EDISP - random signals (former lect9..)
(English) Digital Signal Processing

February 10, 2014

Randomness

We describe as “random” effects that are too complex to precisely analyze in practice, or simply unknown:

- physical noise: thermal, mechanical, acoustic, radio/radar
- somebody’s decisions made from data unknown (to us): aircraft pilotage, human voice

But we know *something* about the constraints: system bandwidth, physical dependencies, vocal tract properties.
We describe:

value constraints as probability

dependencies as conditional probability $\longrightarrow$ correlation

Description of a random variable

ξ is a random variable taking some value $x \in X$; for each value x there is some probability that $\xi = x$.

We may imagine ξ as the ensemble of all possible values together with their probabilities, or as the "set of all possible experiments" (**realizations**.

Imagine the weather:

- what we see this year is **one realization** x
- what may happen during any year is random variable ξ
- it is a variable from a multidimensional space X of "whole yearly weather"

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

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

expectation or probabilistic mean, or mean value:
$\mu_\xi = E[\xi] = \int_{-\infty}^{\infty} x f_\xi(x) dx$

mean square (MS) value $P_\xi = E[\xi^2(t)] = \int_{-\infty}^{\infty} x^2 f_\xi dx$

variance (mean power of variable component)
$\sigma_\xi^2 = E\left[(\xi - \mu_\xi)^2\right] = \int_{-\infty}^{\infty} [x - \mu_\xi(t)]^2 f_\xi(x) dx$

covariance is a measure of relation between *two* variables ξ and η:

Stochastic signal basics (discrete time)

$\xi[n]$ – a sequence of stochastic variables $\xi(n)$ (as a DT signal $x[n]$ is a sequence of numbers $x(n)$)

Imagine the weather again, concentrating on a scalar value of “maximum temperature of a day”:

- what we see this year is **one realization** $x[n]$
- max temperature today is one sample $x(n)$ at $n =$today
- max temperature on 11 January is a random variable $\xi(n)$ at $n =$11 Jan
- max temperature each day of the year is a random signal/process $\xi[n]$

signal or process: $\xi[n]$ – a set of all possible realizations $x[n]$

realization: $x[n]$ one sequence, being particular member of the set $\xi[n]$

process value at the moment n_1, $\xi(n_1)$ is a stochastic variable, described with its PDF $f_\xi(x(n_1); n_1)$

$\longrightarrow$ for the full description of $\xi(n)$ we need all the possible multidimensional (joint) PDF's $f_\xi(x(n_1),\ x(n_2),\ \ldots\ ;\ n_1,\ n_2,\ \ldots)$

practical view: we narrow our interest to the two-dimensional PDF $f_\xi(x(n_1),\ x(n_2);\ n_1,\ n_2)$ to be able to tell the relation between the process values at two points in time.

For the stochastic (random) signal, we use the same description with expectation (called *mean*, or precisely *probabilistic mean*), MS value, variance, covariance... but they are in general dependent on time: $\mu_\xi(n)$, etc.

Complex signal

Complex signal

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

Example: complex filtering (convolution of complex signals)

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

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

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

We describe complex random signals with....

- *Mean value*

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

- *MS value* (mean power)

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

- *Variance*

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

- *Autocorrelation* is a measure of dependency between signal values in different time instants

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

- *Autocovariance*

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

Stationarity and ergodicity

Signal is **stationary** when it fulfills the following

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

Time-domain mean is defined as: $< x[n] > = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} x(n)$
Time-domain correlation is defined as:
$\psi_{xx}(m) = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} x^*(n)x(n+m)$
Signal is **ergodic** when it fulfills the following

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

In other words, *ergodicity* means that we can draw conclusions on *probabilistic* mean, variance, and autocorrelation from *time-domain* mean, mean power, autocorrelation. We usually have to assume the signal is ergodic, no simple test exists ....

Power spectrum density (PSD)

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: $\mu_\xi \ = \ 0$ then $C_{\xi\xi}(m) \ = \ R_{\xi\xi}(m)$
(if not, use $\xi_1[n] \ = \ \xi[n] \ - \ \mu_\xi$).
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) $\longrightarrow S_{\xi\xi}(\theta) \geq 0$, 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

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

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

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

- PSD estimate

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

Estimation accuracy

The actual values μ_ξ, $\xi[n]$, $\hat{\mu}_\xi$, $R_{\xi\xi}(m)$, $S_{\xi\xi}(\theta)$ are constant (=not random)
From different **realizations** $x[n]$ we obtain different estimates.
Estimates $\hat{\sigma}_\xi^2$, $\hat{R}_{\xi\xi}(m)$, $\hat{S}_{\xi\xi}(\theta)$ are random $\longrightarrow$ How to measure the accuracy of estimate?

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

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

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

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

Periodogram is a method to estimate PSD that is faster!
As the ACF is estimated from the convolution

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

we may rewrite $\hat{S}_{\xi\xi}(\theta)$ using transforms (in the following $x_1[m] = x[-m]$)

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

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

finally

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

Further: we can transform N-sample sections of $x^*[n]$, $x_1[n]$ and then average periodograms, reducing variance.

Practical implementations of periodogram

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

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

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

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

Filtering of random signals

$$
\begin{array}{ccc|c|ccc}
\text{process} & x_1[n] \longrightarrow & & & & \longrightarrow y_1[n] & \text{process} \\
\xi[n] & x_2[n] \longrightarrow & & H(z) & & \longrightarrow y_2[n] & \eta[n] \\
 & \ldots \longrightarrow & & & & \longrightarrow \ldots &
\end{array}
$$

For a stationary $\xi[n]$:

- mean value:

$$\mu_\eta = \mu_\xi \sum_{n=-\infty}^{\infty} h(n) = \mu_\xi \, H(e^{j\theta})|_{\theta=0} \tag{1}$$

- autocorrelation

$$R_{\eta\eta}(m) = \sum_{i=-\infty}^{\infty} R_{\xi\xi}(m-i)\, v(i) \quad \text{where} \quad v(i) = \sum_{k=-\infty}^{\infty} h(k)\, h(k+i) \tag{2}$$

- power spectrum density

$$S_{\eta\eta}(\theta) = S_{\xi\xi}(\theta) |H(e^{j\theta})|^2 \tag{3}$$

Applications

- Analysis of AD conversion errors
- Analysis of arithmetic errors in filters
- Signal modelling $\longrightarrow$ compression (LPC):
  - on the compression side, a filter is tuned so that noise passed through it has the properties of the signal being compressed;
  - only filter coefficients are transmitted;
  - signal is reconstructed from noise passed through a filter
- System modelling $\longrightarrow$ identification
- Signal detection $\longrightarrow$ matched filter (presence, time of arrival);
  - a filter matched to signal $x(n)$ has the impulse response $h(n) = x(M - n)$; such a filter maximizes the S/N ratio if the noise is gaussian and white
  - if the noise is not white, we may *whiten* it with another filter (*whitening filter*): we may *model* the noise as if it was white, but then filtered with a filter with transfer function $G(z)$; then the whitening filter will be $1/G(z)$....