Eco 4306 Economic and Business Forecasting

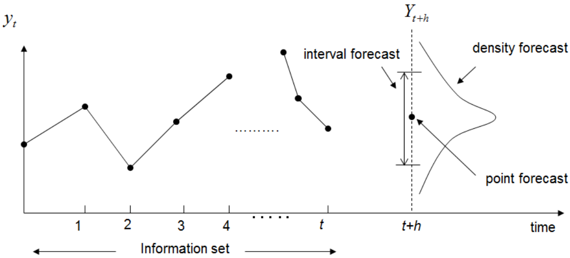
Lecture 6
Chapter 4: Tools of the Forecaster

Introduction

- before constructing a forecast based on a time series model, forecaster needs to decide about three basic elements that guide the production of the forecast
  1. **Information set**
  2. **Forecast horizon**
  3. **Loss function**
- information set will be used to construct conditional density function to be able to evaluate expectations, and the optimal forecast will minimize the expected loss

Introduction

Forecasting Problem



Introduction

- example: to forecast the number of new homes built, we need to

1. construct the **information set**
    - gather relevant up-to-date information for the problem at hand - existing number of houses, state of the local economy, population inflows, . . .
    - this information is used to estimate the time series model and construct the forecast
2. choose **forecast horizon**: how far into the future to forecast
    - 1-month-ahead, 1-quarter-ahead, 1-year-ahead, 10-years-ahead, . . .
    - this depends on the use of the forecast
        - e.g. a policy makers who plans to design or revamp the transportation services of the area or any other infrastructure is likely to be more interested in long-term predictions of new housing (1 year, 2 years, 5 years) than in short-term predictions (1 month, 1 quarter)
    - forecast horizon influences the choice of the frequency of the time series data
        - e.g. if our interest is a 1-month-ahead prediction, we may wish to collect monthly data, or if our interest is a 1-day-ahead forecast, we may collect daily data
3. decide which **loss function** best represents the costs associated with forecast errors
    - forecast errors will happen and more importantly they will be costly
    - costs of underestimation and of overestimation may be of different magnitude
    - we will choose a forecast that minimizes the expected loss

4.1 Information Set

- a **univariate information set** is the historical time series of the process up to time t

$$I_t = \{y_0, y_1, y_2, \ldots, y_t\}$$

- a **multivariate information set** is the collection of several historical time series

$$I_t = \{y_0, y_1, y_2, \ldots, y_t, x_0, x_1, x_2, \ldots, x_t, z_0, z_1, z_2, \ldots, z_t\}$$

- for example, to produce a 1-year-ahead forecast for new houses built
  - univariate information set is the time series of new houses built in previous years
  - multivariate information set may in addition contain the time series for inflows of population, unemployment in the area, ...

4.1 Information Set

- forecast $f_{t,h}$ is constructed as a function of the information set

$$f_{t,h} = g(I_t)$$

  function $g(\cdot)$ represents the time series model that processes the known information up to time t and from which we produce the forecast of the variable of interest at a future date $t+h$

4.1 Information Set

- some examples of 1-step-ahead forecasts of a process $\{Y_t\}$

$$
\begin{aligned}
\text{(i)} \quad & f_{t,1} = 0.8y_t \\
\text{(ii)} \quad & f_{t,1} = 0.2y_t - 0.9y_{t-1} \\
\text{(iii)} \quad & f_{t,1} = \frac{4}{1 + 0.5y_t} \\
\text{(iv)} \quad & f_{t,1} = 1.8y_t - 0.5y_{t-1} + 0.4x_t + 0.3x_{t-1} + 0.6x_{t-2}
\end{aligned}
$$

- in (i), (ii) and (iii) the information set is univariate, in (iv) it is multivariate

4.1 Information Set

- predictability of a time series depends on how useful the information set is
- sometimes univariate information sets are not very helpful, and we need to resort to multivariate information sets
- for example, stock returns are very difficult to predict on the basis of past stock returns alone, but when we add other information such as firm size, price-earnings ratio, cash flows, and so on, we find some predictability
- some time series (e.g. stock returns, interest rates, exchange rates, ...) are inherently very difficult to predict due to
  - lack of understanding of the phenomenon
  - lack of statistical methods
  - high uncertainty making it difficult to separate information from noise

4.2 Forecast Horizon

- we distinguish between a **short-term** forecast and a **long-term** forecast
- in economics up to a 1-year-ahead prediction is a short-term forecast, forecasts between 1 and 10 years are considered short/medium term or medium/long term, and a 10-year-ahead and longer prediction is a long-term forecast
- short-, medium-, and long-term forecast are functions of the frequency of the data and of the properties of the model
- we distinguish between **1-step ahead forecast** $f_{t,1}$ and **multistep forecast** $f_{t,h}$ for $h > 1$

4.2.1 Forecasting Environments

- suppose that we have a time series with T observations, $\{y_1, y_2, \ldots, y_T\}$
- we divide the sample into two parts: **estimation sample** and **prediction sample**
- estimate the model using observations in estimation sample, with $t < T$ observations, $\{y_1, y_2, \ldots, y_t\}$
- we then assess the performance of models in-sample and out-of-sample
- **in-sample assessment** - evaluate goodness of the model (perform specification tests) using observations from 1 to t
- **out-of-sample assessment** - evaluate the forecasting ability of the model using observations from $t + 1$ to T
  - e.g. if we are interested in evaluating accuracy of 1-step-ahead forecasts we first produce a sequence of out-of-sample 1-step-ahead forecasts $f_{t+j,1}$ where $j = 0, 1, \ldots T - t - 1$ for $\{Y_{t+1}, Y_{t+2}, \ldots, Y_T\}$
  - we next compute a sequence of 1-step-ahead forecast errors $e_{t+j,1} = y_{t+j+1} - f_{t+j,1}$ for $j = 0, 1, \ldots, T - t - 1$
  - finally, we assess the accuracy of the forecast by plugging the forecast errors into the loss function and calculating the average or the maximum loss

4.2.1 Forecasting Environments

- three forecasting schemes: **recursive**, **rolling**, and **fixed**

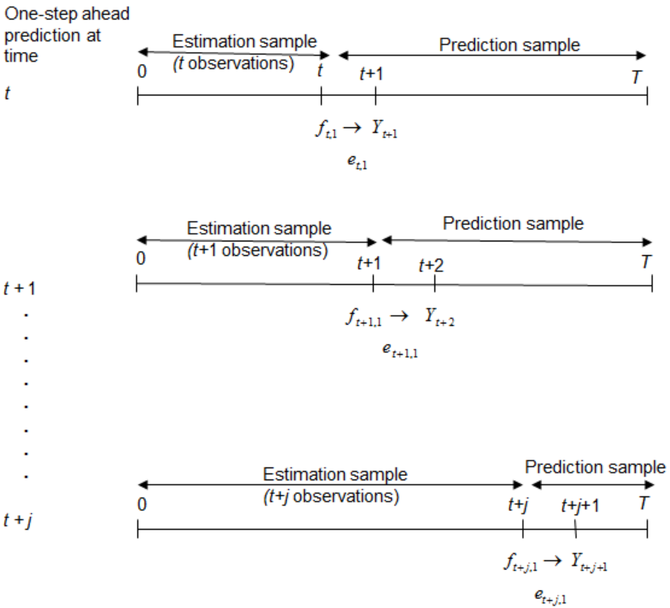
4.2.1 Forecasting Environments

recursive forecasting scheme

- repeatedly increase estimation sample by one observation, reestimate the model with extra observation, and compute a 1-step ahead forecast
- estimation sample keeps expanding until the prediction sample is exhausted
- this yields a sequence of 1-step-ahead forecasting errors $\{e_{t,1}, e_{t+1,1}, \dots e_{T-1,1}\}$
- forecast errors are then used to calculate measures of forecast accuracy based on the loss function

4.2.1 Forecasting Environments

recursive forecasting scheme



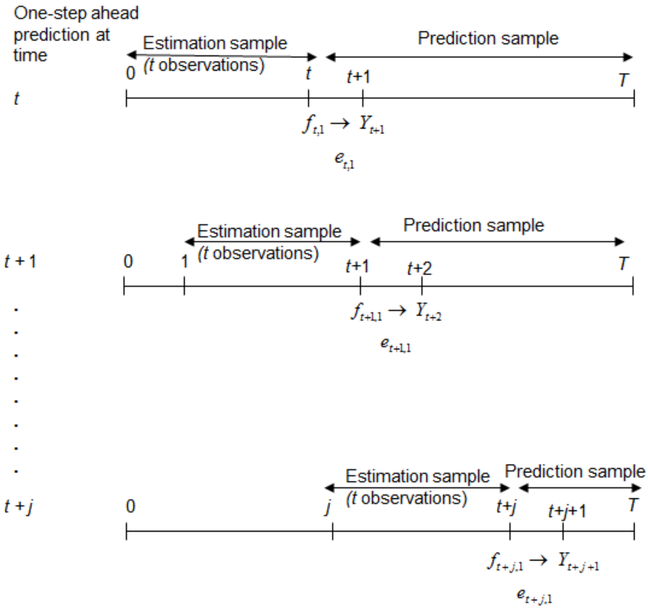
4.2.1 Forecasting Environments

rolling forecasting scheme

- similar to recursive scheme but estimation sample always contains the same number of observations
- thus at t it contains observations 1 to t, at $t+1$ observations 2 to $t+1$, at time $t+2$ observations 3 to $t+2$, ...
- model is reestimated for each rolling sample, and 1-step-ahead forecast is produced
- estimation sample is rolling until the prediction sample is exhausted
- this yields collection of 1-step-ahead forecasting errors $\{e_{t,1}, e_{t+1,1}, \ldots, e_{T-1,1}\}$
- forecast errors are then used to calculate measures of forecast accuracy based on the loss function

4.2.1 Forecasting Environments

rolling forecasting scheme



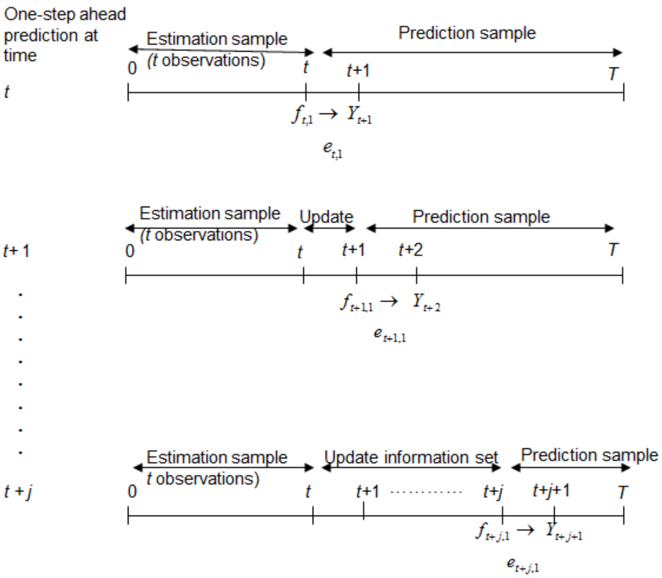
4.2.1 Forecasting Environments

fixed forecasting scheme

- model is estimated only once using the estimation sample that contains the first t observations
- information set is updated but model is not reestimated - each one step ahead forecast is thus constructed using same parameters
- for instance, at time $t+1$, information set contains one more observation, which will contribute to the construction of the 1-step-ahead forecast but will not be used to reestimate model parameters
- information set is updated until the prediction sample is exhausted
- this again yields collection of 1-step-ahead forecasting errors $\{e_{t,1}, e_{t+1,1}, \ldots, e_{T-1,1}\}$
- forecast errors are then used to calculate measures of forecast accuracy based on the loss function

4.2.1 Forecasting Environments

fixed forecasting scheme



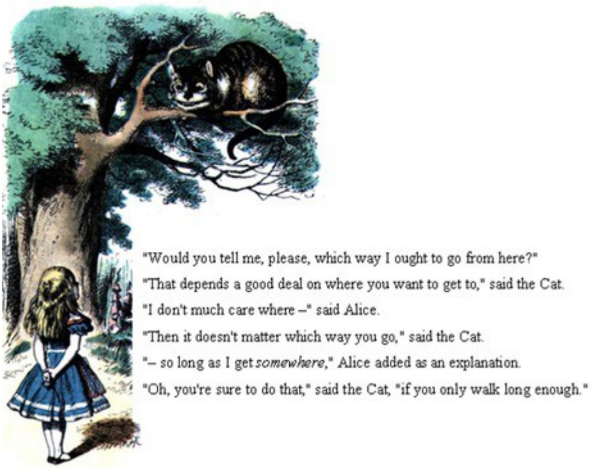
4.2.1 Forecasting Environments

advantages and disadvantages of the three schemes

- recursive scheme
  - incorporates as much information as possible in the estimation of the model
  - advantageous if the model is stable over time
  - if the data have structural breaks, model's stability is in jeopardy and so is the forecast
- rolling scheme
  - avoids the potential problem with the model's stability
  - more robust against structural breaks in the data
  - does not make use of all the data
- fixed scheme
  - fast and convenient because - there is one and only one estimation
  - does not allow for parameter updating, so again problem with structural breaks and model's stability

4.3 Loss function

what the best forecast is depends on the purpose of the forecast, its intended use



4.3 Loss function

- example: suppose you live in Riverside, CA about 90 miles east of Los Angeles
- you are departing on a business trip from Los Angeles International Airport (LAX) to meet with a client in New York
- you need to forecast how many hours it takes to get from Riverside to LAX
- information set I_t will contain the distance between Riverside and LAX, rush hours in the area highways, construction work in the area, time needed for check-in at LAX, time needed for security check at LAX
- suppose the actual time could be either 5 hours or 3 hours with equal probability
- suppose your forecast is the average time needed $f_{t,1} = E(Y_{t+1}|I_t) = 4$ hours

$$f_{t,1} = 4 \qquad y_{t+1} = \begin{cases} 3 \\ 5 \end{cases} \quad \Rightarrow \quad e_{t,1} = y_{t+1} - f_{t,1} = \begin{cases} 1 \\ -1 \end{cases}$$

4.3 Loss function

- suppose that it takes 5 hours to get to LAX and so you miss your flight
- the forecast error is $e_{t,1} = 1$ and the potential costs associated with it are
  - need to wait at the airport to hope to be able to get on the next flight
  - alternatively, purchase another ticket with a different airline
  - need to spend extra money on food, hotel
  - stressed and/or in bad mood for the rest of the day
  - professional reputation might be damaged if you miss the meeting with your client
  - prospective business deal might be lost

- suppose that it takes 3 hours to get to LAX and you thus and an hour spare at LAX
- the forecast error is $e_{t,1} = -1$ and the potential costs associated with it are
  - having to wait in a noisy environment, uncomfortable chairs, crowded space, ...

4.3 Loss function

- note that positive and negative errors are of same magnitude, but costs are not
- your loss function is thus asymmetric
- taking into account your loss function, you decide that it makes sense for you to change your forecast and instead of average time $f_{t,1} = 4$ choose the maximum time thus $f_{t,1} = 5$ hours
- **as this example illustrates, the forecast will depend on the loss function that the forecaster is facing**
- the forecaster thus must know the loss function *before* making the forecast
- note also that in the example if you are avoiding positive forecast errors and always arrive at airport too early, the average forecast errors will be negative, not zero
- **it is rational to consistently make biased forecasts if loss function is asymmetric**

4.3 Loss function

- **loss function** $L(e_{t,h})$ is the evaluation of costs associated with the forecast error
- three properties that loss functions need to satisfy

1. if the forecast error is zero, the loss is zero:
   $L(e_{t,h}) = 0$ when $e_{t,h} = 0$
2. loss function is a non-negative function with minimum value equal to zero:
   $L(e_{t,h}) \geq 0$ for all $e_{t,h}$
3. for positive errors the loss is monotonically increasing, for negative errors it is monotonically decreasing:

$$\begin{array}{llll} \text{if} & e_{t,h}^{(1)} > e_{t,h}^{(2)} > 0 & \text{then} & L(e_{t,h}^{(1)}) > L(e_{t,h}^{(2)}) \\ \text{if} & e_{t,h}^{(1)} < e_{t,h}^{(2)} < 0 & \text{then} & L(e_{t,h}^{(1)}) > L(e_{t,h}^{(2)}) \end{array}$$

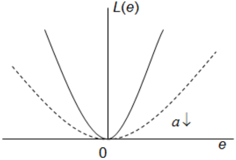
4.3.1 Some Examples of Loss Functions

Symmetric Loss Functions

- sign of the forecast errors is irrelevant, positive or negative errors of the same magnitude have identical costs

Quadratic loss function

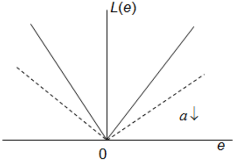
$$L(e) = ae^2, \quad a > 0$$



$$L(e) = L(-e)$$

Absolute value loss function

$$L(e) = a|e|, \quad a > 0$$



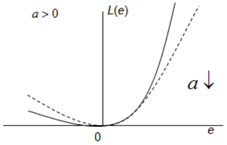
$$L(e) = L(-e)$$

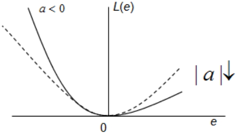
4.3.1 Some Examples of Loss Functions

Asymmetric Loss Functions

Linex function

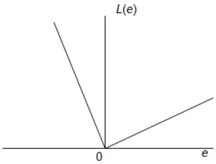
$$L(e) = \exp(ae) - ae - 1, \quad a \neq 0$$





Lin-lin function

$$L(e) = \begin{cases} a\,|e| & e > 0 \\ b\,|e| & e \leq 0 \end{cases}$$



$b > a \rightarrow L(-e) > L(e)$

4.3.1 Some Examples of Loss Functions

- quadratic loss function is the most prevalent in practice - it is mathematically tractable
- most of the time, however economic agents have asymmetric loss functions
  - example with trip to LAX airport - for most people it is less costly to wait at the airport than to miss a flight
  - government planning spending and forecasting tax revenues - deficit and surplus of the same size are not viewed the same by most politicians
  - Fed policymakers deciding about interest rate, facing inflation vs unemployment tradeoff - monetary hawks and inflation doves
  - investment fund managers making predictions of asset returns in their portfolio - underperforming by 5% vs overperforming 5%
  - financial intermediaries are requited to make capital provisions as a preventive measure against insolvency caused by loan defaults

4.3.3 Optimal Forecast: An Introduction

- we now put all three components together - information set I_t, forecast horizon h, and loss function $L(e_{t,h})$
- recall: $e_{t,h} = y_{t+h} - f_{t,h}$ and y_{t+h} is future value unknown at time t, of random variable Y_{t+h}, which has a conditional probability density function $f(y_{t+h}|I_t)$
- because the loss function depends on a random variable, it is also a random variable, thus we can write the expected loss as

$$E(L(y_{t+h} - f_{t,h})) = \int L(y_{t+h} - f_{t,h}) f(y_{t+h}|I_t) dy_{t,h}$$

- the optimal forecast is $f_{t,h}$ which minimizes the above expected loss

$$\min_{f_{t,h}} E(L(y_{t+h} - f_{t,h}))$$

4.3.3 Optimal Forecast: An Introduction

- if the loss function is quadratic, the optimal forecast that is minimizing the expected loss is

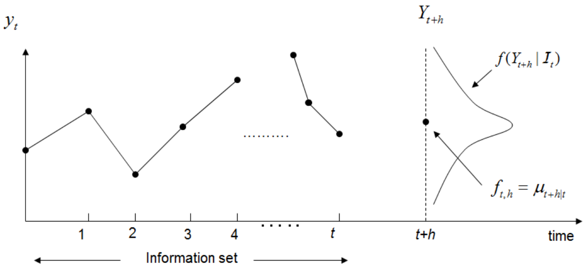
$$f^*_{t,h} = \mu_{t+h|t} = E(y_{t+h}|I_t) = \int y_{t+h} f(y_{t+h}|I_t) dy_{t,h}$$

- we will discuss the optimal forecast under various symmetric and asymmetric loss function in more detail when we get to Chapter 9

4.3.3 Optimal Forecast: An Introduction

Symmetric Loss Functions - Quadratic

$$L(e) = ae^2, \quad a > 0$$



4.3.3 Optimal Forecast: An Introduction

Asymmetric Loss Functions - Linex

$$L(e) = \exp(ae) - ae - 1, \quad a < 0$$