Time Series
Ch4:Estimation in the Time Domain

Author: Ngai Hang Chan

Reporter: 蕭伯任

2017-3-31

Contents

4.2 MOMENT ESTIMATORS

- The simplest type of estimators are the moment estimates. If $EY_t = \mu$, we simply estimate μ by $\bar{Y} = (1/n)\sum_{t=1}^{n} Y_t$ and proceed to analyze the demeaned series $X_t = Y_t - \bar{Y}$. For the covariance and correlation functions, we may use the same idea to estimate $\gamma(k)$ by

$$C_k = (1/n)\sum_{t=1}^{n-k}(Y_t - \bar{Y})(Y_{t+k} - \bar{Y})$$

- Similarly, we can estimate $\rho(k)$ by

$$r_k = C_k/C_0$$

4.3 AUTOREGRESSIVE MODELS

- Given the strong resemblance between an AR(p) model and a regression model, it is not surprising to anticipate that estimation of an AR(p) model is straightforward. Consider an AR(p) process

$$Y_t = \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + Z_t.$$

- This equation bears a strong resemblance to traditional regression models. Rewriting this equation in the familiar regression expression,

$$Y_t = (\phi_1, \cdots, \phi_p) \begin{pmatrix} Y_{t-1} \\ \vdots \\ Y_{t-p} \end{pmatrix} + Z_t = Y'_{t-1}\phi + Z_t,$$

4.3 AUTOREGRESSIVE MODELS

- The least squares estimate (LSE) of ϕ is given by

$$\hat{\phi} = \left(\sum_{t=p+1}^{n} Y_{t-1} Y_{t-1}' \right)^{-1} \left(\sum_{t=p+1}^{n} Y_{t-1} Y_t \right)$$

- Standard regression analysis can be applied here with slight modifications. Furthermore, if $Z_t \sim N(0, \sigma^2)$ i.i.d., then $\hat{\phi}$ is also the maximum likelihood estimate(MLE). In the simple case that $p = 1$, $Y_t = \phi Y_{t-1} + Z_t$, we have $\hat{\phi} = \sum_{t=p+1}^{n} Y_t Y_{t-1} / \sum_{t=p+1}^{n} Y_{t-1}^2$.
- Futher, $\hat{Z}_t = Y_t - Y_{t-1}' \hat{\phi}$ is the fitted residual and almost all techniques concerning residual analysis from classical regression can be carried over. Finally, standard asymptotic results such as consistency and asymptotic normality are available.

4.3 AUTOREGRESSIVE MODELS

- **Theorem 4.1**

$$\sqrt{n}(\hat{\phi} - \phi) \quad \underset{\mathcal{L}}{\rightarrow} \quad N(0, \sigma^2 \Gamma_p^{-1}),$$

- where $\underset{\mathcal{L}}{\rightarrow}$ denotes convergence in distribution of the corresponding random variables as the sample size $n \to \infty$ and

$$\Gamma_p = E \begin{pmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(p-1) \\ * & \gamma(0) & \cdots & \gamma(p-2) \\ \vdots & & & \vdots \\ * & * & \cdots & \gamma(0) \end{pmatrix}$$

4.3 AUTOREGRESSIVE MODELS

- **Example 4.1** For an AR(1) model, we have $\sqrt{n}(\hat{\phi} - \phi) \quad \underset{\mathcal{L}}{\rightarrow} \quad N(0, \sigma^2/\gamma(0))$, with

$$\begin{aligned} \gamma(0) = varY_t &= \phi^2 varY_{t-1} + \sigma^2 \\ &= \phi^2\gamma(0) + \sigma^2 \end{aligned}$$

Thus, $\gamma(0) = \sigma^2/(1 - \phi^2)[i.e., \hat{\phi} \sim AN(\phi, (1 - \phi^2)/n)]$

4.3 AUTOREGRESSIVE MODELS

- From the preceding theorem, usual inference such as constructing approximated confidence intervals or testings for ϕ can be conducted. Alternatively, we can evaluate the Yule-Walker(Y-W) equation via multiplying equation by Y_{t-k} and taking expectations,

$$\begin{aligned} \gamma(k) &= \phi_1 \gamma(k-1) + \cdots + \phi_p \gamma(k-p), \\ \rho(k) &= \phi_1 \rho(k-1) + \cdots + \phi_p \rho(k-p), \quad k = 1, ..., p. \end{aligned}$$

4.3 AUTOREGRESSIVE MODELS

- In matrix notation , these equations become

$$\begin{pmatrix} \rho(1) \\ \vdots \\ \rho(p) \end{pmatrix} = \begin{pmatrix} \rho(0) & \rho(1) & \cdots & \rho(p-1) \\ \rho(1) & \rho(0) & \cdots & \rho(p-2) \\ \vdots & & & \vdots \\ (p-1) & \cdots & \cdots & \rho(0) \end{pmatrix} \begin{pmatrix} \phi_1 \\ \vdots \\ \phi_p \end{pmatrix}$$

Hence, the Yule-Walker estimates are the ϕ such that

$$\hat{\phi} = R^{-1}r = \begin{pmatrix} 1 & r_1 & \cdots & r_{p-1} \\ r_1 & 1 & \cdots & r_{p-2}) \\ \vdots & & & \vdots \\ r_{p-1} & \cdots & \cdots & 1 \end{pmatrix}^{-1} \begin{pmatrix} r_1 \\ \vdots \\ r_p \end{pmatrix}$$

4.3 AUTOREGRESSIVE MODELS

- Again, asymptotic properties of the Y -W estimates can be found. When the sample size n is big and the order p is moderate, computational cost can be enormous for inverting the matrix **R**. In practice, it would be much more desirable to solve these quantities in real time (i.e., in a recursive online manner). The Durbin-Levinson(D-L) algorithm offers such a recursive scheme. We would not pursue the details of this algorithm, but refer the interested reader to the discussion given in Brockwell and Davis(1991). In any case, most of the computer routines, including those of SPLUS programs, use this algorithm to estimate the parameters.

4.3 AUTOREGRESSIVE MODELS

- Roughly speaking, we can classify the estimation steps as follows:
  1. Use the Durbin-Levinson algorithm to evaluate the Yule-Walker estimates.
  2. Use the Yule-Walker estimates as initial values to calculate the maximum likelihood estimates(MLE) of the parameters. Details of the MLE are given in Section 4.6.
  3. To estimate the standard error in the AR equation, use the estimator $\hat{\sigma}^2 = \frac{1}{n}\sum_{t=p+1}^{n}(Y_t - \hat{Y}_t)^2, \quad \hat{Y}_t = \hat{\phi}_1 Y_{t-1} + \cdots + \hat{\phi}_p Y_{t-p}.$

4.4 MOVING AVERAGE MODELS

- Contrary to the AR model, estimation for an MA model is much more tricky. To illustrate this point, consider the simple MA(1) model $Y_t = Z_t - \theta Z_{t-1}$. Suppose that we intend to use a moment estimator for θ. Then

$$\rho_1 = \frac{-\theta}{1+\theta^2}$$

and

$$r_1 = \frac{-\hat{\theta}}{1+\hat{\theta}^2}$$

Thus

$$\hat{\theta} = \frac{-1 \pm \sqrt{1-4r_1^2}}{2r_1}$$

- This estimator is nonlinear in nature. Such a nonlinearity phenomenon is even more prominent for an MA(q) model. In general, it will be very difficult to express the θ_i's of an MA(q) model as functions of r_i's analytically.

4.4 MOVING AVERAGE MODELS

- Alternatively, if $|\theta| < 1$, then

$$Z_t = Y_t + \theta Z_{t-1} = Y_t + \theta Y_{t-1} + \theta^2 Y_{t-2} + \cdots .$$

Let $S(\theta) = \sum_{t=1}^{n} Z_t^2$. We can find the θ such that $S(\theta)$ is minimized, where $Z_t = Z_t(\theta)$ implicitly. Note that even in this simple $MA(1)$ case, $S(\theta)$ cannot be minimized analytically. In particular, for given $Y_i, \ldots, Y_n$ and θ, and conditional on $Z_0 = 0$, set

$$\begin{aligned} Z_1 &= Y_1, \\ Z_2 &= Y_2 + \theta Z_1 = Y_2 + \theta Y_1, \\ &\vdots \\ Z_n &= Y_n + \theta Z_{n-1}, \end{aligned}$$

4.4 MOVING AVERAGE MODELS

- compute $S_*(\theta) = \sum_t^2 Z_t^2$ for the given θ, where we use S_*. to denote that this quantity is evaluated conditional on the initial value $Z_0 = 0$. In general, we can perform a grid search over $(1, -1)$ to find the minimum of $S_*(\theta)$ by means of a numerical method, the Gauss-Newton method, say. This is also known as the conditional least squares(CLS) method. Specifically, consider

$$Z_t(\theta) \cong Z_t(\theta^*) + (\theta - \theta^*)\frac{dZ_t(\theta)}{d\theta}|_{\theta=\theta^*},$$

  from an initial point θ^*. Note that this equation is linear in θ, thus $\sum_{t=1}^n Z_t^2(\theta) = S_*(\theta)$ can be minimized analytically to get a new $\theta_{(1)}$. Substitute $\theta_{(1)}$ forθ^* into equation again and iterate this process until it converges.
- For a general MA(q) model, a multivariate Gauss-Newton procedure can be used to minimize $S_*(\theta)$ via $Z_t = Y_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}$ such that $Z_0 = Z_{-1} = \cdots = Z_{1-q} = 0$, where $\theta = (\theta_1, \cdots, \theta_q)'$.

4.5 ARMA MODELS

- Having seen the intricacies in estimating an MA model, we now discuss the estimation of an ARMA model by means of a simple $ARMA(1,1)$ model.
- **Example 4.2** Let $Y_t - \phi Y_{t-1} = Z_t - \theta Z_{t-1}$ Conditional on $Z_0 = 0 = Y_0$, find (ϕ, θ) that minimizes

$$S_*(\phi, \theta) = \sum_{t=1}^{n} Z_t^2(\phi, \theta),$$

where

$$Z_t = Y_t - \phi Y_{t-1} + \theta Z_{t-1}$$

4.5 ARMA MODELS

- For a general $ARMA(p,q)$, we perform a similar procedure to find the estimates by solving a numerical minimization problem. Let $Z_t = Y_t - \phi_1 Y_{t-1} - \cdots - \phi_p Y_{t-p} + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}$. Compute $Z_t = Z_t(\phi, \theta)$, $t = p+l, \cdots, n$ and find the parameters (ϕ, θ) that minimize

$$S_*(\phi, \theta) = \sum_{t=p+1}^{n} Z_t^2(\phi, \theta).$$

For an invertible MA or ARMA model, the initial values of $Y_0 = Y_{-1} = Y_{1-p} = \cdots = Z_0 = \cdots = Z_{1-q} = 0$ have little effect on the final parameter estimates when the sample size is large.

Thank You for your Attention