6. Methods for Rational Spectra

- It is assumed that signals have *rational spectra*

\[
\phi(\omega) = \frac{\sum_{k=-m}^{m} \gamma_k e^{-j\omega k}}{\sum_{k=-n}^{n} \rho_k e^{-j\omega k}},
\]

where \(\gamma_k = \gamma_k^*\) and \(\rho_k = \rho_k^*\).

- Any continuous PSD can be approximated arbitrary closely by (23), provided that \(m\) and \(n\) are chosen sufficiently large (the Weierstrass theorem from calculus).

- The rational spectra (23) can be factored as

\[
\phi(\omega) = \left| \frac{B(\omega)}{A(\omega)} \right|^2 \sigma^2,
\]

where \(\sigma^2\) is a positive scalar and \(A(\omega)\) and \(B(\omega)\) polynomials

\[
A(\omega) = 1 + a_1 e^{-j\omega} + \ldots + a_n e^{-j\omega} \quad \text{and} \quad B(\omega) = 1 + b_1 e^{-j\omega} + \ldots + b_m e^{-j\omega}.
\]
• Filtered stationary signal with the PSD $\phi(\omega)$ has the PSD $|H(\omega)|^2 \phi(\omega)$, where $H(\omega)$ is the transfer function of the filter.

• Therefore, the rational spectra (24) can be associated with a signal obtained by filtering white noise of power $\sigma^2$ through the rational filter with transfer function $H(\omega) = B(\omega)/A(\omega)$.

• The filtering operation can be written as

$$y(t) = -\sum_{k=1}^{n} a_k y(t-k) + \sum_{l=0}^{m} b_l e(t-l), \quad b_0 = 1 \quad (25a)$$

or

$$\sum_{k=0}^{n} a_k y(t-k) = \sum_{l=0}^{m} b_l e(t-l), \quad a_0 = b_0 = 1. \quad (25b)$$
• This means that the filtered value \( y(t) \) is the weighted sum
\[
\sum_{k=1}^{n} a_k y(t - k)
\]
of \( n \) filtered values added to weighted sum
\[
\sum_{l=0}^{m} b_l e(t - l)
\]
of white noise sequence \( e(t) \).

• Note: In the book, a short notation \( A(q)y(t) \) is used to denote filtering operations in time domain. E.g., in (25b) \( \sum_{k=0}^{n} a_k y(t - k) \) is shortened as \( A(q)y(t) \) and \( \sum_{l=0}^{m} b_l e(t - l) \) as \( B(q)e(t) \). Notation \( q^{-1} \) is the unit delay operator \( (q^{-1}y(t) = y(t - k)) \).

• Signal \( y(t) = -\sum_{k=1}^{n} a_k y(t - k) + e(t) \) is called autoregressive (AR) signal.

• Signal \( y(t) = \sum_{l=0}^{m} b_l e(t - l) \) is called moving average signal.

• Signal (25a) is called ARMA signal.

• Notations \( AR(n) \), \( MA(m) \) and \( ARMA(n, m) \) are used in order to specify the order of the model.
• If the zeros of $A(z) = \sum_{k=0}^{n} a_k z^{-k}$ are inside the unit circle, the model (25) is said to be stable.

• If the zeros of $B(z) = \sum_{l=0}^{m} b_l z^{-l}$ are inside the unit circle, the model (25) is said to be minimum phase (see exercise 3.1 for reason to this name).

• It has been shown that if the signal is assumed to have rational spectra, then the spectral estimation problem is reduced to estimation of ARMA-model parameters.
• The PSD can be computed using (24) with

\[ A(\omega) = \sum_{k=0}^{n} a_k e^{-j\omega k}, \quad a_0 = 1 \]  

(26a)

\[ B(\omega) = \sum_{l=0}^{m} b_l e^{-j\omega l}, \quad b_0 = 1. \]  

(26b)

• The possibly unknown noise power \( \sigma^2 \) needs to be solved if the exact scaling is desired.
• AR part can model narrow peaks (zeros of $A(\omega)$ or poles of $H(\omega)$ near unit circle).

• MA part can model zeros of the spectrum.

• Large orders needed for general cases.
Example

• Let \( a_1 = 0.8e^{-j/(10\pi^{10})} \) and \( a_2 = 0.8e^{j\pi/10} \).

• The roots are inside the unit circle (stable filter). *roots* command in Matlab.

• The normalized PSD is computed using (26) and computing the normalized periodogram of filtered white noise.

• The filtered noise can be produced in Matlab as \( x = \text{filter}(b, a, n) \), where \( b \) contains MA parameters, \( a \) AR parameters and \( n \) is noise.

• Now \( b = 1 \) (pure AR-model) and \( n = \text{rand}(256, 1) \).
b=a,a=1 MA model

true spectrum
filtered noise spectrum
**AR Signals**

- Most often, AR or *all-pole* signals are considered.
  - Signals with narrow peaks quite common.
  - AR parameter estimation well studied problem.
- Methods to find AR coefficients are discussed next.
**Yule-Walker method**

- We want to minimize \( E[|e(t)|^2] = E[|y(t) + \sum_{k=1}^{n} a_k y(t-k)|^2] \).

  Setting the differentiation with respect to \( a_l^*, l = 1, \ldots, N \) equal to zero gives

\[
- \begin{bmatrix}
  r(1) \\
  r(2) \\
  \vdots \\
  r(n)
\end{bmatrix}
= 
\begin{bmatrix}
  r(0) & r(-1) & \ldots & r(1-n) \\
  r(1) & r(0) & \ldots & r(2-n) \\
  \vdots & \vdots & \ddots & \vdots \\
  r(n-1) & r(n-2) & \ldots & r(0)
\end{bmatrix}
\begin{bmatrix}
  a_1 \\
  a_2 \\
  \vdots \\
  a_n
\end{bmatrix}
\]

which, in a compact matrix form is

\[
-r = Ra
\]  

(27)

- The autocorrelation matrix \( R \) is Toeplitz (since \( r(k) = r^*(-k) \)).
• The AR coefficients can be solved from (27) by first estimating the ACS $r(0), \ldots, r(n)$ and then solving $a$ by some linear equation solver.

• Indeed, there are many solutions. The course book presents **Levinson-Durbin** algorithm, which finds $a_k, k = 1, \ldots, n$ iteratively, starting from $a_1$. This is numerically rather efficient way and suited for Toeplitz stuctured autocorrelation matrices $R$. 
• Multiply AR-signal \( y(t) = -\sum_{k=1}^{n} a_k y(t - k) + e(t) \) by \( y^*(t) \) and take expectation. This gives

\[
\begin{align*}
r(0) + \sum_{k=1}^{n} a_k r(-k) &= \mathbb{E}\{e(t)y^*(t)\}. \\
\end{align*}
\]  
(28)

• Since \( y(t) \) is obtained by causal filter, \( y(t) = \sum_{k=0}^{\infty} h_k e(t - k) \). Thus, \( \mathbb{E}\{e(t)y^*(t)\} = \mathbb{E}\{e(t)e^*(t)\} = \sigma^2 \) and we have relationship

\[
\begin{align*}
r(0) + \sum_{k=1}^{n} a_k r(-k) &= \sigma^2. \\
\end{align*}
\]  
(29)

• After finding AR coefficients, the unknown noise power can be estimated using (29).
Using (29) together with (27) we obtain so called Yule-Walker or Normal equations

\[
\begin{bmatrix}
\sigma^2 \\
0 \\
\vdots \\
0
\end{bmatrix} = \begin{bmatrix}
r(0) & r(-1) & \ldots & r(-n) \\
r(1) & r(0) & \ldots & r(1-n) \\
\vdots & \vdots & \ddots & \vdots \\
r(n) & r(n-1) & \ldots & r(0)
\end{bmatrix} \begin{bmatrix}
1 \\
a_1 \\
\vdots \\
a_n
\end{bmatrix}
\]

which, in a compact matrix form is

\[
\begin{bmatrix}
\sigma^2 \\
0
\end{bmatrix} = R_{n+1} \begin{bmatrix}
1 \\
a
\end{bmatrix}.
\]
Least Squares Method

• Least squares (LS) method do not use expectation. Instead, write AR signal sample (length $N$) as

$$
\begin{bmatrix}
    y(t) \\
    \vdots \\
    y(t - N + n + 1)
\end{bmatrix} = \begin{bmatrix}
    y(t - 1) & \ldots & y(t - n) \\
    \vdots & \ddots & \vdots \\
    y(t - N + n) & \ldots & y(t - N + 1)
\end{bmatrix} \begin{bmatrix}
    -a_1 \\
    \vdots \\
    -a_n
\end{bmatrix} + \begin{bmatrix}
    e(t) \\
    \vdots \\
    e(t - N + n + 1)
\end{bmatrix}
$$

which, in a compact matrix form is

$$y = Ya + e. \quad (31)$$
• LS problem (31) solves $\mathbf{a}$ by minimizing $\|e\|^2 = \|\mathbf{y} - \mathbf{Y}\mathbf{a}\|^2$.

• It is easy to verify that LS-solution is

$$
\mathbf{a} = - (\mathbf{Y}^H\mathbf{Y})^{-1}\mathbf{Y}^H\mathbf{y}.
$$

(32)

• This was so called covariance LS method. There are also other ways to write $\mathbf{y}$ and $\mathbf{Y}$, see the course book. These deal how the ends of sample vector are handled. The other often used alternative is autocorrelation LS method, which is equivalent with the Yule-Walker method.

• In Matlab, there are $\text{pcov,pyulear,pmcov}$. 
• It is also possible to use recursive LS algorithms to find the AR coefficients.

• In this case sample is put through an RLS-algorithm (whitening filter), and as a result, the coefficients convergence to desired AR coefficients.

• There exist several RLS-algorithms, see, e.g., S. Haykin ‘Adaptive Filter Theory’, Prentice Hall, 1996.
Example

- In the previous example we used AR signal with \( a_1 = 0.8e^{-j/10\pi} = 0.8 \) and \( a_2 = 0.8e^{j\pi/10} = 0.76 + j0.25 \) and produced filtered signal \( y(t) \).

- Now we estimate the AR coefficients \( a_1 \) and \( a_2 \) by the LS-method described above.

- In Matlab, commands \( y = x(1:254); Y = [x(2:255)x(3:256)]; ahat = \text{inv}(Y' * Y) * Y' * y \) give the solution \( \hat{a}_1 = 0.79 - j0.03 \) and \( \hat{a}_2 = 0.75 + j0.24 \) that are quite close. Note that \text{filter} \ command seems to use complex conjugates of the coefficients, so that you should use \text{filter}(1,\text{conj}(a),n) \ to have correct coefficients.

- The resulting spectrum is compared with the true spectrum in the next slide.
ARMA signals

- MA part forms nonlinear problem, difficult to solve.
- The course books presents two solutions: modified Yule-Walker and two stage-LS methods.
- The first is reasonable if zeros are low (well inside the unit circle).
- The latter one estimates first AR coefficients and then $e(t)$ using obtained AR coefficients (whitening filter). If $e(t)$ is known, MA coefficients are easy to compute. Thus, using the estimated $\hat{e}(t)$, the MA coefficients can be estimated.
- Care must be taken if zeros and nulls are close and near the unit circle.
  - Very large models may be required.