Predictors

Exercise 1

Consider the following process:

$$y(t) = \frac{1}{4} \cdot y(t-1) + e(t) + 2 \cdot e(t-1), \qquad e(t) \sim WN\left(0, \frac{1}{4}\right)$$

1. Derive the expression for the 1-step optimal predictor from the available data.
2. What is the value of the 1-step prediction error variance?
3. Derive the expression for the 2-step optimal predictor from the available data.
4. What is the value of the 2-step prediction error variance?
5. Given the following observations:

$$y(1) = 1, \quad y(2) = \frac{1}{2}, \quad y(3) = -\frac{1}{2}, \quad y(4) = 0, \quad y(5) = -\frac{1}{2}$$

compute $\widehat{y}(6|5)$ and $\widehat{y}(7|5)$.

1) Derive the expression for the 1-step optimal predictor from the available data.

The process is an ARMA$(1,1)$ with the following *operatorial representation:*

$$\begin{aligned} y(t) &= W(z) \cdot e(t) \\ &= \frac{C(z)}{A(z)} \cdot e(t) \\ &= \frac{1 + 2 \cdot z^{-1}}{1 - \frac{1}{4} \cdot z^{-1}} \cdot e(t) \\ &= \frac{z+2}{z - \frac{1}{4}} \cdot e(t) \end{aligned}$$

The pole of $W(z)$ is $z = \frac{1}{4}$ and it is inside the unit circle, so the digital filter $W(z)$ is asymptotically stable. Moreover $e(t)$ is a WSS process. Thus $y(t)$ is a WSS process.

In order to find the predictor, first of all, we have to check if $y(t)$ is a *canonical representation:*

1. Same degree: $\checkmark$;
2. Coprime: $\checkmark$;
3. Monic: $\checkmark$;
4. Roots inside the unitary circle: $\times$ the zeroes (roots of the numerator) are outside!

In order to replace the zero outside the unitary circle, we apply the all pass filter:

$$T(z) = a \cdot \frac{z + \frac{1}{a}}{z + a} = 2 \cdot \frac{z + \frac{1}{2}}{z + 2}$$

So the new dynamic filter is:

$$\begin{aligned} W_1(z) &= W(z) \cdot T(z) \\ &= \frac{\cancel{z+2}}{z - \frac{1}{4}} \cdot 2 \cdot \frac{z + \frac{1}{2}}{\cancel{z+2}} \\ &= 2 \cdot \frac{z + \frac{1}{2}}{z - \frac{1}{4}} \end{aligned}$$

Here the polynomials are not monic. Thus, we define:

$$\eta(t) = 2 \cdot e(t) \implies e(t) = \frac{1}{2} \cdot \eta(t)$$

With the following statistical properties:

$$m_\eta = \mathbb{E}[\eta(t)] = 2 \cdot \mathbb{E}[e(t)] = 2 \cdot m_e = 0$$

$$\lambda_\eta^2 = \mathbb{E}\left[(\eta(t) - m_\eta)^2\right] = \mathbb{E}\left[\eta(t)^2\right] = 4 \cdot \mathbb{E}\left[e(t)^2\right] = 4 \cdot \lambda_e^2 = 1$$

So:

$$\eta(t) \sim WN(0, 1)$$

We can conclude that:

$$\begin{aligned} y(t) &= W_1(z) \cdot e(t) \\ &= 2 \cdot \frac{z + \frac{1}{2}}{z - \frac{1}{4}} \cdot e(t) \\ &= \frac{z + \frac{1}{2}}{z - \frac{1}{4}} \cdot \eta(t) \\ &= \frac{1 + \frac{1}{2} \cdot z^{-1}}{1 - \frac{1}{4} \cdot z^{-1}} \cdot \eta(t) \\ &= \frac{C(z)}{A(z)} \cdot \eta(t) \end{aligned}$$

This is the *canonical representation* of the process $y(t)$.

Now we compute 1 step of the polynomial long division:

$$\begin{array}{cc|cc} 1 & +\frac{1}{2} \cdot z^{-1} & 1 & -\frac{1}{4} \cdot z^{-1} \\ \cline{3-4} -1 & +\frac{1}{4} \cdot z^{-1} & 1 & \\ \hline / & +\frac{3}{4} \cdot z^{-1} & & \end{array}$$

where $C(z) = 1 + \frac{1}{2} \cdot z^{-1}$, $A(z) = 1 - \frac{1}{4} \cdot z^{-1}$, $Q_1(z) = 1$ and $R_1(z) = \frac{3}{4} \cdot z^{-1}$.

Thus we have:

$$\begin{aligned} y(t) &= \left(Q_1(z) + \frac{R_1(z)}{A(z)} \right) \cdot \eta(t) \\ &= Q_1(z) \cdot \eta(t) + \frac{R_1(z)}{A(z)} \cdot \eta(t) \\ &= \eta(t) + \frac{\frac{3}{4} \cdot z^{-1}}{1 - \frac{1}{4} \cdot z^{-1}} \cdot \eta(t) \end{aligned}$$

While the first term is unpredictable with the information at time $t-1$, the second term is totally predictable since it depends on $\eta(t-1)$. Thus the 1-step optimal predictor from the noise is:

$$\begin{aligned} \hat{y}(t|t-1) &= \frac{\frac{3}{4} \cdot z^{-1}}{1 - \frac{1}{4} \cdot z^{-1}} \cdot \eta(t) \\ &= \frac{R_1(z)}{A(z)} \cdot \eta(t) \end{aligned}$$

Whitening filter from an ARMA model:

$$\begin{aligned} y(t) &= \frac{C(z)}{A(z)} \cdot e(t) \\ A(z) \cdot y(t) &= C(z) \cdot e(t) \end{aligned}$$

The whitening filter is then:

$$e(t) = \frac{A(z)}{C(z)} \cdot y(t) = \tilde{W}(z) \cdot y(t)$$

In this particular example we have that:

$$y(t) = \frac{C(z)}{A(z)} \cdot \eta(t) \implies \eta(t) = \frac{A(z)}{C(z)} \cdot y(t)$$

Optimal predictor from an ARMA model:
The predictor frorm the noise is:

$$
\begin{aligned}
y(t) &= \frac{C(z)}{A(z)} \cdot e(t) \\
y(t) &= \underbrace{\left(\frac{R_r(z)}{A(z)} + Q_r(z)\right)}_{\text{r-steps of long division}} \cdot e(t) \\
y(t) &= \underbrace{\frac{R_r(z)}{A(z)} \cdot e(t)}_{\text{r-steps predictor}} + \underbrace{Q_r(z) \cdot e(t)}_{\text{r-steps prediction error}}
\end{aligned}
$$

The predictor from the noise is:

$$
\hat{y}(t|t-r) = \frac{R_r(z)}{A(z)} \cdot e(t)
$$

The prediction error is:

$$
\varepsilon(t) = Q_r(z) \cdot e(t)
$$

The predictor from the available data, using the whitening filter, is:

$$
\begin{aligned}
\hat{y}(t|t-r) &= \frac{R_r(z)}{A(z)} \cdot e(t) \\
\hat{y}(t|t-r) &= \frac{R_r(z)}{A(z)} \cdot \underbrace{\left(\frac{A(z)}{C(z)} \cdot y(t)\right)}_{\text{whitening filter}} \\
&= \frac{R_r(z)}{C(z)} \cdot y(t) \\
&= \frac{\tilde{R}_r(z)}{C(z)} \cdot y(t-r)
\end{aligned}
$$

Thus the 1-step optimal predictor from the available data is:

$$
\begin{aligned}
\hat{y}(t|t-1) &= \frac{R_1(z)}{A(z)} \cdot \overbrace{\left(\frac{A(z)}{C(z)} \cdot y(t)\right)}^{\text{whitening filter}} \\
&= \frac{\frac{3}{4} \cdot z^{-1}}{\cancel{1 - \frac{1}{4} \cdot z^{-1}}} \cdot \frac{\cancel{1 - \frac{1}{4} \cdot z^{-1}}}{1 + \frac{1}{2} \cdot z^{-1}} \cdot y(t) \\
&= \frac{\frac{3}{4} \cdot z^{-1}}{1 + \frac{1}{2} \cdot z^{-1}} \cdot y(t)
\end{aligned}
$$

with the recursive time-domain representation:

$$
\begin{aligned}
\hat{y}(t|t-1) &= \frac{\frac{3}{4}\cdot z^{-1}}{1+\frac{1}{2}\cdot z^{-1}} \cdot y(t) \\
\left(1+\frac{1}{2}\cdot z^{-1}\right)\cdot \hat{y}(t|t-1) &= \left(\frac{3}{4}\cdot z^{-1}\right)\cdot y(t) \\
\hat{y}(t|t-1) + \frac{1}{2}\cdot \hat{y}(t-1|t-2) &= \frac{3}{4}\cdot y(t-1) \\
\hat{y}(t|t-1) &= -\frac{1}{2}\cdot \hat{y}(t-1|t-2) + \frac{3}{4}\cdot y(t-1)
\end{aligned}
$$

Notice that this is obviously equal to the one-step forward shifting predictor due to the stationary properties:

$$
\hat{y}(t+1|t) = -\frac{1}{2}\cdot \hat{y}(t|t-1) + \frac{3}{4}\cdot y(t)
$$

A simpler way to find the 1-step predictor of an ARMA process is given by the theory.

1-step optimal predictor from an ARMA model:

$$
\hat{y}(t|t-1) = \frac{C(z) - A(z)}{C(z)} \cdot y(t)
$$

In this particolar example we have:

$$
\begin{aligned}
\hat{y}(t|t-1) &= \frac{C(z) - A(z)}{C(z)} \cdot y(t) \\
&= \frac{1+\frac{1}{2}\cdot z^{-1} - \left(1-\frac{1}{4}\cdot z^{-1}\right)}{1+\frac{1}{2}\cdot z^{-1}} \cdot y(t) \\
&= \frac{\frac{3}{4}\cdot z^{-1}}{1+\frac{1}{2}\cdot z^{-1}} \cdot y(t)
\end{aligned}
$$

which gives the previously computed predictor.

2) What is the value of the 1-step prediction error variance?

The 1-step prediction error variance is simply given by the unpredictable part:

$$
\begin{aligned}
\mathbb{E}\left[\varepsilon(t)^2\right] &= \mathbb{E}\left[\left(y(t) - \hat{y}(t|t-1)\right)^2\right] \\
&= \mathbb{E}\left[\left(Q_1(z)\cdot \eta(t)\right)^2\right] \\
&= \mathbb{E}\left[\eta(t)^2\right] = \lambda_\eta^2 = 1
\end{aligned}
$$

3) Derive the expression for the 2-step optimal predictor from the available data.

We have already che process in its canonical representation. Let's compute 2 steps of the polynomial long division:

$$
\begin{array}{ccc|cc}
1 & +\frac{1}{2}\cdot z^{-1} & & 1 & -\frac{1}{4}\cdot z^{-1} \\
\cline{4-5}
-1 & +\frac{1}{4}\cdot z^{-1} & & 1 & +\frac{3}{4}\cdot z^{-1} \\
\cline{1-3}
/ & +\frac{3}{4}\cdot z^{-1} & & & \\
 & -\frac{3}{4}\cdot z^{-1} & +\frac{3}{16}\cdot z^{-2} & & \\
\cline{1-3}
 & / & +\frac{3}{16}\cdot z^{-2} & &
\end{array}
$$

where $C(z) = 1 + \frac{1}{2}\cdot z^{-1}$, $A(z) = 1 - \frac{1}{4}\cdot z^{-1}$, $Q_2(z) = 1 + \frac{3}{4}\cdot z^{-1}$ and $R_2(z) = \frac{3}{16}\cdot z^{-2}$.

Notice that:

$$R_2(z) = z^{-r}\cdot \tilde{R}_2(z)$$

Since, in this case, $r = 2$, we have:

$$R_2(z) = z^{-2}\cdot \tilde{R}_2(z) \implies \tilde{R}_2(z) = \frac{3}{16}$$

The 2-step optimal predictor is then given by:

$$
\begin{aligned}
\hat{y}(t|t-2) &= \frac{R_2(z)}{A(z)}\cdot\frac{A(z)}{C(z)}\cdot y(t) \\
&= \frac{\tilde{R}_2(z)}{A(z)}\cdot\frac{A(z)}{C(z)}\cdot y(t-2) \\
&= \frac{\tilde{R}_2(z)}{C(z)}\cdot y(t-2) \\
&= \frac{\frac{3}{16}}{1+\frac{1}{2}\cdot z^{-1}}\cdot y(t-2)
\end{aligned}
$$

with the recursive time-domain representation:

$$
\begin{aligned}
\hat{y}(t|t-2) &= \frac{\frac{3}{16}}{1+\frac{1}{2}\cdot z^{-1}}\cdot y(t-2) \\
\left(1+\frac{1}{2}\cdot z^{-1}\right)\cdot\hat{y}(t|t-2) &= \frac{3}{16}\cdot y(t-2) \\
\hat{y}(t|t-2) + \frac{1}{2}\cdot\hat{y}(t-1|t-3) &= \frac{3}{16}\cdot y(t-2) \\
\hat{y}(t|t-2) &= -\frac{1}{2}\cdot\hat{y}(t-1|t-3) + \frac{3}{16}\cdot y(t-2)
\end{aligned}
$$

4) What is the value of the 2-step prediction error variance?

The 2-step prediction error variance is simply given by the unpredictable part:

$$\begin{aligned}
\mathbb{E}\left[\varepsilon(t)^2\right] &= \mathbb{E}\left[(y(t)-\hat{y}(t|t-2))^2\right] \\
&= \mathbb{E}\left[(Q_2(z)\cdot\eta(t))^2\right] \\
&= \mathbb{E}\left[\left(\eta(t)+\frac{3}{4}\cdot\eta(t-1)\right)^2\right] \\
&= \mathbb{E}\left[\eta(t)^2+\frac{9}{16}\cdot\eta(t-1)^2+\frac{3}{4}\cdot\eta(t)\cdot\eta(t-1)\right] \\
&= 1\cdot\lambda_\eta^2+\frac{9}{16}\cdot\lambda_\eta^2+\frac{3}{4}\cdot\underbrace{\cancel{\mathbb{E}[\eta(t)\cdot\eta(t-1)]}}_{\eta\sim WN} \\
&= 1+\frac{9}{16}=\frac{25}{16}
\end{aligned}$$

Observe the variances:

- Process variance (not explicitly computed):

$$\gamma_y(0)=\frac{8}{5}=1.6$$

- 1-step prediction error variance:

$$\mathbb{E}\left[\varepsilon(t)^2\right]=1$$

- 2-step prediction error variance:

$$\mathbb{E}\left[\varepsilon(t)^2\right]=\frac{25}{16}\approx 1.56$$

The variance of the prediction error tends to the variance of the process, since the r-step predictor tends to the process mean for $r\to\infty$ (the best future prediction when the future is far away is the mean value of the process):

$$\mathbb{E}\left[\varepsilon(t)^2\right]=\mathbb{E}\left[(y(t)-\hat{y}(t|t-r))^2\right]\underset{r\to\infty}{\longrightarrow}\mathbb{E}\left[(y(t)-m_y)^2\right]=\gamma_y(0)$$

4) Computation of $\widehat{y}(6|5)$ and $\widehat{y}(7|5)$.

Using the previous predictors, we can compute $\widehat{y}(6|5)$ and $\widehat{y}(7|5)$. Obviously, for the computation of $\widehat{y}(6|5)$ we will make use of the 1-step predictor:

$$\hat{y}(t|t-1)=-\frac{1}{2}\cdot\hat{y}(t-1|t-2)+\frac{3}{4}\cdot y(t-1)$$

while for the computation of $\widehat{y}(7|5)$ we will adopt the 2-step predictor:

$$\hat{y}(t|t-2)=-\frac{1}{2}\cdot\hat{y}(t-1|t-3)+\frac{3}{16}\cdot y(t-2)$$

We will also use the given observations:

$$y(1)=1,\quad y(2)=\frac{1}{2},\quad y(3)=-\frac{1}{2},\quad y(4)=0,\quad y(5)=-\frac{1}{2}$$

- $\widehat{y}(6|5)$

$$
\begin{aligned}
&\widehat{y}(1|0) = \mathbb{E}[y(t)] = 0 \text{ (initialization)}\\
&\widehat{y}(2|1) = -\frac{1}{2}\cdot\hat{y}(1|0) + \frac{3}{4}\cdot y(1) = -\frac{1}{2}\cdot 0 + \frac{3}{4}\cdot 1 = \frac{3}{4}\\
&\widehat{y}(3|2) = -\frac{1}{2}\cdot\hat{y}(2|1) + \frac{3}{4}\cdot y(2) = -\frac{1}{2}\cdot\frac{3}{4} + \frac{3}{4}\cdot\frac{1}{2} = 0\\
&\widehat{y}(4|3) = -\frac{1}{2}\cdot\hat{y}(3|2) + \frac{3}{4}\cdot y(3) = -\frac{1}{2}\cdot 0 - \frac{3}{4}\cdot\frac{1}{2} = -\frac{3}{8}\\
&\widehat{y}(5|4) = -\frac{1}{2}\cdot\hat{y}(4|3) + \frac{3}{4}\cdot y(4) = +\frac{1}{2}\cdot\frac{3}{8} + \frac{3}{4}\cdot 0 = \frac{3}{16}\\
&\widehat{y}(6|5) = -\frac{1}{2}\cdot\hat{y}(5|4) + \frac{3}{4}\cdot y(5) = -\frac{1}{2}\cdot\frac{3}{16} - \frac{3}{4}\cdot\frac{1}{2} = -\frac{15}{32}
\end{aligned}
$$

The effect of the initialization rapidly vanishes.

- $\widehat{y}(7|5)$

$$
\begin{aligned}
&\widehat{y}(2|0) = \mathbb{E}[y(t)] = 0 \text{ (initialization)}\\
&\widehat{y}(3|1) = -\frac{1}{2}\cdot\hat{y}(2|0) + \frac{3}{16}\cdot y(1) = -\frac{1}{2}\cdot 0 + \frac{3}{16}\cdot 1 = \frac{3}{16}\\
&\widehat{y}(4|2) = -\frac{1}{2}\cdot\hat{y}(3|1) + \frac{3}{16}\cdot y(2) = -\frac{1}{2}\cdot\frac{3}{16} + \frac{3}{16}\cdot\frac{1}{2} = 0\\
&\widehat{y}(5|3) = -\frac{1}{2}\cdot\hat{y}(4|2) + \frac{3}{16}\cdot y(3) = -\frac{1}{2}\cdot 0 - \frac{3}{16}\cdot\frac{1}{2} = -\frac{3}{32}\\
&\widehat{y}(6|4) = -\frac{1}{2}\cdot\hat{y}(5|3) + \frac{3}{16}\cdot y(4) = +\frac{1}{2}\cdot\frac{3}{32} + \frac{3}{16}\cdot 0 = \frac{3}{64}\\
&\widehat{y}(7|5) = -\frac{1}{2}\cdot\hat{y}(6|4) + \frac{3}{16}\cdot y(5) = -\frac{1}{2}\cdot\frac{3}{64} - \frac{3}{16}\cdot\frac{1}{2} = -\frac{15}{128}
\end{aligned}
$$

The effect of the initialization rapidly vanishes.

Exercise 2

Consider the following process:

$$y(t) = \frac{1}{2} \cdot y(t-1) + e(t) - 2 \cdot e(t-1), \qquad e(t) \sim WN(0,2)$$

1. Is the process WSS?
2. Compute the predictors $\hat{y}(t|t-r)$ for $r = 1, 2$
3. Compute the 1-2 step prediction error variances.

1) Is the process WSS?

Let's put the process into its operatorial representation:

$$\left(1 - \frac{1}{2} \cdot z^{-1}\right) \cdot y(t) = \left(1 - 2 \cdot z^{-1}\right) \cdot e(t)$$

So:

$$\begin{aligned} \frac{y(t)}{e(t)} &= W(z) \\ &= \frac{1 - 2 \cdot z^{-1}}{1 - \frac{1}{2} \cdot z^{-1}} \\ &= \frac{z-2}{z-\frac{1}{2}} \end{aligned}$$

The pole ($z = \frac{1}{2}$) is inside the unit circle, so the digital filter is asymptotically stable. Moreover, $e(t)$ is a WSS process. We can conclude that $y(t)$ is a WSS process too.

In order to find the predictor, first of all, we have to check if $y(t)$ is a *canonical representation:*

1. Same degree: $\checkmark$;
2. Coprime: $\checkmark$;
3. Monic: $\checkmark$;
4. Roots inside the unitary circle: $\times$ the zeroes (roots of the numerator) are outside!

Observation: Notice that the pole is the reciprocal of the zero. The transfer function $W(z)$, except from the gain, has the structure of an all pass filter:

$$T(z) = a \cdot \frac{z + \frac{1}{a}}{z + a} = \left(-\frac{1}{2}\right) \cdot \frac{z-2}{z-\frac{1}{2}}$$

The process equation is then an all-pass filter with gain equal to -2:

$$\begin{aligned} y(t) &= W(z) \cdot e(t) \\ &= -2 \cdot \underbrace{\left[\left(-\frac{1}{2}\right) \cdot \frac{z-2}{z-\frac{1}{2}}\right]}_{\text{all-pass filter}} \cdot e(t) \\ &= -2 \cdot T(z) \cdot e(t) \end{aligned}$$

where $T(z)$ is an all-pass filter with gain 1.

Consider the white noise $\eta(t)$, derived from the original noise $e(t)$:

$$\eta(t) = -2 \cdot e(t) \implies e(t) = -\frac{1}{2} \cdot \eta(t)$$

Its mean and variance are given by:

$$m_\eta = \mathbb{E}[\eta(t)] = -2 \cdot \mathbb{E}[e(t)] = -2 \cdot m_e = 0$$

$$\lambda_\eta^2 = \mathbb{E}\left[(\eta(t) - m_\eta)^2\right] = \mathbb{E}\left[\eta(t)^2\right] = 4 \cdot \mathbb{E}\left[e(t)^2\right] = 4 \cdot \lambda_e^2 = 8$$

So:

$$\eta(t) \sim WN(0, 8)$$

If we replace $e(t)$ in the process function, we have that:

$$y(t) = \frac{1}{2} \cdot y(t-1) + e(t) - 2 \cdot e(t-1), \qquad e(t) \sim WN(0, 2)$$

$$\begin{aligned} y(t) &= \frac{1}{2} \cdot y(t-1) + \left(-\frac{1}{2} \cdot \eta(t)\right) - 2 \cdot \left(-\frac{1}{2} \cdot \eta(t-1)\right) \\ &= \frac{1}{2} \cdot y(t-1) - \frac{1}{2} \cdot \eta(t) + \eta(t-1), \qquad \eta(t) \sim WN(0, 8) \end{aligned}$$

The operatorial representation becomes:

$$\begin{aligned} \frac{y(t)}{\eta(t)} &= W_1(z) \\ &= -\frac{1}{2} \cdot \frac{1 - 2 \cdot z^{-1}}{1 - \frac{1}{2} \cdot z^{-1}} \\ &= -\frac{1}{2} \cdot \frac{z - 2}{z - \frac{1}{2}} \\ &= T(z) \end{aligned}$$

Observation: It is an all pass filter!

2) Compute the predictors $\hat{y}(t|t-r)$ for $r = 1, 2$.

Since the process $y(t)$ is the steady-state output of an all-pass filter fed by the white noise $\eta(t)$, we can conclude that $y(t)$ has the same spectrum of the white noise $\eta(t)$. But the white noise is totally unpredictable. So the optimal r-step predictor is the trivial predictor, that is the expected value of the process $y(t)$, which is the expected value of the noise $\eta(t)$:

$$\hat{y}(t|t-r) = \mathbb{E}[y(t)] = \mathbb{E}[\eta(t)] = 0, \qquad \forall r$$

3) Compute the 1-2 step prediction error variances.

Thus, the r-step prediction error variance is simply the variance of the process, i.e. the variance of the white noise $\eta(t)$:

$$\begin{aligned} \mathbb{E}\left[\varepsilon(t)^2\right] &= \mathbb{E}\left[(y(t) - \hat{y}(t|t-2))^2\right] \\ &= \mathbb{E}\left[y(t)^2\right] \\ &= \mathbb{E}\left[\eta(t)^2\right] = \lambda_\eta^2 = 8, \qquad \forall r \end{aligned}$$

Exercise 3

Consider the following WSS process ARMAX model (derived from an exercise from the previous lecture):

$$y(t) = \frac{z^{-2}}{1 - 1.3 \cdot z^{-1} + 0.4 \cdot z^{-2}} \cdot u(t) + \frac{1 - 0.55 \cdot z^{-1}}{1 - 1.3 \cdot z^{-1} + 0.4 \cdot z^{-2}} \cdot e(t), \qquad e(t) \sim WN(0, 16)$$

Thus:

$$\begin{cases} A(z) = 1 - 1.3 \cdot z^{-1} + 0.4 \cdot z^{-2} \\ B(z) = 1 \\ C(z) = 1 - 0.55 \cdot z^{-1} \\ k = 2 \end{cases}$$

$$y(t) = \frac{B(z)}{A(z)} \cdot u(t-k) + \frac{C(z)}{A(z)} \cdot e(t)$$

1. Compute the predictor $\hat{y}(t|t-r)$ for $r = 2$
2. Compute the associated prediction error variance.

1) Compute the predictor $\hat{y}(t|t-r)$ for $r = 2$.

In order to find the predictor, first of all, we have to check if $y(t)$ is a *canonical representation:*

1. Same degree: ✓;
2. Coprime: ✓;
3. Monic: ✓;
4. Roots inside the unitary circle: ✓;

We conclude saying that the process is in its canonical representation. The details of the canonical representation can be looked up in the previous lecture.

Let's compute 2 steps of the polynomial long division between $C(z)$ and $A(z)$:

$$\begin{array}{cccc|ccc}
1 & -0.55 \cdot z^{-1} & & & 1 & -1.3 \cdot z^{-1} & +0.4 \cdot z^{-2} \\ \cline{5-7}
-1 & +1.3 \cdot z^{-1} & -0.4 \cdot z^{-2} & & 1 & +0.75 \cdot z^{-1} & \\ \hline
/ & +0.75 \cdot z^{-1} & -0.4 \cdot z^{-2} & & & & \\
 & -0.75 \cdot z^{-1} & +0.975 \cdot z^{-2} & -0.3 \cdot z^{-3} & & & \\ \cline{1-4}
 & / & +0.575 \cdot z^{-2} & -0.3 \cdot z^{-3} & & &
\end{array}$$

where

$$\begin{cases} A(z) = 1 - 1.3 \cdot z^{-1} + 0.4 \cdot z^{-2} \\ B(z) = 1 \\ C(z) = 1 - 0.55 \cdot z^{-1} \\ k = 2 \\ Q_2(z) = 1 + 0.75 \cdot z^{-1} \\ R_2(z) = 0.575 \cdot z^{-2} - 0.3 \cdot z^{-3} = z^{-2} \cdot \left(0.575 - 0.3 \cdot z^{-1}\right) \end{cases}$$

Notice that:

$$R_2(z) = z^{-r} \cdot \tilde{R}_2(z)$$

Since, in this case, $r = 2$, we have:

$$R_2(z) = z^{-2} \cdot \tilde{R}_2(z) \implies \tilde{R}_2(z) = 0.575 - 0.3 \cdot z^{-1}$$

Whitening filter from an ARMAX model:

$$y(t) = \frac{B(z)}{A(z)} \cdot u(t-k) + \frac{C(z)}{A(z)} \cdot e(t)$$
$$A(z) \cdot y(t) = B(z) \cdot u(t-k) + C(z) \cdot e(t)$$
$$C(z) \cdot e(t) = A(z) \cdot y(t) - B(z) \cdot u(t-k)$$

The whitening filter is then:

$$e(t) = \frac{A(z) \cdot y(t) - B(z) \cdot u(t-k)}{C(z)}$$
$$e(t) = \frac{A(z)}{C(z)} \cdot y(t) - \frac{B(z)}{C(z)} \cdot u(t-k)$$

Optimal predictor from an ARMAX model:
The predictor frorm the noise is:

$$
\begin{aligned}
y(t) &= \frac{B(z)}{A(z)} \cdot u(t-k) + \frac{C(z)}{A(z)} \cdot e(t) \\
y(t) &= \frac{B(z)}{A(z)} \cdot u(t-k) + \underbrace{\left(\frac{R_r(z)}{A(z)} + Q_r(z)\right)}_{\text{r-steps of long division}} \cdot e(t) \\
y(t) &= \underbrace{\frac{B(z)}{A(z)} \cdot u(t-k) + \frac{R_r(z)}{A(z)} \cdot e(t)}_{\text{r-steps predictor}} + \underbrace{Q_r(z) \cdot e(t)}_{\text{r-steps prediction error}}
\end{aligned}
$$

The predictor from the noise is:

$$
\hat{y}(t|t-k) = \frac{B(z)}{A(z)} \cdot u(t-k) + \frac{R_r(z)}{A(z)} \cdot e(t)
$$

The prediction error is:

$$
\varepsilon(t) = Q_r(z) \cdot e(t)
$$

The predictor from the available data, using the whitening filter, is:

$$
\begin{aligned}
\hat{y}(t|t-k) &= \frac{B(z)}{A(z)} \cdot u(t-k) + \frac{R_r(z)}{A(z)} \cdot e(t) \\
\hat{y}(t|t-k) &= \frac{B(z)}{A(z)} \cdot u(t-k) + \frac{R_r(z)}{A(z)} \cdot \underbrace{\left(\frac{A(z)}{C(z)} \cdot y(t) - \frac{B(z)}{C(z)} \cdot u(t-k)\right)}_{\text{whitening filter}} \\
&= \frac{B(z)}{A(z)} \cdot u(t-k) + \frac{R_r(z)}{C(z)} \cdot y(t) - \frac{R_r(z)}{A(z)} \cdot \frac{B(z)}{C(z)} \cdot u(t-k) \\
&= \frac{B(z) \cdot (C(z) - R_r(z))}{A(z) \cdot C(z)} \cdot u(t-k) + \frac{R_r(z)}{C(z)} \cdot y(t) \\
&= \frac{B(z) \cdot Q_r(z) \cdot A(z)}{A(z) \cdot C(z)} \cdot u(t-k) + \frac{R_r(z)}{C(z)} \cdot y(t) \\
&= \frac{B(z) \cdot Q_r(z)}{C(z)} \cdot u(t-k) + \frac{R_r(z)}{C(z)} \cdot y(t) \\
&= \frac{B(z) \cdot Q_r(z)}{C(z)} \cdot u(t-k) + \frac{\tilde{R}_r(z)}{C(z)} \cdot y(t-r), \qquad \forall k \geq r
\end{aligned}
$$

Using the result from the long division:

$$
\begin{aligned}
\frac{C(z)}{A(z)} &= \frac{R_r(z)}{A(z)} + Q_r(z) \\
C(z) &= R_r(z) + Q_r(z) \cdot A(z) \\
C(z) - R_r(z) &= Q_r(z) \cdot A(z)
\end{aligned}
$$

The 2-step optimal predictor is then given by:

$$
\begin{aligned}
\hat{y}(t|t-2) &= \frac{B(z)\cdot Q_2(z)}{C(z)}\cdot u(t-2) + \frac{\tilde{R}_2(z)}{C(z)}\cdot y(t-2) \\
&= \frac{(1)\cdot\left(1+0.75\cdot z^{-1}\right)}{\left(1-0.55\cdot z^{-1}\right)}\cdot u(t-2) + \frac{\left(0.575-0.3\cdot z^{-1}\right)}{\left(1-0.55\cdot z^{-1}\right)}\cdot y(t-2) \\
&= \frac{1+0.75\cdot z^{-1}}{1-0.55\cdot z^{-1}}\cdot u(t-2) + \frac{0.575-0.3\cdot z^{-1}}{1-0.55\cdot z^{-1}}\cdot y(t-2)
\end{aligned}
$$

with the recursive time-domain representation:

$$
\begin{aligned}
\hat{y}(t|t-2) &= \frac{1+0.75\cdot z^{-1}}{1-0.55\cdot z^{-1}}\cdot u(t-2) + \frac{0.575-0.3\cdot z^{-1}}{1-0.55\cdot z^{-1}}\cdot y(t-2) \\
\left(1-0.55\cdot z^{-1}\right)\cdot\hat{y}(t|t-2) &= \left(0.575-0.3\cdot z^{-1}\right)\cdot y(t-2) + \left(1+0.75\cdot z^{-1}\right)\cdot u(t-2) \\
\hat{y}(t|t-2) - 0.55\cdot\hat{y}(t-1|t-3) &= 0.575\cdot y(t-2) - 0.3\cdot y(t-3) + u(t-2) + 0.75\cdot u(t-3) \\
\hat{y}(t|t-2) &= 0.55\cdot\hat{y}(t-1|t-3) + 0.575\cdot y(t-2) - 0.3\cdot y(t-3) + u(t-2) + 0.75\cdot u(t-3)
\end{aligned}
$$

2) Compute the associated prediction error variance.

The 2-step prediction error variance can be simply computed as the variance of the unpredictable part of the process, which is:

$$
\begin{aligned}
\mathbb{E}\left[\varepsilon(t)^2\right] &= \mathbb{E}\left[\left(y(t)-\hat{y}(t|t-2)\right)^2\right] \\
&= \mathbb{E}\left[\left(Q_2(z)\cdot e(t)\right)^2\right] = \mathbb{E}\left[\left(e(t)+0.75\cdot e(t-1)\right)^2\right] \\
&= \mathbb{E}\left[e(t)^2\right] + 0.75^2\cdot\mathbb{E}\left[e(t-1)^2\right] + 1.5\cdot\underbrace{\cancel{\mathbb{E}\left[e(t)\cdot e(t-1)\right]}}_{e\sim WN} \\
&= \lambda_e^2 + 0.75^2\cdot\lambda_e^2 = 16 + 0.75^2\cdot 16 = 25
\end{aligned}
$$