Eco 4306 Economic and Business Forecasting

Lecture 10
Chapter 8: Forecasting Practice I

Motivation

- we learned characteristics of moving average (MA) and autoregressive (AR) processes
- in theory, AC and PAC can serve as basic tool to choose between an MA or an AR process and determine their order
- in practice, there are many time series for which the selection of an AR or an MA process is not straightforward
- the choice among models is not that obvious when we face real time series, forecaster needs to make judgment calls which model(s) to select
- we will introduce new tools, used to evaluate different models and to select or narrow the set of models

Motivation

- AR and MA process can be combined to give rise to a mixed model that we call **autoregressive moving average**, ARMA(p, q)

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q}$$

- the simplest possible is the ARMA(1,1) model

$$Y_t = c + \phi_1 Y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

- AC and PAC functions will display decay toward zero, but there is no clear cutoff to zero at any lag for either of them

Outline

real world application: forecasting San Diego Metropolitan Statistical Area (MSA) house price index

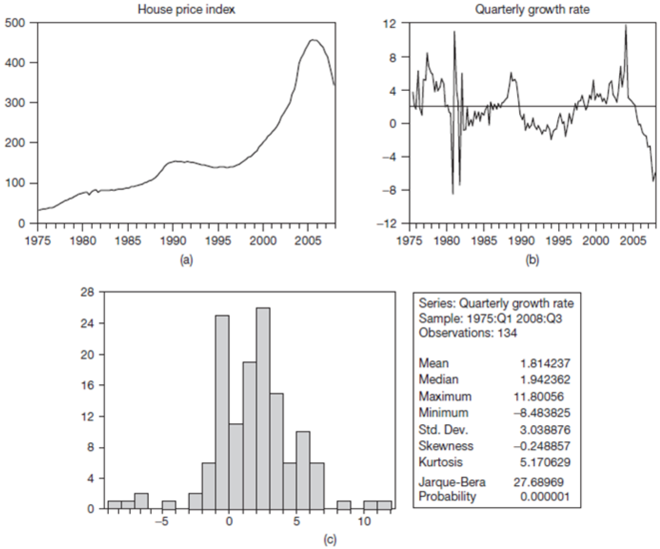
1. Data: source, definition, descriptive statistics, and autocorrelations
2. Model: identification, estimation, evaluation, and selection
3. Forecast: selection of loss function and construction of the forecast

8.1 Data

- house prices data can be obtained from Freddie Mac http://www.freddiemac.com, and from Federal Housing Finance Agency (FHFA) http://www.fhfa.gov
- Freddie Mac database is also available on Quandl https://www.quandl.com/data/FMAC
- we will use quarterly house price index for San Diego MSA from 1975Q1 to 2008Q3: **Figure08_1_SDhouseprices.xls**

8.1 Data

- index has overall upward tendency, seems to come from a nonstationary process
- we will thus model quarterly growth rate of the index instead



8.1 Data

Sample: 1975:Q1 2008:Q4
Included observations: 134

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.487	0.487	32.524	0.000
		2	0.486	0.326	65.135	0.000
		3	0.401	0.121	87.502	0.000
		4	0.464	0.223	117.67	0.000
		5	0.257	-0.140	127.02	0.000
		6	0.276	0.000	137.85	0.000
		7	0.264	0.075	147.86	0.000
		8	0.184	-0.092	152.77	0.000
		9	0.115	-0.040	154.69	0.000
		10	0.049	-0.114	155.04	0.000
		11	0.011	-0.090	155.06	0.000
		12	-0.064	-0.061	155.67	0.000
		13	-0.073	-0.025	156.48	0.000
		14	-0.123	-0.041	158.77	0.000
		15	-0.156	-0.055	162.48	0.000

8.1 Data

- ACF and PACF show large autocorrelation coefficients for several lags - time series has much dependence
- large Q-statistics and p-values practically zero for all lags so we reject H_0 of no autocorrelation
- important to remember that sample AC and PAC functions are estimated functions, subject to sampling error
- this should be taken into account especially when the sample size is not very large - sample ACF and PACF can look like different from their theoretical counterparts

8.2 Model Selection

identification of possible models to be estimated

- option 1: AR model
  - decay toward zero in ACF, limited number non-zero elements in PACF
  - possible candidates are AR(2), AR(4), AR(5)
- Option 2: MA model
  - decay toward zero in PACF, limited number non-zero elements in ACF
  - possible candidates MA(4) or MA(7)
- Option 3: ARMA model
  - decay toward zero in both ACF and PACF, with no clear cutoff in ACF or PACF
  - possible candidates ARMA(2,2) or ARMA(2,4), since first two spikes in the PACF appear most dominant, and the remaining dependence is left to be picked up by either MA(2) or MA(4) component

8.2 Model Selection

identification of possible models to be estimated

- we thus consider six alternative models:

model 1

MA(4) $\quad Y_t = \mu + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \theta_3\varepsilon_{t-3} + \theta_4\varepsilon_{t-4}$

model 2

AR(3) $\quad Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + \varepsilon_t$

model 3

AR(4) $\quad Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + \phi_4 Y_{t-4} + \varepsilon_t$

model 4

AR(5) $\quad Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + \phi_4 Y_{t-4} + \phi_5 Y_{t-5} + \varepsilon_t$

model 5

ARMA(2,2) $\quad Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2}$

model 6

ARMA(2,4) $\quad Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \theta_3\varepsilon_{t-3} + \theta_4\varepsilon_{t-4}$

Model Estimation

- choose **Object** → **New Object** → **Equation** and enter the model specification

model 1	MA(4)	sdg c ma(1) ma(2) ma(3) ma(4)
model 2	AR(3)	sdg c ar(1) ar(2) ar(3)
model 3	AR(4)	sdg c ar(1) ar(2) ar(3) ar(4)
model 4	AR(5)	sdg c ar(1) ar(2) ar(3) ar(4) ar(5)
model 5	ARMA(2,2)	sdg c ar(1) ar(2) ma(1) ma(2)
model 6	ARMA(2,4)	sdg c ar(1) ar(2) ma(1) ma(2) ma(3) ma(4)

Model Evaluation

we next compare the candidate models along several criteria and check whether

- model implies stationarity and invertibility
- residuals are white noise
- parameters of the model are statistically significant
- information criteria

Model Evaluation

- for MA models we need to check invertibility - inverted MA roots should lie inside the unit circle
- for AR models we need to check stationarity - inverted AR roots should lie inside the unit circle
- open the equation object and choose **View** $\rightarrow$ **ARMA Structure** $\rightarrow$ **Roots**

Model Evaluation

- if the model is well specified, residuals should not exhibit any linear dependence and should look like white noise
- recall: for single hypothesis $H_0 : \rho_j = 0$ we can check the 95% confidence interval in ACF plot, if the spike at lag j is outside the dashed lines we reject the null hypothesis at 5% level
- recall: Q-statistic is used test joint hypothesis $H_0 : \rho_1 = \rho_2 = \ldots = \rho_k = 0$, rejecting this hypothesis means that the residuals are not white noise, since there is a $j \leq k$ such that $\rho_j \neq 0$
- open the equation object and choose
  - **Resids** or alternatively **View** $\rightarrow$ **Actual, Fitted, Residual** $\rightarrow$ **Residual Graph**
  - **View** $\rightarrow$ **Residual Diagnostics** $\rightarrow$ **Correlogram - Q-Statistics**

Model Evaluation

- **Akaike information criteria** (AIC) and **Schwarz information criteria** (SIC)
- main idea behind AIC and SIC similar to the adjusted R^2
- objective is to find a model that can explain observed data and at the same time uses is parsimonious enough (with small number of parameters)
- AIC and SIC include a penalty term to capture the trade-off between a large number of parameters and a potential reduction of the residual variance

$$AIC = \log \frac{SSR}{T} + \frac{2m}{T}$$
$$SIC = \log \frac{SSR}{T} + \frac{m \log T}{T}$$

  where SSR is the sum of squared residuals, m is the number of estimated parameters, T is the sample size
- penalty terms $2m/T$ and $(m \log T/T$ increase whenever with number of estimated parameters
- SIC penalizes more heavily than the AIC because $2 < \log T$, SIC thus tends to select more parsimonious models than AIC
- preferred model is found by minimizing AIC or SIC

Model Forecast

house price forecasts useful for property owners, real estate investors, government, and mortgage banks

- property owners: substantial proportion of households' wealth in U.S. is the value of homes, decisions to buy/sell depend on current and future prices
- investors: more likely to invest in housing when they expect capital gains (higher prices in the future)
- government: policy makers may be concerned with the effect of a tighter monetary policy (higher interest rates) on housing prices
- mortgage banks: likelihood of default by borrowers increases when house prices go down

even if data and model used are the same, different agents may have different forecasts because agents may have different loss functions

asymmetric loss function is more likely than a symmetric loss function going to capture the trade-offs for most agents involved