Times Series Analysis (XII) – Parameter Estimation (II)

Jianyong Sun
School of Mathematics and Statistics
Xi'an Jiaotong University

24th Oct., 2017

Recap

- Four methods to do parameter estimation
  - method-of-moment
    - For AR models, Yule-Walker equations
    - For MA models, nonlinear equations
    - For mixed models, difficult. ARMA(1,1) — $\hat{\phi} = \frac{r_2}{r_1}$, nonlinear equation for $\hat{\theta}$
    - To estimate σ_e^2, use sample variance

Recap

- conditional least square
  - For AR(p), minimize $S_c(\phi, \mu) = \sum_{t=p+1}^{n}[(Y_t - \mu) - \sum_{j=1}^{p} \phi_j(Y_{t-j} - \mu)]^2 \rightarrow$ Yule-Walker equations
  - For MA(1) model
    - minimize $S_c(\theta) = \sum[Y_t + \theta Y_{t-1} + \theta^2 Y_{t-2} + \cdots]^2 \rightarrow$ numerical optimization
    - Given θ, through recursion to obtain e_t's, then minimize $S_c(\theta) = \sum e_t^2$
  - For ARMA model, $e_t = Y_t - \phi_1 Y_{t-1} - \cdots - \phi_p Y_{t-p} + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q}$, compute $\sum e_t^2$ results in 'start-up' problems.

Recap

- Maximum likelihood
  - joint probability of $Y_1, \cdots, Y_n$
  - For AR(1), define $S(\phi, \mu) = \sum_{t=2}^{n}[(Y_t - \mu) - \phi(Y_{t-1} - \mu)]^2 + (1 - \phi^2)(Y_1 - \mu)$
  - Define $L(\phi, \mu, \sigma_e^2) = (2\pi\sigma_e^2)^{-n/2}(1 - \phi^2)^{1/2} \exp\left[-\frac{1}{2\sigma_e^2} S(\phi, \mu)\right]$
- Unconditional least square
  - minimize $S(\phi, \mu)$
- Large-sample properties of the MLE and conditional and unconditional least square estimators

Illustrations of Parameter Estimation

1st Example: Consider the simulated MA(1) series with $\theta = -0.9$. We consider the estimates by method-of-moments, maximum likelihood and unconditional sum-of-squares and conditional least-square, listed in the table.

Method	$\hat{\theta}$	*Std.*
MM	-0.554	NA
MLE	-0.915	0.04
Unconditional SS	-0.923	0.04
Conditional SS	-0.879	0.04

where the Std. is the standard error, estimated as

$$\sqrt{Var(\hat{\theta})} \approx \sqrt{\frac{1-\theta^2}{n}} \approx 0.04$$

Illustrations of Parameter Estimation

2nd Example: Consider the simulated MA(1) series with $\theta = 0.9$. We have

Method	$\hat{\theta}$	*Std.*
MM	0.719	NA
MLE	0.958	0.04
Unconditional SS	0.983	0.04
Conditional SS	1.00	0.04

Here the maximum likelihood estimate of $\hat{\theta} = 1$ is a little disconcerting since it corresponds to a noninvertible model.

3rd Example: MA(1) simulation with $\theta = -0.9$. We have

Method	$\hat{\theta}$	*Std.*
MM	-0.719	NA
MLE	-0.894	0.06
Unconditional SS	-0.961	0.06
Conditional SS	-0.979	0.06

The standard error can be computed as

$$\sqrt{Var(\hat{\theta})} \approx \sqrt{\frac{1-\theta^2}{n}} \approx 0.06$$

Illustrations of Parameter Estimation

For our simulated autoregressive models, the results are reported in the following table

Parameter ϕ	MM	Conditional SS	Unconditional SS	MLE	n
0.9	0.831	0.857	0.911	0.892	60
0.4	0.470	0.473	0.473	0.465	60

The standard errors for the estimates are

$$\sqrt{Var(\hat{\phi})} \approx \sqrt{\frac{1-\phi^2}{n}} \approx 0.07$$

and

$$\sqrt{Var(\hat{\phi})} \approx \sqrt{\frac{1-\phi^2}{n}} \approx 0.11$$

Considering the magnitude of these standard errors, all four methods estimate reasonably well for AR(1) models.

Illustrations of Parameter Estimation

Parameters	Method-of-Moments Estimates	Conditional SS Estimates	Unconditional SS Estimates	Maximum Likelihood Estimate	n
$\phi_1 = 1.5$	1.472	1.5137	1.5183	1.5061	120
$\phi_2 = -0.75$	−0.767	−0.8050	−0.8093	−0.7965	120

Figure: Parameter Estimation for a Simulated AR(2) Model

The standard errors for the estimates are

$$\sqrt{Var(\hat{\phi}_1)} \approx \sqrt{Var(\hat{\phi}_2)} \approx \sqrt{\frac{1-\phi_2^2}{n}} \approx 0.06$$

considering the size of the standard errors, all four methods estimate reasonably well for AR(2) models.

Illustrations of Parameter Estimation

Final Example, consider the ARMA(1,1). Here $\phi = 0.6, \theta = -0.3$ and $n = 100$.

Parameters	Method-of-Moments Estimates	Conditional SS Estimates	Unconditional SS Estimates	Maximum Likelihood Estimate	*n*
$\phi = 0.6$	0.637	0.5586	0.5691	0.5647	100
$\theta = -0.3$	–0.2066	–0.3669	–0.3618	–0.3557	100

Figure: Parameter Estimation for a Simulated ARMA(1,1) Model

Illustrations of Parameter Estimation

Look at some real time series. The sample PACF strongly suggested an AR(1) model for the industrial chemical property time series.

Parameter	Method-of-Moments Estimate	Conditional SS Estimate	Unconditional SS Estimate	Maximum Likelihood Estimate	n
ϕ	0.5282	0.5549	0.5890	0.5703	35

Figure: Parameter Estimation for the Color Property Series.

Here the standard error of the estimates is about

$$\sqrt{Var(\hat{\phi})} \approx \sqrt{\frac{1-(0.57)^2}{n}} \approx 0.14$$

so all of the estimates are comparable.

Illustrations of Parameter Estimation

Consider the hare abundance series. We base all modeling on the square root of the original abundance numbers. Based on the partial autocorrelation function, we will estimate an AR(3) model.

Coefficients:	ar1	ar2	ar3	Intercept†
	1.0519	–0.2292	–0.3931	5.6923
s.e.	0.1877	0.2942	0.1915	0.3371

```
sigma^2 estimated as 1.066: log-likelihood = -46.54, AIC = 101.08
```

Figure: Maximum Likelihood Estimates from R Software: Hare Series.

The intercept here is the estimate of the process mean μ not of θ_0.

Noting the standard errors, the estimates of the lag 1 and lag 3 autoregressive coefficients are significantly different from zero, as is the intercept term, but the lag 2 autoregressive parameter estimate is not significant.

The estimated model would be written

$$\sqrt{Y_t} - 5.6923 = 1.0519\left(\sqrt{Y_{t-1}} - 5.6923\right) - 0.2292\left(\sqrt{Y_{t-2}} - 5.6923\right) - 0.3930\left(\sqrt{Y_{t-3}} - 5.6923\right) + e_t$$

or

$$\sqrt{Y_t} = 3.25 + 1.0519\sqrt{Y_{t-1}} - 0.2292\sqrt{Y_{t-2}} - 0.3930\sqrt{Y_{t-3}} + e_t$$

where Y_t is the hare abundance in year t in original terms.

Illustrations of Parameter Estimation

The oil price series. The sample ACF suggested an MA(1) model on the differences of the logs of the prices.

Parameter	Method-of-Moments Estimate	Conditional SS Estimate	Unconditional SS Estimate	Maximum Likelihood Estimate	*n*
θ	–0.2225	–0.2731	–0.2954	–0.2956	241

Figure: Estimation for the Difference of Logs of the Oil Price Series.

The method-of-moments estimate differs quite a bit from the others. The others are nearly equal given their standard errors of about 0.07.

Bootstrapping ARIMA Models

- we summarized some approximate normal distribution results for the estimator $\hat{\gamma}$ where $\hat{\gamma}$ is the vector consisting of all the ARMA parameters.
- These normal approximations are accurate for large samples, and statistical software generally uses those results in calculating and reporting standard errors.
- However, the general theory provides no practical guidance on how large the sample size should be for the normal approximation to be reliable
- **Bootstrap methods** (Efron and Tibshirani, 1993; Davison and Hinkley, 2003) provide an alternative approach to assessing the uncertainty of an estimator and may be more accurate for small samples.

Bootstrapping ARIMA Models

- We shall confine our discussion to the parametric bootstrap that generates the bootstrap time series $Y_1^*, Y_2^*, \cdots, Y_n^*$ by simulation from the fitted ARIMA(p, d, q) model.
  - The bootstrap may be done by fixing the first $p + d$ initial values of Y^* to those of the observed data.
  - For stationary models, an alternative procedure is to simulate stationary realizations from the fitted model, which can be done approximately by simulating a long time series from the fitted model and then deleting the transient initial segment of the simulated data — the so-called **burn-in**.

Bootstrapping ARIMA Models

- If the errors are assumed to be normally distributed, the errors may be drawn randomly and **with replacement** from $\mathcal{N}(0, \hat{\sigma}_e^2)$.
- For the case of an unknown error distribution, the errors can be drawn randomly and with replacement from the residuals of the fitted model.
- For each bootstrap series, let $\hat{\gamma}^*$ be the estimator computed based on the bootstrap time series data using the method of full maximum likelihood estimation assuming stationarity.
- The bootstrap is replicated, say, B times. (For example, $B = 1000$.)

Bootstrapping ARIMA Models

- From the B bootstrap parameter estimates, we can form an empirical distribution and use it to calibrate the uncertainty in $\hat{\gamma}$.
- Suppose we are interested in estimating some function of γ, say $h(\gamma)$, e.g., the AR(1) coefficient.
- Using the percentile method, a 95% bootstrap confidence interval for $h(\gamma)$ can be obtained as the interval from the 2.5 percentile to the 97.5 percentile of the bootstrap distribution of $h(\gamma)$.

Bootstrapping ARIMA Models

We illustrate the bootstrap method with the hare data.

- First bootstrap method: Generate recursively using the equation

$$Y_t^* - \hat{\phi}_1 Y_{t-1}^* - \hat{\phi}_2 Y_{t-2}^* - \hat{\phi}_3 Y_{t-3}^* = \hat{\theta}_0 + e_t^*$$

  for $t = 4, 5, \cdots, 31$, where e_t^* are chosen independently from $\mathcal{N}(0, \sigma_e^2)$ and $Y_1^* = Y_1, \cdots, Y_3^* = Y_3$
- Second, obtained using the same method except that the errors are drawn from the residuals.
- Third, the stationary bootstrap with a normal error distribution
- Fourth, the stationary bootstrap with the empirical residual distribution for the fourth row

Bootstrapping ARIMA Models

Method	ar1	ar2	ar3	intercept	noise var.
I	(0.593, 1.269)	(−0.655, 0.237)	(−0.666, −0.018)	(5.115, 6.394)	(0.551, 1.546)
II	(0.612, 1.296)	(−0.702, 0.243)	(−0.669, −0.026)	(5.004, 6.324)	(0.510, 1.510)
III	(0.699, 1.369)	(−0.746, 0.195)	(−0.666, −0.021)	(5.056, 6.379)	(0.499, 1.515)
IV	(0.674, 1.389)	(−0.769, 0.194)	(−0.665, −0.002)	(4.995, 6.312)	(0.477, 1.530)
Theoretical	(0.684, 1.42)	(−0.8058, 0.3474)	(−0.7684, −0.01776)	(5.032, 6.353)	(0.536, 1.597)

Figure: Bootstrap and Theoretical Confidence Intervals for the AR(3) Model Fitted to the Hare Data

All results are based on about 1000 bootstrap replications, but full maximum likelihood estimation fails for 6.3%, 6.3%, 3.8%, and 4.8% of 1000 cases for the four bootstrap methods I, II, III, and IV, respectively.

- All four methods yield similar bootstrap confidence intervals, although the conditional bootstrap approach generally yields slightly narrower confidence intervals.
- This is expected, as the conditional bootstrap time series bear more resemblance to each other because all are subject to identical initial conditions.
- The bootstrap confidence intervals are generally slightly wider than their theoretical counterparts that are derived from the large-sample results.
- Overall, we can draw the inference that the ϕ_2 coefficient estimate is insignificant, whereas both the ϕ_1 and ϕ_3 coefficient estimates are significant at the 5% significance level.

- The bootstrap method has the advantage of allowing easy construction of confidence intervals for a model characteristic that is a nonlinear function of the model parameters.
- the characteristic AR polynomial of the fitted AR(3) model for the hare data admits a pair of complex roots. Indeed, the roots are $0.84 \pm 0.647i$ and -2.26
- The two complex roots can be written in polar form: $1.06\exp(\pm 0.657i)$.
- As in the discussion of the quasi-period for the AR(2) model, the quasi-period of the fitted AR(3) model can be defined as $2\pi/0.657 = 9.57$. Thus, the fitted model suggests that the hare abundance underwent cyclical fluctuation with a period of about 9.57 years.

Bootstrapping ARIMA Models

- The interesting question of constructing a 95% confidence interval for the quasi-period could be studied using the delta method.
- However, this will be quite complex, as the quasi-period is a complicated function of the parameters.
- the bootstrap provides a simple solution: For each set of bootstrap parameter estimates, we can compute the quasi-period and hence obtain the bootstrap distribution of the quasi-period.
- Confidence intervals for the quasi-period can then be constructed using the percentile method, and the shape of the distribution can be explored via the histogram of the bootstrap quasi-period estimates

Bootstrapping ARIMA Models

- Note that the quasi-period will be undefined whenever the roots of the AR characteristic equation are all real numbers.
- Among the 1000 stationary bootstrap time series obtained by simulating from the fitted model with the errors drawn randomly from the residuals with replacement, 952 series lead to successful full maximum likelihood estimation.
- All but one of the 952 series have well-defined quasi-periods

Bootstrapping ARIMA Models

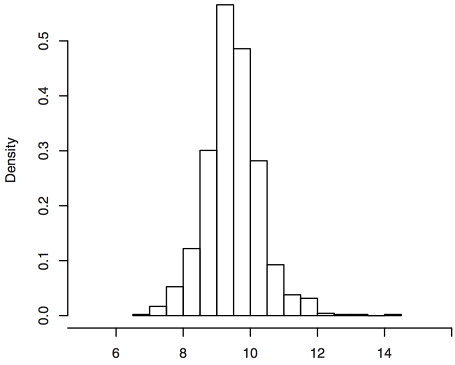

Figure: Histogram of Bootstrap Quasi-period Estimates

- The histogram shows that the sampling distribution of the quasi-period estimate is slightly skewed to the right.

Bootstrapping ARIMA Models

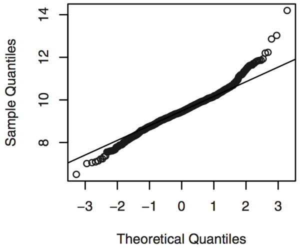

Figure: Q-Q Normal Plot of Bootstrap Quasi-period Estimates

- The Q-Q normal plot suggests that the quasi-period estimator has, furthermore, a thick-tailed distribution.
- The delta method and the normal distribution approximation may be inappropriate
- Finally, using the percentile method, a 95% confidence interval of the quasi-period is found to be (7.84,11.34).

Summary

- we delved into the estimation of the parameters of ARIMA models.
- We considered estimation criteria based on the method of moments, various types of least squares, and maximizing the likelihood function.
- The properties of the various estimators were given, and the estimators were illustrated both with simulated and actual time series data.

Questions?

Delta Method

In statistics, the delta method is a result concerning the approximate probability distribution for a function of an asymptotically normal statistical estimator from knowledge of the limiting variance of that estimator.

Delta Method

if there is a sequence of random variables X_n satisfying

$$\sqrt{n}[X_n - \theta] \to_d \mathcal{N}(0, \sigma^2)$$

Then for any function g satisfying the property that $g'(\theta)$ exists and non-zero, then

$$\sqrt{n}[g(X_n) - g(\theta)] \to_d \mathcal{N}(0, \sigma^2[g'(\theta)]^2)$$