4F7 Spectrum Estimation
Parametric Methods

Sumeetpal Singh
Email : sss40@eng.cam.ac.uk

Modification history:
$\leq$ 2008: S. Godsill, P. Rayner
Revised by S.S. Singh

- We have seen that periodogram-based methods can lead to biased estimators with large variance

- Parametric methods assume a model for the physical process which generated the data, e.g. an ARMA model

- The aim is to estimate the parameters of the assumed model from the observed data

- The choice of the model to be used can be informed by the power spectral density plot (e.g. estimated by the periodogram)

- Recall the result: If a random process $\{X_n\}$ can be modelled as white noise exciting a filter with frequency response $H(e^{j\omega})$ then its spectral density is

$$\boxed{S_X(e^{j\omega}) = \sigma^2 \left|H(e^{j\omega})\right|^2}$$

where σ^2 is the variance of the white noise process. [It is usually assumed that $\sigma^2 = 1$ and the scaling is incorporated as gain in the frequency response]

- The frequency response $H(e^{j\omega})$ of the ARMA model can be represented
by a finite number of parameters which are then to be estimated from the data

- (Example: PSD of the AR(P) process.) Let $X_n = aX_{n-P} + W_n$, $|a| < 1$ and $E\left\{W_n^2\right\} = \sigma^2$.

$$S_X(e^{j\omega}) = \frac{\sigma^2}{\left(1 - ae^{j\omega P}\right)\left(1 - ae^{-j\omega P}\right)}$$
$$= \frac{\sigma^2}{1 + a^2 - 2a\cos(\omega P)}$$

which has period $2\pi/P$.

- (A cautionary note.) Parametric models need to be chosen carefully - an inappropriate model for the data can give misleading results

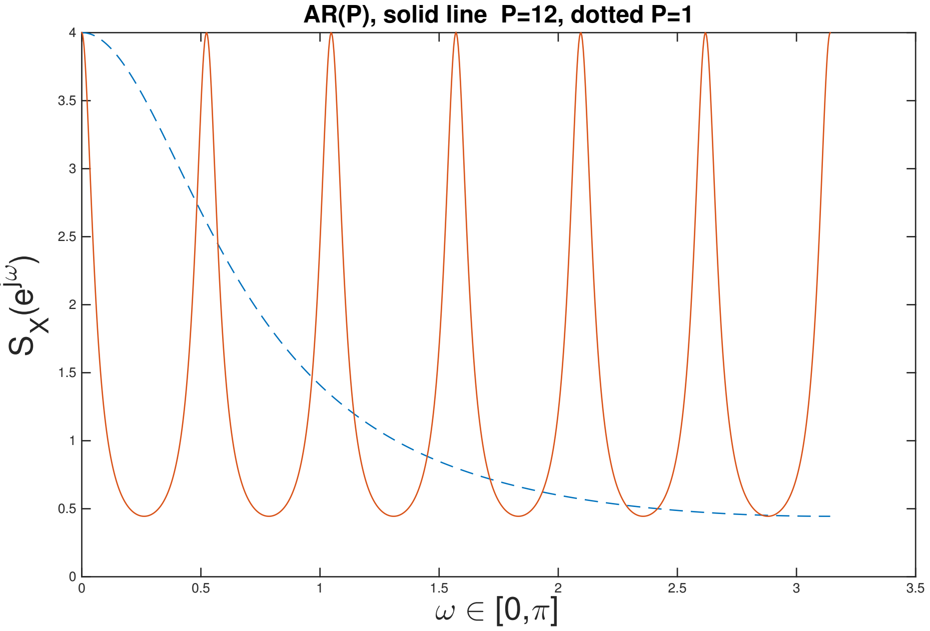

Figure 1: Parameters $\sigma = 1$, $a = 0.5$

1 ARMA Models

A quite general representation is the autoregressive moving-average (ARMA) model:

- The ARMA(P,Q) model difference equation representation is:

$$\boxed{x_n = -\sum_{p=1}^{P} a_p\, x_{n-p} + \sum_{q=0}^{Q} b_q\, w_{n-q}} \qquad (1)$$

where:

$$\begin{array}{ll} a_p & \text{are the AR parameters,} \\ b_q & \text{are the MA parameters} \end{array}$$

and $\{W_n\}$ is white noise with unit variance, $\sigma^2 = 1$.

- Clearly the ARMA model is a pole-zero IIR filter-based model with transfer function

$$H(z) = \frac{B(z)}{A(z)}$$

where

$$A(z) = 1 + \sum_{p=1}^{P} a_p z^{-p}, \quad B(z) = \sum_{q=0}^{Q} b_q z^{-q} \tag{2}$$

- Unless otherwise stated we will always assume that the filter is stable, i.e. the poles (solutions of $A(z) = 0$) all lie within the unit circle to ensure the ARMA process is WSS and has a causal representation

- Using the above result, the power spectrum of the ARMA process is:

$$S_X(e^{j\omega}) = \frac{|B(e^{j\omega})|^2}{|A(e^{j\omega})|^2}$$

- The ARMA model is quite a flexible and general way to model a stationary random process

- The spectrum can be factored as

$$\frac{B(z)}{A(z)} = b_0 \frac{\prod_{i=1}^{Q}(1 - z^{-1}c_i)}{\prod_{i=1}^{P}(1 - z^{-1}d_i)}$$

- The spectrum can be manipulated by choosing Q, P, $\{c_i\}_{i=1}^{Q}$, $\{d_i\}_{i=1}^{P}$ subject to $|d_i| < 1$. (As an exercise, plot $\log_{10} \frac{|B(e^{j\omega})|}{|A(e^{j\omega})|}$ in the interval $\omega \in [0, 2\pi)$ in Matlab.)

- The poles model well the peaks in the spectrum (sharper peaks implies poles closer to the unit circle)

- The zeros model troughs in the spectrum

2 Autocorrelation function of the ARMA Model

- The autocorrelation function $R_{XX}[r]$ for the output x_n of the ARMA model is:

$$R_{XX}[r] = E[x_n x_{n+r}]$$

- Substituting for x_{n+r} from equation 1 gives:

$$\begin{aligned} &R_{XX}[r] \\ &= E\left[x_n \left\{-\sum_{p=1}^{P} a_p\, x_{n+r-p} + \sum_{q=0}^{Q} b_q\, w_{n+r-q}\right\}\right] \\ &= -\sum_{p=1}^{P} a_p\, E[x_n x_{n+r-p}] + \sum_{q=0}^{Q} b_q\, E[x_n w_{n+r-q}] \end{aligned}$$

- The white noise process $\{W_n\}$ is wide-sense stationary so that $\{X_n\}$ is also wide-sense stationary
- Let the system impulse response be

$$x_n = \sum_{m=-\infty}^{\infty} h_m\, w_{n-m}$$

The system is causal, i.e. $h_m = 0$ for $m < 0$

- Therefore

$$E[x_n\, w_{n+k}] = E[w_{n+k} \sum_{m=-\infty}^{\infty} h_m\, w_{n-m}]$$

- For a white process

$$E[w_{n+k}\, w_{n-m}] = \begin{cases} \sigma^2 & \text{if } m = -k \\ 0 & \text{otherwise} \end{cases}$$

and let $\sigma^2 = 1$ (without loosing generality.) Hence $E[x_n\, w_{n+k}]$ is independent of n and let

$$R_{XW}[k] = E[x_n\, w_{n+k}] = h_{-k}$$

- Therefore $R_{XX}[r]$ satisfies the same ARMA difference equation that related x_n and w_n:

$$\boxed{R_{XX}[r] = -\sum_{p=1}^{P} a_p\, R_{XX}[r-p] + \sum_{q=0}^{Q} b_q\, R_{XW}[r-q]} \tag{3}$$

- A more convenient expression for the cross-correlation term $R_{XW}[.]$ is needed.

- Substituting this expression for $R_{XW}[k]$ into equation (3) gives the Yule-Walker Equation for an ARMA process,

$$\boxed{R_{XX}[r] = -\sum_{p=1}^{P} a_p \, R_{XX}[r-p] + \sum_{q=0}^{Q} b_q \, h_{q-r}} \tag{4}$$

- Since the system is causal, or $h_m = 0$ for $m < 0$, equation (4) may be rewritten as:

$$\boxed{R_{XX}[r] = -\sum_{p=1}^{P} a_p \, R_{XX}[r-p] + c_r} \tag{5}$$

where:

$$c_r = \begin{cases} \sum_{q=r}^{Q} b_q \, h_{q-r} & \text{if } 0 \leq r \leq Q \\ 0 & \text{if } r > Q \\ \sum_{q=0}^{Q} b_q \, h_{q+|r|} & \text{if } r < 0 \end{cases} \tag{6}$$

- Note that equation (5) expands to

$$
\begin{aligned}
R_{XX}[0] + \sum_{p=1}^{P} a_p R_{XX}[-p] &= c_0 \\
R_{XX}[1] + \sum_{p=1}^{P} a_p R_{XX}[1-p] &= c_1 \\
&\vdots \\
R_{XX}[Q] + \sum_{p=1}^{P} a_p R_{XX}[Q-p] &= c_Q \\
&\vdots \\
R_{XX}[Q+P] + \sum_{p=1}^{P} a_p R_{XX}[Q+P-p] &= 0
\end{aligned}
$$

(note c_{Q+1} onwards is 0)

- Collect into a matrix form

$$\mathbf{R}_X \underbrace{\begin{bmatrix} 1 \\ a_1 \\ a_2 \\ \vdots \\ a_P \end{bmatrix}}_{\mathbf{a}} = \underbrace{\begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_Q \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{\mathbf{c}} \tag{7}$$

where $\mathbf{R}_X$ is the matrix

$$\begin{bmatrix} R_{XX}[0] & R_{XX}[-1] & \dots & R_{XX}[-P] \\ R_{XX}[1] & R_{XX}[0] & \dots & R_{XX}[1-P] \\ \vdots & \vdots & & \vdots \\ R_{XX}[Q] & R_{XX}[Q-1] & \dots & R_{XX}[Q-P] \\ R_{XX}[Q+1] & R_{XX}[Q] & \dots & R_{XX}[Q-P+1] \\ \vdots & \vdots & & \vdots \\ R_{XX}[Q+P] & R_{XX}[Q+P-1] & \dots & R_{XX}[Q] \end{bmatrix}$$

- This is the matrix version of the Yule-Walker equations

3 Solution of the Yule-Walker Equations

- We would like to solve for the ARMA parameters from estimates of the autocorrelation function:

- In principle, if the auto-correlation function $R_{XX}[r]$ of a discrete random process is specified for sufficient number of values of r then a set of simultaneous equations may be set up and solved for the model parameters $\{a_i\}$ and $\{b_i\}$.

- Unknowns $c_r = \sum_{q=r}^{Q} b_q \, h_{q-r}$ and $\{h_n\}$ are complicated functions of $\{a_i\}$ and $\{b_i\}$

- There are numerous methods for estimating ARMA models, e.g. Prony's method, see Matlab routines. However, a full solution in the general case is difficult.

- We will study the solution of equation 7 for two special cases, namely the AR model and the MA model.

4 The AR Model $(Q = 0)$

- If $Q = 0$ then the ARMA model becomes the AR model:

$$\boxed{x_n = -\sum_{p=1}^{P} a_p\, x_{n-p} + b_0 w_n} \qquad (8)$$

- The AR model is used in numerous applications, including speech, audio, economics, ...

- Equation (6), which was $c_r = \sum_{q=r}^{Q} b_q\, h_{q-r}$ for $r \leq Q$ and $c_r = 0$ for $r > Q$ becomes:

$$c_r = \begin{cases} b_0\, h_0 & \text{if } r = 0 \\ 0 & \text{if } r > 0 \end{cases}$$

- Consideration of the difference equation for the AR model,

$$x_n = h_0 w_n + \sum_{m \geq 1} h_m\, w_{n-m}$$

shows that $h_0 = b_0$.

- The matrix Yule-Walker equation becomes:

$$
\begin{bmatrix}
R_{XX}[0] & R_{XX}[-1] & \dots & R_{XX}[-P] \\
R_{XX}[1] & R_{XX}[0] & \dots & R_{XX}[1-P] \\
\vdots & \vdots & & \vdots \\
R_{XX}[P] & R_{XX}[P-1] & \dots & R_{XX}[0]
\end{bmatrix}
\times \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_P \end{bmatrix} = \begin{bmatrix} b_0^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{9}
$$

and we use this to solve for a_i and b_0

- Solve for a_i and b_0^2 by partitioning the matrix as:

$$
\left[\begin{array}{c|ccc}
R_{XX}[0] & R_{XX}[-1] & \dots & R_{XX}[-P] \\
\hline
R_{XX}[1] & R_{XX}[0] & \dots & R_{XX}[1-P] \\
\vdots & \vdots & & \vdots \\
R_{XX}[P] & R_{XX}[P-1] & \dots & R_{XX}[0]
\end{array}\right]
\times \left[\begin{array}{c} 1 \\ \hline a_1 \\ \vdots \\ a_P \end{array}\right] = \left[\begin{array}{c} b_0^2 \\ \hline 0 \\ \vdots \\ 0 \end{array}\right]
$$

(concentrate on rows 2 onwards only)

- Taking the bottom P elements of the right and left hand sides:

$$\begin{bmatrix} R_{XX}[0] & R_{XX}[-1] & \dots & R_{XX}[1-P] \\ R_{XX}[1] & R_{XX}[0] & \dots & R_{XX}[2-P] \\ \vdots & \vdots & & \vdots \\ R_{XX}[P-1] & R_{XX}[P-2] & \dots & R_{XX}[0] \\ \end{bmatrix} \times \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_P \end{bmatrix} = - \begin{bmatrix} R_{XX}[1] \\ R_{XX}[2] \\ \vdots \\ R_{XX}[P] \end{bmatrix} \tag{10}$$

or

$$\mathbf{R}_{P-1}\mathbf{a} = -\mathbf{r} \tag{11}$$

- Hence $\mathbf{a} = -\mathbf{R}_{P-1}^{-1}\mathbf{r}$ and

$$b_0^2 = \begin{bmatrix} R_{XX}[0] & R_{XX}[-1] & \dots & R_{XX}[-P] \\ \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_P \end{bmatrix} \tag{12}$$

- Thus if we can estimate the autocorrelation function using one of the standard methods described earlier, then we can estimate the AR parameters and hence the spectrum.

- (Relation to MMSE estimation.) The Yule-Walker solution $\mathbf{a} = -\mathbf{R}_{P-1}^{-1}\mathbf{r}$ is also the solution to the following optimization problem:

$$\min_{\mathbf{h}} E\left\{\left(x_n + \sum_{i=1}^{P} h_i x_{n-i}\right)^2\right\}$$

where

$$x_n = -\sum_{i=1}^{P} a_i\, x_{n-i} + b_0 w_n$$

(i.e. the AR model driven by noise w_n.)

Check: Let $e = x_n + \sum_{i=1}^{P} h_i x_{n-i}$.

$$e^2 = x_n^2 + \mathbf{h}^{\mathrm{T}} \begin{bmatrix} x_{n-1} \\ \vdots \\ x_{n-P} \end{bmatrix} [x_{n-1}, \ldots, x_{n-P}]\, \mathbf{h}$$
$$+ 2\mathbf{h}^{\mathrm{T}} \begin{bmatrix} x_{n-1} \\ \vdots \\ x_{n-P} \end{bmatrix} x_n.$$

$$E(e^2) = R_{XX}[0] + \mathbf{h}^{\mathrm{T}} \mathbf{R}_{P-1} \mathbf{h} + 2\mathbf{h}^{\mathrm{T}} \mathbf{r}$$

Using the fact that the minimiser of $E(e^2)$ satisfies $\mathbf{R}_{P-1}\mathbf{h} = -\mathbf{r}$, the minimum value of $E(e^2)$ is

$$R_{XX}[0] + \mathbf{a'}^{\mathrm{T}} \mathbf{r}$$

which is equal to b_0^2 in (12). (Thus $b_0^2 > 0$.)

- Error keeps decreasing until model order is correct:

$$\min_{\mathbf{h}} E\left\{ \left(x_n + \sum_{i=1}^{P} h_i x_{n-i} \right)^2 \right\}$$
$$\text{s.t. } h'_{j+1} = \cdots = h'_P = 0$$

will be greater than or equal to the unconstrained minimisation problem

$$b_0^2 = \min_{\mathbf{h}} E\left\{\left(x_n + \sum_{i=1}^{P} h_i x_{n-i}\right)^2\right\}$$

5 Levinson's method

- Assume that we have an AR(P) process
- The Yule-Walker equation (see (9)) for a j-th order AR process (or AR(j) process) where $j < P$ is

$$\mathbf{R}_j \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_j \end{bmatrix} = \begin{bmatrix} \epsilon_j \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

where $\mathbf{R}_j$ is

$$\begin{bmatrix} R_{XX}[0] & R_{XX}[1] & \dots & R_{XX}[j] \\ R_{XX}[1] & R_{XX}[0] & \dots & R_{XX}[j-1] \\ \vdots & \vdots & \ddots & \vdots \\ R_{XX}[j-1] & \dots & R_{XX}[0] & R_{XX}[1] \\ R_{XX}[j] & R_{XX}[j-1] & \dots & R_{XX}[0] \end{bmatrix}.$$

(All instances of $R_{XX}[i]$ with $i < 0$ replaced with $R_{XX}[|i|]$ since R_{XX} is an even function.)

- Note last row is row one <u>backwards</u>, second last row is row two backwards etc

- Let the solution be $a_1, \ldots, a_j$ and ϵ_j (and we know $\epsilon_j > 0$)

- The Levinson's method is used to extend this solution to an AR$(j+1)$ process. The idea is as follows

- It is clear that

$$\mathbf{R}_{j+1} \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_j \\ 0 \end{bmatrix} = \begin{bmatrix} \epsilon_j \\ 0 \\ \vdots \\ 0 \\ \gamma_j \end{bmatrix}$$

where

$$\gamma_j = R_{XX}[j+1] + R_{XX}[j]a_1 + R_{XX}[j-1]a_2 + \cdots + R_{XX}[1]a_j$$

- Using the fact that the last row of $\mathbf{R}_{j+1}$ is row one <u>backwards</u>, second last row is row two backwards etc, we have that

$$\mathbf{R}_{j+1} \begin{bmatrix} 0 \\ a_j \\ \vdots \\ a_1 \\ 1 \end{bmatrix} = \begin{bmatrix} \gamma_j \\ 0 \\ \vdots \\ 0 \\ \epsilon_j \end{bmatrix}$$

- Thus for any constant c

$$\mathbf{R}_{j+1} \left(\begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_j \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ a_j \\ \vdots \\ a_1 \\ 1 \end{bmatrix} \right) = \begin{bmatrix} \epsilon_j + c\gamma_j \\ 0 \\ \vdots \\ 0 \\ \gamma_j + c\epsilon_j \end{bmatrix}$$

- Setting $c = -\gamma_j/\epsilon_j$ gives the solution to the AR$(j+1)$ model!

$$\mathbf{R}_{j+1} \begin{bmatrix} 1 \\ a'_1 \\ \vdots \\ a'_{j+1} \end{bmatrix} = \begin{bmatrix} \epsilon_{j+1} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

where $a'_i = a_i - (\gamma_j/\epsilon_j)a_{j+1-i}$, $a'_{j+1} = -\gamma_j/\epsilon_j$ and $\epsilon_{j+1} = \epsilon_j - \gamma_j^2/\epsilon_j$

- Computational cost is

γ_j	j multiplications
$\gamma_j/\epsilon_j, \quad \gamma_j^2/\epsilon_j$	2 multiplications, 1 division
$a'_1, \ldots, a'_j$	j multiplications

so $(2j+3)$ multiplications in total

- Cost for solving the AR(P) model recursively is

$$\sum_{j=0}^{P-1} 2j + 3 = P^2 + 2P$$

compared to $O(P^3)$ if we inverted the matrix in (10)