STATISTICAL METHODS WITH APPLICATION TO FINANCE

a.y. 2023-2024

Forecasting

R. Pappadà
rpappada@units.it

Department of Economics, Business, Mathematics and Statistics "B. de Finetti"
University of Trieste

May 11, 2024

Outline

Introduction

Forecasting means predicting future values of a time series using the current *information set* given by past and present values of the time series that is being predicted.

Suppose we have an observed time series $x_1, \ldots, x_n$. Out task is to estimate the next value x_{n+1} or, more generally, the *future values*

$$x_{n+h},\ h \geq 1.$$

The forecast of x_{n+h} made at time n (*forecast origin*) for **h steps ahead** (*forecast horizon*) is denoted by $\hat{x}_n(h)$.

Point forecasts and prediction intervals

Two popular techniques used for prediction of time series are

- Prediction using ARMA models
- Exponential smoothing

Both extend easily to the case of GARCH models, although the latter is regarded as a model-free approach.

As a result, a prediction consists of a *point forecast* related to a particular future time period, and a *prediction interval*, i.e. an interval within which a future value is expected to lie with a prescribed probability

What is exponential smoothing?

Exponential smoothing is a popular technique which is used for both *prediction* of time series and *trend estimation*.

We do not necessarily assume that the data come from a stationary model, although this technique should only be used for non-seasonal time series showing no systematic trend.

In general the method is better suited to undifferenced price or value series rather than return series

General idea

Suppose our data represent realizations of rvs $X_1, \dots X_n$, considered without reference to any concrete parametric model.

The task is to obtain a forecast for X_{n+h}, $h \geq 1$, by weighting the data from most recent to most distant with a sequence of exponentially decreasing weights $\alpha(1-\alpha)^i, i = 0, 1, \dots$

$$\alpha x_n + \alpha(1-\alpha)x_{n-1} + \alpha(1-\alpha)^2 x_{n-2} + \dots \tag{1}$$

where $0 < \alpha < 1$ and the weights $\alpha(1-\alpha)^i$ become smaller as i increases (observations further in the past).

Simple exponential smoothing

As a forecast for X_{n+1} we use a prediction of the form

$$\hat{x}_n(1) = \alpha x_n + \alpha(1-\alpha)x_{n-1} + \alpha(1-\alpha)^2 x_{n-2} + \dots \tag{2}$$

This is written in terms of an infinite number of past observations, but in practice there will only be a finite number:

$$\hat{x}_n(1) = \sum_{i=0}^{n-1} \alpha(1-\alpha)^i x_{n-i} \tag{3}$$

$$= \alpha x_n + (1-\alpha)\underbrace{\left[\alpha x_{n-1} + \alpha(1-\alpha)x_{n-2} + \dots\right]}_{\hat{x}_{n-1}(1)}$$

$$= \alpha x_n + (1-\alpha)\hat{x}_{n-1}(1) \tag{4}$$

the prediction at time $t = n$ is obtained from the prediction at time $n-1$ by a simple recursive scheme.

Simple exponential smoothing

As a forecast for X_{n+1} we use a prediction of the form

$$\hat{x}_n(1) = \alpha x_n + \alpha(1-\alpha)x_{n-1} + \alpha(1-\alpha)^2 x_{n-2} + \dots \tag{2}$$

This is written in terms of an infinite number of past observations, but in practice there will only be a finite number:

$$\hat{x}_n(1) = \sum_{i=0}^{n-1} \alpha(1-\alpha)^i x_{n-i} \tag{3}$$

$$= \alpha x_n + (1-\alpha)\underbrace{\left[\alpha x_{n-1} + \alpha(1-\alpha)x_{n-2} + \dots\right]}_{\hat{x}_{n-1}(1)}$$

$$= \alpha x_n + (1-\alpha)\hat{x}_{n-1}(1) \tag{4}$$

the prediction at time $t = n$ is obtained from the prediction at time $n-1$ by a simple recursive scheme.

Table of Contents

Prediction based on ARMA models

Given the data up to time n, forecasting will involve the observations and the fitted residuals up to and including time n.

Denote our sample of n data $x_1, \ldots, x_n$. As a predictor of X_{t+h} we choose $\hat{x}_n(h)$ such that the expected value of the squared forecast errors

$$E\left((X_{n+h} - \hat{x}_n(h))^2\right)$$

is minimized. It can be shown that the 'best' forecast in the mean square sense is the **conditional expectation** of X_{n+h} based on the history of the process up to time n ($\mathcal{F}_n$):

$$\hat{x}_n(h) = E(X_{n+h}|\mathcal{F}_n) = E(X_{n+h}|X_n, X_{n-1}, \ldots) \tag{5}$$

Prediction based on ARMA models/ 2

The basic idea is that, for $h \geq 1$, the prediction in (5), $E(X_{n+h}|\mathcal{F}_n)$, is recursively evaluated in terms of

$$E(X_{n+h-1}|\mathcal{F}_n)$$

In particular, we use the fact that the 'best' forecast of all future values of the innovations $a_{n+1}, a_{n+2}, \dots$ is its expected value, which is 0.

Point forecasts $\hat{x}_n(h)$ can be computed directly from the ARMA model equation by replacing

i. future values of a_t by zero
ii. future values of x_t by their conditional expectation
iii. present and past values of x_t and a_t by their observed values (*known* at time $t = n$)

Table of Contents

Forecasting using an ARMA(1,1) model

Suppose an ARMA(1, 1) model of the form

$$X_t = \mu + \phi_1(X_{t-1} - \mu) + a_t + \theta_1 a_{t-1}$$

has been fitted to the data and the parameters' estimates are $\hat{\mu}, \hat{\phi}_1, \hat{\theta}_1$ have been found. At time $n+1$ we have

$$X_{n+1} = \mu + \phi_1(X_n - \mu) + a_{n+1} + \theta_1 a_n$$

The **one-step ahead forecast** of X_{n+1} is

$$E(X_{n+1}|\mathcal{F}_n) = \hat{x}_n(1) = \hat{\mu} + \hat{\phi}_1(x_n - \hat{\mu}) + \hat{\theta}_1 \hat{a}_n$$

since $E(a_{n+1}|\mathcal{F}_n) = 0$.

Forecasting using an ARMA(1,1) model / 2

For a **two-step forecast** we get

$$\begin{aligned}\hat{x}_n(2) =& E(\mu + \phi_1(X_{n+1} - \mu) + a_{n+2} + \theta_1 a_{n+1} | \mathcal{F}_n)\\ =& \hat{\mu} + \hat{\phi}_1(E(X_{n+1}|\mathcal{F}_n) - \hat{\mu}) + E(a_{n+2}|\mathcal{F}_n) + \hat{\theta}_1 E(a_{n+1}|\mathcal{F}_n)\\ =& \hat{\mu} + \hat{\phi}_1(\hat{x}_n(1) - \hat{\mu})\end{aligned}$$

where $\hat{x}_n(1)$ is the one-step ahead forecast. Hence, by replacing $\hat{x}_n(1) = \hat{\mu} + \hat{\phi}_1(x_n - \hat{\mu}) + \hat{\theta}_1\hat{a}_n$, we obtain

$$\hat{x}_n(2) = \hat{\mu} + \hat{\phi}_1^2(x_n - \hat{\mu}) + \hat{\phi}_1\hat{\theta}_1\hat{a}_n$$

- a_t is replaced by the model residual $\hat{a}_t$
- $\lim_{h\to\infty} E(X_{n+h}|\mathcal{F}_n) = \mu$, so that the prediction converges to the estimate of the unconditional mean

Forecasting using an ARMA(1,1) model / 3

In general, forecasts of two or more steps ahead are expressed as

$$\hat{x}_n(h) = \hat{\mu} + \hat{\phi}_1(\hat{x}_n(h-1) - \hat{\mu}) = \hat{\mu} + \hat{\phi}_1^h(x_n - \hat{\mu}) + \hat{\phi}_1^{h-1}\hat{\theta}_1\hat{a}_n$$

and the associated forecast error is

$$e_n(h) = x_{n+h} - \hat{x}_n(h)$$

For instance, for $h = 1$,

$$e_n(1) = x_{n+1} - \hat{x}_n(1) \approx a_{n+1}$$

and $V(e_n(1)) = \sigma_a^2$.

Forecasting using an AR(1) model

Consider the AR(1) model with mean μ

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

Suppose that we have data $x_1, \ldots, x_n$ and estimates $\hat{\mu}$ and $\hat{\phi}_1$. We know that

$$X_{n+1} = \mu + \phi_1(X_n - \mu) + a_{n+1}$$

1-Step Ahead Forecast. Given that the best predictor of a_{n+1} is its expected value, which is 0, we predict x_{n+1} by

$$\hat{x}_n(1) = \hat{\mu} + \hat{\phi}_1(x_n - \hat{\mu})$$

or

$$\hat{x}_n(1) = \hat{\phi}_0 + \hat{\phi}_1 x_n, \; \hat{\phi}_0 = \hat{\mu}(1 - \hat{\phi}_1)$$

Forecasting using an AR(1) model

The associated forecast error is $e_n(1) = x_{n+1} - \hat{x}_n(1)$, given by

$$\begin{aligned} e_n(1) &= x_{n+1} - \hat{x}_n(1) \\ &= (\mu + \phi_1(x_n - \mu) + a_{n+1}) - (\hat{\mu} + \hat{\phi}_1(x_n - \hat{\mu})) \\ &= (\mu - \hat{\mu}) + (\phi_1 - \hat{\phi}_1)x_n - (\phi_1\mu - \hat{\phi}_1\hat{\mu}) + a_{n+1} \\ &\approx a_{n+1} \qquad \text{(large-sample approximation)} \end{aligned}$$

Consequently, the variance of the 1-step ahead forecast error is

$$\text{Var}(e_n(1)) = \text{Var}(a_{n+1}) = \sigma_a^2$$

If a_t is normally distributed, then a 95% 1-step ahead interval forecast of x_{n+1} is

$$\hat{x}_n(1) \pm 1.96 \times \sqrt{\text{Var}(e_n(1))}$$

AR(1): Multistep Ahead Forecast

2-Step Ahead Forecast. We forecast x_{n+2} by

$$\begin{aligned}\hat{x}_n(2) &= \hat{\mu} + \hat{\phi}_1(\hat{x}_n(1) - \hat{\mu}) \\ &= \hat{\mu} + \hat{\phi}_1(\hat{\phi}_1(x_n - \hat{\mu})) = \hat{\mu} + \hat{\phi}_1^2(x_n - \hat{\mu})\end{aligned}$$

The general formula for the ***h*-step ahead forecast** from the AR(1) model is

$$\hat{x}_n(h) = \hat{\mu} + \hat{\phi}_1^h(x_n - \hat{\mu})$$

If $|\hat{\phi}_1| < 1$ (stationarity), then as h increases, the forecasts will approach fast its unconditional mean μ.

AR(p): Multistep Ahead Forecast

Forecasting $AR(p)$ processes is similar. For an $AR(2)$ process ($p = 2$)

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

we get the one-step ahead forecast

$$\hat{x}_n(1) = \hat{\mu} + \hat{\phi}_1(x_n - \hat{\mu}) + \hat{\phi}_2(x_{n-1} - \hat{\mu})$$

with associated forecast error

$$e_n(1) = x_{n+1} - \hat{x}_n(1) \approx a_{n+1}$$

and the two-step ahead forecast

$$\hat{x}_n(2) = \hat{\mu} + \hat{\phi}_1(\hat{x}_n(1) - \hat{\mu}) + \hat{\phi}_2(x_n - \hat{\mu})$$

AR(2) Forecast Errors

Using the large-sample approximation again, so $\hat{\mu}$ is replaced by μ and $\hat{\phi}$ by ϕ, the error in the two-step ahead forecast is then

$$e_n(2) = x_{n+2} - \hat{x}_n(2) = \phi_1(x_{n+1} - \hat{x}_n(1)) + a_{n+2} = \phi_1 a_{n+1} + a_{n+2}$$

The variance of the forecast error is

$$\begin{aligned} \text{Var}(e_n(2)) &= \text{Var}(\phi_1 a_{n+1} + a_{n+2}) \\ &= \phi_1^2 \text{Var}(a_{n+1}) + \text{Var}(a_{n+2}) \\ &= (1 + \phi_1^2)\sigma_a^2 \end{aligned}$$

Note that $\text{Var}(e_n(2)) \geq \text{Var}(e_n(1))$, meaning that as the forecast horizon increases the uncertainty in forecast also increases.

Forecasting from MA(q) Models

Consider the MA(1) process, $X_t = \mu + a_t + \theta_1 a_{t-1}$.
When $t = n$, the next observation will be $X_{n+1} = \mu + a_{n+1} + \theta_1 a_n$.

Using estimates $\hat{\mu}$ and $\hat{\theta}_1$, and replacing a_n by the residual $\hat{a}_n$, the **1-step-ahead forecast** is

$$\hat{x}_n(1) = \hat{\mu} + \hat{\theta}_1 \hat{a}_n$$

with error $e_n(1) = a_{n+1}$ and Var$(e_n(1)) = \sigma_a^2$.

The **2-step-ahead forecast** is

$$\hat{x}_n(2) = \hat{\mu}$$

with error $e_n(2) = a_{n+2} + \theta_1 a_{n+1}$ and Var$(e_n(2)) = (1 + \theta_1^2)\sigma_a^2$.

Similarly, it can be proved that

$$\hat{x}_n(h) = \hat{\mu}, \qquad h > 2$$

Forecasting from MA(q) Models/ 2

For an $MA(2)$ model

$$X_t = \mu + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2}$$

we have

$$X_{n+h} = \mu + a_{n+h} + \theta_1 a_{n+h-1} + \theta_2 a_{n+h-2}$$

from which we obtain predictions for $X_{n+h}, h \geq 1$:

$$\hat{x}_n(1) = \hat{\mu} + \hat{\theta}_1 \hat{a}_n + \hat{\theta}_2 \hat{a}_{n-1}$$
$$\hat{x}_n(2) = \hat{\mu} + \hat{\theta}_2 \hat{a}_n$$
$$\hat{x}_n(h) = \hat{\mu} \qquad h > 2$$

In general, for an $MA(q)$ model, multistep ahead forecasts go to the mean after the first q steps.

Prediction Intervals

When making forecasts, the interest is in the uncertainty of the predictions. To this end, we compute the variance of the forecast error $\text{Var}(e_n(h))$. Then, assuming that $a_1, a_2, \ldots$ is Gaussian white noise, a $(1-\alpha)100\%$ prediction interval for x_{n+1} is

$$\hat{x}_n(h) \pm z_{1-\frac{\alpha}{2}} \sqrt{\text{Var}(e_n(h))}$$

where

- $z_{1-\alpha/2}$ is the $(1-\alpha/2)$-quantile of the standard normal distribution
- $\sqrt{\text{Var}(e_n(h))}$ is the forecast error's standard deviation

Note that $\text{Var}(e_n(h)) \to \text{Var}(X_n)$, the variance of the process, as $h \to \infty$

Forecasting using ARIMA models

Suppose that X_t is ARIMA(1,1,0), so that the model for Y_t where

$$Y_t = \nabla X_t = X_t - X_{t-1}$$

is AR(1). To forecast X_{n+k}, $k \geq 1$

- fit an AR(1) model to the Y_t process and forecast Y_{n+k}, $k = 1, 2, \ldots$; denote the h-step ahead forecast by $\hat{y}_n(h)$; for instance, the 1-step-ahead forecast is $\hat{y}_n(1)$
- For $h = 1$, since

$$X_{n+1} = X_n + (X_{n+1} - X_n) = X_n + Y_{n+1}$$

the 1-step ahead point forecast of X_{n+1} is

$$\hat{x}_n(1) = x_n + \hat{y}_n(1)$$

Table of Contents

Problem 1

An AR(3) model has been fit to a time series. The estimates are $\hat{\mu} = 102, \hat{\phi}_1 = 0.5, \hat{\phi}_2 = 0.2, \hat{\phi}_3 = 0.1$. The last four observations were $x_{n-3} = 104$, $x_{n-2} = 101$, $x_{n-1} = 102$, $x_n = 99$.
Forecast x_{n+1} and x_{n+2} using these data and estimates.

[*S*ol: $\hat{x}_n(1) = 100.4$; $\hat{x}_n(2) = 100.6$]

Problem 2

The MA(2) model $X_t = \mu + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2}$ was fit to data and the estimates are

Parameter	Estimate
μ	45
θ_1	0.3
θ_2	-0.15

The last two values of the observed time series and residuals are

t	x_t	$\hat{a}_t$
$n-1$	39.8	-4.3
n	42.7	1.5

Find the forecasts of x_{n+1} and x_{n+2}. What is $\hat{x}_n(h), h \geq 3$?

[Sol: $\hat{x}_n(1) = 46.1$; $\hat{x}_n(2) = 44.78$; $\hat{x}_n(h) = \hat{\mu}, h \geq 3$.]

Problem 3

The following ARMA model has been fit to a time series:

$$x_t = 25 + 0.8x_{t-1} - 0.3x_{t-2} + a_t$$

where $\{a_t\}$ is white noise.

a. Suppose that we are at the end of time period $T = 100$ and we know that $x_{100} = 40$ and $x_{99} = 38$. Determine forecasts for periods 101, 102 from this model at origin 100.

b. Suppose that the observation for time period 101 turns out to be $x_{101} = 35$. Revise your forecasts for period 102 using period 101 as the new origin of time.

[*Sol*: (a) $\hat{x}_{100}(1) = 45.6$; $\hat{x}_{100}(2) = 49.48$ (b) $\hat{x}_{101}(1) = 41$]