Introduction to Time Series Analysis. Lecture 13.

Peter Bartlett

Last lecture:

1. Parameter estimation
2. Maximum likelihood estimator
3. Yule-Walker estimation

Introduction to Time Series Analysis. Lecture 13.

1. Yule-Walker estimators
2. Large-sample distribution of Yule-Walker estimators
3. Yule-Walker example

Review: Yule-Walker estimation

Method of moments: We choose parameters for which the moments are equal to the empirical moments.

In this case, we choose ϕ so that $\gamma = \hat{\gamma}$.

Yule-Walker equations for $\hat{\phi}$:

$$\begin{cases} \hat{\Gamma}_p \hat{\phi} = \hat{\gamma}_p, \\ \hat{\sigma}^2 = \hat{\gamma}(0) - \hat{\phi}' \hat{\gamma}_p. \end{cases}$$

These are the forecasting equations.
We can use the Durbin-Levinson algorithm.

Yule-Walker estimation: Confidence intervals

If $\{X_t\}$ is an AR(p) process,

$$\hat{\phi} \sim AN\left(\phi, \frac{\sigma^2}{n}\Gamma_p^{-1}\right), \qquad \hat{\sigma}^2 \xrightarrow{P} \sigma^2.$$

$$\hat{\phi}_{hh} \sim AN\left(0, \frac{1}{n}\right) \quad \text{for } h > p.$$

Thus, we can use the sample PACF to test for AR order, and we can calculate approximate confidence intervals for the parameters ϕ.

Yule-Walker estimation: Confidence intervals

If $\{X_t\}$ is an AR(p) process, and n is large,

- $\sqrt{n}(\hat{\phi}_p - \phi_p)$ is approximately $N(0, \hat{\sigma}^2 \hat{\Gamma}_p^{-1})$,
- with probability $\approx 1 - \alpha$, ϕ_{pj} is in the interval

$$\hat{\phi}_{pj} \pm \Phi_{1-\alpha/2} \frac{\hat{\sigma}}{\sqrt{n}} \left(\hat{\Gamma}_p^{-1} \right)_{jj}^{1/2},$$

where $\Phi_{1-\alpha/2}$ is the $1 - \alpha/2$ quantile of the standard normal.

Yule-Walker estimation: Confidence intervals

- with probability $\approx 1-\alpha$, ϕ_p is in the ellipsoid

$$\left\{\phi \in \mathbb{R}^p : \left(\hat{\phi}_p - \phi\right)' \hat{\Gamma}_p \left(\hat{\phi}_p - \phi\right) \leq \frac{\hat{\sigma}^2}{n} \chi^2_{1-\alpha}(p)\right\},$$

where $\chi^2_{1-\alpha}(p)$ is the $(1-\alpha)$ quantile of the chi-squared with p degrees of freedom.

To see this, notice that

$$\text{Var}\left(\Gamma_p^{1/2}(\hat{\phi}_p - \phi_p)\right) = \Gamma_p^{1/2}\,\text{var}(\hat{\phi}_p - \phi_p)\Gamma_p^{1/2} = \frac{\sigma_w^2}{n} I.$$

$$\text{Thus,} \quad v = \Gamma_p^{1/2}(\hat{\phi}_p - \phi_p) \sim N(0, \hat{\sigma}_w^2/nI)$$

$$\text{and so} \quad \frac{n}{\hat{\sigma}_w} v'v \sim \chi^2(p).$$

Yule-Walker estimation

It is also possible to define analogous estimators for ARMA(p,q) models with $q > 0$:

$$\hat{\gamma}(j) - \phi_1 \hat{\gamma}(j-1) - \cdots - \phi_p \hat{\gamma}(j-p) = \sigma^2 \sum_{i=j}^{q} \theta_i \psi_{i-j},$$

where $\psi(B) = \theta(B)/\phi(B)$.

Because of the dependence on the ψ_i, these equations are nonlinear in ϕ_i, θ_i.

There might be no solution, or nonunique solutions.

Also, the *asymptotic efficiency* of this estimator is poor: it has unnecessarily high variance.

Efficiency of estimators

Let $\hat{\phi}^{(1)}$ and $\hat{\phi}^{(2)}$ be two estimators. Suppose that

$$\hat{\phi}^{(1)} \sim AN(\phi, \sigma_1^2), \qquad \hat{\phi}^{(2)} \sim AN(\phi, \sigma_2^2).$$

The asymptotic efficiency of $\hat{\phi}^{(1)}$ relative to $\hat{\phi}^{(2)}$ is

$$e\left(\phi, \hat{\phi}^{(1)}, \hat{\phi}^{(2)}\right) = \frac{\sigma_2^2}{\sigma_1^2}.$$

If $e\left(\phi, \hat{\phi}^{(1)}, \hat{\phi}^{(2)}\right) \leq 1$ for all ϕ, we say that $\hat{\phi}^{(2)}$ is a *more efficient* estimator of ϕ than $\hat{\phi}^{(1)}$.

For example, for an AR(p) process, the moment estimator and the maximum likelihood estimator are as efficient as each other.

For an MA(q) process, the moment estimator is less efficient than the innovations estimator, which is less efficient than the MLE.

Yule Walker estimation: Example

AR(1):

$$\gamma(0) = \frac{\sigma^2}{1-\phi_1^2}$$

$$\hat{\phi}_1 \sim AN\left(\phi_1, \frac{\sigma^2}{n}\Gamma_1^{-1}\right) = AN\left(\phi_1, \frac{1-\phi_1^2}{n}\right).$$

AR(2):

$$\begin{pmatrix} \hat{\phi}_1 \\ \hat{\phi}_2 \end{pmatrix} \sim AN\left(\begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}, \frac{\sigma^2}{n}\Gamma_2^{-1}\right)$$

and

$$\frac{\sigma^2}{n}\Gamma_2^{-1} = \frac{1}{n}\begin{pmatrix} 1-\phi_2^2 & -\phi_1(1+\phi_2) \\ -\phi_1(1+\phi_2) & 1-\phi_2^2 \end{pmatrix}.$$

Yule Walker estimation: Example

Suppose $\{X_t\}$ is an AR(1) process and the sample size n is large.

If we estimate ϕ, we have

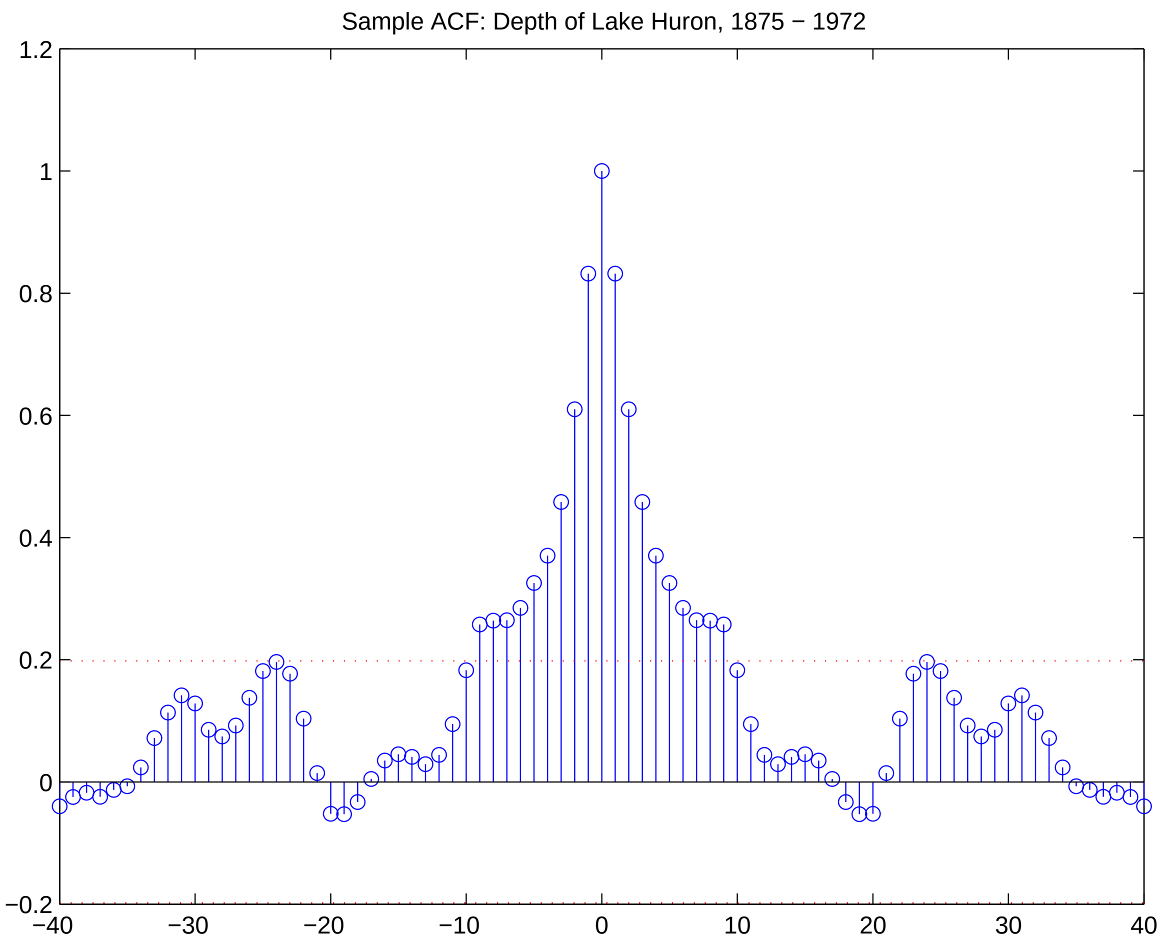
$$\mathrm{Var}(\hat{\phi}_1) \approx \frac{1-\phi_1^2}{n}.$$

If we fit a *larger* model, say an AR(2), to this AR(1) process,

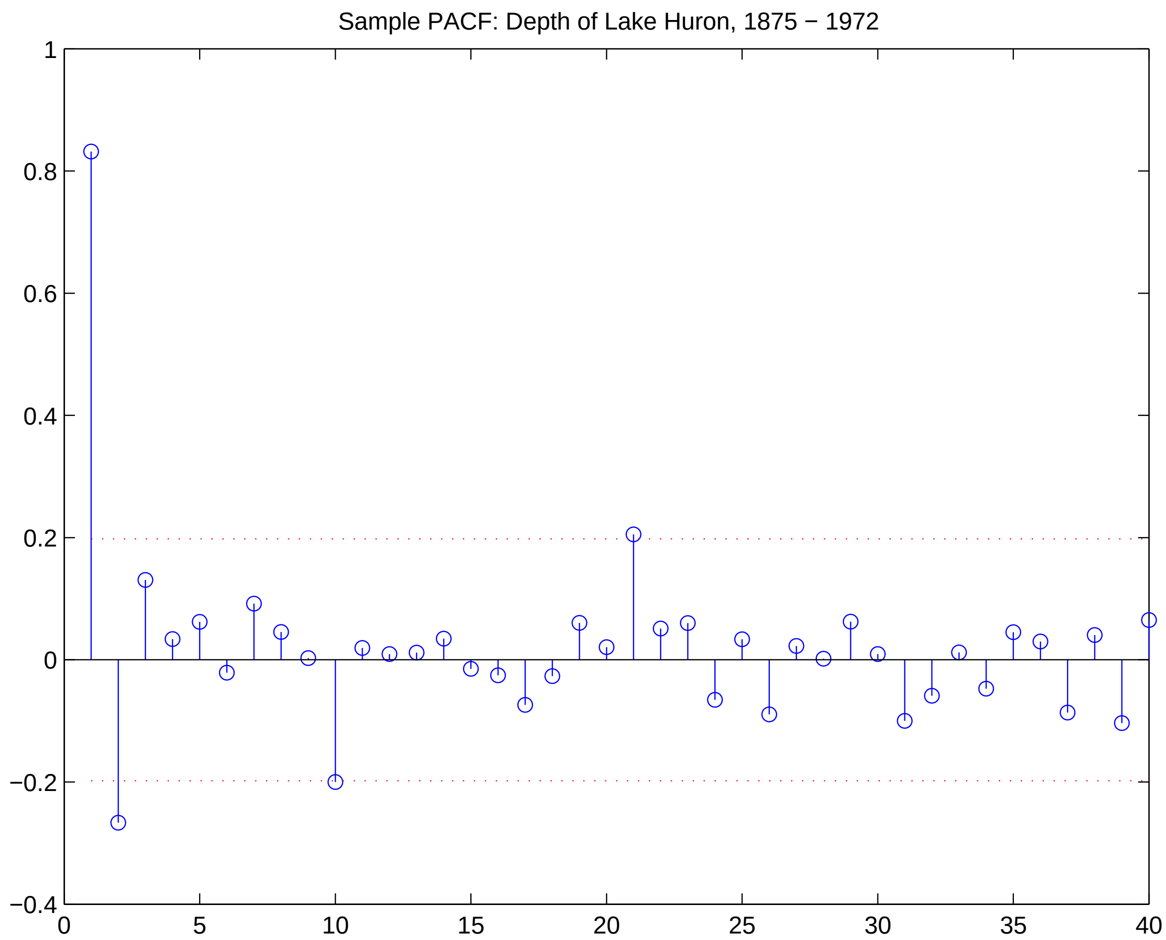
$$\mathrm{Var}(\hat{\phi}_1) \approx \frac{1-\phi_2^2}{n} = \frac{1}{n} \quad > \quad \frac{1-\phi_1^2}{n}.$$

We have lost efficiency.

Yule Walker estimation: Example



Yule Walker estimation: Example



Yule Walker estimation: Example

$$\hat{\Gamma}_2 = \begin{pmatrix} 1.7379 & 1.4458 \\ 1.4458 & 1.7379 \end{pmatrix} \qquad \hat{\gamma}_2 = \begin{pmatrix} 1.4458 \\ 1.0600 \end{pmatrix}$$

$$\hat{\phi}_2 = \hat{\Gamma}_2^{-1}\hat{\gamma}_2 = \begin{pmatrix} 1.0538 \\ -0.2668 \end{pmatrix}$$

$$\hat{\sigma}_w^2 = \hat{\gamma}(0) - \hat{\phi}_2'\hat{\gamma}_2 = 0.4971$$

Yule Walker estimation: Example

Confidence intervals:

$$\hat{\phi}_1 \pm \Phi_{1-\alpha/2} \left(\hat{\sigma}_w^2 \hat{\Gamma}_2^{-1}/n\right)_{11}^{1/2} = 1.0538 \pm 0.1908$$

$$\hat{\phi}_2 \pm \Phi_{1-\alpha/2} \left(\hat{\sigma}_w^2 \hat{\Gamma}_2^{-1}/n\right)_{22}^{1/2} = -0.2668 \pm 0.1908$$