JACOBS UNIVERSITY

Fatemeh Hadaeghi

EEG- Signal Processing

Lecture Notes for BSP, Chapter 5

Master Program Data Engineering

5 Introduction

The complex patterns of neural activity, both in presence and absence of external stimulation, can be recorded by means of EEG recording system with an acceptable temporal resolution. Depending on the functional state of the brain, various types of EEG waveforms, ranging from isolated events (e.g., spikes) to a complex composite signal pattern (i.e., mixture of waveforms) can be distinguished. Different pathological conditions such as physical brain injuries and epileptic seizures can also be characterized by detecting deviations from the normal EEG waves corresponding to alpha or beta activity. For diagnostic purposes, therefore, functional status of the brain needs to be evaluated based on the frequency, amplitude, morphology, and spatial distribution of the observed brain waves. Consequently, in order to design a system for EEG classification, the extraction of relevant signal features is crucially important. The most common technique for extracting characteristic features from an EEG signal is *spectral analysis* which aims at analyzing *spectral power* in different frequency bands. In general, previously developed approaches to spectral analysis of stationary and non-stationary signals can be considered in two branches: model-based (also known as parametric) spectral analysis and non-parametric (mainly, Fourier-based) spectral analysis. In model-based methods for EEG spectral analysis, it is of particular importance to simulate a real EEG signal by artificial models. Therefore, this chapter begins with a brief introduction to linear and non-linear models adopted in EEG signal processing. Then, the characteristics of the most frequently occurring EEG disturbances and artifacts are described and some methods for noise reduction or cancellation are briefly explained. Afterwards, examples of standard methods for spectral analysis of stationary and non-stationary signals are discussed.

The main references for this chapter are:

[1] Cohen, Arnon. Biomedical Signal Processing: Time and frequency domains analysis, Volume I. CRC-Press, 1986.

[2] Sörnmo, Leif, and Pablo Laguna. Bioelectrical signal processing in cardiac and neurological applications. Vol. 8. Academic Press, 2005.

[3] Pardey, James, Stephen Roberts, and Lionel Tarassenko. "A review of parametric modelling techniques for EEG analysis." Medical engineering & physics 18.1 (1996): 2-11.

5.1 Modeling the EEG signal

One of important objectives in biomedical signal processing is mathematical *signal modeling* and *simulation* which provides a critical tool to understand the complex temporal phenomenology of physiological processes. Modeling of the EEG signals is a wide, well-established academic area divided into two branches: linear and non-linear models. In this section, after a brief discussion on stochastic or deterministic nature of EEG recordings, the most frequently used linear and non-linear models previously adopted in EEG signal processing are briefly explained.

5.1.1 Deterministic and Stochastic Signals

As stated in Chapter 1, the objective and constraints of the problem at hand will determine to consider a vital signal as random or deterministic. EEG signal is often viewed as a realization of a stochastic process and is modeled as a random signal. From this point of view, even if the "pure" EEG (which reflects only the cerebral activity) had had deterministic properties, it is always corrupted by random noise, introduced in acquisition system and digitizing process. Therefore, it is reasonable to consider the EEG as a stochastic process.

Under certain conditions such as before and during epileptic seizure, however, techniques developed for characterizing chaotic behavior in the EEG have shown that it is better to model the EEG signal as a chaotic deterministic process. From this perspective, EEG is generated by a nonlinear dynamical system, hence, it can be characterized by a deterministic process which exhibit chaotic behavior, analogous to that of a stochastic process.

Although in this course we model EEG signal as a stochastic signal, it is worth mentioning that the deterministic/stochastic issue is controversial in current research literature. This review paper gives the interested readers a good insight into nonlinear dynamical EEG analysis.

5.1.2 Stochastic Properties

When a stochastic signal description is adopted as the primary approach to EEG modeling and analysis, it is of interest to determine a joint Probability Density Function (PDF) which characterizes this process. Let's assume, $\boldsymbol{x} = (x[0], x[1], \dots, x[N-1])^T$ is column vector representing the EEG signal in an arbitrary observation interval, $[0, N-1]$. Then the probability density function which characterizes this process needs to be estimated, i.e.:

$$p(\boldsymbol{x};\boldsymbol{\theta}) = p(x[0], x[1], \ldots, x[N-1]; \boldsymbol{\theta}). \qquad (5\text{-}1)$$

The elements of parameter vector, $\boldsymbol{\theta}$, are unknown deterministic values which determine the specific shape of the function and provide quantitative information on various signal properties.

In order to estimate the probability density function, $p(\boldsymbol{x};\boldsymbol{\theta})$, a nonparametric approach includes two main steps: 1) to compute the amplitude histogram of the recorded EEG samples and 2) to *guess* the particular structure of $p(\boldsymbol{x};\boldsymbol{\theta})$. In fact, EEG signal is assumed to be a sample function of a stationary and ergodic random process such that it is sufficient to hypothesis about the particular structure of $p(\boldsymbol{x};\boldsymbol{\theta})$ by having only one sample function from the ensemble.

Another approach is to assume that based on some neurophysiological insight, the structure of the PDF is known *a priori* and one only needs to estimate the parameter vector, $\boldsymbol{\theta}$ from the recorded data. However, due to the ever-changing properties of the EEG, a highly complex PDF structure is required to precisely model signals corresponding to normal and abnormal brain states.

From an engineering perspective, the EEG signal is often assumed to be characterized by a *multivariate Gaussian PDF*. But, is this assumption biologically plausible? The EEG is often viewed as the *summation* of signals generated by a large number of individual neural oscillators. Based on the well-known central limit theorem, the sum of *these independent random variables* in the limit has a Gaussian PDF if the number of summed variables are sufficiently large; the original variables themselves are not required to be Gaussian (normally distributed) for this theorem to hold. But, the next question is whether activities generated by individual neural oscillators can be considered as *independent random variables*. In particular, regarding the fact that neural oscillators are organized in groups with significant internal interactions (the group organization is believed to play a major role in producing the synchronous activity that is visible in the EEG). On the other hand, in higher level, groups of neural generators may still act independently, suggesting that the requirements of the central limit theorem may still be valid. This question has been addressed in several experimental studies aimed at investigating how accurately the amplitude of the EEG samples could be described by a Gaussian PDF. These classes of studies revealed that during synchronized activity, such as during the presence of alpha rhythm, the EEG exhibits Gaussian behavior, whereas during the REM sleep or mental arithmetic task it deviates from a Gaussian distribution. In general, when the measurement interval was increased, the amplitude distribution became increasingly non-Gaussian.

Despite the controversial opinions on whether the EEG signal can be characterized by a multivariate Gaussian PDF or not, the PDF is still considered to be Gaussian since, as we will see later, spectral analysis has a natural connection to this particular distribution.

The multivariate Gaussian PDF of a stochastic process $\boldsymbol{x}[n]$ is characterized by its mean value, $m_{x_n} = \mathcal{E}\{\mathbf{x}_n\}$, and the correlation function, $\varphi_{xx}\,[n_1, n_2] = \mathcal{E}\{x[n_1]x[n_2]\}$ which reflects the dependence between the two samples. The correlation function (also known as autocorrelation function) is a symmetric function, i.e., $\varphi_{xx}\,[n_1, n_2] = \varphi_{xx}\,[n_2, n_1]$. Using vector and matrix notations, the mean vector and the correlation matrix are, respectively, defined by:

$$\boldsymbol{m}_x = \mathcal{E}\{\boldsymbol{x}\} \tag{5-2}$$

and

$$\boldsymbol{R}_x = \mathcal{E}\{\boldsymbol{x}\boldsymbol{x}^{\boldsymbol{T}}\} = \begin{bmatrix} \varphi_{xx}\,[0,0] & \varphi_{xx}\,[0,1] & \cdots & \varphi_{xx}\,[0, N-1] \\ \varphi_{xx}\,[1,0] & \varphi_{xx}\,[1,1] & \cdots & \varphi_{xx}\,[1, N-1] \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{xx}\,[N-1,0] & \varphi_{xx}\,[N-1,1] & \cdots & \varphi_{xx}\,[N-1, N-1] \end{bmatrix}. \tag{5-3}$$

Therefore, multivariate Gaussian PDF is described as:

$$p(\boldsymbol{x}; \boldsymbol{m}_x, \boldsymbol{C}_x) = \frac{1}{(2\pi)^{\frac{N}{2}} |\boldsymbol{C}_x|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{m}_x)^T \boldsymbol{C}_x^{-1}((\boldsymbol{x} - \boldsymbol{m}_x))\right) \tag{5-4}$$

where the matrix $\boldsymbol{C}_x$ describes the covariance and is related to the correlation matrix $\boldsymbol{R}_x$:

$$\boldsymbol{C}_x = \mathcal{E}\{(\boldsymbol{x} - \boldsymbol{m}_x)(\boldsymbol{x} - \boldsymbol{m}_x)^T\} = \boldsymbol{R}_x - \boldsymbol{m}_x {\boldsymbol{m}_x}^T. \tag{5-5}$$

When $x[n]$ is a zero-mean process, (i.e., $\boldsymbol{m}_x = 0$), $\boldsymbol{C}_x$ and $\boldsymbol{R}_x$ are identical. With this Gaussian model, it is clear that the correlation matrix, $\boldsymbol{R}_x$, contains the essential information on signal properties and plays an important role in the analysis of EEG signals. However, it is difficult to estimate $\boldsymbol{R}_x$ from a single realization of the data. In fact, as only two samples, i.e., $x[n_1]$ and $x[n_2]$ are considered in $\varphi_{xx}\,[n_1, n_2]$, the resulting estimation will have a large variance, unless several realization are available and can be used for ensemble averaging. In practice, the EEG is assumed to have certain restrictive properties, for example, by viewing the signal as a stationary ergodic process. A stationary process is a stochastic process whose statistical properties are not a function of time. Stationary process are convenient since for such a process we can calculate, for example, the expectation by averaging the values, $x[n]$, overall the ensembles at any time, n. A particular class of stationary signals are called wide-sense stationary for which: i) the mean function is equal to a constant for all time instances such that $\boldsymbol{m}_x[n] = \boldsymbol{m}_x$ and ii) the correlation function, $\varphi_{xx}\,[n_1, n_2]$ is a function of the time lag $k = n_1 - n_2$ between the samples $x[n_1]$ and $x[n_2]$, i.e., $\varphi_{xx}\,[n, n-k] = \varphi_{xx}\,[k]$. The lag-dependent correlation function is denoted by:

$$\varphi_{xx}\,[k] = \mathcal{E}\{x[n]x[n-k]\}\,. \qquad (5\text{-}6)$$

The corresponding correlation matrix is, therefore, defined by:

$$\boldsymbol{R}_x = \begin{bmatrix} \varphi_{xx}\,[0] & \varphi_{xx}\,[-1] & \cdots & \varphi_{xx}\,[-N+1] \\ \varphi_{xx}\,[1] & \varphi_{xx}\,[0] & \cdots & \varphi_{xx}\,[-N+2] \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{xx}\,[N-1] & \varphi_{xx}\,[N-2] & \cdots & \varphi_{xx}\,[0] \end{bmatrix}. \qquad (5\text{-}7)$$

This correlation matrix is symmetric for a real valued process (since, $\varphi_{xx}\,[k] = \varphi_{xx}\,[-k]$) and is Toeplitz (since all elements in a given diagonal are identical, and equal to the correlation function at a certain time lag).

For this class of signals, the first two moments of the process, namely, mean and covariance, (i.e., those which define the Gaussian PDF in Eq. (5-4)) are needed to be taken into account. Spectral analysis is intimately related to the Gaussian distribution since, as stated in Chapter 3, the power spectral density or, more briefly, the power spectrum is defined as the discrete-time Fourier transform (DTFT) of the correlation function $\varphi_{xx}\,[k]$.

In EEG analysis, normal spontaneous activity over relatively short time intervals is stationary and can be subjected to power spectral analysis. However, in long time periods, it is necessary to view the EEG signal as a non-stationary, stochastic process with time-varying mean, correlation function, and higher-order moments. Variation in the degree of wakefulness of the subject over time, opening and closing the eyes, occurrence of transient waveforms such as epileptic spikes can cause the properties of the alpha rhythm to change slowly or rhythmic activity to undergo abrupt changes. Dealing with the non-stationarity of the EEG, specific algorithmic approaches have been presented to process EEG in the above mentioned cases. For instance, in case of slowly time varying properties due to variation in the degree of wakefulness of the subject over time, a method, originally developed for stationary signals, is repeatedly applied to consecutive, overlapping intervals. Time- frequency methods such as Wigner-Ville distribution is an example of these methods. Another approach is to develop a parametric model of the EEG signal and to estimate its parameter values recursively using an adaptive algorithm. In case of non-stationarity due to the abruptly changing activity, EEG signal is firstly decomposed into a series of variable-length, quasi-stationary segments. Each individual segment is then characterized by means of its power spectrum and related spectral parameters.

5.1.3 Linear Stochastic Models

EEG signal is typically modeled by a *parametric* model that is described by a transfer function (or filter) similar to the one explained in Eq. (3-25). That is the sampled EEG signal, $x[n]$, is assumed to be the output of a linear system driven by an unknown, inaccessible input sequence, $u[n]$, which is corrupted by an additive noise, $\zeta[n]$. The block diagram of such system is depicted in Figure 5-1. The sequence $x[n]$ is, therefore, given as the solution to the difference equation

$$x[n] = -\sum_{k=1}^{p} a_k x[n-k] + \sum_{m=0}^{q} b_m u[n-m] + \sum_{k=0}^{p} a_k \zeta[n-k]. \quad (5\text{-}8)$$

It is usually convenient to model the additive noise by means of a white noise sequence, with zero mean or consider it as the output of a noise filter driven by white noise, $w[n]$, such that

$$\sum_{k=0}^{s} c_k \zeta[n-k] = \sum_{k=0}^{t} d_k w[n-k]. \quad (5\text{-}9)$$

If we define the following operations

$$\begin{aligned} A(z^{-1}) &= \sum_{k=0}^{p} a_k z^{-k}\,;\ a_0 = 1 \\ B(z^{-1}) &= \sum_{k=0}^{q} b_k z^{-k} \\ C(z^{-1}) &= \sum_{k=0}^{s} c_k z^{-k} \\ D(z^{-1}) &= \sum_{k=0}^{t} d_k z^{-k} \end{aligned} \quad (5\text{-}10)$$

and transfer Eq. (5-8) and Eq. (5-9) to z-domain, we get:

$$X(z) = \frac{B(z^{-1})}{A(z^{-1})} U(z) + \frac{D(z^{-1})}{C(z^{-1})} W(z). \quad (5\text{-}11)$$

In Eq. (5-11) the sequence is modeled by means of the system parameter vector, $\boldsymbol{\beta}_x$ and the the noise parameter vector, $\boldsymbol{\beta}_w$, where:

$$\boldsymbol{\beta}_x^T = (a_0, a_1, \dots, a_p, b_0, b_1, \dots, b_q)$$

$$\boldsymbol{\beta}_w^T = (c_0, c_1, \dots, c_s, d_0, d_1, \dots, d_t) \tag{5-12}$$

The problem of identifying the parameters of the model when the input is available has been the subject of system identification researches and well covered in the literature. In signal processing modeling, therefore, the input sequence, $u[n]$, is assumed to be a white inaccessible sequence. The parameter vector $\boldsymbol{\beta}_x^T$ is thus describing a linear transformation, converting the white sequence into the signal sequence. We can decouple the transfer function model into the deterministic and noise models:

$A(z^{-1})Y(z) = B(z^{-1})U(z)$	Deterministic system
$C(z^{-1})\zeta(z) = D(z^{-1})W(z)$	Noise model
$X(z) = Y(z) + \zeta(z)$	Observation equation

Here the sequence $y[n]$ is the inaccessible pure (noise-free) output of the system.

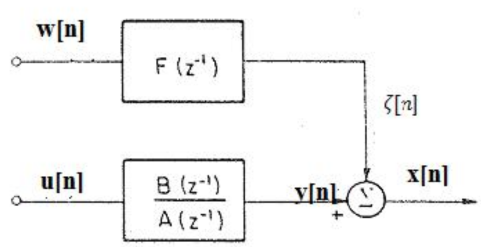

Figure 5-1: The transfer function model [1].

By means of the transfer function model of Eq. (5-11) several time series models have been derived:

- **the Autoregressive moving average exogenous variables (ARMAX) model.** By letting $A(z^{-1}) = C(z^{-1})$, the ARMAX model is written as

$$A(z^{-1})X(z) = B(z^{-1})U(z) + D(z^{-1})W(z) \tag{5-13}$$

or in terms of the difference equation

$$x[n] = -\sum_{k=1}^{p} a_k x[n-k] + \sum_{m=0}^{q} b_m u[n-m] + \sum_{k=0}^{t} d_k w[n-k]. \tag{5-14}$$

- **the autoregressive moving average (ARMA) model.** By assuming there is no external noise, the ARMA model is written as

$$X(z) = \frac{B(z^{-1})}{A(z^{-1})} U(z) \tag{5-15}$$

or in terms of the difference equation

$$x[n] = -\sum_{k=1}^{p} a_k x[n-k] + \sum_{m=0}^{q} b_m u[n-m]. \tag{5-16}$$

It is known as ARMA of order (p, q).

- **the autoregressive (AR) model.** It is an all pole model by assuming $B(z^{-1}) = G$ and there is no external noise considered in the model:

$$X(z) = \frac{G}{A(z^{-1})} U(z) \tag{5-17}$$

and

$$x[n] = -\sum_{k=1}^{p} a_k x[n-k] + G\, u[n]. \tag{5-18}$$

- **the moving average (MA) model.** It is an all zero model by assuming $A(z^{-1}) = 1$ and there is no external noise considered in the model:

$$X(z) = B(z^{-1}) U(z) \tag{5-19}$$

and

$$x[n] = \sum_{m=0}^{q} b_m u[n-m]. \tag{5-20}$$

In Equation (5-14), the current (n-th) sample of the sequence is expressed as linear combination of the past p sequence samples and $(q + 1)$ input samples. Hence, the models discussed here are also known as linear prediction models. Figure 5-2 depicts all these models in the z and in the time domains.

The AR and ARMA models are the most commonly used linear models in EEG modeling and various algorithms for the estimation of system parameters (without the need to have an access to the input) are available. Due to its simplicity, AR model will be discussed here in more derails.

The AR model is commonly used to model EEG signals, because of its simplicity and the effective algorithms which have been developed for the estimation of the AR parameters.

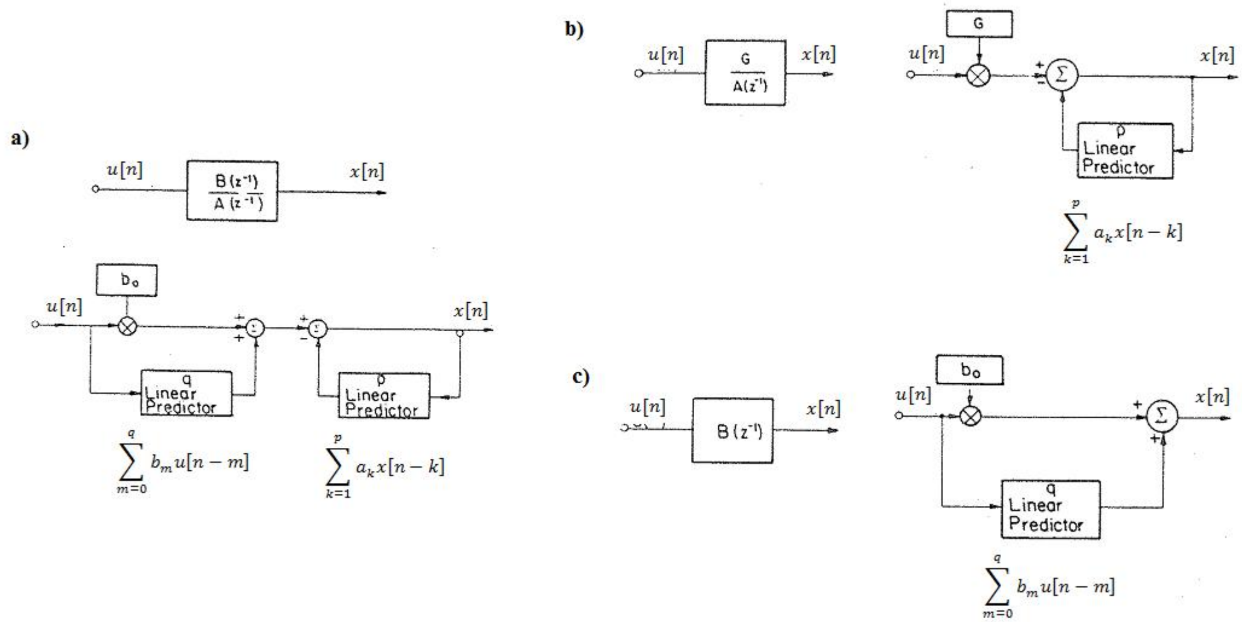

Figure 5-2: Block diagram of a) ARMA model, b) AR model and c) MA model in time domain and their z-transform [1].

When modeling a stationary process, one must make sure that the model is stationary and the system is stable. For a linear process, these are ensured if the complex roots of the characteristic equation

$$A(z^{-1}) = 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_p z^{-p} = 0 \tag{5-21}$$

are all inside the unit circle in the z-plane or outside the unit circle in the z^{-1}-plane.

The inverse filter, $H^{-1}(z^{-1}) = (\frac{1}{A(z^{-1})})^{-1} = A(z^{-1})$, is known as the AR *whitening filter*. Because, when the sequence $x[n]$ serves as the input to the AR whitening filter, the resultant output will have white spectrum.

The simplest AR process is that of the first order, (i.e., $p = 1$). It is known as the *Markov process*, given by the difference equation

$$x[n] = -a_1 x[n-1] + G\, u[n], \tag{5-22}$$

which is stationary for $|a_1| < 1$.

In general, the choice of the order p of the AR model depends on the accuracy required.

As later will be discussed in section 5.3, the AR model parameters contain essential spectral information on a rhythm required for spectral power estimation and finding the dominant frequency of the rhythm. The model order p, determines the number of peaks that can be present in the AR power spectrum.

Figure 5-3 shows AR modeling by presenting a recorded EEG with dominant alpha rhythm and a simulated signal. The simulated signal is obtained by filtering white noise with an all-pole filter whose parameters have been earlier estimated from the EEG signal.

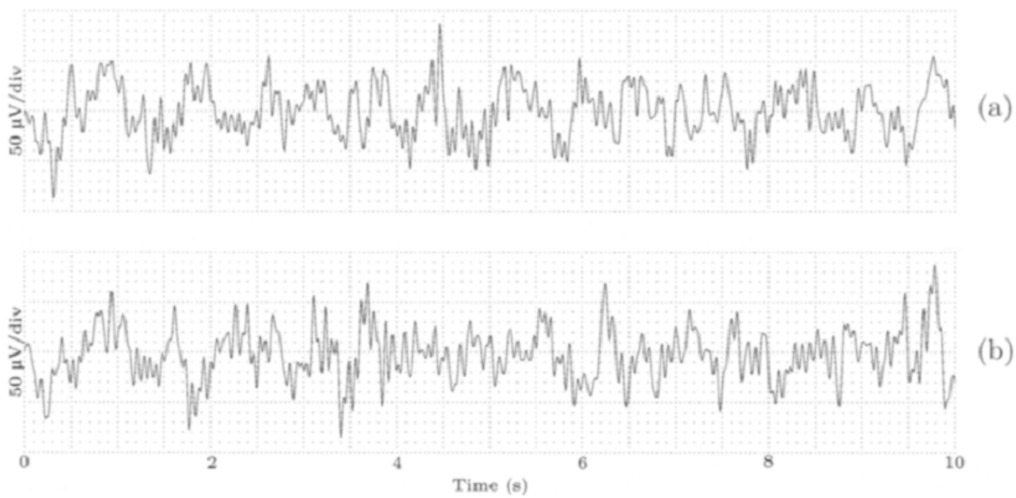

Figure 5-3: An EEG signal with dominant alpha rhythm and its simulated signal by using AR model [2].

Estimation of AR model- Least Square Method

It is required to estimate the order of the process, p, its coefficients, a_k at, $k = 1, 2, \dots, p$, and the gain factor, G. Since the input is inaccessible, the estimation must be performed without the sequence, $u[n]$.

Let's assume we have the samples, $x[j]$, for $j = 0, 1, \dots, (n - 1)$. We can estimate the incoming sample $x[n]$ by the estimator

$$\hat{x}[n] = -\sum_{k=1}^{\hat{p}} \hat{a}_k x[n-k] \qquad (5\text{-}23)$$

where a circumflex ($\hat{\bullet}$) denotes estimated value. For the time being, let's assume, p is given (for example, by guessing). At each time instance, n, we can calculate the error $e[n]$ (known as the "residual") between the actual sequence sample and the predicted one:

$$e[n] = \mathrm{x[n]} - \hat{x}[n] = \mathrm{x[n]} + \sum_{k=1}^{\hat{p}} \hat{a}_k x[n-k] \qquad (5\text{-}24)$$

The residuals, $e[n]$, are, indeed, the estimates of the inaccessible input, i.e., $G\,u[n]$. The least squares method determines the estimated parameters by minimizing the expectation of the squared error:

$$\min_{\hat{a}_k} \mathcal{E}\{e^2[n]\} = \min_{\hat{a}_k} \mathcal{E}\{(\mathrm{x[n]} + \textstyle\sum_{k=1}^{\hat{p}} \hat{a}_k x[n-k])^2\}. \tag{5-25}$$

Performing the minimization by

$$\frac{\partial \mathcal{E}\{e^2[n]\}}{\partial \hat{a}_k}, k = 1,2,\dots,\hat{p} \tag{5-26}$$

and assuming the sequence to be stationary, we get p linear equations:

$$\textstyle\sum_{j=1}^{\hat{p}} \hat{a}_j \varphi_{xx}[k-j] = -\varphi_{xx}[k]\,, k = 1,2,\dots,\hat{p}. \tag{5-27}$$

Where $\varphi_{xx}[k-j] = \mathcal{E}\{x[n-j]x[n-k]\} = \varphi_{xx}[j-k]$.

Equations (5-27) are known as the Yule-Walker equations or the normal equations. They can be solved for the least squares optimal parameters $\hat{a}_j$ for $j = 1, 2, \dots, \widehat{p}$, if the $\hat{p} + 1$ correlations, $\varphi_{xx}[j]$ for $j = 0, 1, \dots, \hat{p}$ are given.

The correlation coefficients are not given; hence, they have to be estimated from the given finite sequence, $\{x[n]\}$. Assume the sequence $\{x[n]\}$ is given for $n = 0, 1, 2, \dots, (N — 1)$. We can estimate the correlation coefficients by

$$\hat{\varphi}_{xx}[k] = \frac{1}{N} \textstyle\sum_{n=0}^{N-1-k} x[n]x[n+k]. \tag{5-28}$$

In Eq. (5-28), it is assumed that all samples of $x[n]$ are zero outside the given range. These estimations (Eq. (5-28)), known as the autocorrelation method, will be used instead of the correlation coefficients of Eq. (5-27). For sake of convenience, we shall continue to use the symbol, $\varphi_{xx}[n]$ where indeed, $\hat{\varphi}_{xx}[n]$ must be used. Eq.(5-27) can be denoted in a matrix form as

$$\begin{bmatrix} \varphi_{xx}[0] & \varphi_{xx}[1] & \varphi_{xx}[2] & \cdots & \dots & \varphi_{xx}[\hat{p}-1] \\ \varphi_{xx}[1] & \varphi_{xx}[0] & \varphi_{xx}[1] & & \dots & \varphi_{xx}[\hat{p}-2] \\ & \vdots & & \ddots & & \vdots \\ \varphi_{xx}[\hat{p}-2] & \dots & \cdots & \cdots & \varphi_{xx}[0] & \varphi_{xx}[1] \\ \varphi_{xx}[\hat{p}-1] & \cdots & \dots & & \varphi_{xx}[1] & \varphi_{xx}[0] \end{bmatrix}. \tag{5-29}$$

or as

$$\boldsymbol{R_x \hat{a}} = \boldsymbol{\varphi}_{xx}. \tag{5-30}$$

where the correlation matrix $\boldsymbol{R_x}$, vector $\boldsymbol{\varphi}_{xx}$. and the AR coefficients vector $\widehat{\boldsymbol{a}}$ are defined in Eq. (5-30). The direct solution of Eq. (5-30) is given by inversion of the correlation matrix

$$\widehat{\boldsymbol{a}} = \boldsymbol{R}^{-1}\boldsymbol{\varphi}_{xx}. \tag{5-31}$$

Note that the correlation matrix is symmetric and in general, positive semi-definite. Hence, efficient algorithms for the solution of Eq. (5-30) have been developed. An efficient recursive procedure, for example, was suggested by Durbin. For more details, interested readers are referred to this paper.

Different variations of AR model can be summarized as follows:

- Time-varying AR modeling

  By replacing the fixed model parameters by their time varying counterparts, the issue of non-stationarity can be handled within the context of AR modeling. To be clearer, the Eq. (5-18) can be re-written as:

  $$x[n] = -a_1[n]x[n-1] - a_2[n]x[n-2] - \cdots - a_p[n]x[n-p] + G\,u[n]. \tag{5-32}$$

  Regarding the fact that the temporal evolution of $a_k[n]$ is unknown, in general, the usefulness of this model extension is limited to the situations where slowly changing spectral properties are expected. In this case, the parameters are estimated by means of adaptive estimation algorithms.
- Multivariate AR modeling

  Another extension of the linear model is the multivariate model which is mainly useful to study the spatial interaction between different regions of the brain. For the AR model in Eq. (5-18) we can write the following multivariate difference equation:

  $$\boldsymbol{x}[n] = -\boldsymbol{A_1}x[n-1] - \boldsymbol{A_2}[n]x[n-2] - \cdots - \boldsymbol{A_p}x[n-p] + G\,\boldsymbol{u}[n]. \tag{5-33}$$

  where $\boldsymbol{x}[n]$ is an M-dimensional column vector that contains the samples at time n of the M channels, and $(\boldsymbol{A_1}, \boldsymbol{A_2}, \ldots, \boldsymbol{A_p})$ are $M \times M$ matrices that together describe temporal as well as spatial correlation properties across the scalp.
- AR modeling with impulse input

  In order to modeling the transient events such as K complexes and vertex waves, observed in sleep recordings, it has been proposed to mix the input white noise with a train of isolated impulses at random occurrence times. Figure 5-4 shows an example of the output of such signal model.

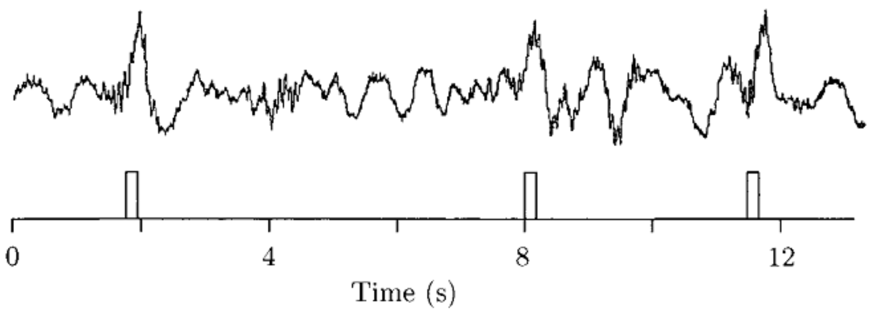

Figure 5-4: Simulation of a "sleep EEG" with transient events by driving the AR model with a mixture of white noise and pulse [2].

5.1.4 Nonlinear EEG modeling

Although the linear models, with either time-invariant or time-varying parameters, have been successfully used in many EEG applications, these models cannot replicate all types of signal patterns, nor do they provide deeper insight into the underlying mechanisms of EEG generation. As a result, nonlinear simulation models have been developed in order to better understand the underlying generation process.

An important way forward in understanding the brain dynamics is to integrate classic neuroscience (imaging and neurotransmitter findings) with recent advances in the mathematics of nonlinear systems. In particular, the neuronal system composed of different neuron populations is exquisitely sensitive to external forcing and internal shifts in functional parameters strongly suggesting that an adequate explanation will include the dynamical properties of chaotic systems.

As this topic is beyond the scope of this course, the interested readers are invited to take a look at this paper.