Spring 2015 Statistics 153 (Time Series) : Lecture Fourteen

Aditya Guntuboyina

10 March 2015

1 ARMA models with non-zero mean

$\{X_t\}$ is said to be ARMA(p, q) with mean μ if $\{X_t - \mu\}$ is ARMA(p, q) i.e., if

$$(X_t - \mu) - \phi_1(X_{t-1} - \mu) - \cdots - \phi_p(X_{t-p} - \mu) = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}.$$

2 Fitting an ARMA(p, q) model to data

We study how to estimate the parameters $\mu, \phi_1, \ldots, \phi_p$ and $\theta_1, \ldots, \theta_q$ and σ_Z^2 of an ARMA(p, q) model from observed data. We assume that it is sensible to fit a stationary model to the data. If obvious forms of nonstationarity such as trend exist, they need to be removed (say, by differencing) before fitting an ARMA model.

We will look at three methods for fitting ARMA models: Method of Moments, Least Squares and Maximum Likelihood.

Each of these methods can be better understood if we first look at them in the special case of fitting AR(p) models.

3 Fitting an AR(p) model to the data

Given data $x_1, \ldots, x_n$ for which we believe than an AR(p) model:

$$(X_t - \mu) - \phi_1(X_{t-1} - \mu) - \cdots - \phi_p(X_{t-p} - \mu) = Z_t.$$

is appropriate, we learn how to fit it i.e., how to estimate the parameters μ, $\phi_1, \ldots, \phi_p$ and σ_Z^2 from the data.

3.1 Method of Moments

Estimate μ by the sample mean: $\bar{x} = (x_1 + \cdots + x_n)/n$.

For estimating the other parameters: $\phi_1, \ldots, \phi_p$ and and σ_Z^2, recall the set of equations (called **Yule-Walker equations**) that we had for the calculation of the autocorrelation function of a causal stationary AR(p) process (this was the second method for calculating the ACF of an ARMA process; discussed in Lectures 7 and 8):

$$\gamma_X(0) - \phi_1\gamma_X(1) - \cdots - \phi_p\gamma_X(p) = \sigma_Z^2, \tag{1}$$

and

$$\gamma_X(k) - \phi_1\gamma_X(k-1) - \cdots - \phi_p\gamma_X(k-p) = 0 \qquad \text{for } k \geq 1. \tag{2}$$

Previously, we considered solving these equations to write $\gamma_X(k)$ in terms of σ_Z^2 and $\phi_1, \dots, \phi_p$. But these same equations can be used to estimate σ_Z^2 and $\phi_1, \dots, \phi_p$ from the data $x_1, \dots, x_n$. Indeed, first estimate the autocovariances $\gamma_X(h)$ by the sample autocovariances $\hat{\gamma}_X(h)$:

$$\hat{\gamma}_X(h) := \frac{1}{n}\sum_{t=1}^{n-h}(x_{t+h} - \bar{x})(x_t - \bar{x})$$

and then solve (1) and (2) for the unknown parameters σ_Z^2 and $\phi_1, \dots, \phi_p$. Note that we have an infinite set of equations in (2). But we only need to estimate $p+1$ parameters. So we will only use (1) the first p of the equations in (2). This gives us $p+1$ equations to solve for the $p+1$ unknowns $\phi_1, \dots, \phi_p$ and σ_Z^2. This estimation method is called Yule-Walker estimation and the resulting estimates are called Yule-Walker Estimates.

Essentially, we are trying to find an AR(p) model whose autocovariance function equals the observed sample autocovariance function at lags $0, 1, \dots, p$. This is why this method is called the method of moments.

For $p = 1$ i.e., the AR(1) case, we just have the two equations:

$$\hat{\gamma}_X(0) - \phi\hat{\gamma}_X(1) = \sigma_Z^2 \quad \text{and} \quad \hat{\gamma}_X(1) = \phi\hat{\gamma}_X(0).$$

This of course gives

$$\hat{\phi} = \frac{\hat{\gamma}_X(1)}{\hat{\gamma}_X(0)} = r_1 \quad \text{and} \quad \hat{\sigma_Z^2} := \hat{\gamma}_X(0)\left(1 - r_1^2\right).$$

When $p = 2$ i.e., AR(2), we get the three equations:

$$\hat{\gamma}_X(0) - \phi_1\hat{\gamma}_X(1) - \phi_2\hat{\gamma}_X(2) = \sigma_Z^2$$

and

$$\hat{\gamma}_X(1) - \phi_1\hat{\gamma}_X(0) - \phi_2\hat{\gamma}_X(1) = 0 \quad \text{and} \quad \hat{\gamma}_X(2) - \phi_1\hat{\gamma}_X(1) - \phi_2\hat{\gamma}_X(0) = 0$$

The last two equations can used to solve for ϕ_1 and ϕ_2 to yield:

$$\hat{\phi}_1 = \frac{r_1(1-r_2)}{1-r_1^2} \quad \text{and} \quad \hat{\phi}_2 = \frac{r_2 - r_1^2}{1-r_1^2}.$$

Plugging these values for ϕ_1 and ϕ_2 into $\hat{\gamma}_X(0) - \phi_1\hat{\gamma}_X(1) - \phi_2\hat{\gamma}_X(2) = \sigma_Z^2$, we get an estimate for σ_Z^2.

3.2 Least Squares

Given data $x_1, \dots, x_n$, one can fit the AR(1) model $X_t - \mu = \phi(X_{t-1} - \mu) + Z_t$ by minimizing the following least squares objective function:

$$S_c(\phi, \mu) = \sum_{i=2}^{n}\left(x_i - \mu - \phi(x_{i-1} - \mu)\right)^2. \tag{3}$$

This is called conditional least squares estimation because, as we will see later, this minimization arises when one maximizes the conditional likelihood of $x_2, \dots, x_n$ given x_1 under the iid gaussian assumption on $\{Z_t\}$.

To minimize (3), let $\beta_0 = \mu(1-\phi)$ and $\beta_1 = \phi$ and rewrite it as:

$$\sum_{i=2}^{n}\left(x_i - \beta_0 - \beta_1 x_{i-1}\right)^2.$$

Minimizing this now is exactly linear regression and the answers are given by

$$\hat{\beta}_1 = \frac{\sum_{i=2}^n (x_i - \bar{x}_{(2)})(x_{i-1} - \bar{x}_{(1)})}{\sum_{i=2}^n (x_{i-1} - \bar{x}_{(1)})^2} \qquad \text{where } \bar{x}_{(1)} := \frac{x_1 + \cdots + x_{n-1}}{n-1} \text{ and } \bar{x}_{(2)} := \frac{x_2 + \cdots + x_n}{n-1}$$

and $\hat{\beta}_0 := \bar{x}_{(2)} - \hat{\beta}_1 \bar{x}_{(1)}$. This will give

$$\hat{\phi} = \frac{\sum_{i=2}^n (x_i - \bar{x}_{(2)})(x_{i-1} - \bar{x}_{(1)})}{\sum_{i=2}^n (x_{i-1} - \bar{x}_{(1)})^2} \quad \text{and} \quad \hat{\mu} := \frac{\bar{x}_{(2)} - \hat{\phi}\bar{x}_{(1)}}{1 - \hat{\phi}}.$$

The parameter σ_Z^2 is estimated by

$$\hat{\sigma_Z^2} := \frac{\sum_{i=2}^n \left(x_i - \hat{\mu} - \hat{\phi}(x_{i-1} - \hat{\mu})\right)^2}{n-1}.$$

It is easily seen that these estimates are very close to those obtained by the Yule-Walker method.

The extension to higher order AR(p) processes is straightforward.

3.3 Maximum Likelihood

Next class.

Book Reading: Section 3.6.