

# **Lecture 7**

# **MULTI-STEP-AHEAD FORECASTING**

# **Multi-step-ahead forecasting**

## **Formalities**

# One-step-ahead forecasting review

MSE-optimal forecast of  $Y_{n+1}$  given data  $\mathcal{Y}_n$  :

$$E(Y_{n+1} \mid \mathcal{Y}_n)$$

One-step-ahead forecast errors:

$$U_{n+1} = Y_{n+1} - E(Y_{n+1} \mid \mathcal{Y}_n)$$

satisfy

$$E(U_{n+1} \mid \mathcal{Y}_n) = 0$$

⇒ errors have no autocorrelation at any lag.

# Multi-step-ahead forecasting

MSE-optimal forecast of  $Y_{n+h}$  given data  $\mathcal{Y}_n$ :

$$E(Y_{n+h} \mid \mathcal{Y}_n)$$

Multi-step-ahead forecast errors:

$$U_{n+h \mid n} = Y_{n+h} - E(Y_{n+h} \mid \mathcal{Y}_n)$$

satisfy

$$E(U_{n+h \mid n} \mid \mathcal{Y}_n) = 0$$

$\Rightarrow$  autocorrelation???

# Multi-step-ahead forecast errors

For any  $t$ :  $E(U_{t+h} | t | \mathcal{Y}_t) = 0$

Unconditional mean:

$$E( U_{t+h} | t ) = E( E( U_{t+h} | t | \mathcal{Y}_t ) ) \quad (\text{LIE})$$

# Multi-step-ahead forecast errors

For any  $t$ :  $E(U_{t+h|t} | \mathcal{Y}_t) = 0$

Unconditional mean:

$$\begin{aligned} E(U_{t+h|t}) &= E(E(U_{t+h|t} | \mathcal{Y}_t)) \quad (\text{LIE}) \\ &= 0 \end{aligned}$$

# Multi-step-ahead forecast errors

For any  $t$ :  $E(U_{t+h|t} | \mathcal{Y}_t) = 0$

Autocovariance at lag  $j \geq h$ :

$$\begin{aligned} & \text{cov}(U_{t+h|t}, U_{t+h-j|t-j}) \\ &= E(U_{t+h|t} U_{t+h-j|t-j}) \quad \text{since } E(U_{t+h|t}) = 0 \text{ for all } t \end{aligned}$$

# Multi-step-ahead forecast errors

For any  $t$ :  $E(U_{t+h|t} | \mathcal{Y}_t) = 0$

Autocovariance at lag  $j \geq h$ :

$$\begin{aligned} & \text{cov}(U_{t+h|t}, U_{t+h-j|t-j}) \\ &= E(U_{t+h|t} U_{t+h-j|t-j}) \quad \text{since } E(U_{t+h|t}) = 0 \text{ for all } t \\ &= E(E(U_{t+h|t} U_{t+h-j|t-j} | \mathcal{Y}_t)) \quad (\text{LIE}) \end{aligned}$$

# Multi-step-ahead forecast errors

For any  $t$ :  $E(U_{t+h|t} | \mathcal{Y}_t) = 0$

Autocovariance at lag  $j \geq h$ :

$$\begin{aligned} & \text{cov}(U_{t+h|t}, U_{t+h-j|t-j}) \\ &= E(U_{t+h|t} U_{t+h-j|t-j}) \quad \text{since } E(U_{t+h|t}) = 0 \text{ for all } t \\ &= E(E(U_{t+h|t} U_{t+h-j|t-j} | \mathcal{Y}_t)) \quad (\text{LIE}) \\ &= E(U_{t+h-j|t-j} E(U_{t+h|t} | \mathcal{Y}_t)) \quad \text{since } t + h - j \leq t \end{aligned}$$

# Multi-step-ahead forecast errors

For any  $t$ :  $E(U_{t+h|t} | \mathcal{Y}_t) = 0$

Autocovariance at lag  $j \geq h$ :

$$\begin{aligned} & \text{cov}(U_{t+h|t}, U_{t+h-j|t-j}) \\ &= E(U_{t+h|t} U_{t+h-j|t-j}) \quad \text{since } E(U_{t+h|t}) = 0 \text{ for all } t \\ &= E(E(U_{t+h|t} U_{t+h-j|t-j} | \mathcal{Y}_t)) \quad (\text{LIE}) \\ &= E(U_{t+h-j|t-j} E(U_{t+h|t} | \mathcal{Y}_t)) \quad \text{since } t + h - j \leq t \\ &= 0 \end{aligned}$$

# Multi-step-ahead forecast errors

For any  $t$ :  $E(U_{t+h|t} | \mathcal{Y}_t) = 0$

Autocovariance at lag  $j < h$  :

$$\begin{aligned} & \text{cov}(U_{t+h|t}, U_{t+h-j|t-j}) \\ &= E(U_{t+h|t} U_{t+h-j|t-j}) \quad \text{since } E(U_{t+h|t}) = 0 \text{ for all } t \\ &= E(E(U_{t+h|t} U_{t+h-j|t-j} | \mathcal{Y}_t)) \quad (\text{LIE}) \\ &\neq E(U_{t+h-j|t-j} E(U_{t+h|t} | \mathcal{Y}_t)) \quad \text{since } t + h - j > t \end{aligned}$$

# Multi-step-ahead forecast errors

For any  $t$ :  $E(U_{t+h|t} | \mathcal{Y}_t) = 0$

$$\text{cov}(U_{t+h|t}, U_{t+h-j|t-j}) = \begin{cases} ? & j = 1, \dots, h-1 \\ 0 & j = h, h+1, \dots \end{cases}$$

# Multi-step-ahead forecast errors

For any  $t$ :  $E(U_{t+h|t} | \mathcal{Y}_t) = 0$

$$\text{cov}(U_{t+h|t}, U_{t+h-j|t-j}) = \begin{cases} ? & j = 1, \dots, h-1 \\ 0 & j = h, h+1, \dots \end{cases}$$

$h$ -step-ahead forecast errors have zero autocorrelation at lags  $h, h+1, \dots$

# Multi-step-ahead forecast errors

For any  $t$ :  $E(U_{t+h|t} | \mathcal{Y}_t) = 0$

$$\text{cov}(U_{t+h|t}, U_{t+h-j|t-j}) = \begin{cases} ? & j = 1, \dots, h-1 \\ 0 & j = h, h+1, \dots \end{cases}$$

$h$ -step-ahead forecast errors *may* have non-zero autocorrelation at lags  $1, \dots, h-1$ .

# **Multi-step-ahead forecasting**

## **Approach 1**

# Multi-step forecasts from 1-step models

The most common approach:

- specify model of  $E( Y_t | \mathcal{Y}_{t-1} )$ , eg ARIMA
- use “extended” LIE

$$E( Y | \mathcal{F} ) = E( E( Y | \mathcal{G} ) | \mathcal{F} ) \quad \text{if } \mathcal{F} \subset \mathcal{G}$$

# Multi-step forecasts from 1-step models

The most common approach:

- specify model of  $E( Y_t | \mathcal{Y}_{t-1} )$ , eg ARIMA
- use “extended” LIE

$$E( Y | \mathcal{F} ) = E( E( Y | \mathcal{G} ) | \mathcal{F} ) \quad \text{if } \mathcal{F} \subset \mathcal{G}$$

to *derive* multi-step-ahead forecasts. Eg:

$$E( Y_{n+2} | \mathcal{Y}_n ) = E( E( Y_{n+2} | \mathcal{Y}_{n+1} ) | \mathcal{Y}_n )$$

since  $\mathcal{Y}_n \subset \mathcal{Y}_{n+1}$

# Multi-step forecasts from 1-step models

The most common approach:

- specify model of  $E( Y_t | \mathcal{Y}_{t-1} )$ , eg ARIMA
- use “extended” LIE

$$Y = Y \quad \text{if } \subset$$

to *derive* multi-step-ahead forecasts. Eg:

$$Y_{n+2} = Y_{n+2} \quad \text{since } \mathcal{Y}_n \subset \mathcal{Y}_{n+1}$$

↑

More data at time  $n + 1$  than time  $n$ .

# Multi-step forecasts from 1-step models

The most common approach:

- specify model of  $E(Y_t | \mathcal{Y}_{t-1})$ , eg ARIMA
- use “extended” LIE

$$Y = Y \quad \text{if } \subset$$

to *derive* multi-step-ahead forecasts. Eg:

$$Y_{n+2} = E(Y_{n+2} | \mathcal{Y}_{n+1})$$

↑  
1-step-ahead forecast

# Example. AR(1) model

$$E(Y_{n+1} | \mathcal{Y}_n) = \phi_1 Y_n$$

2-step-ahead forecast:

$$E(Y_{n+2} | \mathcal{Y}_n) = E(E(Y_{n+2} | \mathcal{Y}_{n+1}) | \mathcal{Y}_n)$$

# Example. AR(1) model

$$E(Y_{n+1} | \mathcal{Y}_n) = \phi_1 Y_n$$

2-step-ahead forecast:

$$\begin{aligned} E(Y_{n+2} | \mathcal{Y}_n) &= E(E(Y_{n+2} | \mathcal{Y}_{n+1}) | \mathcal{Y}_n) \\ &= E(\phi_1 Y_{n+1} | \mathcal{Y}_n) \end{aligned}$$

# Example. AR(1) model

$$E(Y_{n+1} | \mathcal{Y}_n) = \phi_1 Y_n$$

2-step-ahead forecast:

$$\begin{aligned} E(Y_{n+2} | \mathcal{Y}_n) &= E(E(Y_{n+2} | \mathcal{Y}_{n+1}) | \mathcal{Y}_n) \\ &= E(\phi_1 Y_{n+1} | \mathcal{Y}_n) \\ &= \phi_1 E(Y_{n+1} | \mathcal{Y}_n) \end{aligned}$$

# Example. AR(1) model

$$E(Y_{n+1} | \mathcal{Y}_n) = \phi_1 Y_n$$

2-step-ahead forecast:

$$\begin{aligned} E(Y_{n+2} | \mathcal{Y}_n) &= E(E(Y_{n+2} | \mathcal{Y}_{n+1}) | \mathcal{Y}_n) \\ &= E(\phi_1 Y_{n+1} | \mathcal{Y}_n) \\ &= \phi_1 E(Y_{n+1} | \mathcal{Y}_n) \\ &= \phi_1 \phi_1 Y_n \end{aligned}$$

# Example. AR(1) model

$$E(Y_{n+1} | \mathcal{Y}_n) = \phi_1 Y_n$$

2-step-ahead forecast:

$$\begin{aligned} E(Y_{n+2} | \mathcal{Y}_n) &= E(E(Y_{n+2} | \mathcal{Y}_{n+1}) | \mathcal{Y}_n) \\ &= E(\phi_1 Y_{n+1} | \mathcal{Y}_n) \\ &= \phi_1 E(Y_{n+1} | \mathcal{Y}_n) \\ &= \phi_1 \phi_1 Y_n \\ &= \phi_1^2 Y_n \end{aligned}$$

# Example. AR(1) model

$$E(Y_{n+1} | \mathcal{Y}_n) = \phi_1 Y_n$$

$$E(Y_{n+2} | \mathcal{Y}_n) = \phi_1^2 Y_n$$

3-step-ahead forecast:

$$E(Y_{n+3} | \mathcal{Y}_n) = E(E(Y_{n+3} | \mathcal{Y}_{n+1}) | \mathcal{Y}_n)$$

# Example. AR(1) model

$$E(Y_{n+1} | \mathcal{Y}_n) = \phi_1 Y_n$$

$$E(Y_{n+2} | \mathcal{Y}_n) = \phi_1^2 Y_n$$

3-step-ahead forecast:

$$\begin{aligned} E(Y_{n+3} | \mathcal{Y}_n) &= E(E(Y_{n+3} | \mathcal{Y}_{n+1}) | \mathcal{Y}_n) \\ &= E(\phi_1^2 Y_{n+1} | \mathcal{Y}_n) \end{aligned}$$

# Example. AR(1) model

$$E(Y_{n+1} | \mathcal{Y}_n) = \phi_1 Y_n$$

$$E(Y_{n+2} | \mathcal{Y}_n) = \phi_1^2 Y_n$$

3-step-ahead forecast:

$$\begin{aligned} E(Y_{n+3} | \mathcal{Y}_n) &= E(E(Y_{n+3} | \mathcal{Y}_{n+1}) | \mathcal{Y}_n) \\ &= E(\phi_1^2 Y_{n+1} | \mathcal{Y}_n) \\ &= \phi_1^2 E(Y_{n+1} | \mathcal{Y}_n) \end{aligned}$$

# Example. AR(1) model

$$E(Y_{n+1} | \mathcal{Y}_n) = \phi_1 Y_n$$

$$E(Y_{n+2} | \mathcal{Y}_n) = \phi_1^2 Y_n$$

3-step-ahead forecast:

$$\begin{aligned} E(Y_{n+3} | \mathcal{Y}_n) &= E(E(Y_{n+3} | \mathcal{Y}_{n+1}) | \mathcal{Y}_n) \\ &= E(\phi_1^2 Y_{n+1} | \mathcal{Y}_n) \\ &= \phi_1^2 E(Y_{n+1} | \mathcal{Y}_n) \\ &= \phi_1^2 \phi_1 Y_n \end{aligned}$$

# Example. AR(1) model

$$E(Y_{n+1} | \mathcal{Y}_n) = \phi_1 Y_n$$

$$E(Y_{n+2} | \mathcal{Y}_n) = \phi_1^2 Y_n$$

3-step-ahead forecast:

$$\begin{aligned} E(Y_{n+3} | \mathcal{Y}_n) &= E(E(Y_{n+3} | \mathcal{Y}_{n+1}) | \mathcal{Y}_n) \\ &= E(\phi_1^2 Y_{n+1} | \mathcal{Y}_n) \\ &= \phi_1^2 E(Y_{n+1} | \mathcal{Y}_n) \\ &= \phi_1^3 Y_n \end{aligned}$$

## Example. AR(1) model

For any  $h = 1, 2, \dots$

$$E(Y_{n+h} | \mathcal{Y}_n) = \phi_1^h Y_n$$

This process can be applied to derive the  $h$ -step-ahead forecasting formulae for any ARIMA model.

(But R **forecast** handles these formulae for us!)

# **Short- vs long-term forecasting**

# What happens as $h$ increases?

Consider the AR(1) model

$$E(Y_t | \mathcal{Y}_{t-1}) = \phi_1 Y_{t-1}$$

with  $h$ -step-ahead forecast function

$$E(Y_{n+h} | \mathcal{Y}_n) = \phi_1^h Y_n$$

# What happens as $h$ increases?

Consider the AR(1) model

$$E(Y_t | \mathcal{Y}_{t-1}) = \phi_1 Y_{t-1}$$

with  $h$ -step-ahead forecast function

$$E(Y_{n+h} | \mathcal{Y}_n) = \phi_1^h Y_n$$

As  $h \rightarrow \infty$ :

- $|\phi_1| < 1 \Rightarrow \phi_1^h \rightarrow 0$
- $\phi_1 = 1 \Rightarrow \phi_1^h = 1$  for every  $h$
- $|\phi_1| > 1 \Rightarrow \phi_1^h \rightarrow \infty$

# What happens as $h$ increases?

Consider the AR(1) model

$$E(Y_t | \mathcal{Y}_{t-1}) = \phi_1 Y_{t-1}$$

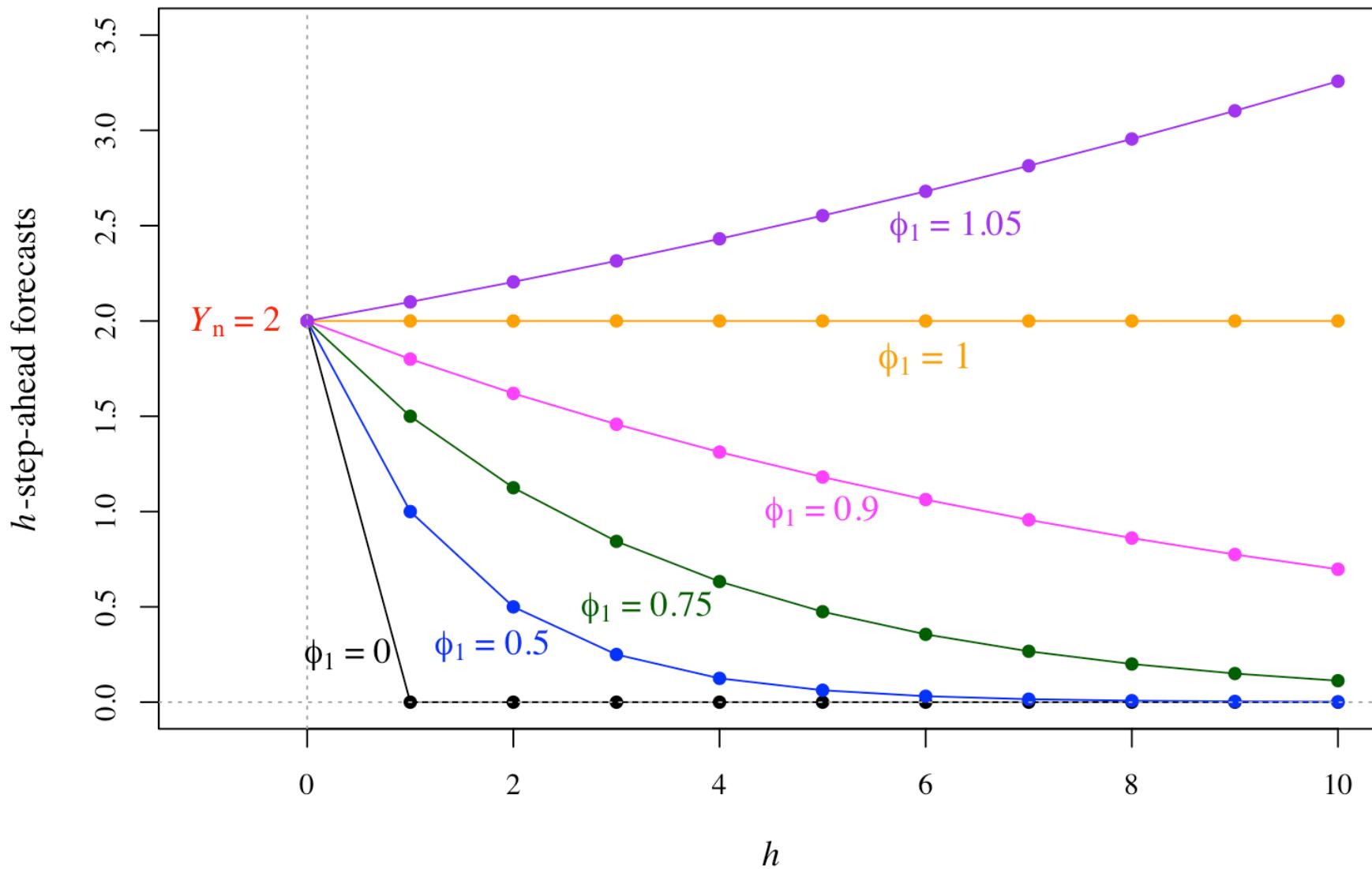
with  $h$ -step-ahead forecast function

$$E(Y_{n+h} | \mathcal{Y}_n) = \phi_1^h Y_n$$

As  $h \rightarrow \infty$ :

- $|\phi_1| < 1 \Rightarrow E(Y_{n+h} | \mathcal{Y}_n) \rightarrow 0$
- $\phi_1 = 1 \Rightarrow E(Y_{n+h} | \mathcal{Y}_n) = Y_n$  for every  $h$
- $|\phi_1| > 1 \Rightarrow E(Y_{n+h} | \mathcal{Y}_n) \rightarrow \pm\infty$

# Numerical illustration with $Y_n = 2$



# Forecasting with a trend function

Suppose  $X_t$  specifies a *deterministic* trend:

$$Y_t = X'_t \beta + Z_t$$

$$E( Z_t | \mathcal{Z}_{t-1} ) = \phi_1 Z_{t-1}$$

# Forecasting with a trend function

Suppose  $X_t$  specifies a *deterministic* trend:

$$Y_t = X'_t \beta + Z_t$$

$$E( Z_t | \mathcal{Z}_{t-1} ) = \phi_1 Z_{t-1}$$

Then

$$E( Y_{n+h} | \mathcal{Y}_n ) = X'_{n+h} \beta + E( Z_{n+h} | \mathcal{Y}_n )$$



Future values of  $X_t$  are *known*.

i.e. no need to forecast these

# Forecasting with a trend function

Suppose  $X_t$  specifies a *deterministic* trend:

$$Y_t = X'_t \beta + Z_t$$

$$E( Z_t | \mathcal{Z}_{t-1} ) = \phi_1 Z_{t-1}$$

Then

$$\begin{aligned} E( Y_{n+h} | \mathcal{Y}_n ) &= X'_{n+h} \beta + E( Z_{n+h} | \mathcal{Y}_n ) \\ &= X'_{n+h} \beta + E( Z_{n+h} | \mathcal{Z}_n ) \end{aligned}$$

$X_t$  is deterministic  $\Rightarrow \mathcal{Y}_n$  and  $\mathcal{Z}_n$  are equivalent.

# Forecasting with a trend function

Suppose  $X_t$  specifies a *deterministic* trend:

$$Y_t = X'_t \beta + Z_t$$

$$E(Z_t | \mathcal{Z}_{t-1}) = \phi_1 Z_{t-1}$$

Then

$$\begin{aligned} E(Y_{n+h} | \mathcal{Y}_n) &= X'_{n+h} \beta + E(Z_{n+h} | \mathcal{Y}_n) \\ &= X'_{n+h} \beta + E(Z_{n+h} | \mathcal{Z}_n) \\ &= X'_{n+h} \beta + \phi_1^h Z_n \end{aligned}$$

# Forecasting with a trend function

Suppose  $X_t$  specifies a *deterministic* trend:

$$Y_t = X'_t \beta + Z_t$$

$$E( Z_t | \mathcal{Z}_{t-1} ) = \phi_1 Z_{t-1}$$

Then

$$\begin{aligned} E( Y_{n+h} | \mathcal{Y}_n ) &= X'_{n+h} \beta + E( Z_{n+h} | \mathcal{Y}_n ) \\ &= X'_{n+h} \beta + E( Z_{n+h} | \mathcal{Z}_n ) \\ &= X'_{n+h} \beta + \phi_1^h Z_n \\ &= X'_{n+h} \beta + \phi_1^h (Y_n - X'_n \beta) \end{aligned}$$

# Forecasting with a trend function

Suppose  $X_t$  specifies a *deterministic* trend:

$$Y_t = \textcolor{cyan}{X}'_t \beta + Z_t$$

$$E(Z_t | \mathcal{Z}_{t-1}) = \phi_1 Z_{t-1}$$

Then

$$E(Y_{n+h} | \mathcal{Y}_n) = \textcolor{cyan}{X}'_{n+h} \beta + \phi_1^h (Y_n - X'_n \beta)$$

forecast for      trend at      forecast for  
time  $n + h$       time  $n + h$       deviation from

trend at time  $n + h$

# Forecasting with a trend function

$$E(Y_{n+h} | \mathcal{Y}_n) = X'_{n+h}\beta + \phi_1^h (Y_n - X'_n\beta)$$

As  $h \rightarrow \infty$ :

$$|\phi_1| < 1 \Rightarrow E(Y_{n+h} | \mathcal{Y}_n) - X'_{n+h}\beta \rightarrow 0$$

Stationarity implies the forecasts converge to the trend function as the forecast horizon  $h$  increases.

# Forecasting with a trend function

$$E(Y_{n+h} | \mathcal{Y}_n) = X'_{n+h}\beta + \phi_1^h (Y_n - X'_n\beta)$$

As  $h \rightarrow \infty$ :

$$\phi_1 = 1 \Rightarrow E(Y_{n+h} | \mathcal{Y}_n) - X'_{n+h}\beta = Y_n - X'_n\beta$$

A **unit root** implies the time series is **not** forecast to return to its **trend**, regardless of length of time horizon.

# Forecasting with a trend function

$$E(Y_{n+h} | \mathcal{Y}_n) = X'_{n+h}\beta + \phi_1^h (Y_n - X'_n\beta)$$

As  $h \rightarrow \infty$ :

$$|\phi_1| > 1 \Rightarrow E(Y_{n+h} | \mathcal{Y}_n) - X'_{n+h}\beta \rightarrow \pm\infty$$

An **explosive** AR model implies the **forecasts** diverge from the **trend function** as the forecast horizon  $h$  increases.

# Forecasting with a trend function + ARMA

$$Y_t = X'_t \beta + Z_t$$

$$\begin{aligned} Z_t = & \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} \\ & + U_t + \theta_1 U_{t-1} + \dots + \theta_q U_{t-q} \end{aligned}$$

- $Z_t$  stationary  $\Rightarrow$  as  $h \rightarrow \infty$  :  
 $E(Y_{t+h} | \mathcal{Y}_n) - X'_{n+h} \beta \rightarrow 0$
- $Y_t$  is called “*trend stationary*” or  
“*mean reverting*” or “*mixingale*”.

# Forecasting with a trend function + ARMA

$$Y_t = X'_t \beta + Z_t$$

$$\begin{aligned} Z_t = & \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} \\ & + U_t + \theta_1 U_{t-1} + \dots + \theta_q U_{t-q} \end{aligned}$$

- $Z_t$  unit root  $\Rightarrow$  as  $h \rightarrow \infty$  :  
 $E(Y_{t+h} | \mathcal{Y}_n) - X'_{n+h} \beta \rightarrow Y_n - X'_n \beta$
- $Z_t$  is called “*difference stationary*”.

# **Multi-step-ahead forecasting**

## **Approach 2**

# Direct forecasting

Usual approach:

- model  $E(Y_{n+1}|\mathcal{F}_n)$
- then derive  $E(Y_{n+h}|\mathcal{F}_n)$  for  $h = 1, 2, \dots$

“Direct forecasting”:

- model  $E(Y_{n+h}|\mathcal{F}_n)$  for  $h = 1, 2, \dots$

# Direct forecasting with AR models

$$E(Y_t | \mathcal{Y}_{t-1}) = \phi_1^{(1)} Y_{t-1} + \phi_2^{(1)} Y_{t-2} + \dots$$

$$E(Y_t | \mathcal{Y}_{t-2}) = \phi_2^{(2)} Y_{t-2} + \phi_3^{(2)} Y_{t-3} + \dots$$

$$E(Y_t | \mathcal{Y}_{t-3}) = \phi_3^{(3)} Y_{t-3} + \phi_4^{(3)} Y_{t-4} + \dots$$

⋮  
⋮  
⋮

i.e. specify and estimate a *separate* model for each forecast horizon  $h$

# Direct forecasting with AR models

$$E(Y_t | \mathcal{Y}_{t-1}) = \phi_1^{(1)} Y_{t-1} + \phi_2^{(1)} Y_{t-2} + \dots$$

$$E(Y_t | \mathcal{Y}_{t-2}) = \phi_2^{(2)} Y_{t-2} + \phi_3^{(2)} Y_{t-3} + \dots$$

$$E(Y_t | \mathcal{Y}_{t-3}) = \phi_3^{(3)} Y_{t-3} + \phi_4^{(3)} Y_{t-4} + \dots$$

⋮  
⋮  
⋮

The *first* AR lag in each model is  $h = 1, 2, 3, \dots$

# Direct forecasting with AR models

$$E( Y_t | \mathcal{Y}_{t-1} ) = \phi_1^{(1)} Y_{t-1} + \dots + \phi_{p^{(1)}}^{(1)} Y_{t-p^{(1)}}$$

$$E( Y_t | \mathcal{Y}_{t-2} ) = \phi_2^{(2)} Y_{t-2} + \dots + \phi_{p^{(2)}}^{(2)} Y_{t-p^{(2)}}$$

$$E( Y_t | \mathcal{Y}_{t-3} ) = \phi_3^{(3)} Y_{t-3} + \dots + \phi_{p^{(3)}}^{(3)} Y_{t-p^{(3)}}$$

⋮  
⋮  
⋮

AR lag order may differ across  $h = 1, 2, 3, \dots$

# Choice of AR lag orders

For each  $h = 1, 2, \dots$

- Check residual autocorrelation in models

$$Y_t = \phi_h^{(h)} Y_{t-h} + \dots + \phi_{p^{(h)}}^{(h)} Y_{t-p^{(h)}} + U_t^{(h)}$$

for  $p^{(h)} = h, \dots, p_{\max}$ .

- Choose  $p^{(h)}$  to minimise  $\text{AIC}_c$  among models that pass the autocorrelation test.

# Ljung-Box test for direct AR models

Correct specification requires  $U_t^{(h)}$  in

$$Y_t = \phi_h^{(h)} Y_{t-h} + \dots + \phi_{p^{(h)}}^{(h)} Y_{t-p^{(h)}} + U_t^{(h)}$$

must satisfy

$$\text{cov}(U_t^{(h)}, U_{t-j}^{(h)}) = 0 \text{ for } j = h, h+1, \dots$$

Autocovariances at  $j = 1, \dots, h$  are unrestricted.

# Ljung-Box test for direct AR( $p$ ) models

Define  $\rho_j^{(h)} = \text{cor}(U_t^{(h)}, U_{t-j}^{(h)})$ .

Standard Ljung-Box test hypotheses:

$$H_0 : \rho_1^{(h)} = \dots = \rho_l^{(h)}$$

$$H_1 : \text{at least one of } \rho_1^{(h)}, \dots, \rho_l^{(h)} \neq 0$$

and statistic:

$$LB_l = n(n+2) \sum_{j=1}^l \frac{\widehat{\rho}_j^2}{n-j}$$

# Ljung-Box test for direct AR( $p$ ) models

Define  $\rho_j^{(h)} = \text{cor}(U_t^{(h)}, U_{t-j}^{(h)})$ .

We want Ljung-Box test hypotheses:

$$H_0 : \rho_{\textcolor{brown}{h}}^{(h)} = \dots = \rho_l^{(h)}$$

$$H_1 : \text{at least one of } \rho_{\textcolor{brown}{h}}^{(h)}, \dots, \rho_l^{(h)} \neq 0$$

and statistic:

$$LB_l^{(\textcolor{brown}{h})} = n(n+2) \sum_{j=\textcolor{brown}{h}}^l \frac{\widehat{\rho}_j^2}{n-j}$$

# Ljung-Box test for direct AR( $p$ ) models

Define  $\rho_j^{(h)} = \text{cor}(U_t^{(h)}, U_{t-j}^{(h)})$ .

We want Ljung-Box test hypotheses:

$$H_0 : \rho_{\textcolor{brown}{h}}^{(h)} = \dots = \rho_l^{(h)}$$

$$H_1 : \text{at least one of } \rho_{\textcolor{brown}{h}}^{(h)}, \dots, \rho_l^{(h)} \neq 0$$

and statistic:

$$LB_l^{(\textcolor{brown}{h})} = LB_l - LB_{h-1}$$

# Ljung-Box test for direct AR( $p$ ) models

Define  $\rho_j^{(h)} = \text{cor}(U_t^{(h)}, U_{t-j}^{(h)})$ .

We want Ljung-Box test hypotheses:

$$H_0 : \rho_{\textcolor{brown}{h}}^{(h)} = \dots = \rho_{\textcolor{magenta}{l}}^{(h)}$$

$$H_1 : \text{at least one of } \rho_{\textcolor{brown}{h}}^{(h)}, \dots, \rho_{\textcolor{magenta}{l}}^{(h)} \neq 0$$

and statistic:

$$LB_{\textcolor{magenta}{l}}^{(\textcolor{brown}{h})} = LB_{\textcolor{magenta}{l}} - LB_{\textcolor{magenta}{h}-1} \sim \chi^2_{\textcolor{magenta}{l}-p} \text{ under } H_0$$

# **Multi-step forecasting summary**

# Approach 1: “Recursive” forecasting

Specify a **single model** for one-step-ahead forecasting:

$$E(Y_t | \mathcal{Y}_{t-1})$$

and *derive/compute* forecasts from **this model**:

$$E(Y_{n+h} | \mathcal{Y}_n), \quad h = 1, 2, \dots$$

## Approach 2: “Direct” forecasting

Specify separate models for  $h$ -step-ahead forecasting:

$$E(Y_t | \mathcal{Y}_{t-1}), E(Y_t | \mathcal{Y}_{t-2}), E(Y_t | \mathcal{Y}_{t-3}), \dots$$

Residual autocorrelation checking in  $E(Y_t | \mathcal{Y}_{t-h})$ :

$$LB_l^{(h)} = LB_l - LB_{h-1}$$

where  $LB_l$  is the usual Ljung-Box statistic.

## Approach 2: “Direct” forecasting

Specify separate models for  $h$ -step-ahead forecasting:

$$E(Y_t | \mathcal{Y}_{t-1}), E(Y_t | \mathcal{Y}_{t-2}), E(Y_t | \mathcal{Y}_{t-3}), \dots$$

Residual autocorrelation checking in  $E(Y_t | \mathcal{Y}_{t-h})$ :

$$LB_l^{(h)} = LB_l - LB_{h-1}$$

Critical/ $p$ -values from  $\chi^2_{l-p}$  distribution.

Eg.  $l = p + 4$  provides 4 d.f. in the test.

# Some comparisons

Recursive forecasting:

- requires only one model search
- forecasts are “model consistent”

Direct forecasting:

- does not require derivations of multi-step formulae
- perceived to be more flexible across  $h$