



7 Forecasting

We briefly summarize what we have done in the previous chapters

1. (Tutorial 1-3) Given the model of $\{Y_t\}$, study its properties, eg. ACVF $\gamma(\cdot)$ and PACF ϕ_{kk} .
2. (Tutorial 5-6) Given the observed data Y_1, \dots, Y_n , select the appropriate order of the ARIMA(p, d, q) model for $\{Y_t\}$ and estimate the AR and MA parameters.

In this tutorial, we discuss how to predict Y_{n+h} ($h \geq 1$) given the observed data Y_1, \dots, Y_n .

Forecasting

Definition 1. The h -step forecasting function of a time series $\{Y_t\}$ given the observed data Y_1, \dots, Y_n is defined as

$$\hat{Y}_{n+h} = Y_{n+h}^n := E(Y_{n+h} | Y_1, \dots, Y_n).$$

and also define

- Forecast error: $e_n(h) := Y_{n+h} - \hat{Y}_{n+h}$.
- Variance of forecasting error: $P_{n+h}^n := \text{Var}(e_n(h) | Y_1, \dots, Y_n)$.

Remark 1. Suppose $\{Y_t\} \sim \text{ARIMA}(p, d, q)$. If we further assume the normality of the white noise sequence, then

$$\left(Y_{n+h}^n - 1.96\sqrt{P_{n+h}^n}, Y_{n+h}^n + 1.96\sqrt{P_{n+h}^n} \right)$$

is a 95% prediction interval (not confidence interval) for Y_{n+h} .

7.1 Box and Jenkins Approach

There are several forecasting methods. For example, the Box and Jenkins Approach and Holt-Winters Approach. In particular, we would study the Box and Jenkins Approach. Before going into detail, we study its assumption and implications.

Assumption of Box and Jenkins Approach

The Box and Jenkins Approach is applicable for giving prediction of the ARIMA(p, d, q) model. Recall that

$$\{Y_t\} \sim \text{ARIMA}(p, d, q) \Rightarrow \{X_t\} \sim \text{ARMA}(p, q),$$

where $X_t = \Delta^d Y_t$. We have to assume $\{X_t\}$ being **causal** and **invertible**.

1. **Invertibility:** Legitimate to assume $Z_k = 0$ for $k \leq p$. It implies that

$$Z_t = \left(\sum_{k=1}^q \theta_k Z_{t-k} \right) + \left(X_t - \sum_{k=1}^p \phi_k X_{t-k} \right)$$

is computable for $t = p+1, \dots, n$.

2. **Causality:** $E(Z_k | Y_1, \dots, Y_n) = 0$ for all $k > n$.

Box and Jenkins Approach

Assume $\{Y_t\} \sim \text{ARMA}(p, q)$, i.e. $Y_t - \sum_{k=1}^p \phi_k Y_{t-k} = Z_t - \sum_{k=1}^q \theta_k Z_{t-k}$ being causal and invertible. It follows that we can write the causal representation as $Y_t = \sum_{i=0}^{\infty} \psi_i Z_{t-i}$. Then we have

1. (Forecast)

$$\hat{Y}_{n+h} = E \left\{ \sum_{i=0}^{\infty} \psi_i Z_{n+h-i} \middle| Y_1, \dots, Y_n \right\} = \sum_{i=h}^{\infty} \psi_i Z_{n+h-i},$$

which is computable once we have the value of Z_{p+1}, \dots, Z_n .

2. (Forecast error with variance)

$$e_n(h) = \sum_{i=0}^{h-1} \psi_i Z_{n+h-i} \Rightarrow P_{n+h}^n = \text{Var} \left\{ \sum_{i=0}^{h-1} \psi_i Z_{n+h-i} \middle| Y_1, \dots, Y_n \right\} = \sigma^2 \sum_{i=0}^{h-1} \psi_i^2.$$

Assume $\{Y_t\} \sim \text{ARIMA}(p, 1, q)$, i.e. $X_t = \Delta Y_t \sim \text{ARMA}(p, q)$ with $\{X_t\}$ being causal and invertible.

- We can obtain $\hat{X}_{n+1}, \dots, \hat{X}_{n+h}$ from above.
- Define $e_{n,X}(h)$ and $e_{n,Y}(h)$ as the forecast error for $\{X_t\}$ and $\{Y_t\}$, respectively.
- $Y_{n+h} = X_{n+h} + Y_{n+h-1} = \dots = X_{n+h} + X_{n+h-1} + \dots + X_{n+1} + Y_n$.

1. (Forecast)

$$\hat{Y}_{n+h} = Y_n + \sum_{j=1}^h \hat{X}_{n+j}.$$

2. (Forecast error with variance)

$$e_{n,Y}(h) = Y_{n+h} - \hat{Y}_{n+h} = \sum_{j=1}^h (X_{n+j} - \hat{X}_{n+j}) = \sum_{j=1}^h e_{n,X}(j).$$

Remark 2. By induction, as for $\{Y_t\} \sim \text{ARIMA}(p, d, q)$, we have $\overbrace{(\Delta \cdots \Delta)}^{(p\text{-th order})}$

$$X_t = \Delta^p Y_t = \Delta^{p-1}(\Delta Y_t) = \Delta^{p-2}(\Delta(\Delta Y_t)) = \dots = \overbrace{(\Delta \cdots \Delta)}^{(p\text{-th order})} Y_t$$

and hence the method works for general $\text{ARIMA}(p, d, q)$ as long as $\{X_t\}$ is causal and invertible.

Operation of Box and Jenkins Approach

In short, the Box and Jenkins Approach tells us to

1. Set $Z_t = 0$ for $t \leq p$. Then solve for Z_{p+1}, \dots, Z_n sequentially by considering for $t = p+1, \dots, n$,

$$Z_t = \left(\sum_{k=1}^q \theta_k Z_{t-k} \right) + \left(Y_t - \sum_{k=1}^p \phi_k Y_{t-k} \right).$$


2. Noticing $E(Z_t | Y_1, \dots, Y_n) = Z_t \mathbb{1}(p+1 \leq t \leq n)$, calculate

$$\hat{Y}_{n+k} = \sum_{i=1}^p \phi_i E \left\{ Y_{n+k-i} \middle| Y_1, \dots, Y_n \right\} + E \left\{ Z_{n+k} - \sum_{i=1}^q \theta_i Z_{n+k-i} \middle| Y_1, \dots, Y_n \right\},$$

i.e. drop away all Z_t except Z_{p+1}, \dots, Z_n in evaluation of Y_{n+k}^n .

3. Express $e_n(h) = Y_{n+h} - \hat{Y}_{n+h}$ in terms of Z_t and find $P_{n+h}^n = \text{Var}(e_n(h) | Y_1, \dots, Y_n)$.
4. Find the 95% Prediction interval $(Y_{n+h}^n - 1.96\sqrt{P_{n+h}^n}, Y_{n+h}^n + 1.96\sqrt{P_{n+h}^n})$.

(★★☆) Forecasting: ARMA(2,2) model

 **Exercise 1.** Consider $Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + Z_t - \theta_1 Z_{t-1} - \theta_2 Z_{t-2}$ to be causal and invertible, where $\{Z_t\} \stackrel{iid}{\sim} N(0, \sigma^2)$. Given $Y_1 = 1$, $Y_2 = 3$, $Y_3 = 1$ and $Y_4 = 6$.

(a) Find \hat{Y}_5 and \hat{Y}_6 in terms of $\phi_1, \phi_2, \theta_1, \theta_2$.

(b) Find the 95% prediction interval for Y_5 and Y_6 in terms of $\phi_1, \phi_2, \theta_1, \theta_2$ and σ^2 .

Solution

(a) Assume $Z_t \leq 0$ for $t \leq 2$ ($p = 2$), i.e., $Z_1 = Z_2 = 0$.

- **(Solve for $\{Z_t\}$)** Consider $Z_t = \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2}$. We have

$$Z_3 = \theta_1 Z_2 + \theta_2 Z_1 + Y_3 - \phi_1 Y_2 - \phi_2 Y_1 = 2 - 3\phi_1 - \phi_2.$$

$$Z_4 = \theta_1 Z_3 + \theta_2 Z_2 + Y_4 - \phi_1 Y_3 - \phi_2 Y_2 = \theta_1(2 - 3\phi_1 - \phi_2) + 6 - \phi_1 - 3\phi_2.$$

- **(Find the forecast)** We have $E(Y_t | Y_1, \dots, Y_n)$ for all $t > n$ and hence

$$\begin{aligned} \hat{Y}_5 &= E(\phi_1 Y_4 + \phi_2 Y_3 + Z_5 - \theta_1 Z_4 - \theta_2 Z_3 | Y_1, \dots, Y_4) = \phi_1 Y_4 + \phi_2 Y_3 + \mathbf{0} - \theta_1 Z_4 - \theta_2 Z_3 \\ &= 6\phi_1 + \phi_2 - \theta_1(\theta_1(2 - 3\phi_1 - \phi_2) + 6 - \phi_1 - 3\phi_2) - \theta_2(2 - 3\phi_1 - \phi_2). \end{aligned}$$

$$\begin{aligned} \hat{Y}_6 &= E(\phi_1 Y_5 + \phi_2 Y_4 + Z_6 - \theta_1 Z_5 - \theta_2 Z_4 | Y_1, \dots, Y_4) = \phi_1 \hat{Y}_5 + \phi_2 Y_4 + \mathbf{0} - \theta_1 \mathbf{0} - \theta_2 Z_4 \\ &= \phi_1 \left[6\phi_1 + \phi_2 - \theta_1(\theta_1(2 - 3\phi_1 - \phi_2) + 6 - \phi_1 - 3\phi_2) - \theta_2(2 - 3\phi_1 - \phi_2) \right] + \\ &\quad 6\phi_2 - \theta_2(\theta_1(2 - 3\phi_1 - \phi_2) + 6 - \phi_1 - 3\phi_2). \end{aligned}$$

(b) We first find the forecast error and its variance

- $e_4(1) = Y_5 - \hat{Y}_5 = Z_5$ and hence $\text{Var}(e_4(1) | Y_1, \dots, Y_4) = \sigma^2$.
- We then evaluate $e_4(2)$:

$$\begin{aligned} e_4(2) &= Y_6 - \hat{Y}_6 = \phi_1(Y_5 - \hat{Y}_5) + Z_6 - \theta_1 Z_5 \\ &= \phi_1 e_4(1) + Z_6 - \theta_1 Z_5 = (\phi_1 - \theta_1)Z_5 + Z_6 \end{aligned}$$

$$\text{and hence } \text{Var}(e_4(2) | Y_1, \dots, Y_4) = (\phi_1^2 - 2\phi_1\theta_1 + \theta_1^2 + 1)\sigma^2.$$

Hence, the 95% prediction intervals for Y_5 and Y_6 are given by


$$\left(\hat{Y}_5 - 1.96\sigma, \hat{Y}_5 + 1.96\sigma \right)$$

and

$$\left(\hat{Y}_6 - 1.96(\phi_1^2 - 2\phi_1\theta_1 + \theta_1^2 + 1)^{1/2}\sigma, \hat{Y}_6 + 1.96(\phi_1^2 - 2\phi_1\theta_1 + \theta_1^2 + 1)^{1/2}\sigma \right),$$

respectively.

(★★☆) Forecasting: ARIMA(2,1,1) model

 **Exercise 2. (Problematic!)** Assume that $\{Y_t\} \sim \text{ARIMA}(2, 1, 1)$, i.e., $X_t = \Delta Y_t$ being causal and invertible

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + Z_t - \theta_1 Z_{t-1},$$

where $\{Z_t\} \stackrel{iid}{\sim} N(0, \sigma^2)$. Suppose that you are given Y_1, \dots, Y_n ($n > 2$).

- Find \hat{Y}_{n+1} and \hat{Y}_{n+2} in terms of (Y_1, Y_2, \dots, Y_n) , (Z_1, Z_2, \dots, Z_n) , ϕ_1 , ϕ_2 , and θ_1 .
- Find 95% prediction intervals of Y_{n+1} and Y_{n+2} in terms of (Y_1, Y_2, \dots, Y_n) , ϕ_1 , ϕ_2 , θ_1 and σ^2 .
- Suppose $Y_1 = 1, Y_2 = 2$ and $Y_3 = 3$, find \hat{Y}_4 and \hat{Y}_5 in terms of ϕ_1, ϕ_2 and θ_1 .

Solution

- Noticing that as $X_t = Y_t - Y_{t-1}$, knowing Y_1, \dots, Y_n gives a value of X_2, \dots, X_n .

$$\begin{aligned} \hat{X}_{n+1} &= E(\phi_1 X_n + \phi_2 X_{n-1} + Z_{n+1} - \theta_1 Z_n | Y_1, \dots, Y_n) = \phi_1 X_n + \phi_2 X_{n-1} + 0 - \theta_1 Z_n \\ &= \phi_1 (Y_n - Y_{n-1}) + \phi_2 (Y_{n-1} - Y_{n-2}) - \theta_1 Z_n. \end{aligned}$$

$$\begin{aligned} \hat{X}_{n+2} &= E(\phi_1 X_{n+1} + \phi_2 X_n + Z_{n+2} - \theta_1 Z_{n+1} | Y_1, \dots, Y_n) = \phi_1 \hat{X}_{n+1} + \phi_2 X_n + 0 - \theta_1 0 \\ &= \phi_1 (\phi_1 (Y_n - Y_{n-1}) + \phi_2 (Y_{n-1} - Y_{n-2}) - \theta_1 Z_n) + \phi_2 (Y_n - Y_{n-1}). \end{aligned}$$

Then we can write

$$\bullet \hat{Y}_{n+1} = Y_n + \hat{X}_{n+1}. \quad \bullet \hat{Y}_{n+2} = Y_n + \hat{X}_{n+1} + \hat{X}_{n+2}.$$

- We then find the forecasting error and its variance. We have

- $e_{n,X}(1) = X_{n+1} - \hat{X}_{n+1} = Z_{n+1}.$
- $e_{n,X}(2) = X_{n+2} - \hat{X}_{n+2} = \phi_1 e_{n,X}(1) + Z_{n+2} - \theta_1 Z_{n+1} = (\phi_1 - \theta_1) Z_{n+1} + Z_{n+2}.$
- $e_{n,Y}(1) = e_{n,X}(1) = Z_{n+1}$ and $e_{n,Y}(2) = e_{n,X}(1) + e_{n,X}(2) = (1 + \phi_1 - \theta_1) Z_{n+1} + Z_{n+2}.$

Hence $\text{Var}(e_{n,Y}(1) | Y_1, \dots, Y_n) = \sigma^2$ and $\text{Var}(e_{n,Y}(2) | Y_1, \dots, Y_n) = [(1 + \phi_1 - \theta_1)^2 + 1] \sigma^2$. Therefore, the 95% prediction interval is given by

$$(\hat{Y}_{n+1} - 1.96\sigma, \hat{Y}_{n+1} + 1.96\sigma) \quad \text{and} \quad (\hat{Y}_{n+2} \pm 1.96[(1 + \phi_1 - \theta_1)^2 + 1]^{1/2} \sigma).$$

- By invertibility, we assume $Y_0 = 0$ and, hence, $X_1 = Y_1 - Y_0 = Y_1$. As $\{X_t\} \sim \text{ARMA}(2, 1)$, assume $Z_t = 0$ for $t \leq 2$. Notice that $Z_t = X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} + \theta_1 Z_{t-1}$ and hence

$$Z_3 = X_3 - \phi_1 X_2 - \phi_2 X_1 + \theta_1 Z_2 = 1 - \phi_1 - \phi_2.$$


Then by (a),

$$\begin{aligned} \hat{X}_4 &= \phi_1 (Y_3 - Y_2) + \phi_2 (Y_3 - Y_2) - \theta_1 Z_3 = \phi_1 + \phi_2 - \theta_1 (1 - \phi_1 - \phi_2) \\ \hat{X}_5 &= \phi_1 \hat{X}_4 + \phi_2 (Y_3 - Y_2) = \phi_1 [\phi_1 + \phi_2 - \theta_1 (1 - \phi_1 - \phi_2)] + \phi_2. \end{aligned}$$

Hence $\hat{Y}_4 = Y_3 + \hat{X}_4 = 3 + \hat{X}_4$ and $\hat{Y}_5 = Y_3 + \hat{X}_4 + \hat{X}_5 = 3 + \hat{X}_4 + \hat{X}_5$.

7.2 Additional Exercise

(★★☆) Forecasting: AR(1) Model

 **Exercise 3.** Consider causal AR(1) model $Y_t = \alpha Y_{t-1} + Z_t$, where $\{Z_t\} \stackrel{iid}{\sim} N(0, \sigma^2)$.

- (a) Evaluate \hat{Y}_{n+1} and \hat{Y}_{n+2} . Also, construct the corresponding 95% prediction intervals.
- (b) By definition, find the explicit form of \hat{Y}_{n+h} for all $h > 0$.
- (c) Find the value of $\text{Cov}\{e_n(h), e_n(\ell)\}$, where $0 < h < \ell$.

Solution

- (a) • For the 1-step forecast, consider $Y_{n+1}^n = \alpha Y_n + E(Z_{n+1}|Y_1, \dots, Y_n) = \alpha Y_n$. Hence, $e_n(1) = Y_{n+1} - Y_{n+1}^n = Z_{n+1}$ and $\text{Var}(e_n(1)|Y_1, \dots, Y_n) = \sigma^2$.
- For the 2-step forecast, consider $Y_{n+2}^n = \alpha Y_{n+1} + E(Z_{n+2}|Y_1, \dots, Y_n) = \alpha Y_{n+1}$. Hence

$$e_n(2) = Y_{n+2} - Y_{n+2}^n = \alpha Y_{n+1} + Z_{n+2} - \alpha Y_{n+1} = \alpha e_n(1) + Z_{n+2} = \alpha Z_{n+1} + Z_{n+2}.$$

$$\text{and } \text{Var}(e_n(2)|Y_1, \dots, Y_n) = (\alpha^2 + 1)\sigma^2.$$

It follows that the 95% prediction intervals for Y_{n+1} and Y_{n+2} are given by

$$(\alpha Y_n - 1.96\sigma^2, \alpha Y_n + 1.96\sigma^2)$$

and

$$(\alpha^2 Y_n - 1.96(\alpha^2 + 1)^{1/2}\sigma^2, \alpha^2 Y_n + 1.96(\alpha^2 + 1)^{1/2}\sigma^2),$$

respectively.

- (b) By induction, we have

$$\begin{aligned} \hat{Y}_{n+h} &= E(Y_{n+h}|Y_1, \dots, Y_n) = E(\alpha Y_{n+h-1} + Z_{n+h}|Y_1, \dots, Y_n) = \alpha E(Y_{n+h-1}|Y_1, \dots, Y_n) \\ &= \dots = E(\alpha^h Y_n|Y_1, \dots, Y_n) = \alpha^h Y_n. \end{aligned}$$


- (c) The MA representation is given by $Y_t = \sum_{k=0}^{\infty} \alpha^k Z_{t-k}$. Then

$$e_n(d) = \sum_{i=0}^{d-1} \alpha^i Z_{n+d-i}$$

and hence for $h < \ell$,

$$\begin{aligned} \text{Cov}\{e_n(h), e_n(\ell)\} &= \text{Cov}\left(\sum_{i=0}^{h-1} \alpha^i Z_{n+h-i}, \sum_{j=0}^{\ell-1} \alpha^j Z_{n+\ell-j}\right) \\ &= \sigma^2 \sum_{i=0}^{h-1} \alpha^i \times \alpha^{\ell-h+i} = \frac{\sigma^2}{1 - \alpha^2} (\alpha^{\ell-h} - \alpha^{\ell+h}). \end{aligned}$$

(★★☆) Forecasting: MA(1) Model

 **Exercise 4.** Consider the invertible MA(1) model $Y_t = Z_t + \theta Z_{t-1}$, where $\{Z_t\} \stackrel{iid}{\sim} \mathcal{N}(0, \sigma^2)$. Given $Y_1 = 2$ and $Y_2 = 3$.

- (a) Evaluate the h -step forecast \hat{Y}_{2+h} in terms of θ for all $h > 0$.
 (b) Find 95% prediction intervals for Y_3 in terms of θ .

Solution

(a) **(Solve for $\{Z_t\}$)** First, assume $Z_0 = 0$. Then $Y_1 = Z_1 + \theta Z_0 = Z_1$

$$Y_2 = Z_2 + \theta Z_1 = Z_2 + \theta Y_1 \Rightarrow Z_2 = Y_2 - \theta Y_1 = 3 - 2\theta.$$

(Find the Forecast) Then consider

- $\hat{Y}_3 = E(Z_3 + \theta Z_2 | Y_1, Y_2) = \theta Z_2 = \theta(3 - 2\theta).$
- $\hat{Y}_{2+h} = E(Z_{2+h} + \theta Z_{1+h} | Y_1, Y_2) = 0$ for $h \geq 1$.

(b) As $Y_3 = Z_3 + \theta Z_2 = Z_3 + \theta(3 - 2\theta)$ and $\hat{Y}_3 = \theta(3 - 2\theta)$. Then $e_2(1) = Z_3$ and hence $P_2^3 = \sigma^2$. It follows that 95% prediction intervals for Y_3 is given by

$$\left(\theta(3 - 2\theta) - 1.96\sigma^2, \theta(3 - 2\theta) + 1.96\sigma^2 \right).$$

7.3 ✂ R-programming

Recall from Tutorial 06, we have conducted some time-series analysis concerning the stock price of Apple company. In this section, we consider the prediction of the stock price of Bank of China Limited.

```
install.packages("tseries") ; install.packages("forecast")
library("tseries") ; library("forecast")
X0 = get.hist.quote(instrument = "3988.HK", start = "2022-10-01", end = "2022-11-13",
quote = "Adjusted", provider = c("yahoo"), compression = "d", retclass = c("ts"))
X = X0[-which(is.na(X0)==1)]

fit = auto.arima(X,max.p=5,max.q=5,ic="aicc")
fit # See the optimal order based on AICC

Series: X
ARIMA(1,0,0) with non-zero mean

Coefficients:
      ar1      mean
    0.5956  2.5984
s.e.  0.1631  0.0135

sigma^2 = 0.001021: log likelihood = 59.54
AIC=-113.07  AICc=-112.11  BIC=-108.97

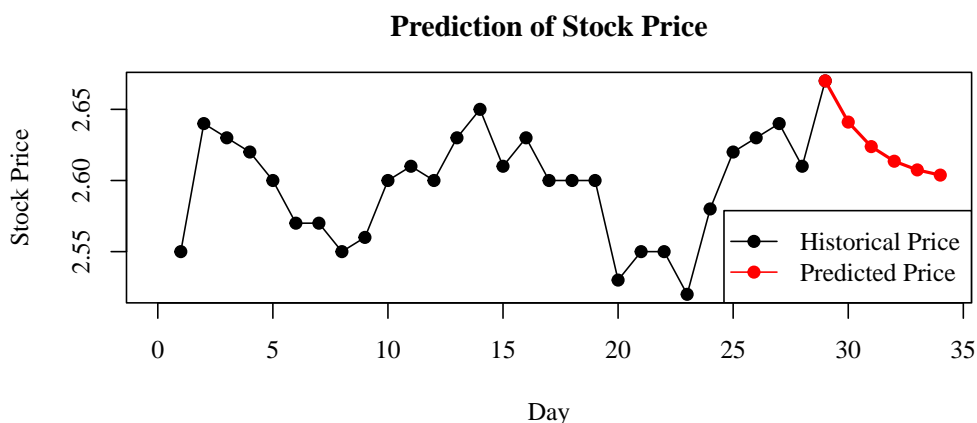
model = arima(X, order = c(1, 0, 0))
BaJ = predict(model,n.ahead=5)

n = length(X)
plot(1:n,X,xlim=c(0,n+5),pch=19,xlab="Day",ylab="Stock Price",main="Prediction of
Stock Price")
lines(1:n,X,xlim=c(0,n+5))
points(n:(n+5),c(X[n],BaJ$pred),col="red",pch=19)
lines(n:(n+5),c(X[n],BaJ$pred),col="red",lwd=2)
legend("bottomright", legend=c("Historical Price","Predicted Price"),
      pch = c(19,19), col=c("black","red"), lty=c(1,1), cex=1)
```

The following are some useful R-commands.

1. `auto.arima`: A function built under the package "forecast", which returns the ARIMA model with the lowest value in the prespecified information criteria.
2. `predict(model,n.ahead=h)`: Provide the value of $\hat{Y}_{n+1}, \dots, \hat{Y}_{n+h}$.

Refer to the following figure for the predicted price.



❗ Remark 3. You should NOT use those two functions in your assignment. Instead, try to derive the forecasting function and write the code without relying on the forecast package.