Financial Econometrics I

Lecture 3

Fu Ouyang

November 29, 2018

Outline

Extended Autocorrelation Function (EACF)

Recall that

- ACF cuts off at q for MA(q) model, i.e., $\rho(k) = 0$ when $k > q$.
- PACF cuts of at p for AR(p) model, i.e., $\pi(k) = 0$ when $k > q$.

Hence, one can use ACF to select q for pure MA(q) model and use PACF to select p for pure AR(p) model.

Now consider a general ARMA(p, q)

$$X_t = \beta_1 X_{t-1} + \cdots + \beta_p X_{t-p} + \epsilon_t + \alpha_1 \epsilon_{t-1} + \cdots + \alpha_q \epsilon_{t-q}$$

To determine the order (p, q) of this model, one can use the *extended autocorrelation function* (EACF).

Extended Autocorrelation Function (EACF)

Properties of EACF

For each fixed AR order $l = 0, 1, 2, \ldots$, let $Z_t \equiv X_t - b_1 X_{t-1} - \cdots - b_l X_{t-l}$.

- When $l = p$, $Z_t \sim \mathrm{MA}(q)$, then for each MA order $k = 0, 1, 2, \ldots$

  (1) One have l Yule-Walker equations for Z_t:

  $$\rho(k+j) = b_1^{(k)} \rho(k+j-1) + \cdots + b_l^{(k)} \rho(k+j-l), \; j = 1, \ldots, l \tag{1}$$

  (2) Solve $b_1^{(k)}, b_2^{(k)}, \ldots, b_l^{(k)}$ from (1), and define

  $$Z_t^{(k)} = X_t - b_1^{(k)} X_{t-1} - \cdots - b_l^{(k)} X_{t-l} \tag{2}$$

  (3) Note that if $l = p$, $b_1^{(k)} = b_1, \ldots, b_p^{(k)} = b_p$ and $Z_t^{(k)} \sim \mathrm{MA}(q)$ for all $k > q$.

  (4) The l-th EACF at lag k, $\rho(k; l)$, is defined as the ACF of $Z_t^{(k)}$ at lag k.

Extended Autocorrelation Function (EACF)

Properties of EACF

- The p-th EACF, $\rho(\cdot; p)$, cuts off at q for stationary ARMA(p, q) processes, i.e., $\rho(k; p) = 0$ for all $k > q$.
- When $l < p$, $\rho(k; l) \neq 0$ as $Z_t^{(k)} \sim \text{MA}(\infty)$.
- When $l > p$, the l-th EACF cuts off at $q + (l - p)$.

Calculating EACF

- Replace $\rho(\cdot)$ in (1) by $\widehat{\rho}(\cdot)$, which leads to the estimator $\widehat{b}_j^{(k)}$ of $b_j^{(k)}$.
- Replace $b_j^{(k)}$ with $\widehat{b}_j^{(k)}$ in (2), which gives $\widehat{Z}_t^{(k)}$.
- The sample ACF of $\widehat{Z}_t^{(k)}$ at lag k is taken as the l-th sample EACF at lag k, and denoted by $\widehat{\rho}(l; k)$.

Extended Autocorrelation Function (EACF)

The list of $\widehat{\rho}(l;k)$ for $l = 0, 1, 2...$ and $k = 0, 1, 2, ...$ can be organized in the following table, where the (p, q)-th entry is $\widehat{\rho}(q+1;p)$ (as $k > q$ is required) and $\widehat{\rho}(k;0) = \widehat{\rho}(k)$ by construction.

	$q=0$	$q=1$	$q=2$	$q=3$	$\cdots$
$p=0$	$\widehat{\rho}(1)$	$\widehat{\rho}(2)$	$\widehat{\rho}(3)$	$\widehat{\rho}(4)$	$\cdots$
$p=1$	$\widehat{\rho}(1;1)$	$\widehat{\rho}(2;1)$	$\widehat{\rho}(3;1)$	$\widehat{\rho}(4;1)$	$\cdots$
$p=2$	$\widehat{\rho}(1;2)$	$\widehat{\rho}(2;2)$	$\widehat{\rho}(3;2)$	$\widehat{\rho}(4;2)$	$\cdots$
$p=3$	$\widehat{\rho}(1;3)$	$\widehat{\rho}(2;3)$	$\widehat{\rho}(3;3)$	$\widehat{\rho}(4;3)$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Extended Autocorrelation Function (EACF)

The ideal (theoretical) pattern of the EACF table for a ARMA(2,1) model should look like the following table, where the "x" symbol indicates that the corresponding $\widehat{\rho}(k;l)$ is significantly nonzero, while "o" means insignificant.

	$q=0$	$q=1$	$q=2$	$q=3$	$\cdots$
$p=0$	x	x	x	x	$\cdots$
$p=1$	x	x	x	x	$\cdots$
$p=2$	x	o	o	o	$\cdots$
$p=3$	x	x	o	o	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Model Identification using ACF, PACF, and EACF

Practical Guidelines

1. If sample ACF $\widehat{\rho}(k)$ are close to 1 for all small k and decay slowly, consider $X_t - X_{t-1}$.
2. If sample ACF $\widehat{\rho}(k)$ are bounded by $\pm 1.96(1 + 2\sum_{j=1}^{k-1} \widehat{\rho}(j)^2)^{1/2}/\sqrt{T}$ for all $k > q$, fit an MA(q) model.
3. If sample PACF $\widehat{\pi}(k)$ are bounded by $\pm 1.96\sqrt{T}$ for all $k > p$, fit an AR(p) model.
4. If sample EACF $\widehat{\rho}(q + j; p + i)$ are bounded by $\pm 1.96\sqrt{T - i - j}$ for any $i \geq 0$ and $j > i$, fit an ARMA(p, q) model (the upper-left vertex of the triangle of all "o").
5. If more than one criteria among 2, 3, and 4 are satisfied, apply the *parsimony principle*.

Outline

Model Diagnostics

Model Diagnostics is important as it can help check if the specification of a time series model is appropriate.

The key idea is straightforward: if a ARMA(p, q) model is adequate for an observed data, the residuals

$$\widehat{\epsilon}_t = X_t - \sum_{j=1}^{p} \widehat{b}_j X_{t-j} - \sum_{l=1}^{q} \widehat{a}_l \widehat{\epsilon}_{t-l}, t = p+1, ..., T$$

assuming $\widehat{\epsilon}_{p+1-q} = 0, ..., \widehat{\epsilon}_p = 0$ should behave like white noise.

Two classes of the methods for residual analysis: visual diagnostics and tests for white noise.

Model Diagnostics

Residual Analysis

Visual diagnostics

- Residual plot, i.e., plotting $\widehat{\epsilon}_t$ (more often standardized, $\widehat{\epsilon}_t / SE(\widehat{\epsilon}_t)$) against time t. (no "pattern")
- Plot $\widehat{\epsilon}_t$ against the fitted values $\widehat{X}_t$. (no "pattern")
- ACF, PACF, EACF of the residuals (no significant $\rho(k), \pi(k), \rho(k, l)$).
- QQ plots for diagnosing normality assumption. (heavy tails?)

Statistical tests for white noise

- Perform the Ljung-Box tests for white noise ($df = m - p - q$).

Outline

Information Criteria

Model diagnostics via residual analysis tends to favor complex models as more regressors can often improve the goodness of fit (GoF). But this may give an *overfitting* model. Including irrelevant regressors may:

- inflate the SE of the estimated parameters, and/or
- make the interpretation of the fitted model difficult.

A good strategy for model selection is to find a balance between the GoF and the complexity of the model (the *parsimony principle*).

Akaike's information criterion (AIC) is proposed for (nested) model selection based on the trade-off between these two factors:

$$AIC = -2l(\widehat{\theta}) + 2K$$

where $l(\cdot)$ is the log-likelihood function, $\widehat{\theta}$ is the MLE, and K is the number of estimated parameters.

Information Criteria

An alternative to AIC is the *Bayesian information criterion* (BIC):

$$BIC = -2l(\widehat{\theta}) + K \log T$$

For a stationary ARMA(p, q) model with $\epsilon_t \sim N(0, \sigma^2)$,

$$AIC(p, q) = T \log(\widehat{\sigma}^2) + 2(p + q + 1)$$

$$BIC(p, q) = T \log(\widehat{\sigma}^2) + (p + q + 1) \log T$$

- The first term is a measure of the GoF.
- The second term is a penalty for the complexity of the model.

In practice, one can select the order (p, q) minimizing AIC(p, q) or BIC(p, q).

Information Criteria and Model Identification

Model Selection via AIC/BIC

- AIC tends to overestimate the orders, while BIC tends to give oversimplified models ($\log T > 2$).
- For interpretation purpose, one can use BIC to select a simple model. For forecasting purpose, one can try AIC first as a slightly more complex model causes little harm.
- AIC is asymptotically efficient, while BIC is not.
- Compute AIC and select a bunch of competitive candidates (with AIC close to to the minimum AIC).
- Select optimal model from these candidates based on interpretation, simplicity, diagnostic checking among other considerations.

Let's see an illustrating example using real-world data. [R markdown].