Univariate Time Series Modelling and Forecasting

Noé Debrois

June 16, 2024

1 Introduction

Univariate time series modelling aims to predict returns using only past values.

2 Notation and Concepts

- **Strictly Stationary Process**: The probability measure for the sequence $\{y_t\}$ is the same as that for $\{y_{t+m}\} \forall m$.
- **Weakly Stationary Process**: A series is weakly stationary if:
  1. $E(y_t) = \mu, \forall t$
  2. $\mathrm{Var}(y_t) = \sigma^2 < \infty$
  3. $\mathrm{Cov}(y_{t_1}, y_{t_2}) = \gamma_{t_2 - t_1}, \forall t_1, t_2$

3 Univariate Time Series Models

3.1 Covariance and Autocorrelation

- For a covariance stationary process, all variances are the same and covariances depend on the difference $t_1 - t_2$.
- **Autocovariances** γ_s. Covariances normalized by variance: since the value of autocovariances depend on the units of measurement of y_s.
- **Autocorrelation Function (ACF)**: Plotting autocorrelations $\tau_s = \frac{\gamma_s}{\gamma_0}$ against s yields the correlogram.

3.2 White Noise Process : (virtually) no discernible structure.

- Defined by $E(y_t) = 0$ and $\mathrm{Var}(y_t) = \sigma^2$ for all t, & $\gamma_{t-r} = \sigma^2$ if $t = r$, 0 otherwise.
- The ACF will be zero except for a peak of 1 at $s = 0$.
- Confidence intervals for significance testing can be constructed, e.g., for 95% CI: $\pm 1.96/\sqrt{T}$.

3.3 Joint Hypothesis Tests

T: sample size.

m: maximum lag length.

- **Box-Pierce Q-statistic**: $Q = T\sum_{k=1}^{m} \tau_k^2$, asymptotically χ^2_m distributed.

Test the joint hypothesis that all m of the τ_k corr. coeffs. are simultaneously $= 0$.

- **Ljung-Box statistic**: $\widetilde{Q} = T(T+2)\sum_{k=1}^{m} \frac{\tau_k^2}{T-k}$, used as a general test of linear dependence in time series. $\sim \chi^2_m$.

3.4 Moving Average (MA) Processes

- MA(q) model: $y_t = \mu + u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + ... + \theta_q u_{t-q}$, where u_t are iid with $E(u_t) = 0$ and $\text{Var}(u_t) = \sigma^2$.

$\begin{cases} \mathbb{E}[y_t] = \mu. \\ \text{Var}(y_t) = \gamma_0 = (1+\theta_1^2+\dots+\theta_q^2)\sigma^2. \end{cases} \longrightarrow \gamma_s = \begin{cases} (\theta_s + \theta_{s+1}\theta_1 + \dots + \theta_q\theta_{q-s})\sigma^2 & \text{for } s \in [1,q] \\ 0 & \text{for } s > q \end{cases}$

- Example: For $y_t = \theta_1 u_{t-1} + \theta_2 u_{t-2} + u_t$, calculate mean, variance, and ACF.

MA(2)

ACF: $\tau_1 = \frac{\gamma_1}{\gamma_0} = \frac{\text{Cov}(y_t, y_{t+1})}{\gamma_0} = \frac{\mathbb{E}[(y_t - 0)(y_{t+1} - 0)]}{\gamma_0} = \frac{\mathbb{E}[y_t y_{t+1}]}{\gamma_0}$.

3.5 Autoregressive (AR) Processes

- AR(p) model: $y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + u_t = \mu + \sum_{i=1}^{p} \phi_i y_{t-i} + u_t = \mu + u_t + \sum_{i=1}^{p} \phi_i L^i y_t$
- Stationarity condition: Roots of $\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - ... - \phi_p L^p$ lie outside the unit circle.

$\boxed{\phi(L) y_t = \mu + u_t}$

ACF decays exponentially to zero.

- Mean and ACF derived using Yule-Walker equations.

$\mathbb{E}[y_t] = \frac{\mu}{1 - \phi_1 - \dots - \phi_p}$

Yule Walker eqs:

$\begin{cases} \tau_1 = \phi_1 + \tau_1\phi_2 + \dots + \tau_{p-1}\phi_p \\ \tau_2 = \tau_1\phi_1 + \phi_2 + \dots + \tau_{p-2}\phi_p \\ \vdots \\ \tau_p = \tau_{p-1}\phi_1 + \tau_{p-2}\phi_2 + \dots + \phi_p \end{cases}$

- Wold's Decomposition Thm: any stationary series can be decomposed into the Σ of 2 unrelated processes: a purely det. part & a purely stoch. part, which will be MA(∞).

Example: for an AR(p) model with $\mu = 0$: $\phi(L) y_t = u_t$. Wold decomposition is: $y_t = \psi(L) u_t$ where: $\psi(L) = (1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p)^{-1}$.

3.6 Partial Autocorrelation Function (PACF)

τ_{kk}: corr. between y_t & y_{t-k} after removing the effects of $y_{t-k+1}, \dots, y_{t-1}$.

- PACF measures correlation between observations k periods apart, after controlling for intermediate lags (i.e. all lags $< k$). and the current observation

acf = pacf @ lag 1 always.

- For AR(p), PACF will be zero after lag p. For MA(q), PACF will decline geometrically. → So PACF is useful for telling the difference between AR and MA processes.

3.7 ARMA Processes

- ARMA(p,q) model: Combines AR(p) and MA(q) processes: $y_t = \mu + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j u_{t-j} + u_t$. $\mathbb{E}[y_t] = \frac{\mu}{1 - \phi_1 - \dots - \phi_p}$.
- Stationarity and invertibility conditions similar to AR and MA processes.

3.8 Summary of ACF and PACF Behaviour

- AR process: Geometrically decaying ACF, number of PACF spikes equals AR order.
- MA process: Number of ACF spikes equals MA order, geometrically decaying PACF.

4 Examples and Exercises

- Exercises on calculating mean, variance, and ACF for given AR and MA processes.
- Example ACF and PACF plots for standard processes.

- Likelihood fct: of a sample $y_1, \dots, y_T$ w/ joint pdf $f(y|\theta)$: $L(\theta|y) = f(y|\theta)$.
  [see all about MLE]

- Information criteria for model selection: AIC vs BIC vs HQIC.
  $AIC = 2k - 2\ln(L)$; $\underbrace{BIC = k\ln(T) - 2\ln(L)}_{\text{the best according to M.A. (STRONGLY CONSISTENT)}}$; $HQIC = 2k\ln(\ln(T)) - 2\ln(L)$
  L: likelihood of the model.
  k = sample size.

  SELECT MODEL W/ LOWER IC (parsimonious) → VAR of estimator $\sim \frac{1}{\text{nb d.o.f}}$.
  Don't choose with acf!

Box & Jenkin method to estimate ARMA:

1. Identification: select model w/ lowest IC.
2. Estimation of params.: least squares OR MLE.
3. Model checking: residual diagnostics.

- ARIMA: → $(D_t)_t = (I_t - I_{t-1})_t$ & build ARMA on $(D_t)_t$.
  "integrated" → ARMA(p,q) in the variable diff. d times.
  ⇕
  ARIMA(p,d,q) on the original data.

- ARMAX: when you have a target time series & r exogenous variables $x_{t,k}$:
$$y_t = \mu + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j u_{t-j} + \underline{\sum_{k=1}^{r} \beta_k x_{t,k}} + u_t$$
MLE

→ see the article about election cycle [3] in american stock returns.

- Trend & seasonality:
  - DETERMINISTIC TREND: $y_t = z_t + \beta \times t$ (z_t: stat. process) — Detrend: fit a linear regression to find $\hat{\beta}$ and $y(t) - \hat{\beta}$ is the detrended time-series.

If non linear: $y_t = z_t + \beta_1 t^2$ OR $y_t = z_t + \beta_2 \sqrt{t}$ — find the right regression

⚠ Consider simple functions. ⚠ ALWAYS DETREND before starting your econometrics analysis: bc. a trendprocess is NON-STATIONARY.

- SEASONALITY: periodic behavior w/ known period.

Univariate Time series estimation: 1) Detrend & Deseasonalize w/ linear reg. → Plot ACF & PACF & fit ARMA/X w/ ≠ lags to find best model w/ IC. → Make the residuals diagnostics: ARMAX residuals should satisfy 4 CLRM.

- Forecasting in econometrics:
  - with ARMA: $y_{t+s} = f_{t,s} = \mu + \sum_{i=1}^{p} \phi_i f_{t,s-i} + \sum_{j=1}^{q} \theta_j u_{t+s-j}$ — General formula.
  - with MA: let $y_t = \mu + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \theta_3 u_{t-3} + u_t$ MA(3)

$$\downarrow$$

$$y_{t+1} = f_{t,1} \quad y_{t+2} = f_{t,2} \quad \cdots \quad y_{t+s} = f_{t,s}.$$

$$f_{t,1} = \mathbb{E}[y_{t+1} | t] = \mathbb{E}_t[\mu + \theta_1 u_t + \theta_2 u_{t-1} + \theta_3 u_{t-2} + u_{t+1}]$$
$$= \mu + \theta_1 u_t + \theta_2 u_{t-1} + \theta_3 u_{t-2}$$

$$\vdots$$

$$\left.\begin{aligned} f_{t,4} &= \mu \\ f_{t,5} &= \mu \\ &\vdots \end{aligned}\right\} \forall s \geq 4.$$

- with AR(2): $y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + u_t$

$$f_{t,1} = \mathbb{E}[y_{t+1}|t] = \mu + \phi_1 y_t + \phi_2 y_{t-1}$$

$$f_{t,2} = \mu + \phi_1 f_{t,1} + \phi_2 y_t$$

$$f_{t,3} = \mu + \phi_1 f_{t,2} + \phi_2 f_{t,1}$$

$\vdots$

Generic formula: $f_{t,s} = \mu + \phi_1 f_{t,s-1} + \phi_2 f_{t,s-2}$

- Criteria to assess accuracy of t-s forecasting:

$$MSE = \frac{1}{N}\sum_{t=1}^{N}(y_{t+s} - f_{t,s})^2$$

$$MAE = \frac{1}{N}\sum_{t=1}^{N}|y_{t+s} - f_{t,s}|$$

A measure more closely correlated with trading profitability:

% correct sign pred. $= \frac{1}{N}\sum_{t=1}^{N} z_{t+s}$. $z_{t+s} = 1$ if $x_{t+s} f_{t,s} > 0$, $= 0$ otherwise.

Methodology:
1. decide what kind of forecast.
2. Select a model & compute $f_{t,s}$.
3. Evaluate the goodness of forecast w/ relevant metrics.

- Probabilistic forecasting: informs us about the expected distribution.
  - example of density forecast of MA(1):

$$y_t = \mu + u_t + \theta_1 u_{t-1}$$

$$\left.\begin{aligned} \mathbb{E}_t[y_{t+1}] &= \mu + \theta_1 u_t \\ Var_t(y_{t+1}) &= \sigma^2 \end{aligned}\right\} y_{t+1}|t \sim N(\mu + \theta_1 u_t, \sigma^2).$$

  - AR(1): $y_{t+1}|t \sim N(\mu + \phi_1 y_t, \sigma^2)$.

  - Test for accuracy: pinball loss; winkler score.