T-61.3040 Statistical Signal Modeling Lendasse, Miche, Eirola
Exercise 6, October 20, 2011, Solutions

1. The matrix form for the equations of the Padé approximation is:

$$\begin{bmatrix} x(0) & 0 & 0 \\ x(1) & x(0) & 0 \\ x(2) & x(1) & x(0) \\ x(3) & x(2) & x(1) \\ x(4) & x(3) & x(2) \end{bmatrix} \begin{bmatrix} 1 \\ a(1) \\ a(2) \end{bmatrix} = \begin{bmatrix} b(0) \\ b(1) \\ b(2) \\ 0 \\ 0 \end{bmatrix}$$

Inserting the known values for $x(n)$:

$$\begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \\ -1 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ a(1) \\ a(2) \end{bmatrix} = \begin{bmatrix} b(0) \\ b(1) \\ b(2) \\ 0 \\ 0 \end{bmatrix}$$

It is immediately seen that $a(1) = 0$, $a(2) = 1$ and $b(0) = 1$. From the second and third lines it is then trivial to calculate that $b(1) = 1$ and $b(2) = 2$.

2. (a) The coefficients of the linear predictor

$$\hat{x}(n) = a_5 x(n-5) + a_{10} x(n-10)$$

are chosen so that the squared prediction error $\mathrm{E}[e(n)e^*(n)]$ is minimised. This allows one to apply the principle of orthogonality:

$$\begin{aligned} \mathrm{E}\left[\big(x(n) - \hat{x}(n)\big)x^*(n-5)\right] &= 0 \\ \mathrm{E}\left[\big(x(n) - \hat{x}(n)\big)x^*(n-10)\right] &= 0, \end{aligned}$$

that is

$$\begin{aligned} \mathrm{E}\left[\big(a_5 x(n-5) + a_{10} x(n-10)\big)x^*(n-5)\right] &= \mathrm{E}[x(n)x^*(n-5)] \\ \mathrm{E}\left[\big(a_5 x(n-5) + a_{10} x(n-10)\big)x^*(n-10)\right] &= \mathrm{E}[x(n)x^*(n-10)], \end{aligned}$$

resulting in the normal equations

$$\begin{pmatrix} r_x(0) & r_x(5) \\ r_x(5) & r_x(0) \end{pmatrix} \begin{pmatrix} a_5 \\ a_{10} \end{pmatrix} = \begin{pmatrix} r_x(5) \\ r_x(10) \end{pmatrix}$$

where we have used that $r_x(-k) = r_x^*(k) = r_x(k)$ when the autocorrelations are real valued.
Inserting the given numerical values, we get

$$\begin{pmatrix} 1 & 0.9 \\ 0.9 & 1 \end{pmatrix} \begin{pmatrix} a_5 \\ a_{10} \end{pmatrix} = \begin{pmatrix} 0.9 \\ 0.7 \end{pmatrix}.$$

Solving the pair of equations gives $a_5 = \frac{27}{19} \approx 1.42$ and $a_{10} = -\frac{11}{19} \approx -0.58$.

The squared error is given by the formula

$$\mathrm{E}[ee^*] = \mathrm{E}[e(x^*(n) - a_5 x^*(n-5) - a_{10}x^*(n-10)]$$

(orthogonality principle)

$$\begin{aligned}
&= \mathrm{E}[ex^*(n)] \\
&= \mathrm{E}[x(n)x^*(n) - a_5 x(n-5)x^*(n) - a_{10}x(n-10)x^*(n)] \\
&= r_x(0) - a_5 r_x^*(5) - a_{10} r_x^*(10) = \frac{12}{95} \approx 0.126.
\end{aligned}$$

(b) Following the same procedure is in part (a), one obtains

$$\begin{pmatrix} r_x(0) & r_x(1) \\ r_x(1) & r_x(0) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} r_x(1) \\ r_x(2) \end{pmatrix}$$

The solution is $a_1 = a_2 = \frac{6}{21} \approx 0.29$. The squared error is $\mathrm{E}(ee^*) = r_x(0) - a_1 r_x(1) - a_2 r_x(2) = \frac{27}{35} \approx 0.77$ We notice that the error is much larger than in part (a).

(c) This time the orthogonality principle results in one normal equation: $r_x(0)a_N = r_x(N)$, so that $a_N = r_x(N)/r_x(0)$. The minimal error is

$$\mathrm{E}(ee^*) = r_x(0) - a_N r_x(N) = r_x(0) - \frac{r_x^2(N)}{r_x(0)}$$

Checking all the errors for $N = 1$ to 10:

$$0.84 \quad 0.84 \quad 0.91 \quad 0.96 \quad 0.19 \quad 0.84 \quad 0.84 \quad 0.96 \quad 0.99 \quad 0.51$$

This is minimised for $N = 5$ In this case, $r_x(N) = r_x(5) = 0.9$ and $\mathrm{E}(ee^*) = 1 - (0.9)^2/1 = 0.19$.

The error is larger than in (a) but still smaller than in (b).

The considered linear predictors have a connection with the autoregressive parts of AR models. The squared prediction error in turn has a straightforward connection with the coefficient b_0 of AR models (assuming unit noise variance): the term $b_0 v(n)$ in AR models explains the part of the variance of the observations that is left unexplained by the autoregressive part. The expression for this residual variance $|b_0|^2$ is the first Yule-Walker equation of the AR model:

$$|b_0|^2 = r_x(k) + \sum_{l=1}^{p} a(l) r_x(n-l)$$

Here the sign of the coefficients $a(l)$ has been defined to be opposite of the predictor coefficients a_l above. We may notice that the expression is the same as the formula that was used for evaluating the squared prediction error. However, this is true only if the predictor coefficients are chosen so that the squared prediction error is minimised. This gives the insight that linear predictors correspond to AR models only if the predictor coefficients are optimal (w.r.t. the process that is to be predicted). Otherwise the prediction error would not be orthogonal with the previous values of the process, i.e. it would not be white. On the other hand, if coefficients $a(i)$ of a (stable) linear predictor are given, one always can find (via the Yule-Walker equations) autocorrelations of a process that can be optimally predicted with that particular predictor.

3. The coefficients of an AR(2) process

$$x(n) + \sum_{l=1}^{2} a(l)x(n-l) = b(0)v(n)$$

can be solved using the Yule-Walker equations

$$r_x(k) + \sum_{l=1}^{2} a(l)r_x(n-l) = \sigma_v^2 |b(0)|^2 \delta(k) \quad k \geq 0$$

which involve only those autocorrelations that are known. These equations are

$$\begin{pmatrix} r_x(0) & r_x^*(1) & r_x^*(2) \\ r_x(1) & r_x(0) & r_x^*(1) \\ r_x(2) & r_x(1) & r_x(0) \end{pmatrix} \begin{pmatrix} 1 \\ a(1) \\ a(2) \end{pmatrix} = \begin{pmatrix} |b(0)|^2 \\ 0 \\ 0 \end{pmatrix}$$

where we have used the symmetry $r_x(k) = r_x^*(-k)$ and the assumption $\sigma_v^2 = 1$. The coefficients $a(1)$ and $a(2)$ are solved using the two lowermost equations (the column consisting of constants is moved to the right hand side of equations):

$$\begin{pmatrix} r_x(0) & r_x^*(1) \\ r_x(1) & r_x(0) \end{pmatrix} \begin{pmatrix} a(1) \\ a(2) \end{pmatrix} = - \begin{pmatrix} r_x(1) \\ r_x(2) \end{pmatrix}.$$

This pair of equations could be solved by inverting the matrix that multiplies the unknown vector:

$$\begin{pmatrix} a(1) \\ a(2) \end{pmatrix} = - \begin{pmatrix} r_x(0) & r_x^*(1) \\ r_x(1) & r_x(0) \end{pmatrix}^{-1} \begin{pmatrix} r_x(1) \\ r_x(2) \end{pmatrix}.$$

The inverse can be found by the conventional formula for 2×2 matrices. The determinant of the matrix is $(r_x(0))^2 - r_x(1)r_x(1)^* = 4 - (1-j)(1+j) = 2$. Then:

$$\begin{aligned} \begin{pmatrix} a(1) \\ a(2) \end{pmatrix} &= -\frac{1}{2} \begin{pmatrix} r_x(0) & -r_x^*(1) \\ -r_x(1) & r_x(0) \end{pmatrix} \begin{pmatrix} r_x(1) \\ r_x(2) \end{pmatrix} \\ &= -\frac{1}{2} \begin{pmatrix} 2 & -1-j \\ -1+j & 2 \end{pmatrix} \begin{pmatrix} 1-j \\ \frac{2}{3} - j \end{pmatrix} = \begin{pmatrix} -\frac{1}{6} + \frac{5}{6}j \\ -\frac{2}{3} \end{pmatrix} \end{aligned}$$

From the first Yule-Walker equation

$$|b(0)|^2 = r_x(0) + a(1)r_x^*(1) + a(2)r_x^*(2) = \frac{5}{9}.$$

Thus

$$b(0) = \frac{\sqrt{5}}{3} e^{j\phi},$$

where ϕ is an arbitrary phase angle.

4. Let's approximate the observed process $x(n)$ with some ARMA process. An ARMA process can be interpreted as filtering white noise with an LSI filter $H(z)$: $v(n) \rightarrow H(z) \rightarrow x(n)$. We may select the noise variance to be one. $x(n)$ can be whitened for example by filtering it with the inverse filter of $H(z)$, resulting in the original noise. When we want

the inverse filter to be an FIR filter of order 2, we select $H(z)$ to be a second order all-pole filter, i.e. the observations are approximated with an AR(2) process. So,

$$H(z) = \frac{b(0)}{1 + a(1)z^{-1} + a(2)z^{-2}}$$

To fit an AR(2) model to the observations and solve the three parameters, we use the three first Yule-Walker equations

$$\begin{pmatrix} r_x(0) & r_x(1) & r_x(2) \\ r_x(1) & r_x(0) & r_x(1) \\ r_x(2) & r_x(1) & r_x(0) \end{pmatrix} \begin{pmatrix} 1 \\ a(1) \\ a(2) \end{pmatrix} = \begin{pmatrix} |b(0)|^2 \\ 0 \\ 0 \end{pmatrix}$$

where we have used the symmetry $r_x(k) = r_x(-k)$ (the process is real valued). The two lowermost YW equations yield $a(1)$ and $a(2)$, giving the numerical values $a(1) = -0.5$ and $a(2) = 0.25$. Inserting these values into the first equations, one gets $|b(0)|^2 = 1.575$ and $b(0) \approx 1.255$ if $b(0)$ is chosen to be a positive real number.

Then $H(z) = 1.255/(1 - 0.5z^{-1} + 0.25z^{-2})$. Inverting this gives

$$B(z) = \frac{1}{H(z)} = c - c \cdot 0.5z^{-1} + c \cdot 0.25z^{-2},$$

where $c = 1/1.255 \approx 0.7968$. This is an FIR filter corresponding to an MA(2) model that approximately whitens $x(n)$. The approximateness of the comes from the fact that the MA(2) was fitted to the first three values of the autocorrelation sequence. However, one can not be certain that the MA(2) model describes the observations perfectly. Some further approximation is caused by the fact that the autocorrelations are estimated from the observations.

5. (a) The difference equation for an ARMA(1,1)-model is

$$x(n) + a(1)x(n-1) = b(0)v(n) + b(1)v(n-1)$$

The impulse response could easily be found calculating $x(n)$ with the input being the impulse signal $v(n) = \delta(n)$ and stating $x(n) = 0$ for $n < 0$.

Alternatively, it can be found by by taking the inverse transform of the transfer function

$$H(z) = \frac{b(0) + b(1)z^{-1}}{1 + a(1)z^{-1}}$$

which leads to

$$h(n) = b(n) * \big((-a(1))^n u(n)\big) = b(0)(-a(1))^n u(n) + b(1)(-a(1))^{n-1} u(n-1)$$

where $u(n)$ is the unit step, or as the coefficients of the transfer function $H(z)$ when expanded to an infinite polynomial of z^{-1}:

$$H(z) = \frac{b(0) + b(1)z^{-1}}{1 + a(1)z^{-1}} = (b(0) + b(1)z^{-1}) \sum_{k=0}^{\infty} \big(-a(1)z^{-1}\big)^k$$

$$= b(0) + \sum_{k=1}^{\infty} \big(b(1) - a(1)b(0)\big)(-a(1))^{k-1} z^{-k}$$

Either way, we obtain that $h(0) = b(0)$ and $h(1) = b(1) - a(1)b(0)$

(b) Let's write down the Yule-Walker equations for the case $p = 1$, $q = 1$:

$$\begin{pmatrix} r_x(0) & r_x(-1) \\ r_x(1) & r_x(0) \\ r_x(2) & r_x(1) \end{pmatrix} \begin{pmatrix} 1 \\ a(1) \end{pmatrix} = \begin{pmatrix} 0.2 & 0.08 \\ 0.08 & 0.2 \\ 0.064 & 0.08 \end{pmatrix} \begin{pmatrix} 1 \\ a(1) \end{pmatrix} = \begin{pmatrix} c(0) \\ c(1) \\ 0 \end{pmatrix}$$

where $c(k) = b(k) * h^*(-k) = \sum_{l=0}^{1-k} b(l+k)h^*(l)$. Thus

$$c(0) = \sum_{l=0}^{1} b(l)h^*(l) = b(0)h^*(0) + b(1)h^*(1) = |b(0)|^2 + |b(1)|^2 - b(1)a^*(1)b^*(0)$$

and

$$c(1) = \sum_{l=0}^{0} b(l+1)h^*(l) = b(1)h^*(0) = b(1)b^*(0)$$

The coefficient $a(1)$ for the AR part of the model can be solved from the lowermost row of the equations: $0.064 + a(1)0.08 = 0 \implies a(1) = -0.8$.
Obtaining $b(0)$ and $b(1)$ requires solving the non-linear equations

$$\begin{pmatrix} 0.2 & 0.08 \\ 0.08 & 0.2 \end{pmatrix} \begin{pmatrix} 1 \\ a(1) \end{pmatrix} = \begin{pmatrix} c(0) \\ c(1) \end{pmatrix} \quad \Longrightarrow \quad \begin{pmatrix} 0.136 \\ -0.08 \end{pmatrix} = \begin{pmatrix} |b(0)|^2 + |b(1)|^2 + 0.8b(1)b^*(0) \\ b(1)b^*(0) \end{pmatrix}$$

Inserting the lower equation into the top expression, one get the requirements

$$\begin{cases} |b(0)|^2 + |b(1)|^2 = 0.2 \\ b(1)b^*(0) = -0.08 \end{cases}$$

The lower of the equations results the condition

$$|b(1)| = \frac{0.08}{|b(0)|}$$

for the absolute values. Inserting these to the first condition gives

$$|b(0)|^2 + 0.0064|b(0)|^{-2} = 0.2$$

Multiplication with $|b(0)|^2$ results in a second order polynomial equation for $|b(0)|^2$:

$$\left(|b(0)|^2\right)^2 - 0.2|b(0)|^2 + 0.0064 = 0.$$

The solution to this is $|b(0)|^2 = 0.1 \pm 0.06$, that is $|b(0)| = 0.2$ or $|b(0)| = 0.4$. In the first case $|b(1)| = \frac{0.08}{|b(0)|} = 0.4$.
From the equation $b(1)b^*(0) = -0.08$ we also get the condition that the coefficients are opposite phase, i.e., $\arg(b(1)) = \arg(b(0) + \pi$ for the phase angles, but otherwise the phase is arbitrary.
Summarising the results in this case,

$$a(1) = -0.8, \qquad \begin{pmatrix} b(0) \\ b(1) \end{pmatrix} = e^{j\phi} \begin{pmatrix} 0.2 \\ -0.4 \end{pmatrix}$$

where ϕ is an arbitrary phase factor.
For the other solution of the second order equation, $|b(1)| = \frac{0.08}{|b(0)|} = 0.2$. Then the coefficients are given by

$$a(1) = -0.8, \qquad \begin{pmatrix} b(0) \\ b(1) \end{pmatrix} = e^{j\phi} \begin{pmatrix} 0.4 \\ -0.2 \end{pmatrix}.$$