Lab 8: Parametric Spectrum Estimation

DSP Practical
Signal Processing Group
Technische Universität Darmstadt

December 16, 2014

1 Introduction

In the last practical, we introduced the problem of spectrum estimation. In order to determine the spectral content of a random signal, i.e., the distribution of power over frequency, we used non-parametric techniques. We have seen that these rely mainly on a smoothed version of the squared magnitude of the Fourier transform of the random observations.

Another approach is to use a parametric model to describe the random process and therewith the spectrum. This reduces the spectral estimation problem to that of only estimating the parameters of the assumed model. The estimated spectrum is then formed by inserting the estimated parameters into the parameterized spectrum. If the model assumption is correct, parametric modeling usually achieves better accuracy and resolution compared with non-parametric methods.

In the following, we will review the basics of parametric modeling, especially autoregressive (AR) and moving average (MA) random processes. We will discover how the model order can be determined using information theoretic criteria. Finally, we will have a look at speech signal analysis, which is a typical application of AR modeling.

1.1 Parametric Modeling

Let $X(n)$ be a zero-mean, wide-sense stationary random process, i.e.,

$$\mathrm{E}\{X(n)\} = 0 \quad \text{and} \quad c_{XX}(k) = \mathrm{E}\{X(n+k)X^*(n)\}$$

where $c_{XX}(k)$ is the covariance function. The power spectrum is defined as the Fourier transform of the covariance function:

$$C_{XX}\left(e^{j\omega}\right) = \sum_{k=-\infty}^{\infty} c_{XX}(k) e^{-j\omega k}.$$

When using a parametric approach, we model our stochastic process $X(n)$ as the output of a system $H(z)$, where the input $Z(n)$ is a white noise process. This is depicted in Figure 1.

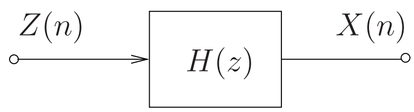

Figure 1: Parametric modeling of $X(n)$, input $Z(n)$ is white noise.

Recall that a white noise process has a covariance function $c_{ZZ}(k) = \sigma_Z^2 \delta(k)$, and a constant power spectral density $C_{ZZ}(e^{j\omega}) = \sigma_Z^2$. The transfer function of our system model is assumed to be

$$H(z) = \frac{b_0 + b_1 z^{-1} + \cdots + b_q z^{-q}}{a_0 + a_1 z^{-1} + \cdots + a_p z^{-p}}$$

where the a_k coefficients are usually normalized such that $a_0 = 1$. Depending on the number of a_k and b_k coefficients, we distinguish between

- $q = 0$, the filter is recursive, or so-called all-pole; $X(n)$ is modeled as an AR process.
- $p = 0$, the filter is a non-recursive, or so-called all-zero; $X(n)$ is modeled as an MA process.
- q and p are greater than zero, the filter is recursive, pole-zero; $X(n)$ is modeled as an ARMA process.

The spectrum of $X(n)$ is then generally given by

$$C_{XX}\left(e^{j\omega}\right) = \left|H\left(e^{j\omega}\right)\right|^2 C_{ZZ}\left(e^{j\omega}\right) = \left|\frac{b_0 + b_1 e^{-j\omega} + \cdots + b_q e^{-j\omega q}}{1 + a_1 e^{-j\omega} + \cdots + a_p e^{-j\omega p}}\right|^2 \sigma_Z^2. \quad (1)$$

We observe that the parametric spectrum so obtained is fully described by the filter coefficients b_k for $k = 0, \ldots, q$; a_k for $k = 1, \ldots, p$; and the variance of the white noise process σ_Z^2.

2 Parameter Estimation

We continue by further characterizing AR and MA processes, and how to estimate the filter parameters based on the estimated covariance function.

2.1 Autoregressive Process

The input-output relationship of an AR model can be expressed as

$$X(n) + \sum_{k=1}^{p} a_k X(n-k) = Z(n)$$

where $Z(n)$ is a white noise process with variance σ_Z^2. The spectrum of the AR process $X(n)$ is then given by

$$C_{XX}\left(e^{j\omega}\right) = \frac{\sigma_Z^2}{\left|1 + \sum_{k=1}^{p} a_k e^{-j\omega k}\right|^2}.$$

Using the Yule-Walker equations, it is possible to estimate the AR parameters a_k for $k = 1, \ldots, p$ from the covariance function

$$c_{XX}(l) + \sum_{k=1}^{p} a_k c_{XX}(l-k) = \begin{cases} \sigma_Z^2, & l = 0 \\ 0, & l = 1, \ldots, p \end{cases}.$$

In MATLAB, the function `aryule` does this job.

2.2 Moving Average Process

If the input-output relationship were instead

$$X(n) = Z(n) + \sum_{k=1}^{q} b_k Z(n-k),$$

then $X(n)$ is a moving average (MA) process. The spectrum becomes

$$C_{XX}\left(e^{j\omega}\right) = \sigma_Z^2 \left|1 + \sum_{k=1}^{q} b_k e^{-j\omega k}\right|^2. \tag{2}$$

The covariance function can be expressed in terms of the MA parameters

$$c_{XX}(l) = \begin{cases} \sigma_Z^2 \sum_{k=0}^{q-|l|} b_{k+|l|} b_k, & l = 0, \ldots, q \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

with $b_0 = 1$. It can be shown that (3) results in a system of non-linear equations and that there is no unique solution.

However, it can be shown that, given certain assumptions, it is possible to estimate the MA parameters. Say we are given estimates of the covariance function $\hat{c}_{XX}(-q), \ldots, \hat{c}_{XX}(q)$. The estimated spectrum is then given by

$$\hat{C}_{XX}(z) = \sum_{n=-q}^{q} \hat{c}_{XX}(n) z^{-n}. \tag{4}$$

Based on (2), the estimated spectrum can be decomposed as follows:

$$\hat{C}_{XX}(z) = \sigma_Z^2 \hat{B}(z) \hat{B}\left(\frac{1}{z}\right) \quad \text{with} \quad \hat{B}(z) = 1 + \sum_{k=1}^{q} \hat{b}_k z^{-k}.$$

We now assume the filter with polynomial $B(z)$ to be minimum phase, i.e., all the zeros of $z^q \hat{B}(z)$ lie inside the unit circle. The zeros of $\hat{B}(1/z)$, on the other hand, are mirrored versions with respect to the unit circle. Using this assumption, we can estimate the MA parameters $b_1, \ldots, b_q$ as the coefficients of the polynomial $B(z)$. Specifically, this can be done as follows:

1. Find the coefficients of polynomial $P(z) = z^q \hat{C}_{XX}(z)$.
2. Determine the roots of $P(z)$, denoted z_i for $i = 1, \ldots, 2q$.
3. Select all z_i which satisfy $|z_i| < 1$ and assign them to $\hat{B}(z)$.

2.3 Estimation of the Covariance Function

The covariance function of a zero-mean, stationary random process $X(n)$ can be estimated using a realization of $X(n)$ for $n = 0, 1, \ldots, N-1$:

$$\hat{c}_{XX}(m) = \begin{cases} \dfrac{1}{N} \displaystyle\sum_{n=0}^{N-m-1} X(n+m) X^*(n), & m \geq 0 \\ \hat{c}_{XX}^*(-m), & m < 0 \end{cases}.$$

The scaling with $1/N$ results in a biased estimate. By replacing $1/N$ with $1/(N - |m|)$, we obtain an unbiased estimate but with higher variance for large lags m. Therefore, the estimator presented here is usually preferred. The estimation of the covariance function or the correlation function can be done in MATLAB using `xcorr`.

3 Order Selection

For AR processes, we have presented the Yule-Walker equations, which relate the covariance function to the process parameters. By using the sample autocorrelation function, estimates of these parameters can be found. However, evaluating these expressions requires knowledge of the number of system parameters, i.e., the model order.

In order to estimate the model order, a number of criterion functions can be used, such as Akaike's Information Criterion (AIC) or the Minimum Description Length (MDL) criterion. These are defined as follows:

$$\begin{aligned}
\mathrm{AIC}(m) &= \ln \hat{\sigma}^2_{Z,m} + m \cdot \frac{2}{N} \\
\mathrm{MDL}(m) &= \ln \hat{\sigma}^2_{Z,m} + m \cdot \frac{\ln N}{N}
\end{aligned}$$

where $\hat{\sigma}^2_{Z,m}$ is the estimated variance of the residuals and N is the size of the data sample used. The variance of the residuals for an AR process of order m can be estimated as follows

$$\hat{\sigma}^2_{Z,m} = \frac{1}{N-m} \sum_{n=m}^{N} \left\{ X(n) + \sum_{k=1}^{m} \hat{a}_k X(n-k) \right\}^2$$

where $\hat{a}_1, \ldots, \hat{a}_m$ are the estimated AR parameters and the term in curly braces is referred to as the residual process. By evaluating the AIC or MDL criterion for different orders m, the order that minimizes the criterion is selected as the model order p.

Generally, as m increases, the variance of the residuals will decrease, that means fitting larger and larger models will usually lead to a smaller $\sigma^2_{Z,m}$, regardless of the true model order. To offset this, penalty functions (the second terms in AIC and MDL) are added. These functions increase with m. The combined effect is that an order is chosen that has a small residual variance, while considering the order-complexity tradeoff.

4 Application

A typical application for AR modeling is in the area of speech signal analysis, where an all-pole model is assumed to describe the behavior of the human

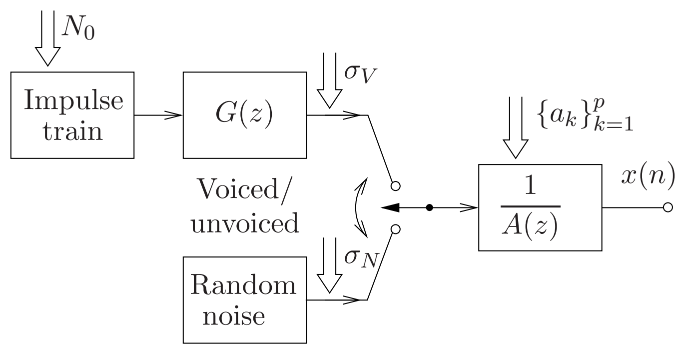

Figure 2: Discrete-time model for speech.

vocal tract. More specifically, we use the discrete-time model for speech, as displayed in Figure 2.

Our speech signal $x(n)$ is regarded as the output of an autoregressive filter, with parameters $\{a_k\}_{k=1}^p$, modeling the vocal tract. A switch is used to model the change of "character" of speech segments:

- For unvoiced speech, the excitation of the AR filter is simply random noise.
- Voiced speech, on the other hand, is a result of the quasi-periodic opening and closing of the vocal cords (glottis). This is modeled by a combination of the impulse train generator with pitch period N_0 and the glottal pulse filter $G(z)$.

The resulting spectral envelope of a speech signal is given by the frequency response of the vocal tract, which is

$$H\left(e^{j\omega}\right) = \frac{1}{1 + \sum_{k=1}^{p} a_k e^{-j\omega k}}.$$

The spectral envelope describes the structure of the spectrum for unvoiced speech quite well. When looking at the spectrum of voiced speech, we have additionally a fine comb structure which is due to the quasiperiodic excitation. We note that the described speech parameters, especially AR parameters $\{a_k\}_{k=1}^p$, are time-varying, i.e., non-stationary, and standard methods do not apply. However, it is common in speech processing to assume quasi-stationarity for short segments with a duration of approximately $20 - 40$ ms. Sometimes, for carefully picked examples, the duration of stationary intervals can be extended.

5 Preparation

1. Try to describe the general principle of parametric spectrum estimation in simple words.

2. Assume an AR(p) random process $X(n)$, whose covariance function has values $c_{XX}(k)$ for $k = 0, 1$ and 2. The white noise variance is $\sigma_Z^2 = 1$.

   (a) Take $p = 1$ and use the Yule-Walker equations to calculate a_1.

   (b) Now take $p = 2$, and calculate a_1 and a_2.

3. Assume an MA(1) random process $X(n)$, whose covariance function has values $c_{XX}(k)$ for $k = 0, 1$. For σ_Z^2, calculate b_1 directly using (3). Describe the problems that will occur for higher orders.

4. Explain the concept behind model order selection.

6 Experiments

6.1 AR process

The following AR(3) process is considered

$$X(n) = Z(n) - 0.5X(n-1) - 0.7X(n-2) - 0.2X(n-3).$$

1. Use commands `filter` and `randn` to generate 1024 samples of $X(n)$.

2. From the generated process, estimate the AR parameters using the built-in AR-Yule-Walker function, `aryule`. Assume the correct model order is known.

3. Use the estimated parameters and plot the spectrum using `freqz`.

4. Compare these results with those obtained through `pyulear` and the spectrum computed from the true AR parameters.

6.2 MA process

You are given 1024 samples of the MA(3) process contained in `ma3.mat`, generated according to

$$X(n) = Z(n) + 0.4Z(n-1) - 0.2Z(n-2) + 0.15Z(n-3).$$

1. Write a routine `maparam` that can be used to estimate the parameters of the MA process. You may use MATLAB commands `xcorr` to estimate the covariance function; and `roots` and `poly` to convert between roots and polynomial representation.

2. Plot the spectrum estimate using the estimated parameters from above, as well as the true spectrum. Comment on your results.

6.3 Order Selection

We wish to determine the order of the AR process contained in the file `arunknown.mat`. Note that the order is between 1 and 10.

1. Use the AIC and MDL to estimate the order of the process. For this, evaluate both criteria for orders $m = 1, 2, \ldots, 10$.

2. Plot the AIC and MDL against the order.

3. Where are these functions minimized? Discuss the plots and results.

6.4 Speech Signal Analysis

The file `s5.mat` contains the utterance "Oak is strong and also gives shade" sampled at 8 kHz. Two segments are to be analyzed: the phonemes SH and AA in "shade" from sample 15600-16300 and 16800-17500, respectively.

1. Load the signal `s5.mat` into your workspace. Extract the two segments SH and AA in "shade". Listen to the signals and display them.

2. Assume that the phonemes of speech segments can be modeled by an autoregressive random process. Estimate the parametric spectrum using an AR(10) model. Compare the result with a windowed averaged periodogram. For this, use command `pwelch` with default settings, i.e., $L = 8$ segments with 50% overlap and a Hamming window.

3. Based on Section 4, how could you determine whether your speech segment is voiced or unvoiced?

4. Now assume the model order is unknown. Use the AIC method to determine the model and re-estimate the parametric spectrum. Comment on your results.