Signal Analysis Lectures

Lecture II: Stochastic process, Time series, Autcorrelation and Power Spectral Density

Elena Cuoco

European Gravitational Observatory
and
Scuola Normale Superiore

May 12, 2020

Summary

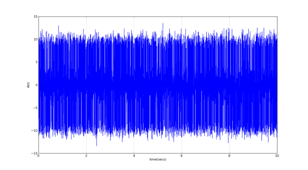
1. Stochastic process: time series
   - Time series, Autocorrelation function, Spectral Density
   - Not parametric modeling, Classical estimators
   - The Periodiogram

Stochastic process: time series

Time series, Autocorrelation function, Spectral Density

Stochastic process and Time Series.

A stochastic discrete process $x[n]$ is a sequence of random variable for each value of n. If n represents the time, we call it time series. A time series is a sequence of data points measuring a physical quantity at successive times spaced at uniform time intervals.



We say that $x[n]$ is a stationary process, if its statistical description does not depend on n.
That is the moments $E(x^{k_0}[n_0], x^{k_1}[n_1], ...x^{k_M}[n_M])$ do not depend on the value of x_M but only on the distance L between the value of $x[n]$ and $x[n+L]$.

Gaussian process

We define a Gaussian stocastich process if $\{x[n_0], x[n_1] \ldots x[n_{N-1}]\}$ has a multivariate distribution. If we assume that the process is stationary and with zero mean, the covariance matrix is the same to the correlation matrix $\mathbf{r}_{xx}$

$$\mathbf{r}_{xx} = \begin{bmatrix} r_{xx}[0] & r_{xx}[-1] & \ldots & r_{xx}[-(N-1)] \\ r_{xx}[1] & r_{xx}[0] & \ldots & r_{xx}[-(N-2)] \\ \vdots & \vdots & \ddots & \vdots \\ r_{xx}[N-1] & r_{xx}[N-2] & \ldots & r_{xx}[0] \end{bmatrix}, \tag{1}$$

where

$$r_{xx}[k] = \mathcal{E}\{x^*[n]x[n+k]\} \ . \tag{2}$$

If samples are taken at a later time to generate the vector $\mathbf{x} = (\mathbf{x}[\mathbf{n_0}], \mathbf{x}[\mathbf{n_1}] \ldots \mathbf{x}[\mathbf{n_{N-1}}])^{\mathbf{T}}$, we can write the density function of probability of a Gaussian random process and real as

$$P[\mathbf{x}] = \frac{\mathbf{1}}{(\mathbf{2}\pi)^{\mathbf{N}/\mathbf{2}} |\mathbf{r_{xx}}|^{\mathbf{1}/\mathbf{2}}} \mathbf{e}^{\mathbf{x}^{\mathbf{T}} \mathbf{r_{xx}^{-1}} \mathbf{x}} . \qquad (3)$$

Any linear operation that acts on a Gaussian stochastic process produces a still gaussianly distributed process.
For a Gaussian stochastic process the statistics of the second order, i.e. moment one and moment two, is a full description of the process. If we then consider real gaussian stochastic stationary processes with zero mean the correlation function provides all the information necessary to describe the process.

Autocorrelation function

The correlation function is connected to the power spectrum $S(f)$ through the Wiener–Khintchine theorem:

$$S_{xx}(f) = \sum_{k=-\infty}^{k=\infty} r_{xx}[k] \exp(-2i\pi f k) \,, \tag{4}$$

where the spectrum power is defined as

$$S_{xx}(f) = \lim_{M \to \infty} \mathcal{E} \left[\frac{1}{2M+1} \left| \sum_{-M}^{M} x[n] \exp(-i2\pi f n) \right|^2 \right] \,. \tag{5}$$

So a zero mean stationary gaussian stochastic process can be described indifferently with the correlation function or with the spectral density.

Many discrete stochastic processes are achieved as a result of a sampling of a continuous stochastic process over time $x(t)$, where in order to avoid problems with *aliasing*[1] we sample the continuous process $x(t)$ with a period $T_s = \frac{1}{f_s}$, where f_s is the maximum frequency present in the spectrum of $x(t)$.
The relationship between the continuous stochastic process $x(t)$ and the series stochastic sampled discreetly $x[n]$ and given by the formula of Shannon's interpolation

$$x(t) = \sum_{n=-\infty}^{\infty} x[n] \frac{\sin \pi f_s (t - nT_s)}{\pi f_s (t - nT_s)}\,, \tag{6}$$

dove

$$x[n] = x(nT_s) \qquad \text{per} \qquad -\infty < n < \infty\,. \tag{7}$$

[1] aliasing is a distortion that occurs in spectrum reconstruction that has not been sampled at the maximum frequency present in the spectrum and therefore presents spectral frequencies that do not belong to the original spectrum

Power Spectral Density

If the process spectrum $x(t)$ is limited bandwidth i.e. if $S(f) = 0$ for $|f| > \frac{f_s}{2}$ then the function of autocorrelation over time of the discrete stochastic process is given by

$$r_{xx}[n] = r(nT_s) = \int_{-f_s/2}^{f_s/2} S(f) e^{i2\pi f T_s n} df \, . \tag{8}$$

The discrete power spectrum is periodic

$$S_{xx}(f) = S_{xx}(f + kf_s) \qquad k = 0, \pm 1, \pm 2, \ldots \tag{9}$$

and equal to the continuous power spectrum in the range of Nyquist

$$S_{xx}(f) = S(f) \qquad \frac{-f_s}{2} \leq f \frac{f_s}{2}, . \tag{10}$$

Stochastic process: time series

Not parametric modeling, Classical estimators

Not parametric methods

We want to resolve now the issues related to the search for the process $x[n]$ if we know the power spectrum $S(f)$ and vice versa the estimate of $S(f)$ if we know the process $x[n]$, using not parametric methods. To this end it is convenient to introduce the complex spectrum of the discrete time series

$$\mathcal{S}(z) = T_s \sum_{n=-\infty}^{\infty} r_{xx}[n] z^{-n} . \tag{11}$$

The one that we wrote in the transformed z of the autocorrelation function $r_{xx}[n]$. The transformed z is the equivalent of Laplace's transformation for discrete timing sequences, where $z = e^{-i2\pi f}$.

If we can write $\mathcal{S}(z)$ in the form

$$\mathcal{S}(z) = T_s \sum_{n=-\infty}^{\infty} r_{xx}[n] z^{-n} = T_s \mathcal{H}(z) \mathcal{H}(1/z^*)\,, \tag{12}$$

then we can generate the discrete time series $x[n]$ filtering white noise as follows

$$x[n] = \sum_{m=0}^{\infty} h[m] w[n-m]\,, \tag{13}$$

where $h[m]$ is the filter impulse function (z):

$$h[m] = \int_{-f_s/2}^{f_s/2} H(f) e^{i2\pi f T_s m} df \tag{14}$$

and $w[n]$ a gaussian white noise at zero mean and variance T_s

$$\mathcal{E}\{w[n]\} = 0 \qquad \mathcal{E}\{w[n]w[n']\} = T_s \delta[n-n']\,. \tag{15}$$

Spectral Factorization

Let's consider the problem of finding the $\mathcal{H}(z)$ filter that let us to simulate the noise data $x[n]$, starting with the knowledge of the spectrum $\mathcal{S}(z)$. Assuming that the signal in input is a white noise at medium zero and unitary variance, the $\mathcal{H}(z)$ transfer function must satisfy

$$\mathcal{S}(z) = \mathcal{H}(z)\mathcal{H}(1/z^*)\,. \tag{16}$$

The problem of the factoring of $\mathcal{S}(z)$ does not have a single solution because if $\mathcal{H}(z)$ is a solution, then also

$$\mathcal{H}'(z) = \pm z^{-k}\mathcal{H}(z) \tag{17}$$

$$\mathcal{H}''(z) = \mathcal{H}(1/z^*) \tag{18}$$

are solutions.

It is possible to show that, apart from an ambiguity in the sign, we can find a spectral factorization in which the filter $\mathcal{H}(z)$ and its inverse $1/\mathcal{H}(z)$ are causal and stable. For this to be possible, it is necessary that the $S(f)$ spectrum meets Paley-Wiener's condition:

$$\int_{-1/2}^{1/2} |\ln S(f)| df < \infty \,. \quad (19)$$

This condition does not allow the spectrum $S(f)$ to have extended regions along the f axis where $S(f)$ is zero. So when we perform filtering operations on the spectrum we have to be careful we don't send any part of it to zero if we want to then build the causal and stable filter over time that reproduces the spectrum.

The Paley-Wiener condition is automatically satisfied if we do the spectral factoring in the following way. Starting from the complex spectrum $\mathcal{S}(z)$ we call *cepstrum* complex the quantity

$$\hat{\mathcal{S}}(z) = \ln(\mathcal{S}(z)) = \sum_{n=-\infty}^{\infty} \hat{r}[n] z^{-n} \,, \tag{20}$$

where

$$\hat{r}[n] = \int_{-1/2}^{1/2} \hat{S}(f) e^{i2\pi f n} df = \int_{-1/2}^{1/2} \hat{S}(-f) e^{i2\pi f n} df = \hat{r}[-n] \,. \tag{21}$$

Introducing the causal part of $\hat{\mathcal{S}}(z)$ defined as

$$\hat{\mathcal{S}}_+(z) = \frac{1}{2}\hat{r}[0] + \sum_{n=1}^{\infty} \hat{r}[n] z^{-n} \,, \tag{22}$$

we can write

$$\hat{\mathcal{S}}(z) = \hat{\mathcal{S}}_+(z) + \hat{\mathcal{S}}_+(1/z^*) \,. \tag{23}$$

Then the canonical factorization of $S(z)$ is given by

$$\mathcal{H}(z) = \exp \hat{\mathcal{S}}_+(z) \,, \tag{24}$$

where $\mathcal{H}(z)$ has no zeros or poles outside the unitary circle in the complex plan.
This ensures that the $\mathcal{H}(z)$ impulse response function both a stable and causal filter and in the time domain given by

$$h[n] = \int_{-1/2}^{1/2} H(f) e^{i2\pi f n} df \ . \tag{25}$$

The periodogram

When we have the discrete experimental data available in the domain of time $x[n]$ we want to estimate the power spectral density (PSD) of the process.
A non-parametric estimate of the PSD is obtained through the periodogram defined as

$$P_{\text{PER}} = \frac{1}{N} \left| \sum_{n=0}^{N-1} x[n] \exp(-i2\pi f n) \right|^2 . \quad (26)$$

The periodogram satisfies

$$P_{\text{PER}}(f) \to P(f)$$

for a N number of data long enough ($N \to \infty$), but its variance doesn't tend to zero for $N \to \infty$; in particular the variance is constant regardless of the value of N. So the periodogram is a flimsy estimation of the spectrum of the process, being the standard deviation of the estimator large as its mean.

In order to improve the quality of the statistics of the periodogram, we introduce an averaged periodoram defined as

$$P_{\text{averaged}} = \frac{1}{K}\sum_{m=0}^{K-1} P_{\text{PER}}^{(m)}(f)\,, \tag{27}$$

where the periodogram is calculated on a data length $L = N/K$. In this way the variance decreases by a factor $1/K$. In this operation, however, you pay the price of increasing the *bias*, since you use shorter data set. In fact, it can be seen that the average value of the periodogram is given by

$$\mathcal{E}\{P_{\text{PER}}\} = \int_{-1/2}^{1/2} W_B(f-\nu)P(\nu)d\nu\,, \tag{28}$$

where P is the PSD of the process and $W_B(f)$ is the Fourier transform of Bartlett's window

$$w_B[k] = \begin{cases} 1 - \frac{|k|}{L} & \text{per} \quad |k| \leq L - 1 \\ 0 & \text{per} \quad |k| > L \end{cases} , \tag{29}$$

that is

$$W_B(f) = \frac{1}{L} \left(\frac{\sin \pi f L}{\sin \pi f} \right) . \tag{30}$$

A value of L shorter corresponds to a window of Bartlett narrower and therefore at a main peak in $W_B(f)$ wider. So as the value of L decreases the spectral resolution because we are not able to to resolve PSD details smaller than $1/L$. It is not possible therefore to have both a small variance and a small bias; what we can get is only a compromise between the two requests.

References

Kay S 1988 *Modern spectral estimation:Theory and Application* Prentice Hall Englewood Cliffs

Haykin S 1996*Adaptive Filter Theory* (Upper Saddle River: Prentice Hall)

Alexander S T 1986 *Adaptive Signal Processing* (Berlin: Springer-Verlag)

Hayes M H 1996*Statistical Digital Signal Processing and Modeling* Wiley

Therrien C W 1992*Discrete Random Signals and Statistical Signal Processing* (Englewood Cliffs: Prentice Hall)

Zubakov L A Wainstein V D 1962 *Extraction of signals from noise* (Englewood Cliffs: Prentice Hall)