Random Variables

- Let X be a random variable on $\mathbb{R}$, then
  - X is usually denoted by an upper case letter.
  - The cummulative distribution function is given by

$$P\{X \leq x\} = F_X(x)$$

  - If the probability density function exists, it is given by

$$p_X(x) = \frac{dF_X(x)}{dx}$$

    so that

$$\begin{aligned} P\{x_1 < X \leq x_2\} &= F_X(x_2) - F_X(x_1) \\ &= \int_{x_1}^{x_2} p_X(\tau) d\tau \end{aligned}$$

  - The expectation of X is given by

$$E[X] = \int_{-\infty}^{\infty} \tau p_X(\tau) d\tau$$

    or more precisely by the Riemann-Stieltjes integral

$$E[X] = \int_{-\infty}^{\infty} \tau dF_X(\tau)$$

    if it exists.

Deterministic versus Random

- Let X and Z be random variables, and let $f(\cdot)$ be a function from $\mathbb{R}$ to $\mathbb{R}$
  - Is Y a random variable

$$Y = f(X)$$

  - Is μ a random variable

$$\mu = E[X]$$

  - Is $\hat{X}$ a random variable

$$\hat{X} = E[X|Z]$$

Properties of Expectation

- Expectation is linear

$$E[X+Y]=E[X]+E[Y]$$

- What is $E[E[X|Y]]$ equal to?

$$E[E[X|Y]]=E[X]$$

- What is $E[X|X,Y]$ equal to?

$$E[X|X,Y]=X$$

- When X, Y, and Z are (jointly) Gaussian

$$E[X|Y,Z]=aY+bZ+c$$

for some scalar values a, b, and c.

2-D Discrete Space Random Processes

- Notation
  - X_s is a pixel at position $s = (s_1, s_2) \in \mathcal{Z}^2$
  - S denotes the set of 2-D Lattice points where $S \subset \mathcal{Z}^2$
- Definitions
  - Mean $\mu_s = E[X_s]$
  - Autocorrelation $R_{sr} = E[X_s X_r]$
  - Autocovariance $C_{sr} = E[(X_s - \mu_s)(X_r - \mu_r)]$
  - A process is said to be **second order** if $E[X_s]$ and $E[X_s X_r]$ exist for all $s \in S$ and $r \in S$.
  - A second order random process is said to be **wide sense stationary** if for all $s \in \mathcal{Z}^2$

$$\mu_s = \mu_{(0,0)}$$

$$C_{r,r+s} = C_{(0,0),s}$$

2-D Power Spectral Density

Let X_s be a zero mean wide sense stationary random process.

Define

$$\hat{X}_N(e^{j\mu}, e^{j\nu}) = \sum_{m=-N}^{N} \sum_{n=-N}^{N} X_{(m,n)} e^{j(m\mu+n\nu)}$$

- Then the power spectrum (i.e. energy spectrum per unit sample) is

$$\frac{1}{(2N+1)^2} \left| \hat{X}_N(e^{j\mu}, e^{j\nu}) \right|^2$$

The following limit does not converge!!

$$\lim_{N\to\infty} \frac{1}{(2N+1)^2} \left| \hat{X}_N(e^{j\mu}, e^{j\nu}) \right|^2$$

Intuition - The spectral estimate remains noisy as the window size increases.

Definition of Power Spectral Density

- Definition of **Power Spectral Density**

$$S_x(e^{j\mu}, e^{j\nu}) \stackrel{\triangle}{=} \lim_{N \to \infty} \frac{1}{(2N+1)^2} E\left[\left| \hat{X}_N(e^{j\mu}, e^{j\nu}) \right|^2 \right]$$

Expectation removes the noise.

Weiner-Khintchine Theorem

- For a wide sense stationary random process, the power spectral density equals the Fourier transform of the autocorrelation

$$S_x(e^{j\mu}, e^{j\nu}) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} R(m,n) e^{-j(m\mu + n\nu)}$$

where

$$R(m,n) = E[X_{(0,0)} X_{(m,n)}]$$

Filtered Random Processes

- Consider the 2-D linear system

$$Y(m,n) = h(m,n) * X(m,n)$$

  where $X(m,n)$ is a 2-D wide sense stationary random process.

- It may be easily shown that

$$\begin{aligned} R_y(m,n) &= h(m,n) * h(-m,-n) * R_x(m,n) \\ S_y(e^{j\mu}, e^{j\nu}) &= \left| H(e^{j\mu}, e^{j\nu}) \right|^2 S_x(e^{j\mu}, e^{j\nu}) \end{aligned}$$

White Noise

- Definition:
  - $X(m,n)$ - independent identically distributed (i.i.d.) Gaussian random variables with distribution $N(0,1)$.
- Then
  - $X(m,n)$ is wide sense stationary with

$$\begin{aligned}
\mu(m,n) &= 0 \\
R_x(k,l) &= E[X(0,0)X(k,l)] \\
&= \delta(k,l) \\
S_x(e^{j\mu}, e^{j\nu}) &= DSFT\{R_x(k,l)\} \\
&= 1
\end{aligned}$$

Filtered White Noise

- Definitions:
  - $X(m,n)$ - independent identically distributed (i.i.d.) Gaussian random variables with distribution $N(0,1)$.
  - $Y(m,n) = h(m,n) * X(m,n)$.
- Then
  - $Y(m,n)$ is wide sense stationary with

$$
\begin{aligned}
S_y(e^{j\mu}, e^{j\nu}) &= \left|H(e^{j\mu}, e^{j\nu})\right|^2 S_x(e^{j\mu}, e^{j\nu}) \\
&= \left|H(e^{j\mu}, e^{j\nu})\right|^2 \cdot 1 \\
R_y(k,l) &= h(m,n) * h(-m,-n) \\
&= \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} h(m,n) h(m+k, n+l)
\end{aligned}
$$

  - $R_y(k,l)$ is the autocorrelation of $h(m,n)$ with itself.

Causal Prediction

- Y_s is a 2-D wide sense stationary Gaussian random process.
- Define
  - The past values are $Y_{<s} = \{Y_r : r < s\}$.
  - The minimum mean squared error (MMSE) predictor of Y_s given the past is

$$\hat{Y}_s = E[Y_s | Y_{<s}]$$

  - The prediction error is $X_s = Y_s - \hat{Y}_s$.

Properties of Causal Predictors

- Fact 1: (WLOG, assume $r < s$.)

$$\begin{aligned}
E[X_s X_r] &= E[E[X_s X_r | Y_{<s}]] \\
&= E[E[(Y_s - \hat{Y}_s)(Y_r - \hat{Y}_r) | Y_{<s}]] \\
&= E[E[(Y_s - \hat{Y}_s) | Y_{<s}](Y_r - \hat{Y}_r)] \\
&= E[(E[Y_s | Y_{<s}] - \hat{Y}_s)(Y_r - \hat{Y}_r)] \\
&= E[(\hat{Y}_s - \hat{Y}_s)(Y_r - \hat{Y}_r)] \\
&= E[0(Y_r - \hat{Y}_r)] = 0
\end{aligned}$$

- Fact 2: $\sigma^2 \triangleq E[X_s^2]$ is the prediction variance.
- Fact 3: The predictor must be space invariant and linear.

$$\hat{Y}_s = \sum_{r > (0,0)} h_r Y_{s-r}$$

Autoregressive (AR) Processes

- Definitions:
  - Y_s - 2-D wide sense stationary Gaussian random process.
  - h_s - MMSE linear predictor for Y_s.
  - $X_s = Y_s - h_s * Y_s$ - predictor error.

- If h_s is FIR, then Y_s is known as an autoregressive (AR) process.

Properites of AR Processes

- Remember that

$$X_s = Y_s - h_s * Y_s$$

- Then
  - We know that

$$Y(e^{j\mu}, e^{j\nu}) = \frac{1}{1 - H(e^{j\mu}, e^{j\nu})} X(e^{j\mu}, e^{j\nu})$$

  - Since X_s is white noise,

$$\begin{aligned} R_x(s) &= \sigma^2 \delta(s) \\ S_x(e^{j\mu}, e^{j\nu}) &= \sigma^2 \end{aligned}$$

  - So the power spectrum of Y_s is given by

$$S_y(e^{j\mu}, e^{j\nu}) = \frac{\sigma^2}{\left|1 - H(e^{j\mu}, e^{j\nu})\right|^2}$$

Spectral Estimate Using AR Processes

- Compute MMSE linear predictor $\hat{h}_s$ for Y_s.
- Compute the prediction variance

$$\hat{\sigma}^2 = \frac{1}{|S|} \sum_{s \in S} |Y_s - h_s * Y_s|^2$$

  where S is a finite set of points in plain, and $|S|$ is the number of points in S.

- Estimate the power spectrum

$$S_y(e^{j\mu}, e^{j\nu}) = \frac{\hat{\sigma}^2}{\left|1 - \hat{H}(e^{j\mu}, e^{j\nu})\right|^2}$$

- Can produce a more accurate estimate of the power spectrum.

Generating AR Processes

- Select a causal prediction filter h_s.
- Apply IIR filter to white noise random process X_s

$$Y(e^{j\mu}, e^{j\nu}) = \frac{1}{1 - H(e^{j\mu}, e^{j\nu})} X(e^{j\mu}, e^{j\nu})$$

- Y_s is sometimes referred to as a white noise driven process.
- Do linear FIR prediction filters $\hat{h}_s$ always form a stable IIR filter?
  - In 1-D, yes.
  - In 2-D, not always!
- Other problems:
  - Causal ordering of points may cause asymmetric artifacts in results.
  - Complexity increases rapidly with IIR filter order P.