COMPUTATIONAL METHODS FOR EVALUATING COVARIANCE FUNCTIONS

Torsten Söderström

Systems and Control Group
Uppsala University
Box 27, S-751 03 Uppsala, Sweden

January 14, 2003

1 Introduction

There are several (good) methods for computing covariance functions for a given model. In a general setting, the problem can be stated as follows. Consider two ARMA processes

$$y(t) = \frac{C(q)}{A(q)}e(t) \qquad w(t) = \frac{F(q)}{D(q)}e(t) \tag{1.1}$$

driven by the *same white* noise sequence $\{e(t)\}$ of zero mean and variance λ^2. Then compute the cross-covariance function

$$r_{yw}(\tau) = Ey(t+\tau)w(t) \tag{1.2}$$

for one or more (problem-specific) time arguments τ.

Many problems of covariance calculations can be reduced to the above form (cf Example 2, to follow).

Occasionally, one has instead to evaluate a "deterministic expectation" such as

$$r_{yu}(\tau) \triangleq \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} y(t+\tau)u(t) \tag{1.3}$$

where $u(t)$ is a given deterministic signal and $y(t)$ a filtered version of $u(t)$. We will consider such a case in Example 5 below.

2 Examples

In this section, we introduce some examples of covariance function evaluations. In sections to come, we will describe some different methods for covariance computations and the same examples will be used for illustration and comparison.

Example 1. Let $y(t)$ be an AR(1) process

$$y(t) + ay(t-1) = e(t) \qquad |a| < 1$$

We seek the (auto-)covariance function $r_y(\tau)$ for $\tau = 0, 1, 2$. Hence, in (1.1), we have $C = F = q$, $A = D = (q + a)$. □

Example 2. Let $u(t)$ be an AR(1) process

$$u(t) + du(t-1) = v(t) \qquad |d| < 1$$

where $v(t)$ is white noise and let $u(t)$ be the input and $y(t)$ the output of a first order stable system

$$y(t) + ay(t-1) = bu(t-1) \qquad |a| < 1$$

Seek the cross-covariance $r_{yu}(\tau)$ for $\tau = 0, 1$. To phrase this example in the setting of (1.1), we set

$$\frac{C(q)}{A(q)} = \frac{b}{(q+a)(q+d)} \qquad \frac{F(q)}{D(q)} = \frac{1}{q+d}$$

We further have $e(t) = v(t+1)$, which of course is a white noise sequence. Note that it is important to have $e(t)$ *white* in (1.1). □

Example 3. Let $y(t)$ be an MA(2) process

$$y(t) = e(t) + c_1 e(t-1) + c_2 e(t-2)$$

and seek the autocovariance function of $y(t)$, i.e. $r_y(\tau)$, for all τ. □

Example 4. Let $y(t)$ and $w(t)$ be two correlated AR(1) processes,

$$\begin{aligned} y(t) + ay(t-1) &= e(t) \\ w(t) + dw(t-1) &= e(t) \end{aligned}$$

and seek the cross-covariance function $r_{yw}(\tau)$ for all τ. Note that Example 1 will be a special case (namely, corresponding to $d = a$). □

Example 5. Here comes a "deterministic" problem. Let $u(t)$ be a step of size σ and $y(t)$ the output of the asymptotically stable system

$$A(q)y(t) = B(q)u(t)$$

The output will then converge to $\overline{y} = S\sigma$, where $S = B(1)/A(1)$ is the static gain of the system. The deviation $y(t) - \overline{y}$ will be a transient that decays exponentially and it will give no contribution to covariance functions such as $r_{yu}(\tau)$ and $r_y(\tau)$. In this example, we get for any finite τ

$$\begin{aligned} r_u(\tau) &= \sigma^2 \\ r_{yu}(\tau) &= S\sigma^2 \\ r_y(\tau) &= S^2\sigma^2 \end{aligned}$$

□

3 Method 1 - Division

This approach is difficult to apply for examples of order 2 or higher. The basic idea is as follows. Divide the polynomials in (1.1), or expressed differently, rewrite the models into weighting function form. Let us assume

$$\begin{aligned}
y(t) &= \frac{C(q)}{A(q)}e(t) \\
&= h_o e(t) + h_1 e(t-1) + h_2 e(t-2) + \ldots \\
&= \sum_{j=0}^{\infty} h_j e(t-j) \\
w(t) &= \frac{F(q)}{D(q)}e(t) \\
&= k_o e(t) + k_1 e(t-1) + k_2 e(t-2) + \ldots \\
&= \sum_{j=0}^{\infty} k_j e(t-j)
\end{aligned}$$

Noting that $e(t-i)$ and $e(t-j)$ are uncorrelated if $i \neq j$, we get

$$\begin{aligned}
Ey(t)w(t) &= E[h_o k_o e^2(t)] + E[h_1 k_1 e^2(t-1)] + \ldots \\
&= \sum_{j=0}^{\infty} h_j k_j E e^2(t-j) = \lambda^2 \sum_{j=0}^{\infty} h_j k_j
\end{aligned}$$

Example 1. We get in this case

$$\begin{aligned}
\frac{C(q)}{A(q)} = \frac{q}{q+a} = \frac{1}{1+aq^{-1}} &= 1 - aq^{-1} + a^2 q^{-2} - a^3 q^{-3} + \ldots \\
&= \sum_{j=0}^{\infty} (-a)^j q^{-j}
\end{aligned}$$

Note that due to the stability assumption, $|a| < 1$, and the series converges. We have also

$$\frac{F(q)}{D(q)} = \sum_{j=0}^{\infty} (-a)^j q^{-j}$$

For the case $\tau = 0$, we get

$$\begin{aligned}
r_y(0) &= E[e(t) - ae(t-1) + a^2 e(t-2) + \ldots][e(t) - ae(t-1) + \ldots] \\
&= \lambda^2 (1 + a^2 + a^4 + \ldots) = \frac{\lambda^2}{1-a^2}
\end{aligned}$$

while the cases $\tau = 1$ and $\tau = 2$ are handled as

$$\begin{aligned}
r_y(1) &= E[e(t+1) - ae(t) + a^2 e(t-1) + \ldots][e(t) - ae(t-1) + \ldots] \\
&= \lambda^2 (-a - a^3 - a^5 + \ldots) = \frac{-a\lambda^2}{1-a^2}
\end{aligned}$$

$$\begin{aligned} r_y(2) &= E[e(t+2) - ae(t+1) + a^2e(t) - a^3e(t-1) + \ldots] \\ &\quad \times[e(t) - ae(t-1) + ae(t-2) + \ldots] \\ &= \lambda^2(a^2 + a^4 + \ldots) = \frac{a^2\lambda^2}{1-a^2} \end{aligned}$$

□

Example 2. As this case corresponds to a second order filter, the division approach is a bit cumbersome. We have

$$\begin{aligned} \frac{b}{(q+a)(q+d)} &= bq^{-2}[1 - aq^{-1} + a^2q^{-2} - a^3q^{-3} + \ldots] \\ &\quad \times[1 - dq^{-1} + d^2q^{-2} - d^3q^{-3} + \ldots] \\ &= bq^{-2}[1 + h_1q^{-1} + h_2q^{-2} + h_3q^{-3} + \ldots] \qquad \text{(say)} \end{aligned}$$

Then the weighting function coefficients (normalized with bq^{-2} as above) are given by

$$h_k = (-a)^k + (-a)^{k-1}(-d) + \ldots + (-d)^k = \frac{(-a)^{k+1} - (-d)^{k+1}}{(-a) - (-d)}$$

We then get

$$\begin{aligned} r_{yu}(0) &= E[be(t-2) + bh_1e(t-3) + bh_2e(t-4) + \ldots] \\ &\quad \times[e(t-1) - de(t-2) + d^2e(t-3) + \ldots] \\ &= \lambda^2(-bd)[1 - dh_1 + d^2h_2 + \ldots] \\ &= -bd\lambda^2[1 - \frac{d}{d-a}\{(-a)^2 - (-d)^2\} + \frac{d^2}{d-a}\{(-a)^3 - (-d)^3\} + \ldots] \\ &= -bd\lambda^2[1 + \frac{1}{d-a}\{\sum_{i=1}^{\infty}(-d)^i(-a)^{i+1}\} + \frac{1}{d-a}\sum_{i=1}^{\infty} d^{2i+1}] \\ &= -bd\lambda^2[1 + \frac{-a}{d-a}\,\frac{ad}{1-ad} + \frac{d^3}{(d-a)(1-d^2)}] \\ &= \frac{-bd\lambda^2}{d-a}\,\frac{1}{(1-ad)(1-d^2)}[(d-a)(1-ad)(1-d^2) \\ &\quad -a^2d(1-d^2) + d^3(1-ad)] \\ &= \frac{-bd\lambda^2}{(1-ad)(1-d^2)} \end{aligned}$$

$$\begin{aligned} r_{yu}(1) &= E[be(t-1) + bh_1e(t-2) + bh_2e(t-3) + \ldots] \\ &\quad \times[e(t-1) - de(t-2) + d^2e(t-3) + \ldots] \\ &= b\lambda^2[1 - dh_1 + d^2h_2 + \ldots] \\ &= r_{yu}(0)/(-d) \\ &= \frac{b\lambda^2}{(1-ad)(1-d^2)} \end{aligned}$$

□

Example 3. The model is already in a weighting function format. As $e(t)$ and $e(s)$ are independent when $t \neq s$, we get

$$\begin{aligned}
r_y(0) &= E[e(t) + c_1 e(t-1) + c_2 e(t-2)]^2 \\
&= \lambda^2(1 + c_1^2 + c_2^2) \\
r_y(1) &= E[e(t+1) + c_1 e(t) + c_2 e(t-1)][e(t) + c_1 e(t-1) + c_2 e(t-2)] \\
&= \lambda^2(c_1 + c_2 c_1) \\
r_y(2) &= E[e(t+2) + c_1 e(t+1) + c_2 e(t)][e(t) + c_1 e(t-1) + c_2 e(t-2)] \\
&= \lambda^2 c_2
\end{aligned}$$

It is clear that we also get

$$r_y(\tau) = 0 \quad \text{if} \;\; \tau > 2$$

□

Example 4. As in Example 1, we have in this case

$$\begin{aligned}
y(t) &= e(t) - ae(t-1) + a^2 e(t-2) - a^2 e(t-3) + \ldots \\
w(t) &= e(t) - de(t-1) + d^2 e(t-2) - d^2 e(t-3) + \ldots
\end{aligned}$$

Hence

$$r_{yw}(0) = Ey(t)w(t) = \lambda^2(1 + ad + a^2 d^2 + \ldots) = \frac{\lambda^2}{1 - ad}$$

If $\tau > 0$

$$\begin{aligned}
r_{yw}(\tau) &= E[e(t+\tau) - ae(t+\tau-1) + \ldots + (-a)^\tau e(t) + (-a)^{\tau+1} e(t-1) + \ldots] \\
&\quad \times [e(t) - de(t-1) + d^2 e(t-2) + \ldots] \\
&= \lambda^2(-a)^\tau[1 + ad + a^2 d^2 + \ldots] = (-a)^\tau \frac{\lambda^2}{1 - ad}
\end{aligned}$$

$$\begin{aligned}
r_{yw}(-\tau) &= E[e(t) - ae(t-1) + a^2 e(t-2) + \ldots] \\
&\quad \times [e(t+\tau) - de(t+\tau-1) + \ldots + (-d)^\tau e(t) + (-d)^{\tau+1} e(t-1) + \ldots] \\
&= \lambda^2(-d)^\tau[1 + ad + a^2 d^2 + \ldots] = (-d)^\tau \frac{\lambda^2}{1 - ad}
\end{aligned}$$

□

4 Method 2 - Yule Walker equations

4.1 The basic approach

The basic idea is to multiply the model by delayed signals $y(t-\tau)$, $w(t-\tau)$. Taking expectations, one will then get some linear equations in the covariance equation. We omit the details for the general case. The procedure is simple for pure AR processes and more involved for full ARMA processes.

Example 1. Multiplying the model by $y(t)$ gives

$$r_y(0) + ar_y(1) = Ey(t)e(t)$$

Noting that $y(t) = e(t) +$ sum of *older* $e(t)$'s gives $Ey(t)e(t) = \lambda^2$. Multiplying the model by $y(t-1)$ gives

$$r_y(1) + ar_y(0) = 0$$

as $y(t-1)$ is independent of $e(t)$. We now solve the linear system of equations

$$\begin{pmatrix} 1 & a \\ a & 1 \end{pmatrix} \begin{pmatrix} r_y(0) \\ r_y(1) \end{pmatrix} = \begin{pmatrix} \lambda^2 \\ 0 \end{pmatrix}$$

with the solution

$$r_y(0) = \frac{\lambda^2}{1-a^2} \qquad r_y(1) = \frac{-a\lambda^2}{1-a^2}$$

Multiplying the model with also $y(t-2)$ gives finally

$$r_y(2) + ar_y(1) = 0$$

from which we easily get

$$r_y(2) = -ar_y(1) = a^2 \frac{\lambda^2}{1-a^2}$$

$\square$

Example 2. The models can be written as

$$\begin{aligned} y(t) + (a+d)y(t-1) + ady(t-2) &= be(t-2) \\ w(t) + dw(t-1) &= e(t-1) \end{aligned}$$

Multiplying the first equation with $w(t-1)$ and the second with $y(t)$ and $y(t-1)$ generates the equations

$$\begin{aligned} r_{yw}(1) + (a+d)r_{yw}(0) + adr_{yw}(-1) &= bEw(t-1)e(t-2) \\ r_{yw}(0) + dr_{yw}(1) &= Ey(t)e(t-1) \\ r_{yw}(-1) + dr_{yw}(0) &= Ey(t-1)e(t-1) \end{aligned}$$

Next, we must determine the right hand sides of these equations. Then recall that

$$\begin{aligned} y(t) &= be(t-2) + \text{ older } e(t)'\text{s} \\ w(t) &= e(t-1) + \text{ older } e(t)'\text{s} \end{aligned}$$

and we get

$$Ew(t-1)e(t-2) = \lambda^2 \qquad Ey(t)e(t-1) = 0 \qquad Ey(t-1)e(t-1) = 0$$

Hence, we have

$$\begin{pmatrix} ad & a+d & 1 \\ 0 & 1 & d \\ 1 & d & 0 \end{pmatrix} \begin{pmatrix} r_{yw}(-1) \\ r_{yw}(0) \\ r_{yw}(1) \end{pmatrix} = \begin{pmatrix} b\lambda^2 \\ 0 \\ 0 \end{pmatrix}$$

with the solution

$$\begin{aligned} r_{yw}(-1) &= \frac{bd^2\lambda^2}{(1-ad)(1-d^2)} \\ r_{yw}(0) &= \frac{-bd\lambda^2}{(1-ad)(1-d^2)} \\ r_{yw}(1) &= \frac{b\lambda^2}{(1-ad)(1-d^2)} \end{aligned}$$

□

Example 3. The Yule-Walker approach offers no advantage in this case and is omitted. □

Example 4. First note that

$$\begin{aligned} y(t) &= e(t) + \text{ "old noise"} \\ w(t) &= e(t) + \text{ "old noise"} \end{aligned}$$

Next, multiply

$$y(t) + ay(t-1) = e(t)$$

with $w(t-\tau)$, $\tau \geq 0$ to get

$$r_{yw}(\tau) + ar_{yw}(\tau - 1) = \begin{cases} \lambda^2 & \tau = 0 \\ 0 & \tau > 0 \end{cases} \tag{4.1}$$

Similarly, multiply

$$w(t) + dw(t-1) = e(t)$$

with $y(t-\tau)$, $\tau \geq 0$ to get

$$r_{yw}(-\tau) + dr_{yw}(-\tau + 1) = \begin{cases} \lambda^2 & \tau = 0 \\ 0 & \tau > 0 \end{cases} \tag{4.2}$$

We next need to combine one equation from (4.1) with another one from (4.2). The choice is not unique. Taking (4.1) with $\tau = 0$ and (4.2) with $\tau = 1$, gives

$$\begin{pmatrix} 1 & a \\ d & 1 \end{pmatrix} \begin{pmatrix} r_{yw}(0) \\ r_{yw}(-1) \end{pmatrix} = \begin{pmatrix} \lambda^2 \\ 0 \end{pmatrix}$$

and hence

$$r_{yw}(0) = \frac{\lambda^2}{1-ad} \qquad r_{yw}(-1) = \frac{-d\lambda^2}{1-ad}$$

From (4.1) with $\tau = 1, 2, \ldots$

$$r_{yw}(\tau) = -ar_{yw}(\tau - 1) = \ldots = (-a)^\tau \frac{\lambda^2}{1-ad}$$

From (4.2) with $\tau = 1, 2, \ldots$

$$r_{yw}(-\tau) = -dr_{yw}(-\tau + 1) = \ldots = (-d)^\tau \frac{\lambda^2}{1-ad}$$

We could also have started by combining (4.1) with $\tau = 1$ and (4.2) with $\tau = 0$. This alternative lead to

$$\begin{pmatrix} 1 & a \\ d & 1 \end{pmatrix} \begin{pmatrix} r_{yw}(1) \\ r_{yw}(0) \end{pmatrix} = \begin{pmatrix} 0 \\ \lambda^2 \end{pmatrix}$$

and

$$r_{yw}(0) = \frac{\lambda^2}{1-ad} \qquad r_{yw}(1) = \frac{-a\lambda^2}{1-ad}$$

From which we can proceed as with the first alternative. □

4.2 A systematic approach based on a Diophantine equation

This approach can be considered as a way to apply the Yule-Walker equations systematically. It is primarily based on the relations

$$\phi_{yw}(\tau) = \lambda^2 \frac{C(e^{i\omega})}{A(e^{i\omega})} \frac{F(e^{-i\omega})}{D(e^{-i\omega})} = \sum_{k=-\infty}^{\infty} r_{yw}(k) e^{-ik\omega} \tag{4.3}$$

In order to simplify the derivation, assume

$$\begin{aligned} \deg A &= \deg C = n \\ \deg D &= \deg F = m \end{aligned}$$

Introduce the polynomials $G(z)$ and $H(z^{-1})$ by the *Diophantine equation*

$$C(z)F(z^{-1}) \equiv G(z)D(z^{-1}) + zA(z)H(z^{-1}) \tag{4.4}$$

where

$$\begin{aligned} G(z) &= g_o z^n + q_1 z^{n-1} + \ldots + g_n \\ H(z^{-1}) &= h_o z^{-m} + \ldots + h_{m-1} z^{-1} \end{aligned}$$

Then, from (4.3), by substituting $e^{i\omega} = z$:

$$\lambda^2 \frac{G(z)}{A(z)} + \lambda^2 \frac{zH(z^{-1})}{D(z^{-1})} = \sum_{k=-\infty}^{\infty} r_{yw}(k) z^{-k} \tag{4.5}$$

Here, we can identify

$$\lambda^2 \frac{G(z)}{A(z)} = \sum_{k=0}^{\infty} r_{yw}(k) z^{-k} \tag{4.6}$$

$$\lambda^2 \frac{zH(z^{-1})}{D(z^{-1})} = \sum_{k=-\infty}^{-1} r_{yw}(k) z^{-k} \tag{4.7}$$

Now, (4.6) can be rewritten as

$$\lambda^2 G(z) = A(z) \sum_{k=0}^{\infty} r_{yw}(k) z^{-k} \tag{4.8}$$

Equating (4.8) for different powers of z gives

$$\begin{array}{r} z^n: \\ z^{n-1}: \\ \vdots \\ z^o: \\ z^{-1}: \end{array} \left\{ \begin{array}{rcl} \lambda^2 g_o & = & 1 \cdot r_{yw}(0) \\ \lambda^2 g_1 & = & 1 \cdot r_{yw}(1) + a_1 r_{yw}(0) \\ & \vdots & \\ \lambda^2 g_n & = & 1 \cdot r_{yw}(n) + \ldots + a_n r_{yw}(0) \\ 0 & = & 1 \cdot r_{yw}(n+1) + \ldots + a_n r_{yw}(1) \end{array} \right.$$

and hence

$$\left\{ \begin{array}{rcll} r_{yw}(0) & = & \lambda^2 g_o & \\ r_{yw}(1) & = & \lambda^2 g_1 - a_1 r_{yw}(0) & \\ \vdots & & & \\ r_{yw}(k) & = & \lambda^2 g_k - \sum_{j=1}^{k} a_j r_{yw}(k-j) & (\text{for} \quad k = 2, \ldots, n) \\ r_{yw}(k) & = & -\sum_{j=1}^{n} a_j r_{yw}(k-j) & (\text{for} \quad \text{all} \quad k > n) \end{array} \right. \tag{4.9}$$

Note that the last equation is nothing but the Yule-Walker equation.

Similarly, we get from (4.7)

$$\lambda^2 z H(z^{-1}) = D(z^{-1}) \sum_{k=-\infty}^{-1} r_{yw} z^{-k} \tag{4.10}$$

and after equating the powers in (4.10), we get

$$\begin{array}{r} z^{-m+1}: \\ z^{-m+2}: \\ \vdots \\ z^o: \\ z^1: \end{array} \left\{ \begin{array}{rcl} \lambda^2 h_o & = & 1 \cdot r_{yw}(-1) \\ \lambda^2 h_1 & = & 1 \cdot r_{yw}(-2) + d_1 r_{yw}(-1) \\ & \vdots & \\ \lambda^2 h_{m-1} & = & 1 \cdot r_{yw}(-m) + \ldots + d_{m-1} r_{yw}(-1) \\ 0 & = & 1 \cdot r_{yw}(-m-1) + \ldots + d_m r_{yw}(-1) \end{array} \right.$$

This leads to

$$\left\{ \begin{array}{rcll} r_{yw}(-1) & = & \lambda^2 h_o & \\ r_{yw}(-k) & = & \lambda^2 h_{k-1} - \sum_{j=1}^{k-1} d_j r_{yw}(-k+j) & (k = 2, \ldots, m) \\ r_{yw}(-k) & = & -\sum_{j=1}^{m} d_j r_{yw}(-k+j) & (\text{for} \quad \text{all} \quad k > m) \end{array} \right. \tag{4.11}$$

Example 1. In this case, $C = F = q$, $A = D = q + a$, $n = m = 1$. The Diophantine equation (4.4) becomes

$$(z)(z^{-1}) \equiv (g_o z + g_1)(z^{-1} + a) + z(z+a) h_o z^{-1}$$

Equating powers of z gives

$$\left\{ \begin{array}{rrcl} z: & 0 & = & g_o a + h_o \\ z^o: & 1 & = & g_o + g_1 a + a h_o \\ z^{-1}: & 0 & = & g_1 \end{array} \right.$$

leading to

$$g_o = \frac{1}{1-a^2} \qquad h_o = \frac{-a}{1-a^2}$$

Next, (4.9) gives

$$\begin{aligned} r_y(0) &= \lambda^2 g_o = \frac{\lambda^2}{1-a^2} \\ r_y(1) &= \lambda^2 g_1 - a r_y(0) = \frac{-a\lambda^2}{1-a^2} \\ r_y(2) &= (-a) r_y(1) = \frac{-a^2\lambda^2}{1-a^2} \end{aligned} \tag{4.12}$$

□

Example 2. In this case

$$\begin{aligned} C &= b \qquad F = 1 \qquad A = (q+a)(q+d) \qquad D = q+d \\ n &= 2 \qquad m = 1 \end{aligned}$$

The Diophantine equation (4.4) becomes

$$b \cdot 1 \equiv (g_o z^2 + g_1 z + g_2)(z^{-1} + d) + z(z+a)(z+d)(h_o z^{-1})$$

Equating powers of z gives

$$\left\{ \begin{array}{rcl} z^2 : & 0 = & g_o d + h_o \\ z : & 0 = & g_o + g_1 d + a h_o + d h_o \\ z^o : & b = & g_1 + g_2 d + a d h_o \\ z^{-1} : & 0 = & g_2 \end{array} \right.$$

leading to

$$g_o = \frac{-bd}{(1-ad)(1-d^2)} \qquad g_1 = \frac{b(1-ad-d^2)}{(1-ad)(1-d^2)} \qquad h_o = \frac{bd^2}{(1-ad)(1-d^2)}$$

Next, (4.9) gives

$$\begin{aligned} r_{yu}(0) &= \lambda^2 g_o = \frac{-bd\lambda^2}{(1-ad)(1-d^2)} \\ r_{yu}(1) &= \lambda^2 [g_1 - (a+d) g_o] \\ &= \frac{b\lambda^2}{(1-ad)(1-d^2)} [1 - ad - d^2 + ad + d^2] = \frac{b\lambda^2}{(1-ad)(1-d^2)} \end{aligned}$$

□

Example 3. Here

$$C = F = q^2 + c_1 q + c_2 \qquad A = D = q^2 \qquad n = m = 2$$

The Diophantine equation will now be

$$\begin{aligned} (z^2 + c_1 z + c_2)(z^{-2} + c_1 z^{-1} + c_2) \equiv\ & (g_o z^2 + g_1 z + g_2) z^{-2} \\ & + z^3 (h_o z^{-2} + h_1 z^{-1}) \end{aligned}$$

Equating powers of z gives

$$\begin{cases} z^2: & c_2 & = & h_1 \\ z: & c_1 + c_1 c_2 & = & h_o \\ z^o: & 1 + c_1^2 + c_2^2 & = & g_o \\ z^{-1}: & c_1 + c_1 c_2 & = & g_1 \\ z^{-2}: & c_2 & = & g_2 \end{cases}$$

and from (4.9) gives

$$\begin{cases} r_y(0) & = & \lambda^2 g_o = \lambda^2(1 + c_1^2 + c_2) \\ r_y(1) & = & \lambda^2 g_1 = \lambda^2(c_1 + c_1 c_2) \\ r_y(2) & = & \lambda^2 g_2 = \lambda^2 c_2 \end{cases}$$

□

Example 4. In this case

$$C = 1 \qquad F = 1 \qquad A = q + a \qquad D = q + d \qquad n = m = 1$$

The Diophantine equation (4.4) gives

$$1 \cdot 1 = (g_o z + g_1)(z^{-1} + d) + z(z + a)h_o z^{-1}$$

and hence

$$\begin{cases} z^1: & 0 & = & g_o d + h_o \\ z^o: & 1 & = & g_o + g_1 d + a h_o \\ z^{-1}: & 0 & = & g_1 \end{cases}$$

leading to

$$g_o = \frac{1}{1 - ad} \qquad h_o = \frac{-d}{1 - ad}$$

Next, the use of (4.9) yields

$$\begin{cases} r_{yw}(0) & = & \lambda^2 g_o = \frac{\lambda^2}{1-ad} \\ r_{yw}(1) & = & \lambda^2 g_1 - a r_{yw}(0) = \frac{-a}{1-ad}\lambda^2 \\ r_{yw}(k) & = & -a r_{yw}(k-1) = \ldots = \frac{(-a)^k}{1-ad}\lambda^2 & \qquad k \geq 1 \end{cases}$$

Further, (4.11) gives

$$\begin{cases} r_{yw}(-1) & = & \lambda^2 h_o = \frac{-d}{1-ad}\lambda^2 \\ r_{yw}(-k) & = & -d r_{yw}(-k+1) = \frac{(-d)^k}{1-ad}\lambda^2 & \qquad k \geq 1 \end{cases}$$

□

5 Method 3 - Integrals around the unit circle

This approach is based on the relations

$$\begin{aligned} \phi(\omega) &= \sum_{k=-\infty}^{\infty} r(k) e^{-ik\omega} \\ r(k) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi(\omega) e^{ik\omega} d\omega \end{aligned}$$

that hold between the spectral density and the covariance function.

The cross-spectrum $\phi_{yw}(\omega)$ for the processes of (1.1) can be written as

$$\phi_{yw}(\omega) = \frac{C(e^{i\omega})}{A(e^{i\omega})} \frac{F(e^{-i\omega})}{D(e^{-i\omega})} \lambda^2$$

The cross-covariance function is then obtained from the integral formula. By changing the independent variable from ω to $z = e^{i\omega}$, the integration path becomes the unit circle counterclockwise. Noting that $dz = izd\omega$, we get

$$\begin{aligned} r_{yw}(k) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{C(e^{i\omega})}{A(e^{i\omega})} \frac{F(e^{-i\omega})}{D(e^{-i\omega})} \lambda^2 e^{ik\omega} d\omega \\ &= \frac{1}{2\pi i} \oint \frac{C(z)}{A(z)} \frac{F(z^{-1})}{D(z^{-1})} \lambda^2 z^k \frac{dz}{z} \end{aligned} \tag{5.1}$$

Sometimes, it is more convenient to make the transformation $z = e^{-i\omega}$, which leads to

$$r_{yw}(k) = \frac{1}{2\pi i} \oint \frac{C(z^{-1})}{A(z^{-1})} \frac{F(z)}{D(z)} \lambda^2 z^{-k} \frac{dz}{z} \tag{5.2}$$

where $\oint$ still denotes integration around the unit circle counterclockwise. The integrals (5.1), (5.2) can be evaluated using residue calculus. This is a feasible approach, in particular for low order models with distinct poles inside the unit circle.

Example 1. In this case, we have

$$\begin{aligned} r_y(\tau) &= \frac{\lambda^2}{2\pi i} \oint \frac{z}{z+a} \frac{z^{-1}}{z^{-1}+a} z^\tau \frac{dz}{z} \\ &= \frac{\lambda^2}{2\pi i} \oint \frac{z^\tau}{(z+a)(1+az)} dz \end{aligned}$$

The only pole inside the unit circle (as $\tau \geq 0$ due to the state problem formulation) is $z = -a$. By residue calculus,

$$r_y(\tau) = \lambda^2 \frac{(-a)^\tau}{1-a^2}$$

Should we want to evaluate $r_y(\tau)$ for negative τ, the alternative form (5.2) (corresponding to the transformation $\tilde{z} = -z^{-1}$) would lead to the integral

$$r_y(-\tau) = \frac{\lambda^2}{2\pi i} \oint \frac{\tilde{z}^{-\tau}}{(\tilde{z}+a)(1+a\tilde{z})} d\tilde{z}$$

and we are back to the above calculation. □

Example 2. We get

$$\begin{aligned} r_{yu}(\tau) &= \frac{\lambda^2}{2\pi i} \oint \frac{b}{(z+a)(z+d)} \frac{1}{z^{-1}+d} z^\tau \frac{dz}{z} \\ &= \frac{\lambda^2}{2\pi i} \oint \frac{bz^\tau}{(z+a)(z+d)(1+dz)} dz \end{aligned}$$

The integrand has poles in $z = -a$ and $z = -d$ inside the unit circle. Residue calculus gives

$$\begin{aligned} r_{yu}(\tau) &= \lambda^2 \left[\frac{b(-a)^\tau}{(d-a)(1-ad)} + \frac{b(-d)^\tau}{(a-d)(1-d^2)} \right] \\ &= \frac{b\lambda^2}{(d-a)} \frac{(-a)^\tau(1-d^2) - (-d)^\tau(1-ad)}{(1-ad)(1-d^2)} \end{aligned}$$

As the numerator will vanish if $a = d$, we can cancel a factor $(d-a)$. We consider the values $\tau = 0$ and 1 separately:

$$\begin{aligned} r_{yw}(0) &= \frac{b\lambda^2}{d-a} \frac{1-d^2-1+ad}{(1-ad)(1-d^2)} = \frac{-bd\lambda^2}{(1-ad)(1-d^2)} \\ r_{yw}(1) &= \frac{b\lambda^2}{d-a} \frac{-a+ad^2+d-ad^2}{(1-ad)(1-d^2)} = \frac{b\lambda^2}{(1-ad)(1-d^2)} \end{aligned}$$

□

Example 3. Here, we have

$$\begin{aligned} r_{yu}(\tau) &= \frac{\lambda^2}{2\pi i} \oint (z^2 + c_1 z + c_2)(z^{-2} + c_1 z^{-1} + c_2) z^\tau \frac{dz}{z} \\ &= \frac{\lambda^2}{2\pi i} \oint [(1 + c_1^2 + c_2^2) + (z + z^{-1})(c_1 + c_1 c_2) + (z^2 + z^{-2}) c_2] z^\tau \frac{dz}{z} \end{aligned}$$

Recalling that

$$\frac{1}{2\pi i} \oint z^k \frac{dz}{z} = \delta_{k,0} = \begin{cases} 1 & \text{if} \quad k = 0 \\ 0 & \text{if} \quad k \neq 0 \end{cases}$$

we get

$$\begin{aligned} r_{yu}(\tau) = &\ (1 + c_1^2 + c_2^2)\delta_{\tau,0} + (c_1 + c_1 c_2)[\delta_{\tau,1} + \delta_{\tau,-1}] \\ &+ c_2[\delta_{\tau,2} + \delta_{\tau,-2}] \end{aligned}$$

□

Example 4. Following the recipe, we get

$$r_{yw}(\tau) = \frac{\lambda^2}{2\pi i} \oint \frac{1}{z+a} \frac{1}{z^{-1}+d} z^\tau \frac{dz}{z}$$

Assume first that $\tau \geq 0$. We then get ($z = -a$ will be the only pole inside the unit circle),

$$\begin{aligned} r_{yw}(\tau) &= \frac{\lambda^2}{2\pi i} \oint \frac{1}{z+a} \frac{1}{1+dz} z^\tau dz \\ &= \frac{\lambda^2}{1-ad} (-a)^\tau \end{aligned}$$

Next, consider the case $\tau = -2$. We then get

$$r_{yu}(-2) = \frac{\lambda^2}{2\pi i} \oint \frac{1}{z^2(z+a)(1+dz)} dz \tag{5.3}$$

We have a distinct zero in $z = -a$ and a *double* pole in $z = 0$. Multiple poles imply that use of residue calculus requires some extra effort. In this example, this difficulty can be circumvented by a change of variables $z \to \tilde{z} = z^{-1}$.

Should we insist to proceed with the integral (5.3), we get

$$\begin{aligned}
r_{yw}(-2) &= \lambda^2 \mathrm{Res}_{z=0}\left[\frac{1}{(z+a)(1+dz)}\right] + \lambda^2 \frac{1}{(-a)^2(1-ad)} \\
&= \lambda^2 \frac{d}{dz}\left[\frac{1}{(z+a)(1+dz)}\right]_{|z=0} + \frac{\lambda^2}{a^2(1-ad)} \\
&= -\lambda^2 \frac{(1+dz)+d(z+a)}{(z+a)^2(1+dz)^2}_{|z=0} + \frac{\lambda^2}{a^2(1-ad)} \\
&= -\lambda^2 \frac{1+ad}{a^2} + \frac{\lambda^2}{a^2(1-ad)} = \frac{\lambda^2}{a^2(1-ad)}[-(1-a^2d^2)+1] = \frac{\lambda^2 d^2}{1-ad}
\end{aligned}$$

□

6 Method 4 - State space form

The basic idea of this method is to rewrite the model into state space form

$$\begin{aligned}
x(t+1) &= Fx(t) + Ge(t) \\
z(t) &= Hx(t)
\end{aligned} \tag{6.1}$$

The variables of interest ($y(t+\tau)$ and $w(t)$) should if possible appear in the state vector $x(t)$, but otherwise in the output vector $z(t)$.

We first derive an equation for the stationary state covariance matrix $P = Ex(t)x^T(t)$ $= Ex(t+1)x^T(t+1)$. Noting that $x(t)$ depends on $e(s)$ for $s < t$, we find that $x(t)$ and $e(t)$ are uncorrelated. Using this fact, we get

$$\begin{aligned}
P &= Ex(t+1)x^T(t+1) = E[Fx(t) + Ge(t)][x^T(t)F^T + e(t)G^T] \\
&= F[Ex(t)x^T(t)]F^T + GEe^2(t)G^T \\
&= FPF^T + \lambda^2 GG^T
\end{aligned} \tag{6.2}$$

This is a so-called *Lyapunov equation.* Noting that P is a symmetric matrix, it can be formulated elementwise as a linear system of equations with $\frac{n(n+1)}{2}$ unknowns.

Next, we consider the covariance *function* of $x(t)$. Let $\tau > 0$. Then

$$\begin{aligned}
r_x(\tau) &= Ex(t+\tau)x^T(t) \\
&= E[Fx(t+\tau-1) + Ge(t+\tau-1)]x^T(t) \\
&= Fr_x(\tau-1) = \ldots \\
&= F^\tau r_x(0) = F^\tau P
\end{aligned}$$

The covariance function of the outputs is then readily obtained as

$$\begin{aligned}
r_y(\tau) &= Ez(t+\tau)z^T(t) \\
&= EHx(t+\tau)x^T(t)H^T \\
&= HF^\tau PH^T
\end{aligned}$$

Example 1. The simplest state space model in this case is

$$x(t+1) = -ax(t) + e(t)$$

(corresponding to $F = -a$, $G = 1$, $H = 1$). The Lyapunov equation becomes $P = a^2 P + \lambda^2$ giving $P = \lambda^2/(1-a^2)$. The covariance function of $x(t) \equiv y(t)$ is

$$r_y(\tau) = (-a)^\tau \frac{\lambda^2}{1-a^2} \qquad \tau \geq 0$$

A more complicated alternative would be to set

$$x(t) = \begin{pmatrix} y(t) \\ y(t-1) \\ y(t-2) \end{pmatrix}$$

as state vector. The corresponding P matrix will then be related to the covariance function as

$$P = \begin{pmatrix} r_y(0) & r_y(1) & r_y(2) \\ r_y(1) & r_y(0) & r_y(1) \\ r_y(2) & r_y(1) & r_y(0) \end{pmatrix}$$

The state space model will be

$$x(t+1) = \begin{pmatrix} -a & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} x(t) + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} e(t)$$

Without imposing more than symmetry on P, we set

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{12} & p_{22} & p_{23} \\ p_{13} & p_{23} & p_{33} \end{pmatrix}$$

and the Lyapunov equation becomes

$$\begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{12} & p_{22} & p_{23} \\ p_{13} & p_{23} & p_{33} \end{pmatrix} = \begin{pmatrix} -a & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{12} & p_{22} & p_{23} \\ p_{13} & p_{23} & p_{33} \end{pmatrix} \begin{pmatrix} -a & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} + \lambda^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Evaluating the various matrix elements (and again using the symmetry) gives

$$\begin{aligned} p_{11} &= a^2 p_{11} + \lambda^2 \\ p_{12} &= -a p_{11} \\ p_{13} &= -a p_{12} \\ p_{22} &= p_{11} \\ p_{23} &= p_{12} \\ p_{33} &= p_{22} \end{aligned}$$

leading quickly to

$$P = \frac{\lambda^2}{1-a^2}\begin{pmatrix} 1 & -a & a^2 \\ -a & 1 & -a \\ a^2 & -a & 1 \end{pmatrix}$$

$\square$

Example 2. It is natural to use a state vector

$$x(t) = \begin{pmatrix} u(t) \\ y(t) \end{pmatrix}$$

This leads to

$$x(t+1) = \begin{pmatrix} -d & 0 \\ b & -a \end{pmatrix} x(t) + \begin{pmatrix} 1 \\ 0 \end{pmatrix} v(t)$$

The Lyapunov equation becomes

$$P = \begin{pmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{pmatrix} = \begin{pmatrix} -d & 0 \\ b & -a \end{pmatrix}\begin{pmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{pmatrix}\begin{pmatrix} -d & b \\ 0 & -a \end{pmatrix} + \lambda^2 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Equating the elements gives

$$\begin{aligned} p_{11} &= d^2 p_{11} + \lambda^2 \\ p_{12} &= -bdp_{11} + adp_{12} \\ p_{22} &= b^2 p_{11} + a^2 p_{22} - 2abp_{12} \end{aligned}$$

with the solution

$$p_{11} = \frac{\lambda^2}{1-d^2} \qquad p_{12} = \frac{-bd\lambda^2}{(1-d^2)(1-ad)} \qquad p_{22} = \frac{b^2(1+ad)\lambda^2}{(1-a^2)(1-ad)(1-d^2)}$$

Here, we can by construction identify $r_{yu}(0) = p_{12}$. We also have

$$r_x(1) = E\begin{pmatrix} u(t+1) \\ y(t+1) \end{pmatrix}(u(t) \;\; y(t)) = \begin{pmatrix} r_u(1) & r_{uy}(1) \\ r_{yu}(1) & r_y(1) \end{pmatrix}$$

Hence

$$\begin{aligned} r_{yu}(1) &= (0 \;\; 1) r_x(1) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = (0 \;\; 1) FP \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= (0 \;\; 1) \begin{pmatrix} -d & 0 \\ b & -a \end{pmatrix}\begin{pmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= (b \;\; -a)\begin{pmatrix} p_{11} \\ p_{12} \end{pmatrix} = \frac{b\lambda^2}{(1-d^2)(1-ad)} \end{aligned}$$

$\square$

Example 3. We take in this case

$$x(t) = \begin{pmatrix} e(t) \\ e(t-1) \\ e(t-2) \end{pmatrix}$$

giving

$$\begin{aligned} x(t+1) &= \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} x(t) + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} e(t) \\ y(t) &= (1 \;\; c_1 \;\; c_2) x(t) \end{aligned}$$

The Lyapunov equation gives

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{12} & p_{22} & p_{23} \\ p_{13} & p_{23} & p_{33} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{12} & p_{22} & p_{23} \\ p_{13} & p_{23} & p_{33} \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} + \lambda^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

which leads to the fairly natural result $P = \lambda^2 I$!

We next get

$$\begin{aligned} r_y(0) &= HPH^T = (1 \;\; c_1 \;\; c_2)\lambda^2 I \begin{pmatrix} 1 \\ c_1 \\ c_2 \end{pmatrix} = \lambda^2(1 + c_1^2 + c_2^2) \\ r_y(1) &= HFPH^T = (1 \;\; c_1 \;\; c_2)\lambda^2 \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ c_1 \\ c_2 \end{pmatrix} = \lambda^2(c_1 + c_1 c_2) \\ r_y(2) &= HF^2PH^T = \lambda^2(1 \;\; c_1 \;\; c_2) \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ c_1 \\ c_2 \end{pmatrix} = \lambda^2 c_2 \end{aligned}$$

As $F^3 = 0$, we have $r_y(\tau) = 0$ for $\tau \geq 3$. $\square$

Example 4. In this case, it is natural to choose $y(t)$ and $w(t)$ as state variables. With

$$x(t) = \begin{pmatrix} y(t) \\ w(t) \end{pmatrix}$$

we have

$$x(t+1) = \begin{pmatrix} -a & 0 \\ 0 & -d \end{pmatrix} x(t) + \begin{pmatrix} 1 \\ 1 \end{pmatrix} e(t)$$

The Lyapunov equation will be

$$P = \begin{pmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{pmatrix} = \begin{pmatrix} -a & 0 \\ 0 & -d \end{pmatrix} \begin{pmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{pmatrix} \begin{pmatrix} -a & 0 \\ 0 & -d \end{pmatrix} + \lambda^2 \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

and hence

$$\begin{aligned} p_{11} &= a^2 p_{11} + \lambda^2 \\ p_{12} &= ad p_{12} + \lambda^2 \\ p_{22} &= d^2 p_{22} + \lambda^2 \end{aligned} \implies P = \begin{pmatrix} \lambda^2/(1-a^2) & \lambda^2/(1-ad) \\ \lambda^2/(1-ad) & \lambda^2/(1-d^2) \end{pmatrix}$$

Let $\tau \geq 0$. Then

$$\begin{aligned} r_x(\tau) &= \begin{pmatrix} r_y(\tau) & r_{yw}(\tau) \\ r_{wy}(\tau) & r_w(\tau) \end{pmatrix} = \begin{pmatrix} r_y(\tau) & r_{yw}(\tau) \\ r_{yw}(-\tau) & r_w(\tau) \end{pmatrix} \\ &= \begin{pmatrix} -a & 0 \\ 0 & -d \end{pmatrix}^\tau P = \begin{pmatrix} (-a)^\tau & 0 \\ 0 & (-d)^\tau \end{pmatrix} \begin{pmatrix} \lambda^2/(1-a^2) & \lambda^2/(1-ad) \\ \lambda^2/(1-ad) & \lambda^2/(1-d^2) \end{pmatrix} \end{aligned}$$

Inspecting the nondiagonal element, we have

$$
\begin{aligned}
r_{yw}(\tau) &= \frac{(-a)^{\tau}\lambda^2}{1-ad} \\
r_{yw}(-\tau) &= \frac{(-d)^{\tau}\lambda^2}{1-ad}
\end{aligned}
$$

□

7 Exercises

1. Determine the autocovariance function $r(\tau)$ [for all τ] for the ARMA(1,1) process
$$y(t) + ay(t-1) = e(t) + ce(t-1)$$

2. Determine the autocovariance function $r(\tau)$ for $\tau = 0, 1, 2$ for the AR(2) process
$$y(t) + a_1 y(t-1) + a_2 y(t-2) = e(t)$$

The following problems in

> T Söderström and P Stoica: *System Identification* (Prentice Hall International, Hemel Hempstead, UK, 1989)

do all involve covariance calculations:
2.4, 2.5, 7.14, 8.3, 8.6, 8.7, 10.1, 10.2, 11.7, 11.8, 12.4, 12.7.