Stochastic signal basics

CDF stochastic variable ξ is described by $F_\xi(x)$ – the probability that $\xi > x$; $F_\xi(x)$ is a Cumulative Density Function

PDF Probability Density Function is more intuitive, defined as $f_\xi(x) = \frac{dF_\xi(x)}{dx}$

signal or process: $\xi(t)$ – a set of all possible realizations $x(t)$

process value at the moment t_1, $\xi(t_1)$ is a stochastic variable, described with its PDF $f_\xi(x(t_1);t_1)$

$\longrightarrow$ for the full description of $\xi(t)$ we need all the possible multidimensional PDF's $f_\xi(x(t_1),\, x(t_2),\, \ldots\, ;\, t_1,\, t_2,\, \ldots)$

practical view: we narrow our interest to the two-dimensional PDF $f_\xi(x(t_1),\, x(t_2);\, t_1,\, t_2)$ to be able to tell the relation between the process values at two points in time.

- *mean value* (expectation)

$$\mu_\xi(t) \;=\; E\left[\xi(t)\right] \;=\; \int_{-\infty}^{\infty} x f_\xi(x;t)dx$$

- *mean square (MS) value* (mean power)

$$P_\xi(t) \;=\; E\;\left[\xi^2(t)\right] \;=\; \int_{-\infty}^{\infty} x^2 f_\xi(x;t)dx$$

- *variance* (mean power of variable component)

$$\sigma_\xi^2(t) \;=\; E\;left\{[(\xi \;-\; \mu_\xi(t)]^2 right\} \;=\; \int_{-\infty}^{\infty}[x \;-\; \mu_\xi(t)]^2\; f_\xi(x;t)dx$$

- *autocorrelation function (ACF)*

$$R_{\xi\xi}(t_1,t_2) \;=\; E[\xi(t_1)\xi(t_2)] \;=\; \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1x_2\; f_{\xi_1\xi_2}(x_1,x_2;t_1,t_2)dx_1dx_2$$

- *autocovariance function*

$$C_{\xi\xi}(t_1,t_2) \;=\; E\;\left\{[\xi(t_1) \;-\; \mu_\xi(t_1)]\,[\xi(t_2) \;-\; \mu_\xi(t_2)]\right\}$$

Stationarity (wide sense)

$$\begin{array}{rcl} \mu_\xi(t) &=& \mu_\xi \;=\; const \\ R_{\xi\xi}(t_1,\,t_2) &=& R_{\xi\xi}(\tau), \quad \tau \;=\; t_2 - t_1 \end{array} \quad scalesym200\}$$

Ergodicity *hypothesis*

$$\mu_\xi \;=\; <x> \;=\; \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} x(t)dt$$

$$R_{\xi\xi}(\tau) \;=\; <x(t)x(t+\tau)> \;=\; \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} x(t)x(t+\tau)dt$$

Power spectrum density (PSD)

$$S_{\xi\xi}(\omega) \;=\; \int_{-\infty}^{\infty} R_{\xi\xi}(\tau)e^{-j\omega\tau}d\tau$$

$$R_{\xi\xi}(\tau) \;=\; \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{\xi\xi}(\omega)\, e^{\,j\omega\tau}d\omega$$

Discrete stochastic signal

$\xi[n]$ – a sequence of stochastic variables $\xi(n)$

Complex signal

$$\xi[n] = \xi_R[n] + j\,\xi_I[n]$$

Example: complex filtering (convolution of complex signals)

$$h[n] = h_R(n) + jh_I[n]$$

$$\eta[n] = h[n] * \xi[n] \longrightarrow \eta(n) = \sum_{m=0}^{n} h(m)\xi(n-m)$$

$$\eta_R[n] = h_R[n] * \xi_R[n] - h_I[n] * \xi_I[n]$$
$$\eta_I[n] = h_I[n] * \xi_R[n] + h_R[n] * \xi_I[n]$$

- *Mean value*

$$\mu_\xi(n) = E\left[\xi(n)\right] = E\left[\xi_R(n)\right] + jE\left[\xi_I(n)\right] = \mu_{\xi_R}(n) + j\mu_{\xi_I}(n)$$

- *MS value* (mean power)

$$P_\xi(n) = E\left[\xi(n)\,\xi^*(n)\right] = E\left[|\xi(n)|^2\right]$$

- *Variance*

$$\sigma_\xi^2(n) = E\left\{\left[\xi(n) - \mu_\xi(n)\right]\left[\xi^*(n) - \mu_\xi^*(n)\right]\right\} = E\left[|\xi(n)|^2 - |\mu_\xi(n)|^2\right]$$

- *Autocorrelation*

$$R_{\xi\xi}(n_1, n_2) = E\left[\xi^*(n_1)\,\xi(n_2)\right]$$

- *Autocovariance*

$$C_{\xi\xi}(n_1, n_2) = E\left\{\left[\xi^*(n_1) - \mu_\xi^*(n_1)\right]\left[\xi(n_2) - \mu_\xi(n_2)\right]\right\}$$

Stationarity

$$\mu_\xi(n) = \mu_\xi = const$$
$$R_{\xi\xi}(n_1, n_2) = R_{\xi\xi}(m), \quad m = n_2 - n_1$$

Time-domain mean

$$< x[n] > = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} x(n)$$

Time-domain correlation

$$\psi_{xx}(m) = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} x^*(n)x(n+m)$$

Ergodicity

$$< x[n] > = \mu_\xi = const$$
$$\psi_{xx}(m) = R_{\xi\xi}(m), \; m = n_2 - n_1$$

Power spectrum density

For a stationary $\xi[n]$, $C_{\xi\xi}(n1,\ n2) = C_{\xi\xi}(m), \quad m = n_2 - n_1$.

$$C_{\xi\xi}(m) \ = \ E\ \left\{ [\xi^*(n) \ - \ \mu_\xi^*] \ [\xi(n \ + \ m) \ - \ \mu_\xi] \right\}$$

If our signal $\xi[n]$ is zero-mean: $\mu_\xi \ = \ 0$ (if not, use $\xi_1[n] \ = \ \xi[n] \ - \ \mu_\xi$), then $C_{\xi\xi}(m) \ = \ R_{\xi\xi}(m)$.

Power spectrum density of a stationary discrete signal $\xi[n]$ (MS convergent if σ_ξ^2 is bounded)

$$S_{\xi\xi}(\theta) \ = \ \sum_{m=-\infty}^{\infty} C_{\xi\xi}(m) e^{-jm\theta}$$

- Periodic (over 2π)
- (if $\xi[n]$ is real) $\longrightarrow S_{\xi\xi}(\theta) \leq 0$, symmetric
- $C_{\xi\xi}(m) \ = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_{\xi\xi}(\theta)\, e^{\,jm\theta} d\theta$

Autocovariance and PSD estimation

$\xi[n]$ (stationary, ergodic) $\longrightarrow$ estimation of properties from N (finite number) of samples of $x[n]$.

- mean value estimate

$$\hat{\mu}_\xi = \frac{1}{N}\sum_{n=0}^{N-1} x(n)$$

- Variance estimate

$$\hat{\sigma}_\xi^2 = \frac{1}{N}\sum_{n=0}^{N-1} [x^*(n) - \hat{\mu}_\xi^*]\,[x(n) - \hat{\mu}_\xi]$$

- Autocovariance estimate (equal to autocorrelation with $\mu_\xi = 0$)

$$\hat{R}_{\xi\xi}(m) = \frac{1}{N}\sum_{n=0}^{N-|m|-1} x^*(n)x(n+m), \quad -(N-1) \leq m \leq N-1$$ *(or, maybe $\frac{1}{N-|m|}$?)*

- PSD estimate

$$\hat{S}_{\xi\xi}(\theta) = \sum_{m=-(N-1)}^{N-1} \hat{R}_{\xi\xi}(m)e^{-jm\theta}$$

Estimation accuracy

Different **realizations** $x[n] \longrightarrow$ different estimates.
$\xi[n]$, $\hat{\mu}_\xi$, $\hat{\sigma}^2_\xi$, $\hat{R}_{\xi\xi}(m)$, $\hat{S}_{\xi\xi}(\theta)$ are stochastic $\longrightarrow$ How to measure the accuracy of estimate?

bias $B = \alpha - E[\hat{\alpha}]$

variance $var\,[\hat{\alpha}] = \sigma^2_{\hat{\alpha}} = E\,\{[\hat{\alpha}^* - E\,(\hat{\alpha})^*]\,[\alpha - E\,(\hat{\alpha})]\,\}$

MS error $E\,[|\hat{\alpha} - \alpha|^2] = B^2 + \sigma^2_{\hat{\alpha}}$

consistency $\lim_{N\to\infty} var\,[\hat{\alpha}] \to 0$ and $\lim_{N\to\infty} B\,[\hat{\alpha}] \to 0$

If $\xi[n]$ is stationary and gaussian . . .

- mean value estimate $\hat{\mu}_\xi = \frac{1}{N}\sum_{n=0}^{N-1} x(n) \longrightarrow$ unbiased, with variance σ^2_ξ/N
- variance estimate $\hat{\sigma}^2_\xi = \frac{1}{N}\sum_{n=0}^{N-1}\,[x^*(n) - \hat{\mu}^*_\xi]\,[x(n) - \hat{\mu}_\xi]$
  $\longrightarrow$ bias $B[\hat{\sigma}^2_\xi] = \sigma^2_\xi/N$, variance $var[\hat{\sigma}^2_\xi] \sim 1/N$ (consistent)

Autocovariance and PSD estimate properties

$\hat{R}_{\xi\xi}(m) = \frac{1}{N} \sum_{n=0}^{N-|m|-1} x^*(n)x(n+m)$ is biased:

$$E[\hat{R}_{\xi\xi}(m)] = \frac{N-|m|}{N} R_{\xi\xi}(m)$$

$$var[\hat{R}_{\xi\xi}(m)] \approx \frac{1}{N} \sum_{r=-\infty}^{\infty} \left[R^2_{\xi\xi}(r) + R_{\xi\xi}(r+m)R_{\xi\xi}(r-m) \right], \quad N \gg m$$

$\hat{S}_{\xi\xi}(\theta) = \sum_{m=-(N-1)}^{N-1} \hat{R}_{\xi\xi}(m)e^{-jm\theta}$:

$$E[\hat{S}_{\xi\xi}(\theta)] = \sum_{m=-(N-1)}^{N-1} \frac{N-|m|}{N} R_{\xi\xi}(m)e^{-jm\theta}$$

$$var[\hat{S}_{\xi\xi}(\theta)] = S^2_{\xi\xi}(\theta) \left\{ 1 + \left[\frac{\sin N\theta}{N \sin \theta} \right]^2 \right\}$$ very large, estimate not consistent!

Periodogram

$$\hat{R}_{\xi\xi}[m] = \begin{cases} \frac{1}{N}\, x^*[m] * x[-m], & |m| \le N-1 \\ 0, & |m| > N-1 \end{cases}$$

(in the following $x_1[m] = x[-m]$, and we transform N-sample sections of $x^*[n]$, $x_1[n]$)

$$\hat{S}_{\xi\xi}(\theta) = \frac{1}{N} \cdot X^*(e^{j\theta}) \cdot X_1(e^{j\theta})$$

$$X_1(\mathrm{e}^{j\theta}) = \sum_{n=-(N-1)}^{0} x_1(n)\mathrm{e}^{-jn\theta} = \sum_{n=0}^{N-1} x_1(-n)\mathrm{e}^{jn\theta} = \sum_{n=0}^{N-1} [x^*(n)\mathrm{e}^{-jn\theta}]^* = [X^*(e^{j\theta})]^*$$

finally

$$\hat{S}_{\xi\xi}(\theta) = \frac{1}{N} \cdot |X(\mathrm{e}^{j\theta})|^2$$

Practical implementations

- Choose FFT length to avoid cyclic efects
- Average K segments (of length $M = N/K$) to reduce variance ath the cost of bias (*Bartlett procedure*).
- Add overlapping of segments and use non-rectangular window (*Welch procedure*).

$$\hat{S}_W^i(\theta) \;=\; \frac{1}{MF}\left|\sum_{n=0}^{M-1} x^i(n)\, g(n) \mathrm{e}^{-\mathrm{j}n\theta}\right|^2, \quad (i \;=\; 1,2,\ldots,K \text{ is a segment number})$$

$$F \;= \frac{1}{M}\sum_{n=0}^{M-1} g^2(n) \;\; (\text{energetic normalizing factor})$$

$$\hat{S}_{W\xi\xi}(\theta) \;=\; \frac{1}{K}\sum_{i=1}^{K} \hat{S}_W^i(\theta)$$