Least Square Method of Estimation

Least Square Method of Estimation

- Since method-of-moments performs poorly for some models, we examine another method of parameter estimation: *Least Squares.*
- We first consider autoregressive models.
- We assume our time series is stationary (or that the time series has been transformed so that the transformed data can be modeled as stationary).
- To account for the possibility that the mean is nonzero, we subtract μ from each observation and treat μ as a parameter to be estimated.

LS Estimation for the AR(1) Model

- Consider the mean-centered $AR(1)$ model:

$$Y_t - \mu = \phi(Y_{t-1} - \mu) + \overset{\varepsilon_t}{e_t}$$

$$\Rightarrow \varepsilon_t = (Y_t - \mu) - \phi(Y_{t-1} - \mu)$$

- The least squares method seeks the parameter values that minimize the sum of squared differences:

$$\sum_{t=2}^{n} \varepsilon_t^2$$

$$\boxed{S_c(\phi, \mu) = \sum_{t=2}^{n} [(Y_t - \mu) - \phi(Y_{t-1} - \mu)]^2}$$

- This criterion is called the *conditional sum-of-squares function* (CSS).

$$\frac{\partial S_c}{\partial \mu} = 2 \sum_{t=2}^{n} (Y_t - \mu) - \phi(Y_{t-1} - \mu)\Big[[0-1] - \phi[0-1]\Big] = 0$$

LS Estimation of μ for the AR(1) Model

- Taking the derivative of CSS with respect to μ, setting equal to 0 and solving for μ, we obtain the LS estimator of μ:

$$\hat{\mu} = \frac{1}{(n-1)(1-\phi)}\left[\sum_{t=2}^{n} Y_t - \phi \sum_{t=2}^{n} Y_{t-1}\right]$$

- For large n, this $\hat{\mu} \approx \bar{Y}$, regardless of the value of ϕ.

LS Estimation of ϕ for the AR(1) Model

- Taking the derivative of CSS with respect to ϕ, setting equal to 0 and solving for ϕ, we obtain the LS estimator of ϕ:

$$\hat{\phi} = \frac{\sum_{t=2}^{n}(Y_t - \bar{Y})(Y_{t-1} - \bar{Y})}{\sum_{t=2}^{n}(Y_{t-1} - \bar{Y})^2}$$

$\hat{\rho}_1$

- This estimator is almost identical to r_1: it's just missing one term in the denominator, $(Y_n - \bar{Y})^2$.
- So, especially for large n, the LS and MOM estimators are nearly identical in the $AR(1)$ model.
- In the general $AR(p)$ model, the LS estimators of μ and of $\phi_1, \ldots, \phi_p$ are approximately equal to the MOM estimators, especially for large samples.

LS Estimation for Moving Average Models

- Consider now the $MA(1)$ model:

$$Y_t = e_t - \theta e_{t-1}$$ or $\varepsilon_t - \theta \varepsilon_{t-1}$

- Recall that this can be written as

$$Y_t = -\theta Y_{t-1} - \theta^2 Y_{t-2} - \theta^3 Y_{t-3} - \cdots + e_t.$$

- So a least squares estimator of θ can be obtained by finding the value of θ that minimizes

$$S_c(\theta) = \sum [Y_t + \theta Y_{t-1} + \theta^2 Y_{t-2} + \theta^3 Y_{t-3} + \cdots]^2$$

- But this is nonlinear in θ, and the infinite series causes technical problems.

LS Estimation for Moving Average Models

- Instead, we proceed by conditioning on one previous value of e_t. Note that
$$e_t = Y_t + \theta e_{t-1}$$
- If we set $e_0 = 0$, then we have the set of recursive equations $e_1 = Y_1$, $e_2 = Y_2 + \theta e_1, \ldots, e_n = Y_n + \theta e_{n-1}$.
- Since we know $Y_1, Y_2, \ldots, Y_n$ (these are the observed data values) and can calculate the $e_1, e_2, \ldots, e_n$ recursively, the only unknown quantity here is θ.
- We can do a numerical search for the value of θ (within the invertible range between -1 and 1) that minimizes $\sum(e_t)^2$, conditional on $e_0 = 0$.
- A similar approach works for higher-order $MA(q)$ models, except that we assume $e_0 = e_{-1} = \cdots = e_{-q} = 0$ and the numerical search is multidimensional, since we are estimating $\theta_1, \ldots, \theta_q$.

LS Estimation for ARMA Models

- With the $ARMA(1,1)$ model:

$$Y_t = \phi Y_{t-1} + e_t - \theta e_{t-1},$$

  we note that

$$e_t = Y_t - \phi Y_{t-1} + \theta e_{t-1}$$

  and minimize $S_c(\phi, \theta) = \sum_{t=2}^{n} e_t^2$; note that the sum starts at $t = 2$ to avoid having to choose an "initial" value Y_0.
- With the general $ARMA(p, q)$ model, the procedure is similar, except that we assume $e_p = e_{p-1} = \cdots = e_{p+1-q} = 0$, and we estimate $\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q$.
- For *large samples*, when the parameter sets yield invertible models, the initial values for $e_p, e_{p-1}, \ldots, e_{p+1-q}$ have little effect on the final parameter estimates.