Supplementary Exercises:

Spectral Analysis and Optimal Filtering

Random Processes With Applications (MVE 135)

Mats Viberg
Department of Signals and Systems
Chalmers University of Technology
412 96 Göteborg, Sweden
Email: viberg@chalmers.se

Oct. 2007

Problems

1. We wish to estimate the spectrum of a stationary stochastic process using the method of averaged periodograms. The analog signal is sampled at a sampling frequency of 1 kHz. Suppose a frequency resolution of 1 Hz and a relative variance of 1% is desired for the estimated spectrum. How long time (in seconds) does it take to collect the required data?

2. It is desired to estimate the spectrum of a stationary continuous-time stochastic process $\{x_a(t)\}$. An ideal lowpass filter with $F_p = 5$ kHz is first applied to $\{x_a(t)\}$. The signal is then sampled, and a total of 10 s of data is collected. The spectrum is estimated by averaging non-overlapping periodograms. The periodogram lengths are selected to give a frequency resolution of approximately 10 Hz. Emilia uses a sampling frequency of $F_s = 10$ kHz, whereas Emil suggests that $F_s = 100$ kHz should give a lower variance of the spectrum estimate, since more data is obtained. Determine the normalized variance for the two choices of F_s, and investigate if Emil is right!

3. Let $x[n]$ be a stationary stochastic process. Suppose the spectrum is estimated as

$$\hat{P}_x\left(e^{j\omega}\right) = \sum_{n=-100}^{100} \hat{r}_x[n] w[n] e^{-j\omega n}$$

where

$$w[n] = \begin{cases} e^{-0.1|n|} & ; \quad |n| \le 100 \\ 0 & ; \quad n > 100 \end{cases}$$

and $\hat{r}_x[n]$ is based on $N = 10000$ samples.

Determine the approximate normalized variance

$$\nu = \frac{\mathrm{Var}\left(\hat{P}_x(e^{j\omega})\right)}{P_x^2(e^{j\omega})}$$

4. Consider the following three signal models:

$$x[n] + 0.6x[n-1] - 0.2x[n-2] = e[n]$$
$$x[n] = e[n] + 0.8e[n-1] + 0.2e[n-2]$$
$$x[n] + 0.8x[n-1] = e[n] - 0.2e[n-1]$$

where $e[n]$ is a zero-mean white noise with variance $\sigma_e^2 = 1$. Determine the autocorrelation function $r_x[k]$ and the spectrum $P_x(e^{j\omega})$ for each of the cases!

5. The spectrum of a stationary stochastic process is to be estimated from the data:

$$x[n] = \{0.6, -0.7, 0.2, 0.3\}\,.$$

Due to the small sample support, a simple AR(1)-model is exploited:

$$x[n] + a_1 x[n-1] = e[n]\,.$$

Determine estimates of the AR-parameter a_1 and the white noise variance σ_e^2. Based on these, give a parametric estimate of the spectrum $P_x(e^{j\omega})$!

6. We wish to estimate an Auto-Regressive model for a measured signal $x[n]$. The covariance function is estimated based on $N = 1000$ data points as

$$\hat{r}_x(0) = 7.73, \quad \hat{r}_x(1) = 6.80, \quad \hat{r}_x(2) = 4.75, \quad \hat{r}_x(3) = 2.36, \quad \hat{r}_x(4) = 0.23$$

Use the Yule-Walker method to estimate a model

$$x[n] + a_1 x[n-1] + a_2 x[n-2] = e[n]$$

for the signal, where $e[n]$ is white noise. Also give an estimate of the noise variance σ_e^2!

7. We are given a noise-corrupted measurement $x[n]$ of a desired signal $d[n]$. The observed signal is thus modeled by

$$x[n] = d[n] + w[n]\,.$$

Suppose $w[n]$ is white noise with variance $\sigma_w^2 = 1$. The desired signal is represented by the low-pass AR-model

$$d[n] = 0.9d[n-1] + e[n]\,,$$

where $e[n]$ is white noise with variance σ_e^2. An estimate of the desired signal is generated by filtering the measurement as

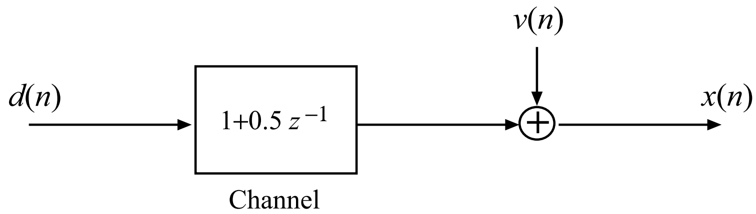
$$\hat{d}[n] = h_0 x[n] + h_1 x[n-1]\,.$$

Determine h_0 and h_1 such that $E[(d[n] - \hat{d}[n])^2]$ is minimized. Then sketch the amplitude characteristics of the resulting filter $h_0 + h_1 z^{-1}$ for σ_e^2 "small", "medium" and "large", respectively. Explain the result!

8. The goal of this problem is to equalize a communication channel using an FIR Wiener filter. The desired signal $d[n]$ is described as

$$d[n] - 0.8d[n-1] = e[n],$$

where $e[n]$ is zero-mean white noise with variance $\sigma_e^2 = 1$. The measured signal $x[n]$ is given by



where $v[n]$ is a zero-mean white noise with variance $\sigma_v^2 = 0.1$. Determine a filter

$$\hat{d}[n] = w_0\, x[n] + w_1\, x[n-1]$$

so that $E\left[\left(d[n] - \hat{d}[n]\right)^2\right]$ is minimized!

9. A 2:nd order digital FIR notch filter can be designed by placing the zeros at $e^{\pm j\omega}$. The resulting transfer function is

$$H(z) = 1 - 2\cos\omega z^{-1} + z^{-2}.$$

Thus, a natural approach for estimating the frequency of a sine-wave is to apply a constrained FIR filter

$$W(z) = 1 - wz^{-1} + z^{-2}$$

to the measured signal, and select w such that the output power is minimized. The frequency could then be estimated as $\omega = \arccos\frac{w}{2}$, provided $|w| < 2$. Now, suppose the measured signal is

$$x[n] = A\cos(\omega_0 n + \varphi) + v[n],$$

where φ is a $U(0, 2\pi)$ random phase and $v[n]$ is a zero-mean stationary white noise. The filter output is given by

$$y[n] = x[n] - wx[n-1] + x[n-2]\,.$$

Find the w that minimizes $E[y^2[n]]$, and compute $\hat{\omega} = \arccos\frac{w}{2}$. Then, prove that $\hat{\omega} = \omega_0$ for $\sigma_v^2 = 0$, but that the estimate is otherwise biased.

Solutions to Selected Problems

1. In the method of periodogram averaging (Bartlett's method), the available N samples are split into $K = N/M$ non-overlapping blocks, each of length M samples. The periodogram for each sub-block is computed, and the final estimate is the average of these. The variance of each sub-periodogram is

$$Var\{\hat{P}_{per}^{(i)}(e^{j\omega})\} \approx P_x^2(e^{j\omega})$$

Thus, the normalized variance is

$$\frac{Var\{\hat{P}_{per}^{(i)}(e^{j\omega})\}}{P_x^2(e^{j\omega})} \approx 1$$

or 100%. Assuming the sub-periodograms to be uncorrelated, averaging reduces the normalized variance to

$$\frac{Var\{\hat{P}_B(e^{j\omega})\}}{P_x^2(e^{j\omega})} \approx 1/K$$

To achieve a normalized variance of 0.01 (1%) thus requires averaging $K \geq 100$ sub-periodograms. Each sub-periodogram has the approximate frequency resolution $\Delta\omega \approx 2\pi/M$, or $\Delta f \approx 1/M$. In un-normalized frequency this becomes $\Delta F \approx F_s/M$ Hz. With $F_s = 1000$ and a desired frequency resolution of $\Delta F \leq 1$ Hz, the length of each sub-block must be $M \geq 1000$. Together with $K \geq 100$, this shows that the number of available samples must be $N = KM \geq 10^5$. At 1 kHz, the data collection takes at least 100 seconds.

3. For Blackman-Tukey's method, the normalized variance is given by

$$\nu = \frac{Var\{\hat{P}_{BT}(e^{j\omega})\}}{P_x^2(e^{j\omega})} \approx \frac{1}{N}\sum_{k=-M}^{M} w_{lag}^2[k]\,,$$

where $w_{lag}[k]$ is the lag window. In this case we get

$$\sum_{k=-M}^{M} w_{lag}^2[k] = \sum_{k=-100}^{100} e^{-0.2|k|} = 2\sum_{k=0}^{100} e^{-0.2k} - 1 = 2\,\frac{1-e^{-200}}{1-e^{-0.2}} - 1 \approx 10\,.$$

Thus, the normalized variance is $\nu = 10/N = 10^{-3}$.

4. The first model is an AR(2)-process:

$$x[n] + 0.6x[n-1] - 0.2x[n-2] = e[n]$$

Multiplying by $x[n-k]$, $k \geq 0$, and taking expectation leads to the relation

$$r_x[k] + 0.6r_x[k-1] - 0.2r_x[k-2] = \sigma_e^2\,\delta[k]$$

(Yule-Walker). With $k = \{0, 1, 2\}$, we get three equations in three unknowns (see also the theory complement):

$$\begin{bmatrix} 1 & 0.6 & -0.2 \\ 0.6 & 1-0.2 & 0 \\ -0.2 & 0.6 & 1 \end{bmatrix} \begin{bmatrix} r_x[0] \\ r_x[1] \\ r_x[2] \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Solving this, for example using Matlab, leads to $r_x[0] = 2.4$, $r_x[1] = -1.8$ and $r_x[2] = 1.6$. We can then continue for $k = 3, 4, \ldots$ to get the autocorrelation at any lag, for example,

$$r_x[3] = -0.6r_x[2] + 0.2r_x[1] = -1.3\,, \quad r_x[4] = -0.6r_x[3] + 0.2r_x[2] = 1.1\,.$$

It is of course also possible to solve the homogenous difference equation

$$r_x[k] + 0.6r_x[k-1] - 0.2r_x[k-2] = 0\,,$$

with the given "initial conditions", to give the general solution. The spectrum of the AR model is given by

$$P_x(e^{j\omega}) = \frac{\sigma_e^2}{|A(e^{j\omega})|^2} = \frac{1}{(1+0.6e^{-j\omega} - 0.2e^{-j2\omega})(1+0.6e^{j\omega} - 0.2e^{j2\omega})} = \frac{1}{1.4 + 0.96\cos\omega - 0.4\cos 2\omega}$$

The second model is an MA(2) process:

$$x[n] = e[n] + 0.8e[n-1] + 0.2e[n-2]$$

In this case, the autocorrelation function is finite. First, multiply by $x[n]$ and take expectation, which leads to

$$r_x[0] = E\{(e[n] + 0.8e[n-1] + 0.2e[n-2])^2\} = (1 + 0.8^2 + 0.2^2)\sigma_e^2 = 1.68\,.$$

Next, multiply by $x[n-1]$ and take expectation:

$$r_x[1] = E\{(e[n]+0.8e[n-1]+0.2e[n-2])(e[n-1]+0.8e[n-2]+0.2e[n-3])\} = (0.8+0.2\times 0.8)\sigma_e^2 = 0.96\,.$$

Finally, multiplying by $x[n-2]$ and taking expectation leads to

$$r_x[1] = E\{(e[n] + 0.8e[n-1] + 0.2e[n-2])(e[n-2] + 0.8e[n-3] + 0.2e[n-4])\} = 0.2\,.$$

Multiplying by $x[n-k]$, $k > 2$ and taking expectation wee see that $r_x[k] = 0$, $|k| > 2$. The spectrum of the MA process is given by

$$P_x(e^{j\omega}) = \sigma_e^2|B(e^{j\omega})|^2 = (1+0.8e^{-j\omega}+0.2e^{-j2\omega})(1+0.8e^{j\omega}+0.2e^{j2\omega}) = 1.68+1.92\cos\omega+0.4\cos 2\omega$$

The third model is the ARMA(1,1) process:

$$x[n] + 0.8x[n-1] = e[n] - 0.2e[n-1]$$

In general, computing the autocorrelation function for an ARMA model is quite difficult. However, in the ARMA(1,1) case the following procedure is easiest. First, multiply by $x[n-1]$ and take expectation, which gives:

$$r_x[1] + 0.8r_x[0] = -0.2\sigma_e^2\,.$$

Second, square the ARMA model equation and again take expectation:

$$E\{(x[n] + 0.8x[n-1])^2\} = E\{(e[n] - 0.2e[n-1])^2\}\,.$$

This leads to

$$1.64r_x[0] + 1.6r_x[1] = 1.04\sigma_e^2\,.$$

Now we have two equations in the two unknowns $r_x[0]$ and $r_x[1]$:

$$\begin{bmatrix} 0.8 & 1 \\ 1.64 & 1.6 \end{bmatrix} \begin{bmatrix} r_x[0] \\ r_x[1] \end{bmatrix} = \begin{bmatrix} -0.2 \\ 1.04 \end{bmatrix}$$

Applying the usual formula for inverting a 2×2 matrix, the solution is easily obtained as

$$\begin{bmatrix} r_x[0] \\ r_x[1] \end{bmatrix} = \frac{1}{0.8 \times 1.6 - 1 \times 1.64} \begin{bmatrix} 1.6 & -1 \\ -1.64 & 0.8 \end{bmatrix} \begin{bmatrix} -0.2 \\ 1.04 \end{bmatrix} \approx \begin{bmatrix} 3.8 \\ -3.2 \end{bmatrix}.$$

The remaining auto-correlation parameters can be obtained by multiplying the ARMA equation by $x[n-k]$, $k \geq 2$ and taking expectation:

$$r_x[k] + 0.8r_x[k-1] = 0\,,\ k \geq 2.$$

Thus, $r_x[2] = -0.8r_x[1] \approx 2.6$ etc. In general, $r_x[k] = (-0.8)^{k-1}r_x[1] = (-0.8)^{k-1}(-3.2)$ for $k \geq 2$. The spectrum of the ARMA model is given by

$$P_x(e^{j\omega}) = \sigma_e^2 \frac{|B(e^{j\omega})|^2}{|A(e^{j\omega})|^2} = \frac{(1-0.2e^{j\omega})(1-0.2e^{-j\omega})}{(1+0.8e^{j\omega})(1+0.8e^{-j\omega})} = \frac{1-0.4\cos\omega}{1+1.6\cos\omega}$$

5. The Yule-Walker method gives the estimate

$$\hat{a}_1 = -\hat{r}_x[0]^{-1}\hat{r}_x[1]\,.$$

With the given data, the sample autocorrelation function is calculated as

$$\hat{r}_x[0] = \frac{1}{4}(0.6^2 + 0.7^2 + 0.2^2 + 0.3^2) = 0.245$$

and

$$\hat{r}_x[1] = \frac{1}{4}(0.6 \times (-0.7) + (-0.7) \times 0.2 + 0.2 \times 0.3) = -0.125$$

Thus, we get

$$\hat{a}_1 = \frac{0.125}{0.245} \approx 0.51\,.$$

The noise variance estimate follows as

$$\hat{\sigma}_e^2 = \hat{r}_x[0] + \hat{a}_1\hat{r}_x[1] \approx 0.18\,.$$

7. The optimal filter (in the LMMSE sense) is given by the Wiener-Hopf (W-H) equations:

$$\sum_k h[k]r_x[l-k] = r_{dx}[l]\,.$$

In this case, the filter is a causal FIR filter with two taps (first-order filter):

$$\hat{d}[n] = h_0x[n] + h_1x[n-1]\,.$$

Thus, the W-H equations are

$$\sum_{k=0}^{1} h_k r_x[l-k] = r_{dx}[l]\,,\ l = 0, 1.$$

In matrix form, this becomes

$$\begin{bmatrix} r_x[0] & r_x[1] \\ r_x[1] & r_x[0] \end{bmatrix} \begin{bmatrix} h_0 \\ h_1 \end{bmatrix} = \begin{bmatrix} r_{dx}[0] \\ r_{dx}[1] \end{bmatrix}$$

The measured signal is given by

$$x[n] = d[n] + w[n] .$$

The auto-correlation function is then given by

$$r_x[k] = E\{(d[n] + w[n])(d[n-k] + w[n-k])\} = r_d[k] + r_w[k]$$

(since $d[n]$ and $w[n]$ are uncorrelated). Now, $w[n]$ is white noise, so $r_w[k] = \sigma_w^2 \delta[k] = \delta[k]$. The desired signal is the AR(1) process:

$$d[n] - 0.9d[n-1] = e[n] .$$

To compute the autocorrelation function we use the Yule-Walker equations, which on matrix form become

$$\begin{bmatrix} 1 & -0.9 \\ -0.9 & 1 \end{bmatrix} \begin{bmatrix} r_d[0] \\ r_d[1] \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \sigma_e^2.$$

Solving this gives

$$\begin{bmatrix} r_d[0] \\ r_d[1] \end{bmatrix} \approx \begin{bmatrix} 5.26 \\ 4.74 \end{bmatrix} \sigma_e^2.$$

Inserting the autocorrelation functions into the W-H equations now gives

$$\left(\begin{bmatrix} 5.26 & 4.74 \\ 4.74 & 5.26 \end{bmatrix} \sigma_e^2 + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} h_0 \\ h_1 \end{bmatrix} = \begin{bmatrix} 5.26 \\ 4.74 \end{bmatrix} \sigma_e^2 ,$$

where we have also used that $r_{dx}[k] = r_d[k]$, since $d[n]$ and $w[n]$ are uncorrelated. It is possible, but a bit cumbersome, to solve this 2×2 system for a general σ_e^2. Instead we take the extreme cases $\sigma_e^2 \to 0$ and $\sigma_e^2 \to \infty$ first, and then the intermediate case $\sigma_e^2 = 1$. For σ_e^2 very small we have

$$\begin{bmatrix} h_0 \\ h_1 \end{bmatrix} \to \begin{bmatrix} 5.26 \\ 4.74 \end{bmatrix} \sigma_e^2 \to \begin{bmatrix} 0 \\ 0 \end{bmatrix} .$$

Thus, when the SNR is very low, the estimation is $\hat{d}[n] = 0$. For very large σ_e^2 we have instead

$$\begin{bmatrix} 5.26 & 4.74 \\ 4.74 & 5.26 \end{bmatrix} \begin{bmatrix} h_0 \\ h_1 \end{bmatrix} = \begin{bmatrix} 5.26 \\ 4.74 \end{bmatrix} ,$$

which gives the optimal FIR coefficients as

$$\begin{bmatrix} h_0 \\ h_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Thus, in the high SNR case we take $\hat{d}[n] = x[n]$, i.e. no filtering is necessary. For the intermediate case $\sigma_e^2 = 1$ we get

$$\begin{bmatrix} 6.26 & 4.74 \\ 4.74 & 6.26 \end{bmatrix} \begin{bmatrix} h_0 \\ h_1 \end{bmatrix} = \begin{bmatrix} 5.26 \\ 4.74 \end{bmatrix} ,$$

which leads to

$$\begin{bmatrix} h_0 \\ h_1 \end{bmatrix} = \begin{bmatrix} 0.63 \\ 0.28 \end{bmatrix} .$$

The estimate is then $\hat{d}[n] = 0.62x[n] + 0.28x[n-1]$, corresponding to the transfer function

$$H(e^{j\omega}) = 0.62 + 0.28e^{-j\omega} .$$

This is a low-pass filter with DC gain $H(e^{j0}) = 0.62 + 0.28 = 0.9$, and 3 dB cutoff frequency $f_c \approx 0.25$ "per samples". The filter is an optimal compromise for removing as much noise (which has a flat spectrum) as possible, without distorting the signal (which is low-pass) too much.