Chapter 6 - Forecasting with Box-Jenkins models

AMA4650 Forecasting and Applied Time Series Analysis



6.1 Forecast of ARMA(p, q)

- Recall a (*well-established*) ARMA model in general expressed in the form

$$\Phi(B)X_t = \Theta(B)a_t, \quad a \sim WN(0, \sigma_a^2).$$

- By causality,

$$X_t = \Phi(B)^{-1}\Theta(B)a_t,$$

i.e. X_t always has a white noise representation in general

$$X_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots,$$

where ψ_i can be determined by $\Phi(B)^{-1}\Theta(B)$; see tutorials for Chapter 4.

- Let $\hat{X}_i(\ell)$ denote the forecast at i for ℓ period ahead.[1]
- To determine $\hat{X}_i(\ell)$, assume

$$\hat{X}_i(\ell) = \psi_\ell^* a_i + \psi_{\ell+1}^* a_{i-1} + \psi_{\ell+2}^* a_{i-2} + \cdots.$$

can only use data a_t, $t \le i$

The ψ^* are to be determined by *least square:* $\min \mathbb{E}[(X_{i+\ell} - \hat{X}_i(\ell))^2]$, where the observation at $i+\ell$ is

$$X_{i+\ell} = a_{i+\ell} + \psi_1 a_{i+\ell-1} + \psi_2 a_{i+\ell-2} + \cdots + \psi_{\ell-1} a_{i+1} + (\psi_\ell a_i + \psi_{\ell+1} a_{i-1} + \ldots)$$

$\phi_a a_b \Rightarrow a+b = i+\ell$

单独拎出来的比

[1] Note that the forecast made at time i can only depend on the information up to time i: $\{a_j,\ j \le i\}$.

Introduction & general theory, *cont'd*

- Taking the expression of $X_{i+\ell}$ and $\hat{X}_i(\ell)$ into the mean square error:

$$\begin{aligned}
&\mathbb{E}\Big[\big(X_{i+\ell} - \hat{X}_i(\ell)\big)^2\Big] \\
&= \mathbb{E}\Big[\big(a_{i+\ell} + \psi_1 a_{i+\ell-1} + \psi_2 a_{i+\ell-2} + \cdots + \psi_{\ell-1} a_{i+1} \\
&\quad + (\psi_\ell - \psi_\ell^*) a_i + (\psi_{\ell+1} - \psi_{\ell+1}) a_{i-1} + \ldots\big)^2\Big] \\
&= (1 + \psi_1^2 + \cdots + \psi_{\ell-1}^2)\sigma_a^2 + \sum_{j=\ell}^{\infty} (\psi_j - \psi_j^*)^2 \sigma_a^2.
\end{aligned}$$

- Thus the mean square forecast error is at minimum when

$$\psi_j^* = \psi_j, \quad \text{for } j = \ell, \ell+1, \ldots.$$

- So the **forecast made at time i for ℓ period ahead** is

$$\hat{X}_i(\ell) = \psi_\ell a_i + \psi_{\ell+1} a_{i-1} + \ldots. \tag{1}$$

- The **forecast error** at i for lead ℓ is

$$\begin{aligned}
e_i(\ell) &= X_{i+\ell} - \hat{X}_i(\ell) \\
&= a_{i+\ell} + \psi_1 a_{i+\ell-1} + \cdots + \psi_{\ell-1} a_{i+1}.
\end{aligned} \tag{2}$$

Forecast by conditional expectation

For any r.v. ξ, let $\mathbb{E}_i[\xi] = \mathbb{E}[\xi | a_i, a_{i-1}, a_{i-2} \cdots]$ be the conditional expectation conditioned on the information up to i.

Then clearly

$$\mathbb{E}_i[a_{i+\ell}] = \begin{cases} 0, & \text{if } \ell > 0; \\ a_{i+\ell}, & \text{if } \ell \le 0. \end{cases}$$

Recall
$$\hat{X}_i(\ell) = \psi_\ell a_i + \psi_{\ell+1} a_{i-1} + \cdots \quad (1)$$

Equation (1) gives

$$\mathbb{E}_i[\hat{X}_i(\ell)] = \hat{X}_i(\ell), \quad \ell > 0.$$

Recall
$$e_i(\ell) = X_{i+\ell} - \hat{X}_i(\ell) = a_{i+\ell} + \psi_1 a_{i+\ell-1} + \cdots + \psi_{\ell-1} a_{i+1} \quad (2)$$

Equation (2) gives

$$\mathbb{E}_i[e_i(\ell)] = 0, \quad \ell > 0,^2 \quad (3)$$

$$\mathbb{E}_i[X_{i+\ell}] = \hat{X}_i(\ell), \quad \ell > 0.$$

Clearly $\mathbb{E}_i[X_{i+\ell}] = X_{i+\ell}, \quad \ell \le 0.$

[2] Thus, the forecast is *unbiased*.

Confidence intervals of forecasts

Recall again the forecast error

$$e_i(\ell) = a_{i+\ell} + \psi_1 a_{i+\ell-1} + \cdots + \psi_{\ell-1} a_{i+1} \tag{2}$$

Equation (2) shows that the variance of the forecast error at lead ℓ is

$\mathrm{Var}(a+b) = \mathrm{Var}(a) + \mathrm{Var}(b)$, $a \perp b$

$$\mathrm{Var}\big(e_i(\ell)\big) = (1 + \psi_1^2 + \cdots + \psi_{\ell-1}^2)\sigma_a^2.$$

Hence $(1-\alpha)\times 100\%$ confidence intervals for the forecasts[3] are obtained as

$$\hat{X}_i(\ell) \pm k_\alpha \left[\Big(\sum_{j=0}^{\ell-1} \psi_j^2 \Big) \sigma_a^2 \right]^{1/2}, \tag{4}$$

where k_α is the $(1-\alpha/2)\times 100\%$ quantile of $\mathcal{N}(0,1)$ and $\psi_0 = 1$. When the data is available, we can replace ψ and σ_a^2 by estimators.

[3] Also called *forecast interval, or interval forecast.*

Forecasting procedure

The difference equation form of an ARMA(p, q) process is

$$\begin{aligned} X_{i+\ell} &= \phi_1 X_{i+\ell-1} + \cdots + \phi_p X_{i+\ell-p} \\ &\quad + a_{i+\ell} + \theta_1 a_{i+\ell-1} + \cdots + \theta_q a_{i+\ell-q}. \end{aligned}$$

So, taking conditional expectation at time i, for $\ell > 0$, the forecast made at i for ℓ period ahead can be written as follows.

For $1 \le \ell \le \min\{p, q\}$, the forecast is

$$\begin{aligned} \hat{X}_i(\ell) = \mathbb{E}_i[X_{i+\ell}] &= \left(\phi_1 \mathbb{E}_i[X_{i+\ell-1}] + \cdots + \phi_{\ell-1}\mathbb{E}_i[X_{i+1}]\right) \\ &\quad + \left(\phi_\ell \mathbb{E}_i[X_i] + \cdots + \phi_p \mathbb{E}_i[X_{i+\ell-p}]\right) \\ &\quad + \left(\theta_1 \mathbb{E}_i[a_{i+\ell-1}] + \theta_2 \mathbb{E}_i[a_{i+\ell-2}] + \cdots + \theta_{\ell-1}\mathbb{E}_i[a_{i+1}]\right) \\ &\quad + \left(\theta_\ell \mathbb{E}_i[a_i] + \cdots + \theta_q \mathbb{E}_i[a_{i+\ell-q}]\right) \\ &= \phi_1 \hat{X}_i(\ell-1) + \cdots + \phi_{\ell-1}\hat{X}_i(1) \\ &\quad + \phi_\ell X_i + \cdots + \phi_p X_{i+\ell-p} \\ &\quad + \theta_\ell a_i + \cdots + \theta_q a_{i+\ell-q}. \end{aligned} \tag{5}$$

Forecasting procedure, *cont'd*

For $1 \leq \ell \leq q$ but $\ell > p$, the forcast made at time i for ℓ period ahead is

$$\begin{aligned}
\hat{X}_i(\ell) &= \mathbb{E}_i[X_{i+\ell}] \\
&= \phi_1 \mathbb{E}_i[X_{i+\ell-1}] + \cdots + \phi_p \mathbb{E}_i[X_{i+\ell-p}] \\
&\quad + \Big(\theta_1 \mathbb{E}_i[a_{i+\ell-1}] + \theta_2 \mathbb{E}_i[a_{i+\ell-2}] + \cdots + \theta_{\ell-1} \mathbb{E}_i[a_{i+1}]\Big) \\
&\quad + \Big(\theta_\ell \mathbb{E}_i[a_i] + \cdots + \theta_q \mathbb{E}_i[a_{i+\ell-q}]\Big) \\
&= \phi_1 \hat{X}_i(\ell-1) + \cdots + \phi_p \hat{X}_i(\ell-p) \\
&\quad + \theta_\ell \mathbb{E}_i[a_i] + \cdots + \theta_q \mathbb{E}_i[a_{i+\ell-q}].
\end{aligned} \tag{6}$$

→ all X_{i+k} is after i

For $1 \leq \ell \leq p$ but $\ell > q$, the forcast made at time i for ℓ period ahead is

$$\begin{aligned}
\hat{X}_i(\ell) &= \mathbb{E}_i[X_{i+\ell}] \\
&= \Big(\phi_1 \mathbb{E}_i[X_{i+\ell-1}] + \cdots + \phi_{\ell-1} \mathbb{E}_i[X_{i+1}]\Big) \\
&\quad + \Big(\phi_\ell \mathbb{E}_i[X_i] + \cdots + \phi_p \mathbb{E}_i[X_{i+\ell-p}]\Big) \\
&= \phi_1 \hat{X}_i(\ell-1) + \cdots + \phi_{\ell-1} \hat{X}_i(1) \\
&\quad + \phi_\ell X_i + \cdots + \phi_p X_{i+\ell-p}.
\end{aligned} \tag{7}$$

Example 6.1



$X_t(1)$ 和 $X_{t+1}(2)$ 一样吗

Assume a dataset with 121 observations fits an An AR(2) model i.e.,

$$X_t - \mu = \phi_1(X_{t-1} - \mu) + \phi_2(X_{t-2} - \mu) + a_t,$$

which is equivalent to

$$\begin{aligned} X_t &= \mu(1 - \phi_1 - \phi_2) + \phi_1 X_{t-1} + \phi_2 X_{t-2} + a_t \\ &= \theta_0 + \phi_1 X_{t-1} + \phi_2 X_{t-2} + a_t, \end{aligned}$$

where μ is the mean, and $\theta_0 = \mu(1 - \phi_1 - \phi_2)$ is the constant.

Assume we have estimated $\hat{\phi}_1 = 1.561, \hat{\phi}_2 = -0.656, \hat{\theta}_0 = 0.486, \hat{\mu} = 5.082$. Moreover, in the observations, $X_{121} = 6.20$ and $X_{120} = 6.63$

Thus at $t = 121$,

$$\hat{X}_t(1) = \hat{\theta}_0 + \hat{\phi}_1 X_t + \hat{\phi}_2 X_{t-1} = 5.80$$

and

$$\hat{X}_t(2) = \hat{\theta}_0 + \hat{\phi}_1 \hat{X}_t(1) + \hat{\phi}_2 X_t = 5.45.$$

$\hat{X}_t(l) = E(\phi_1 X_{t+l-1} + \phi_2 X_{t+l-2} \cdots)$
$+ E(a_{t+l} + a_{t+l-1} + a_{t+2} \cdots)$

Example 6.1, *cont'd*

Recall the confidence interval of the forecast is

$$\hat{X}_i(\ell) \pm k_\alpha \left[\left(\sum_{j=0}^{\ell-1} \psi_j^2 \right) \sigma_a^2 \right]^{1/2},$$

and assume we have obtained $\hat{\sigma}_a^2 = 0.1308$, and [4]

$$\begin{aligned}
&\hat{\psi}_1 = \hat{\phi}_1 = 1.561, \\
&\hat{\psi}_2 = \hat{\phi}_1 \hat{\psi}_1 + \hat{\phi}_2 = 1.781, \\
&\hat{\psi}_3 = \hat{\phi}_1 \hat{\psi}_2 + \hat{\phi}_2 \hat{\psi}_1 = 1.756. \\
&\cdots
\end{aligned}$$

We want to calculate the forecast interval of $\hat{X}_{121}(2)$: $\ell = 2$. So, the 95% confidence interval for $\hat{X}_{121}(2)$ is

$$5.45 \pm 1.96\sqrt{0.1308}(1 + 1.561^2)^{1/2} = \cdots$$

[4] Recall Chapter 4, Page 11; this is one reason why we are interested in the white noise representation $X_t = \sum_{j=0}^{\infty} \psi_j a_{t-j}$.

Example 6.2

Consider a causal ARMA(1,1) process

$$(1-\phi B)X_t = (1-\theta B)a_t$$

$$\Rightarrow \quad X_{t+\ell} = \phi X_{t+\ell-1} + a_{t+\ell} - \theta a_{t+\ell-1}.$$

Forecasts are

$$\hat{X}_t(1) = \mathbb{E}_t[X_{t+1}] = \mathbb{E}[X_{t+1}|a_t, a_{t-1}, \cdots]$$
$$= \mathbb{E}[\phi X_t + a_{t+1} - \theta a_t | a_t, a_{t-1}, \cdots] = \phi X_t - \theta a_t.$$

unknown
① if invertable
$a_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}$
② $a_0 = X_0 - \hat{X}_{-1}(1)$

Moreover, for $\ell \geq 2$,

$$\hat{X}_t(\ell) = \mathbb{E}_t[X_{t+\ell}] = \mathbb{E}[X_{t+\ell}|a_t, a_{t-1}, \cdots]$$
$$= \mathbb{E}\left[X_{t+\ell} = \phi X_{t+\ell-1} + a_{t+\ell} - \theta a_{t+\ell-1} | a_t, a_{t-1}, \cdots\right]$$
$$= \phi \hat{X}_t(\ell-1).$$

$\hat{X}_t(\ell+h-1) = \phi^h \hat{X}_t(\ell-1)$

The forecasts decay geometrically.

Updating ARMA forecasts

- Let the observations be $\{X_1, \cdots, X_N\}$. Based on the observations and equation (5)-(7), the forecast $\hat{X}_i(\ell)$, $\ell = 1, 2, \ldots, L$, $i = 1, \cdots, N$ can be obtained recursively.
- When the next observation X_{N+1} becomes available, we can calculate the forecast $\hat{X}_{N+1}(\ell)$ by update from $\hat{X}_N(\ell)$, without dependence on the forecast made at other times.
- Recall WN representation $X_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots$. Let $t = N + 1$:

must be in ∞ representation

$$X_{N+1} = a_{N+1} + \psi_1 a_N + \psi_2 a_{N-1} + \cdots$$
$$\hat{X}_N(1) = \mathbb{E}_N[X_{N+1}] = \psi_1 a_N + \psi_2 a_{N-1} + \cdots$$

now, only $a_{N+1}, a_{N-1} \cdots$ is known

The difference is

$$a_{N+1} = X_{N+1} - \hat{X}_N(1).$$

- Recall the forecast formula Equation (1): $\hat{X}_i(\ell) = \psi_\ell a_i + \psi_{\ell+1} a_{i-1} + \cdots$

$$\hat{X}_{N+1}(\ell) = \psi_\ell a_{N+1} + \psi_{\ell+1} a_N + \psi_{\ell+2} a_{N-1} + \cdots$$
$$\hat{X}_N(\ell+1) = \psi_{\ell+1} a_N + \psi_{\ell+2} a_{N-1} + \cdots,$$

The difference is $\hat{X}_{N+1}(\ell)$ & $\hat{X}_N(\ell+1)$ please refer to P5

$$\hat{X}_{N+1}(\ell) = \hat{X}_N(\ell+1) + \psi_\ell a_{N+1}. \tag{8}$$

Updating ARMA forecasts, *cont'd*

			Lead ℓ					
i	X_i	a_i	1	2	$\cdots$	$L-1$	L	
N	X_N		$\hat{X}_N(1)$	$\hat{X}_N(2)$		$\hat{X}_N(L-1)$	$\hat{X}_N(L)$	$(*)$
			$+\psi_1 a_{N+1}$	$\cdots$	$\cdots$		$+\psi_{L-1} a_{N+1}$	
$N+1$	X_{N+1}	a_{N+1}	$\hat{X}_{N+1}(1)$	$\hat{X}_{N+1}(2)$		$\hat{X}_{N+1}(L-1)$	$\hat{X}_{N+1}(L)$	$(\dagger)$

from (5)-(7)

Table 6.1

Summary of updating ARMA forecasts.

- Based on observations $\{X_1, X_2, \cdots, X_N\}$, calculate forecast $\hat{X}_i(\ell)$, $i = 1, \cdots, N$, $\ell = 1, \cdots, L$, by Equations (5)-(7).
- If the next observations X_{N+1} is available, then the forecast $\hat{X}_{N+1}(\ell)$, $\ell = 1, \cdots, L-1$ can be obtained by (8). 只用于 update 未来同一时间点预测
- The last update $\hat{X}_{N+1}(L)$ has to be calculated from (5)-(7).

6.2 Forecast of ARIMA(1,1,1)

Forecast of ARIMA(1,1,1), *cont'd*

Consider forecasting an ARIMA(1,1,1) process

$$(1-\phi B)(1-B)X_t = (1-\theta B)a_t$$

$$\Rightarrow \quad (1-\phi B)(1-B)X_{t+\ell} = (1-\theta B)a_{t+\ell}.$$

We will consider two of the most commonly used approaches:

1. As discussed earlier, the first approach is based on the white noise representation:

$$(1-\phi B)(1-B)X_{t+\ell} = (1-\theta B)a_{t+\ell}$$

$$\Rightarrow \quad X_{t+\ell} = \sum_{i=0}^{\infty} \psi_i B^i a_{t+\ell} = \sum_{i=0}^{\infty} \psi_i a_{t+\ell-i}$$

$$= a_{t+\ell} + \psi_1 a_{t+\ell-1} + \psi_2 a_{t+\ell-2} + \ldots,$$

where the weights ψ_i can be calculated from

$$(\psi_0 + \psi_1 B + \ldots)(1-\phi B)(1-B) = (1-\theta B).$$

Forecast of ARIMA(1,1,1), *cont'd*

Let the ℓ-step ahead forecast made at time t be

$$\hat{X}_t(\ell) = \psi_\ell^* a_t + \psi_{\ell+1}^* a_{t-1} + \dots,$$

where ψ_j^* is to be determined by least square:

$$\min_{\psi_\ell^*, \psi_{\ell+1}^*, \cdots} \mathbb{E}[(X_{t+\ell} - \hat{X}_t(\ell))^2].$$

Note that

$$\begin{aligned}\mathbb{E}[(X_{t+\ell} - \hat{X}_t(\ell))^2] &= \mathbb{E}\Big[\big(a_{t+\ell} + \psi_1 a_{t+\ell-1} + \psi_2 a_{t+\ell-2} + \cdots + \psi_{\ell-1} a_{t+1} \\ &\qquad + (\psi_\ell - \psi_\ell^*) a_t + (\psi_{\ell+1} - \psi_{\ell+1}) a_{t-1} + \dots\big)^2\Big] \\ &= (1 + \psi_1^2 + \cdots + \psi_{\ell-1}^2)\sigma_a^2 + \sum_{j=\ell}^{\infty} (\psi_j - \psi_j^*)^2 \sigma_a^2.\end{aligned}$$

Thus, $\psi_j^* = \psi_j$ and the ℓ-step ahead forecast made at time t follows

$$\hat{X}_t(\ell) = \psi_\ell a_t + \psi_{\ell+1} a_{t-1} + \dots$$

Unobservable a_t be determined by $a_t = X_t - \hat{X}_{t-1}(1) := e_{t-1}(1)$: $\curvearrowright$

Forecast of ARIMA(1,1,1), *cont'd*

Recall

$$X_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots$$

$$\hat{X}_{t-1}(1) = \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots$$

Then $a_t = X_t - \hat{X}_{t-1}(1)$.

The forecast error is

$$X_{t+\ell} - \hat{X}_t(\ell) = a_{t+\ell} + \psi_1 a_{t+\ell-1} + \psi_2 a_{t+\ell-2} + \cdots + \psi_{\ell-1} a_{t+1},$$

which implies

$$\mathbb{E}_t[X_{t+\ell}] = \hat{X}_t(\ell), \qquad \ell \geq 1.$$

2. (*Very simple and most useful*!) The second approach is to use difference equations as given by

$$X_{t+\ell} = (1+\phi)X_{t+\ell-1} - \phi X_{t+\ell-2} + a_{t+\ell} - \theta a_{t+\ell-1}.$$

For $\ell = 1$, the one step ahead forecast made at time t is

$$\hat{X}_t(1) = \mathbb{E}_t(X_{t+1}) = (1+\phi)X_t - \phi X_{t-1} - \theta e_{t-1}(1).$$

We can further show that for lead times $\ell \geq 2$, the forecast is

$$\hat{X}_t(\ell) = (1+\phi)\hat{X}_t(\ell-1) - \phi \hat{X}_t(\ell-2),$$

which is a *linear difference equation*; refer to Chapter 4.

Example 6.3

Consider an ARIMA(1,1,0) process

$$(1-0.8B)(1-B)X_t = a_t,$$

i.e.

$$(1-1.8B+0.8B^2)X_t = a_t$$

$$\Rightarrow \quad X_{t+\ell} = 1.8X_{t+\ell-1} - 0.8X_{t+\ell-2} + a_{t+\ell}.$$

Thus, the forecasts at time t are given by

$$\hat{X}_t(1) = \mathbb{E}_t[X_{t+1}] = \mathbb{E}_t[1.8X_t - 0.8X_{t-1} + a_{t+1}] = 1.8X_t - 0.8X_{t-1},$$

$$\hat{X}_t(2) = \mathbb{E}_t[X_{t+2}] = \mathbb{E}_t[1.8X_{t+1} - 0.8X_t + a_{t+2}] = 1.8\hat{X}_t(1) - 0.8X_t,$$

$$\hat{X}_t(\ell) = \mathbb{E}_t[X_{t+\ell}] = 1.8\hat{X}_t(\ell-1) - 0.8\hat{X}_t(\ell-2), \quad \text{for } \ell \geq 3.$$

If observations are given, all forecasts can be calculated.

Example 6.4

Consider an ARIMA(0,1,1) process

$$X_t - X_{t-1} = a_t + \theta a_{t-1}.$$

Since

$$X_{t+\ell} = X_{t+\ell-1} + a_{t+\ell} + \theta a_{t+\ell-1},$$

we have

$$\hat{X}_t(1) = \mathbb{E}_t[X_{t+1}] = X_t + \theta a_t,$$
$$\hat{X}_t(\ell) = \mathbb{E}_t[X_{t+\ell}] = \hat{X}_t(\ell-1), \quad \text{for } \ell \geq 2.$$

It remains to determine a_t: either by

(1) $a_t = X_t - \hat{X}_{t-1}(1)$

or by infinite sum representation if X is invertible

(2) $$a_t = \frac{1}{1+\theta B}(X_t - X_{t-1}) = \sum_{n=0}^{\infty}(-\theta)^n B^n (X_t - X_{t-1})$$
$$= \sum_{n=0}^{\infty}(-\theta)^n X_{t-n} - \sum_{n=0}^{\infty}(-\theta)^n X_{t-n-1} = \sum_{n=0}^{\infty} b_n X_{t-n},$$

where $b_0 = 1$ and $b_n = (-\theta)^n - (-\theta)^{n-1}$, $n \geq 2$.

Example 6.5

Consider an ARIMA(0,1,1) process with constant (drift):

$$X_{t+\ell} = \mu + X_{t+\ell-1} + a_{t+\ell} + \theta a_{t+\ell-1}$$

$$\Rightarrow \quad \begin{cases} \hat{X}_t(1) = \mu + X_t + \theta a_t, \\ \hat{X}_t(\ell) = \mu + \hat{X}_t(\ell-1) \\ \qquad\qquad = \hat{X}_t(1) + (\ell-1)\mu, \quad \text{for } \ell \geq 2. \end{cases}$$

Example 6.6

A time series of 100 observations satisfies the following model

$$X_t = 20.62 + 0.777X_{t-1} - 0.492X_{t-2} + a_t.$$

Its last 5 observations are shown below:

t	96	97	98	99	100
X_t	31.68	29.10	23.15	26.74	32.44

Forecast its 101th observation at $t = 100 \Rightarrow$ one step ahead forecast made at time $t = 100$.

Solution.

$$\begin{aligned} & X_{101} = 20.62 + 0.777X_{100} - 0.492X_{99} + a_{101} \\ \Rightarrow \quad & \hat{X}_{100}(1) = \mathbb{E}_{100}[X_{101}] = 20.62 + 0.777X_{100} - 0.492X_{99} \\ & = 20.62 + 0.777(32.44) - 0.492(26.74) \\ & = 32.67. \end{aligned}$$

Example 6.7

For the model $(1-0.2B)(1-B)X_t = (1-0.5B)a_t$, find its forecast for one and two steps ahead at t, if $a_t = 1, X_t = 4, X_{t-1} = 3$.

Solution.

$$
\begin{aligned}
& (1-0.2B)(1-B)X_t = (1-0.5B)a_t \\
\Rightarrow\ & (1-1.2B+0.2B^2)X_t = (1-0.5B)a_t \\
\Rightarrow\ & X_t - 1.2X_{t-1} + 0.2X_{t-2} = a_t - 0.5a_{t-1} \\
\Rightarrow\ & X_t = 1.2X_{t-1} - 0.2X_{t-2} + a_t - 0.5a_{t-1}.
\end{aligned}
$$

For one-step ahead forecast at time t, we have

$$
\begin{aligned}
& X_{t+1} = 1.2X_t - 0.2X_{t-1} + a_{t+1} - 0.5a_t \\
\Rightarrow\quad & \hat{X}_t(1) = \mathbb{E}_t[X_{t+1}] = 1.2X_t - 0.2X_{t-1} - 0.5a_t \\
& \qquad = 1.2(4) - 0.2(3) - 0.5(1) \\
& \qquad = 3.7.
\end{aligned}
$$

Example 6.7, *cont'd*

For two-step ahead forecast at time t, we have

$$
\begin{aligned}
& X_{t+2} = 1.2X_{t+1} - 0.2X_t + a_{t+2} - 0.5a_{t+1} \\
\Rightarrow \quad & \hat{X}_t(2) = \mathbb{E}_t[X_{t+2}] = 1.2\hat{X}_t(1) - 0.2X_t \\
& \qquad = 1.2(3.7) - 0.2(4) \\
& \qquad = 3.64.
\end{aligned}
$$