XX}[k]| \leq r_{XX}[0]$
  - ACS is a measure of predictability of the random process.
  - Autocorrelation matrix $\mathbf{R} = \mathbb{E}[\mathbf{X}\mathbf{X}^{\mathsf{T}}]$ is positive definite, where $\mathbf{X} = [X[0]\ X[1]\ \cdots\ X[N-1]]^{\mathsf{T}}$

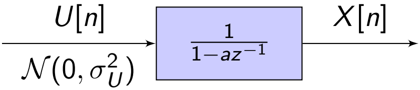
Moving Average & Difference Processes



Moving Average & Difference Processes

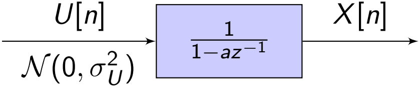


Autoregressive Process



$$
\begin{aligned}
X[n] &= aX[n-1] + U[n] \\
r_{XX}[k] &= \mathbb{E}[X[n]X[n+k]] \\
&= \mathbb{E}[X[n](aX[n+k-1] + U[n+k])] \\
&= a\,\mathbb{E}[X[n]X[n+k-1]] \qquad k \geq 1 \\
&= ar_{XX}[k-1] \\
&= r_{XX}[0]a^k
\end{aligned}
$$

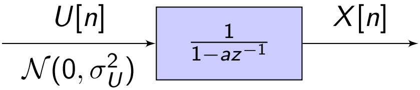
Autoregressive Process



$$\begin{aligned}
X[n] &= aX[n-1] + U[n] \\
r_{XX}[k] &= \mathbb{E}[X[n]X[n+k]] \\
&= \mathbb{E}[X[n](aX[n+k-1] + U[n+k])] \\
&= a\,\mathbb{E}[X[n]X[n+k-1]] \qquad k \geq 1 \\
&= a r_{XX}[k-1] \\
&= r_{XX}[0]a^k
\end{aligned}$$

$$\begin{aligned}
r_{XX}[0] &= \mathbb{E}[(aX[n-1] + U[n])^2] \\
&= a^2 r_{XX}[0] + \sigma_U^2 \\
&= \frac{\sigma_U^2}{1-a^2}
\end{aligned}$$

Autoregressive Process

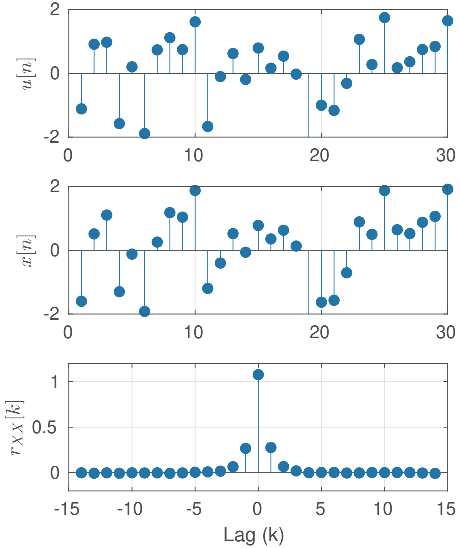


$$
\begin{aligned}
X[n] &= aX[n-1] + U[n] \\
r_{XX}[k] &= \mathbb{E}[X[n]X[n+k]] \\
&= \mathbb{E}[X[n](aX[n+k-1] + U[n+k])] \\
&= a\,\mathbb{E}[X[n]X[n+k-1]] \qquad k \geq 1 \\
&= a r_{XX}[k-1] \\
&= r_{XX}[0]a^k
\end{aligned}
$$

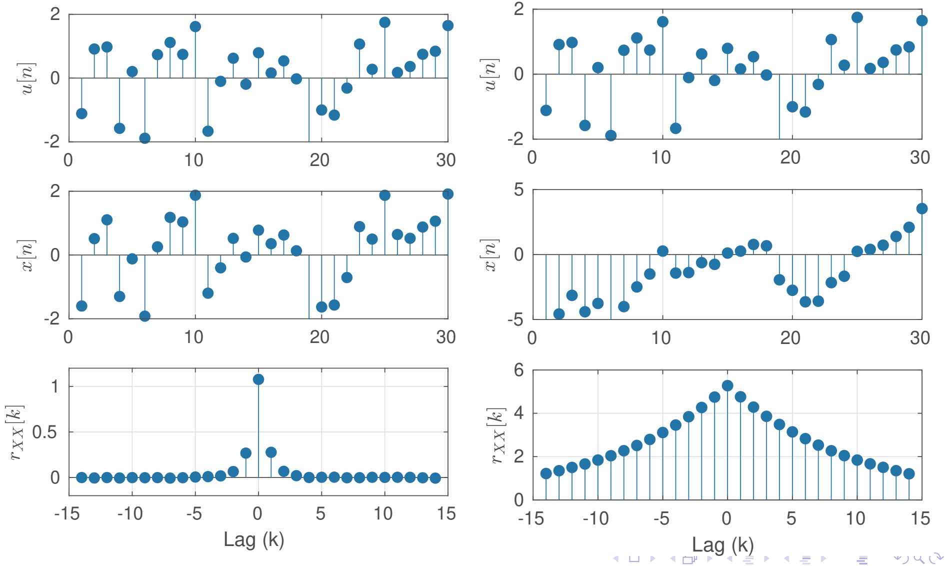
$$
\begin{aligned}
r_{XX}[0] &= \mathbb{E}[(aX[n-1] + U[n])^2] \\
&= a^2 r_{XX}[0] + \sigma_U^2 \\
&= \frac{\sigma_U^2}{1-a^2}
\end{aligned}
$$

$$
r_{XX}[k] = \frac{\sigma_U^2}{1-a^2} a^{|k|} \quad \forall k
$$

Autoregressive Process



Autoregressive Process



Predict $X[n+k]$ from $X[n]$

- In general $X[n_2]$ can be predicted from $X[n_1]$ as

$$\hat{X}[n_2] = \frac{c_x[n_1, n_2]}{c_x[n_1, n_1]}\left(X[n_1] - \mu_X[n_1]\right) + \mu_X[n_2]$$

- For a WSS process, letting $n_1 = n$ and $n_2 = n_1 + k$:

$$\hat{X}[n+k] = \frac{r_{XX}[k] - \mu^2}{r_{XX}[0] - \mu^2}\left(X[n] - \mu\right) + \mu$$

- For a zero-mean WSS process

$$\hat{X}[n+k] = \frac{r_{XX}[k]}{r_{XX}[0]} X[n]$$

- How to get an ensemble of sample functions to estimate μ and r_{XX}?

Ergodicity

- Moments of a RP are supposed to be computed across the ensemble.
- For a WSS process, mean does not depend on time.

Ergodicity

- Moments of a RP are supposed to be computed across the ensemble.
- For a WSS process, mean does not depend on time.
- When we observe one realization of a RP, we pretend as though we are observing multiple realizations of a RV with that mean.

Ergodicity

- Moments of a RP are supposed to be computed across the ensemble.
- For a WSS process, mean does not depend on time.
- When we observe one realization of a RP, we pretend as though we are observing multiple realizations of a RV with that mean.
- Thus, we may be able to determine the mean from a single infinite length realization.

Ergodicity

- Moments of a RP are supposed to be computed across the ensemble.
- For a WSS process, mean does not depend on time.
- When we observe one realization of a RP, we pretend as though we are observing multiple realizations of a RV with that mean.
- Thus, we may be able to determine the mean from a single infinite length realization.
- A random process is said to be *ergodic in mean* if the temporal average converges to ensemble average

$$\mu_X = \lim_{M\to\infty} \frac{1}{M} \sum_{m=1}^{M} x_m[16] = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} x_1[n]$$

Ensemble vs Temporal Averages

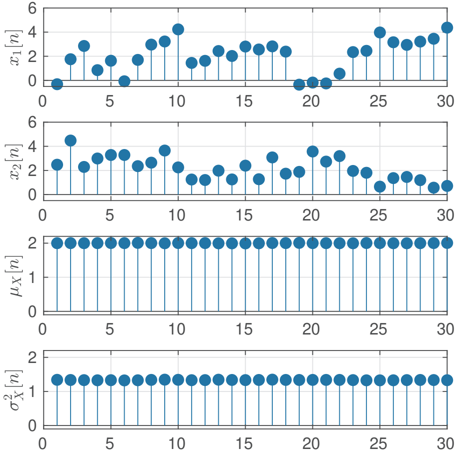
Ensemble Averages

$U[n] \sim \mathcal{N}(1,1) \quad H(z) = \frac{1}{1-0.5z^{-1}}$

Ensemble vs Temporal Averages

Ensemble Averages

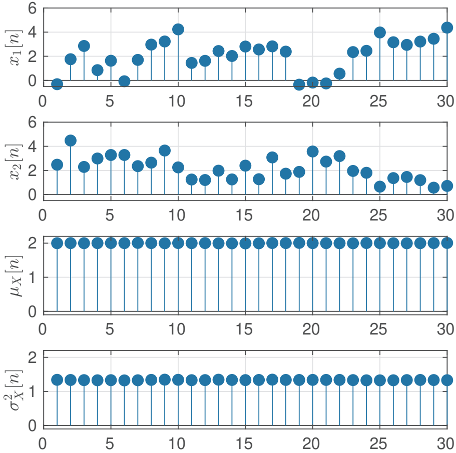
$$U[n] \sim \mathcal{N}(1,1) \quad H(z) = \frac{1}{1-0.5z^{-1}}$$



Ensemble vs Temporal Averages

Ensemble Averages

$U[n] \sim \mathcal{N}(1,1) \quad H(z) = \frac{1}{1-0.5z^{-1}}$



Temporal Averages

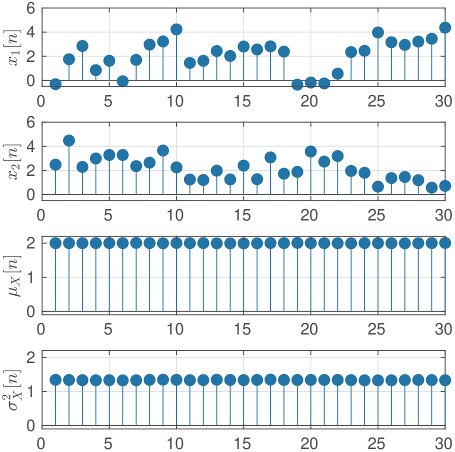
- Calculated on a single sample function

$$\hat{\mu}_X[n] = \frac{1}{n}\sum_{k=1}^{n} x_1[n]$$

Ensemble vs Temporal Averages

Ensemble Averages

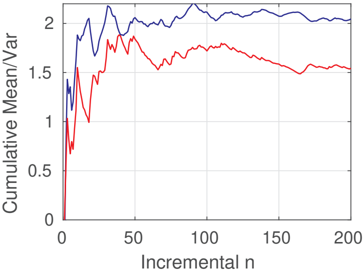
$U[n] \sim \mathcal{N}(1,1) \quad H(z) = \frac{1}{1-0.5z^{-1}}$



Temporal Averages

- Calculated on a single sample function

$$\hat{\mu}_X[n] = \frac{1}{n}\sum_{k=1}^{n} x_1[n]$$

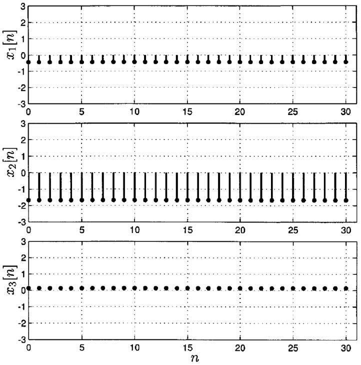


WSS doesn't imply Ergodicity

Random DC level process

$$X[n] = A, \quad \forall n$$

where $A \sim \mathcal{N}(0, 1)$

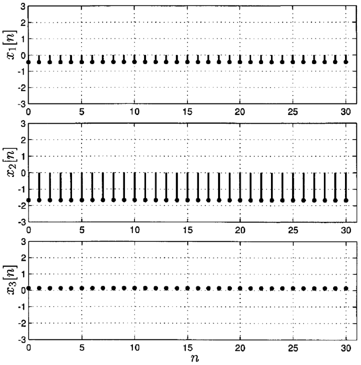


WSS doesn't imply Ergodicity

Random DC level process

$$X[n] = A, \quad \forall n$$

where $A \sim \mathcal{N}(0, 1)$



First and second order moments:

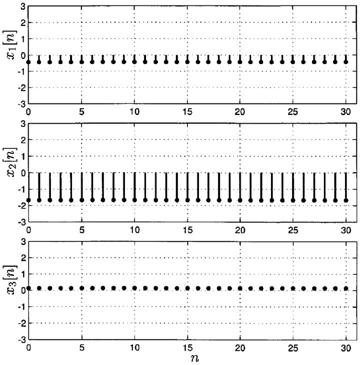
$$\mu_X[n] = \mathbb{E}[X[n]] = \mathbb{E}[A] = 0$$
$$r_{XX}[k] = \mathbb{E}[X[n]X[n+k]] = \mathbb{E}[A^2] = 1$$

WSS doesn't imply Ergodicity

Random DC level process

$$X[n] = A, \quad \forall n$$

where $A \sim \mathcal{N}(0, 1)$



First and second order moments:

$$\mu_X[n] = \mathbb{E}[X[n]] = \mathbb{E}[A] = 0$$
$$r_{XX}[k] = \mathbb{E}[X[n]X[n+k]] = \mathbb{E}[A^2] = 1$$

$$\mathbb{E}[x_1[n]] \neq \mathbb{E}[x_2[n]]$$

The process is WSS, but not ergodic

Note on WSS & Ergodicity

- All the traditional systems routinely *"assume"* WSS and ergodicity.
  - *L*inear models with *Gaussian* distributed errors
  - Model parameters, typically, depend on 1^{st} and 2^{nd} order statistics.
  - WSS: Assume that $\mu_X[n]$ and $r_{XX}[k]$ are independent of n
  - Ergodicity: Assume that a single realization is enough to estimate.
  - Offers, solid mathematical analysis
  - Performance is limited by linearity and Gaussianity assumptions

Note on WSS & Ergodicity

- All the traditional systems routinely *"assume"* WSS and ergodicity.
  - *L*inear models with *Gaussian* distributed errors
  - Model parameters, typically, depend on 1st and 2nd order statistics.
  - WSS: Assume that $\mu_X[n]$ and $r_{XX}[k]$ are independent of n
  - Ergodicity: Assume that a single realization is enough to estimate.
  - Offers, solid mathematical analysis
  - Performance is limited by linearity and Gaussianity assumptions
- Moving towards nonlinear-nongaussian systems
  - Example: $\hat{X}[n_2] = \tanh(aX[n_1] + b)$
  - The parameters a and b depends on higher order moments of $X[n]$
  - It is not meaningful to extend WSS and ergodicity to estimate $\mathbb{E}[X^p[n_1]X^q[n_2] \quad p, q > 1$
  - This explains the *data hungry* nature of the modern DNN models

Power Spectral Density

- ACS measures the correlation between the samples of a WSS process.
- ACS can be related to rate of change of the random process
  - Large fluctuations in ACS $\rightarrow$ Realization are rapidly varying in time
  - Slowly decaying ACS $\rightarrow$ Realizations are slowly varying in time
- For deterministic signals, the FT is used to analyze rate of change
- In the case of random process, each realization is slightly different. However, ACS be used as their common representative.
- FT of ACS is referred to as the power spectral density of the RP

$$P_{XX}(\omega) = \sum_{k=-\infty}^{\infty} r_{XX}[k] e^{-j\omega k} \qquad -\pi \leq \omega < \pi$$

Evaluating PSD

Properties of PSD

- PSD is a real function: $P_{XX}(\omega) = \sum_k r_{XX}[k] \cos(\omega k)$

Properties of PSD

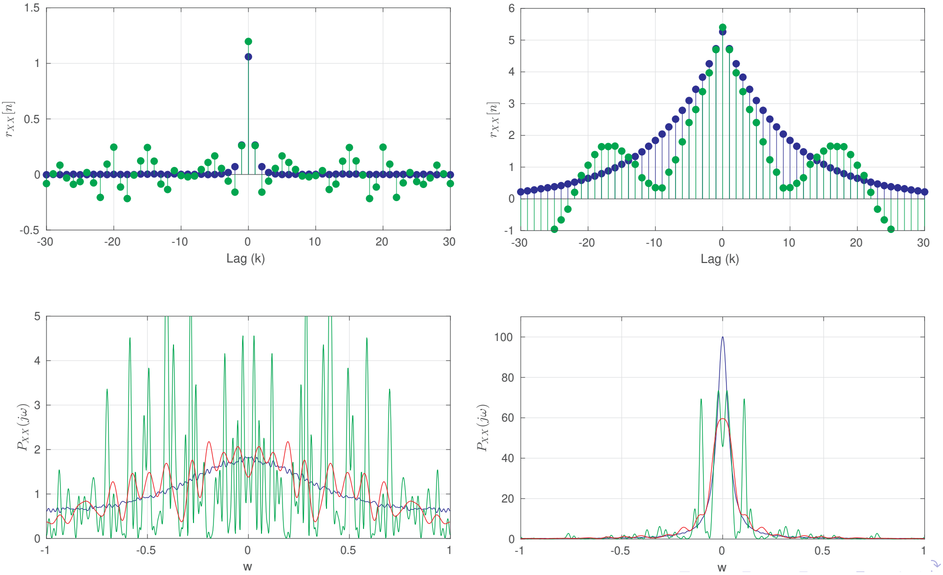
- PSD is a real function: $P_{XX}(\omega) = \sum_k r_{XX}[k] \cos(\omega k)$
- PSD is nonnegative: $P_{XX}(\omega) \geq 0$

Properties of PSD

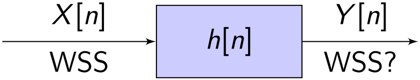
- PSD is a real function: $P_{XX}(\omega) = \sum_k r_{XX}[k]\cos(\omega k)$
- PSD is nonnegative: $P_{XX}(\omega) \geq 0$
- PSD is an even function of ω: $P_{XX}(-\omega) = P_{XX}(\omega)$
- PSD is periodic with period 2π: $P_{XX}(\omega + 2\pi) = P_{XX}(\omega)$
- ACS can be recovered from PSD using inverse FT

$$r_{XX}[k] = \int_{-\pi}^{\pi} P_{XX}(\omega)e^{j\omega k}d\omega \qquad -\infty < k < \infty$$

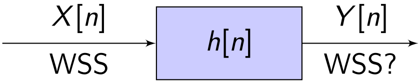
Ergodic Estimates of ACS and PSD



Filtering WSS Process through LTI System



Filtering WSS Process through LTI System



- The output of the system:

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

- The mean sequence of the output:

$$\mathbb{E}[Y[n]] = \left(\sum_{k=-\infty}^{\infty} h[k] \right) \mu_X = H(j0)\mu_X$$

- ACS and PSD of the output:

$$r_{YY}[k] = h[-k] * h[k] * r_{XX}[k] \qquad P_{YY}(\omega) = |H(j\omega)|^2 P_{XX}(\omega)$$

Input & Output Power Calculations

- Average input power: $r_{XX}[0] = E[X^2[n]]$
- The output of the system can be written as $Y[n] = \mathbf{h}^{\mathsf{T}}\mathbf{X}[n]$ where $\mathbf{h} = [h[0]\ h[1] \cdots\ h[N]]^{\mathsf{T}}$ & $\mathbf{X}[n] = [X[n]\ X[n-1]\ \cdots\ X[n-N]]^{\mathsf{T}}$

Input & Output Power Calculations

- Average input power: $r_{XX}[0] = E[X^2[n]]$
- The output of the system can be written as $Y[n] = \mathbf{h}^{\mathsf{T}}\mathbf{X}[n]$ where $\mathbf{h} = [h[0]\ h[1] \cdots\ h[N]]^{\mathsf{T}}$ & $\mathbf{X}[n] = [X[n]\ X[n-1]\ \cdots\ X[n-N]]^{\mathsf{T}}$
- Average output power:

$$
\begin{aligned}
r_{YY}[0] &= \mathbb{E}[Y^2[n]] \\
&= \mathbb{E}[\mathbf{h}^{\mathsf{T}}\mathbf{X}[n]\mathbf{X}^{\mathsf{T}}[n]\mathbf{h}] = \mathbf{h}^{\mathsf{T}}\,\mathbb{E}[\mathbf{X}[n]\mathbf{X}[n]^{\mathsf{T}}]\mathbf{h} \\
&= \mathbf{h}^{\mathsf{T}}\mathbf{R}_{XX}\mathbf{h} > 0
\end{aligned}
$$

Input & Output Power Calculations

- Average input power: $r_{XX}[0] = E[X^2[n]]$
- The output of the system can be written as $Y[n] = \mathbf{h}^\mathsf{T}\mathbf{X}[n]$ where $\mathbf{h} = [h[0]\ h[1] \cdots\ h[N]]^\mathsf{T}$ & $\mathbf{X}[n] = [X[n]\ X[n-1]\ \cdots\ X[n-N]]^\mathsf{T}$
- Average output power:

$$\begin{aligned} r_{YY}[0] &= \mathbb{E}[Y^2[n]] \\ &= \mathbb{E}[\mathbf{h}^\mathsf{T}\mathbf{X}[n]\mathbf{X}^\mathsf{T}[n]\mathbf{h}] = \mathbf{h}^\mathsf{T}\,\mathbb{E}[\mathbf{X}[n]\mathbf{X}[n]^\mathsf{T}]\mathbf{h} \\ &= \mathbf{h}^\mathsf{T}\mathbf{R}_{XX}\mathbf{h} > 0 \end{aligned}$$

- The autocorrelation matrix $\mathbf{R}_{XX}$ is a positive semidefinite matrix

Questions?