Handout 7

Forecasting for ARMA and ARIMA Models

Class notes for Statistics 451: Applied Time Series
Iowa State University

Copyright 2004 W. Q. Meeker.

March 2, 2006
12h 42min

Forecasting Future Values From an ARIMA Model

The ARIMA model for Z_t is

$$\phi_p(\mathrm{B})Z_t = \theta_0 + \theta_q(\mathrm{B})a_t$$

or, in the unscrambled form

$$Z_t = \theta_0 + \phi_1 Z_{t-1} + \cdots + \phi_p Z_{t-p} - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q} + a_t$$

Assume that we have data $Z_1, Z_2, \ldots, Z_n$ and we want to forecast Z_{n+l}, (i.e., l steps ahead from forecast origin n). Then the actual value is

$$Z_{n+l} = \theta_0 + \phi_1 Z_{n+l-1} + \cdots + \phi_p Z_{n+l-p} - \theta_1 a_{n+l-1} - \cdots - \theta_q a_{n+l-q} + a_{n+l}$$

The "minimum mean square error" forecast for Z_{n+l} is

$$\widehat{Z}_n(l) = \theta_0 + \phi_1[Z_{n+l-1}] + \cdots + \phi_p[Z_{n+l-p}] - \theta_1[a_{n+l-1}] - \cdots - \theta_q[a_{n+l-q}]$$

For quantities inside [ ], substitute value if known, forecast if unknown: $\widehat{Z}_n(l-k)$ for Z_{n+l-k} and 0 for a_{n+l-k}.

ARMA Model Forecast Equation in Infinite MA Form

The ARMA model for Z_t is

$$Z_t = \theta_0' + \psi_\infty(\mathrm{B})a_t$$

where $\theta_0' = \mathrm{E}(Z_t) = \theta_0/(1 - \phi_1 - \cdots - \phi_p)$. Expanding $\psi_\infty(\mathrm{B})$ gives

$$Z_t = \theta_0' + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \psi_3 a_{t-3} + \cdots + a_t$$

Assume that we have data $Z_1, Z_1, \ldots, Z_n$ and we want to forecast Z_{n+l}, (i.e., l steps ahead from forecast origin n). Then the actual value is

$$Z_{n+l} = \theta_0' + \psi_1 a_{n+l-1} + \psi_2 a_{n+l-2} + \psi_3 a_{n+l-3} + \cdots + a_{n+l}$$

The "minimum mean square error" forecast for Z_{n+l} is

$$\widehat{Z}_n(l) = \theta_0' + \psi_1[a_{n+l-1}] + \psi_2[a_{n+l-2}] + \psi_3[a_{n+l-3}] + \cdots$$

This form is not very useful for computing forecasts, but is useful in finding a simple expression for the forecast error.

Forecast Error and Forecast Error Variances

One step-ahead ($l = 1$):

$$\begin{aligned}
Z_{n+1} &= \theta_0' + \psi_1 a_n + \psi_2 a_{n-1} + \psi_3 a_{n-2} + \cdots + a_{n+1} \\
\widehat{Z}_n(1) &= \theta_0' + \psi_1 a_n + \psi_2 a_{n-1} + \psi_3 a_{n-2} + \cdots \\
e_n(1) &= Z_{n+1} - \widehat{Z}_n(1) = a_{n+1}. \\
\mathrm{Var}[e_n(1)] &= \mathrm{Var}(a_{n+1}) = \sigma_a^2
\end{aligned}$$

Two steps-ahead ($l = 2$):

$$\begin{aligned}
Z_{n+2} &= \theta_0' + \psi_1 a_{n+1} + \psi_2 a_n + \psi_3 a_{n-1} + \cdots + a_{n+2} \\
\widehat{Z}_n(2) &= \theta_0' + \psi_1[a_{n+1}] + \psi_2 a_n + \psi_3 a_{n-1} + \cdots \\
&= \theta_0' + \quad 0 \quad + \psi_2 a_n + \psi_3 a_{n-1} + \cdots \\
e_n(2) &= Z_{n+2} - \widehat{Z}_n(2) = a_{n+2} + \psi_1 a_{n+1} \\
\mathrm{Var}[e_n(2)] &= \mathrm{Var}(a_{n+2} + \psi_1 a_{n+1}) = \mathrm{Var}(a_{n+2}) + \psi_1^2 \mathrm{Var}(a_{n+1}) \\
&= \sigma_a^2(1 + \psi_1^2).
\end{aligned}$$

Forecast Error and Forecast Error Variances

Three steps-ahead ($l = 3$):

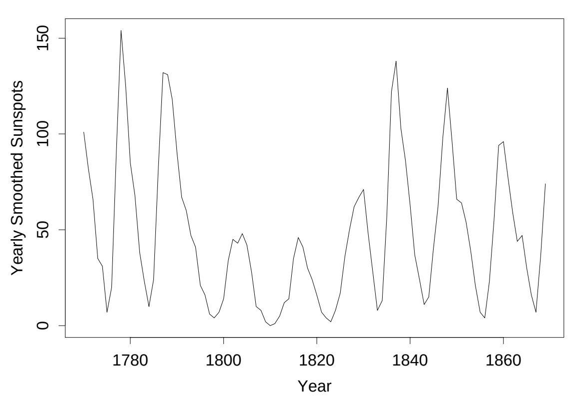
$$\begin{aligned}
Z_{n+3} &= \theta_0' + \psi_1 a_{n+2} + \psi_2 a_{n+1} + \psi_3 a_n + \cdots + a_{n+3} \\
\widehat{Z}_n(3) &= \theta_0' + \psi_1[a_{n+2}] + \psi_2[a_{n+1}] + \psi_3 a_n + \cdots \\
&= \theta_0' + \quad 0 \quad + \quad 0 \quad + \psi_3 a_n + \cdots \\
e_n(3) &= Z_{n+3} - \widehat{Z}_n(3) = a_{n+3} + \psi_1 a_{n+2} + \psi_2 a_{n+1} \\
\mathrm{Var}[e_n(3)] &= \mathrm{Var}(a_{n+3} + \psi_1 a_{n+2} + \psi_2 a_{n+1}) = \sigma_a^2(1 + \psi_1^2 + \psi_2^2).
\end{aligned}$$

In general, l steps-ahead:

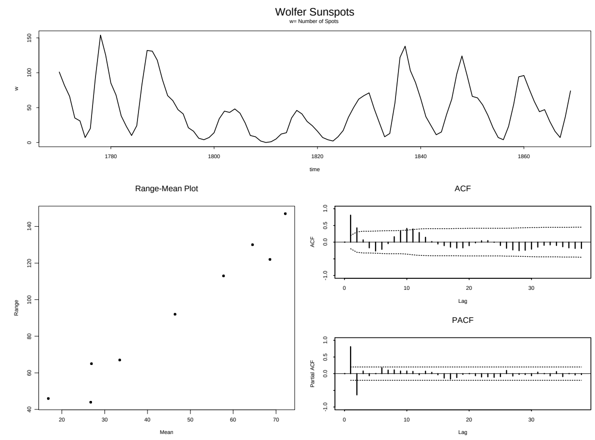
$$\begin{aligned}
\widehat{Z}_n(l) &= \theta_0' + 0 + 0 + \cdots + \psi_l a_n + \psi_{l+1} a_{n-1} + \cdots \\
e_n(l) &= Z_{n+l} - \widehat{Z}_n(l) = a_{n+l} + \psi_1 a_{n+l-1} + \cdots + \psi_{l-1} a_{n+1}
\end{aligned}$$

$$\begin{aligned}
\mathrm{Var}[e_n(l)] &= \sigma_a^2(1 + \psi_1^2 + \cdots + \psi_{l-1}^2) \\
&= \sigma_a^2 \sum_{i=0}^{l-1} \psi_i^2, \quad \text{where } \psi_0 \equiv 1
\end{aligned}$$

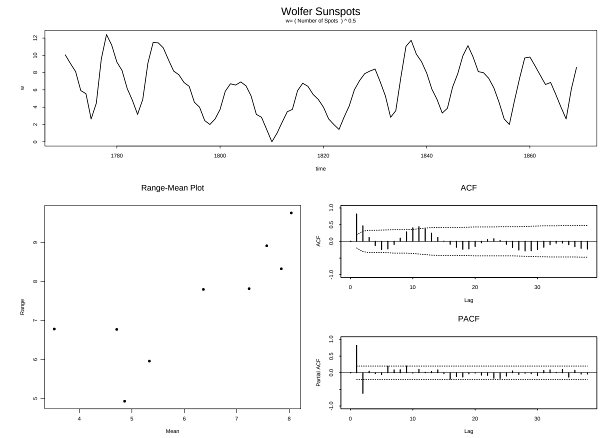
Wolfer Sunspot Numbers 1770-1869



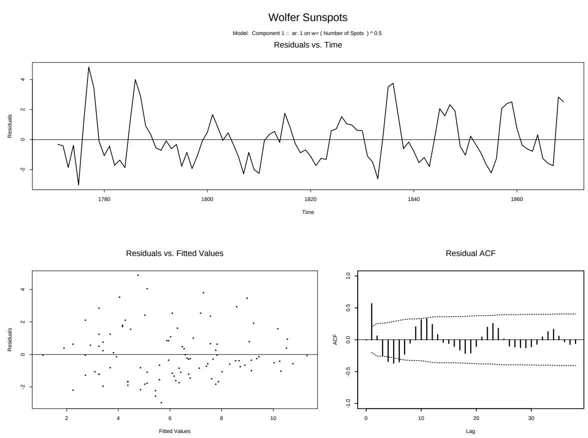
Function `iden` Output for the Wolfer Sunspot Numbers 1770-1869



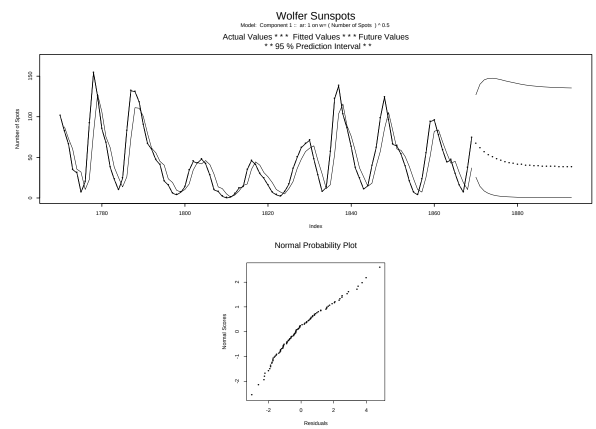
Function `iden` Output based on the Square Roots of the Wolfer Sunspot Numbers 1770-1869



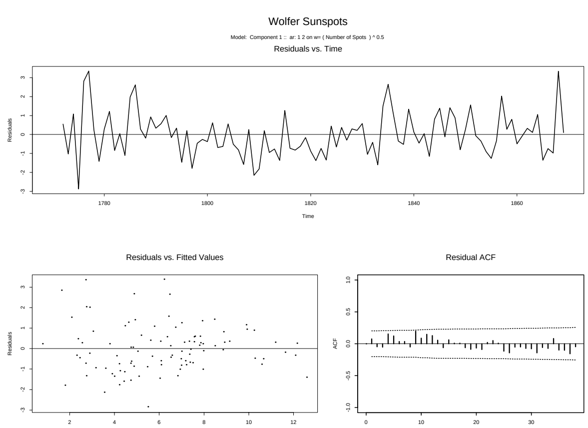
Function `esti` Output based on the Square Roots of the Wolfer Sunspot Numbers 1770-1869 AR(1) Model—Part 1



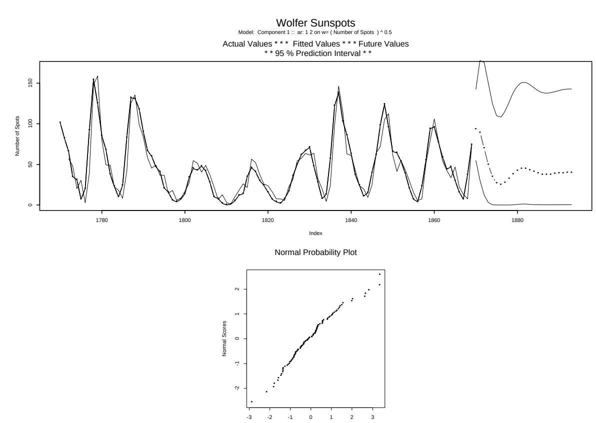
Function `esti` Output based on the Square Roots of the Wolfer Sunspot Numbers 1770-1869 AR(1) Model—Part 2



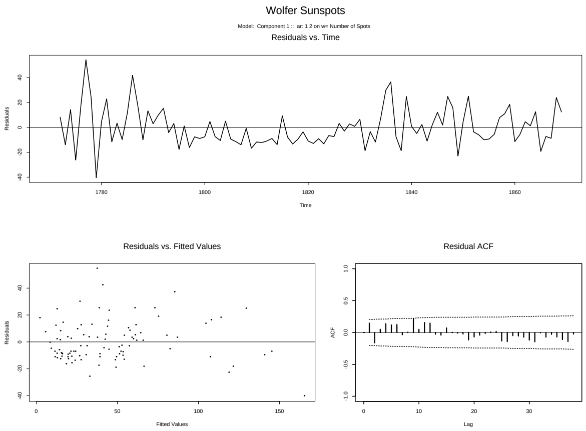
Function `esti` Output based on the Square Roots of the Wolfer Sunspot Numbers 1770-1869 AR(2) Model—Part 1



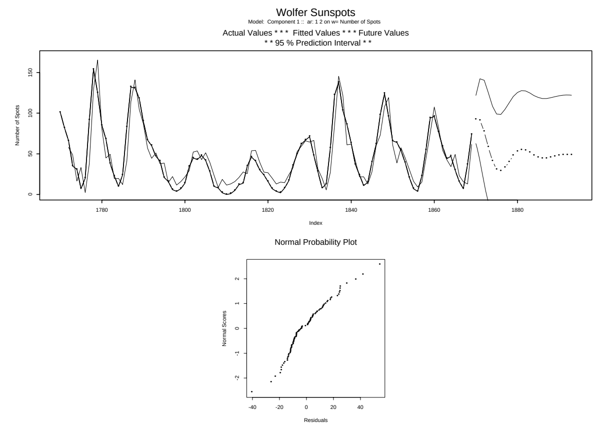
Function `esti` Output based on the Square Roots of the Wolfer Sunspot Numbers 1770-1869 AR(2) Model—Part 2



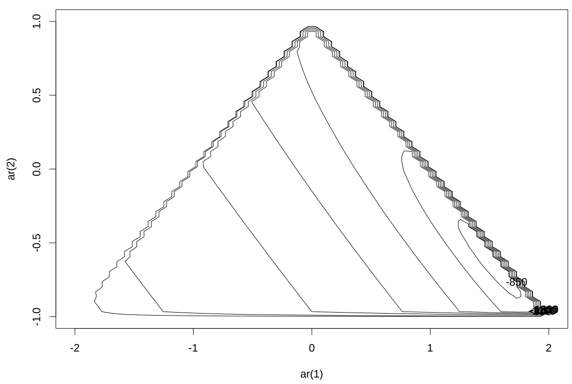
Function esti Output for the Wolfer Sunspot Numbers 1770-1869 AR(2) Model—Part 1



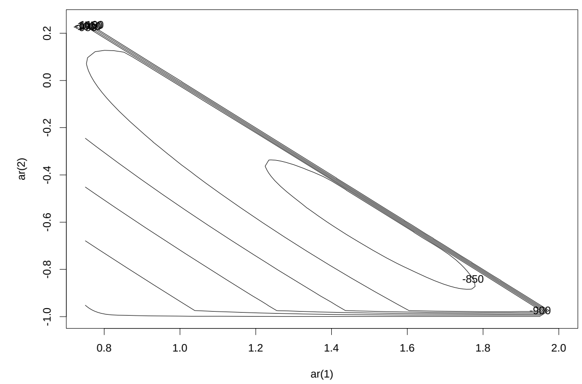
Function esti Output for the Wolfer Sunspot Numbers 1770-1869 AR(2) Model—Part 2



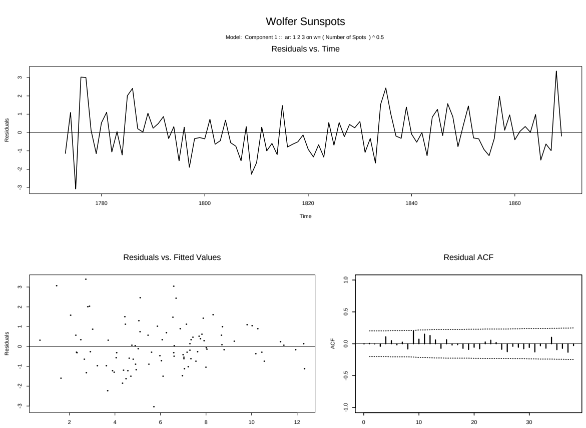
Plot of AR(2) Model Log-likelihood Surface for the Square Root Wolfer Sunspot Data



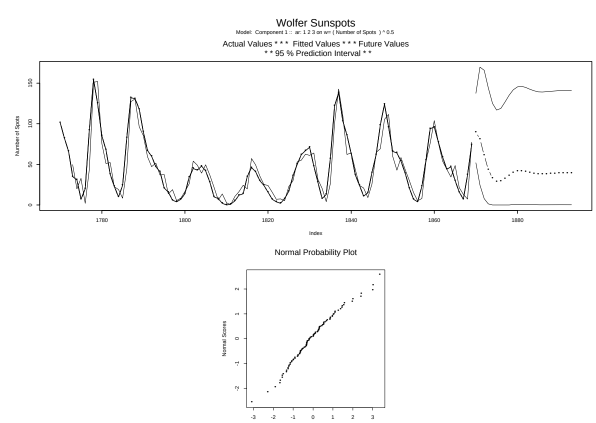
Plot of AR(2) Model Log-likelihood Surface for the Square Root Wolfer Sunspot Data



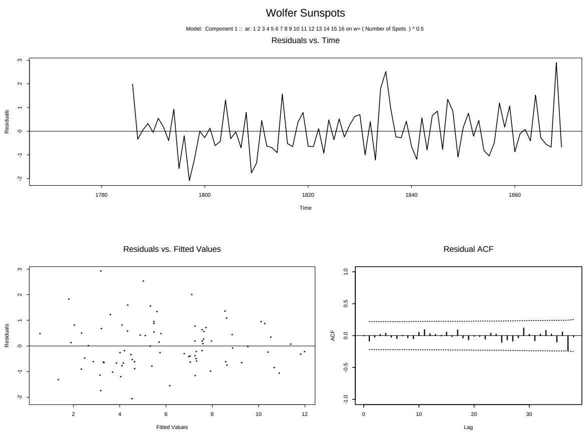
Function esti Output based on the Square Roots of the Wolfer Sunspot Numbers 1770-1869 AR(3) Model—Part 1



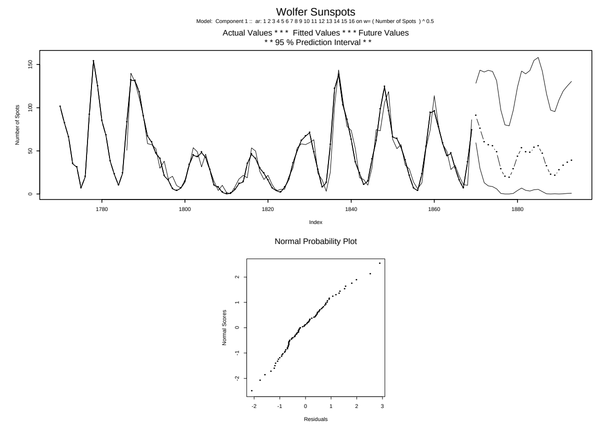
Function esti Output based on the Square Roots of the Wolfer Sunspot Numbers 1770-1869 AR(3) Model—Part 2



Function `esti` Output based on the Square Roots of the Wolfer Sunspot Numbers 1770-1869 AR(16) Model—Part 1



Function `esti` Output based on the Square Roots of the Wolfer Sunspot Numbers 1770-1869 AR(16) Model—Part 2



Prediction Interval for Z_{n+l}

A 95% prediction interval for Z_{n+l} (l steps ahead) is

$$\widehat{Z}_n(l) \pm 1.96\sqrt{\text{Var}[e_n(l)]}$$

For one step-ahead the simplifies to

$$\widehat{Z}_n(1) \pm 1.96\sigma_a$$

For two steps-ahead the simplifies to

$$\widehat{Z}_n(2) \pm 1.96\sigma_a\sqrt{1+\psi_1^2}$$

For three steps-ahead the simplifies to

$$\widehat{Z}_n(3) \pm 1.96\sigma_a\sqrt{1+\psi_1^2+\psi_2^2}$$

and so on. When computing prediction intervals from data, we substitute estimates for parameters, giving approximate prediction intervals (i.e., substitute $\widehat{\phi}_1$ for $\phi_1, \ldots, \widehat{\phi}_p$ for $\phi_p, \widehat{\theta}_1$ for $\theta_1, \ldots \widehat{\theta}_q$ for θ_q, and $\widehat{\sigma}_a = S$ for σ_a).

Eventual (Long-run) Forecasts

For stationary time series, from

$$\widehat{Z}_n(l) = \theta_0' + 0 + 0 + \cdots + \psi_l a_n + \psi_{l+1}a_{n-1} + \psi_{l+2}a_{n-2} + \cdots$$

because the ψ weights die down, the long-run forecast is

$$\lim_{l\to\infty} \widehat{Z}_n(l) = \theta_0' = \text{E}(Z_t)$$

and from

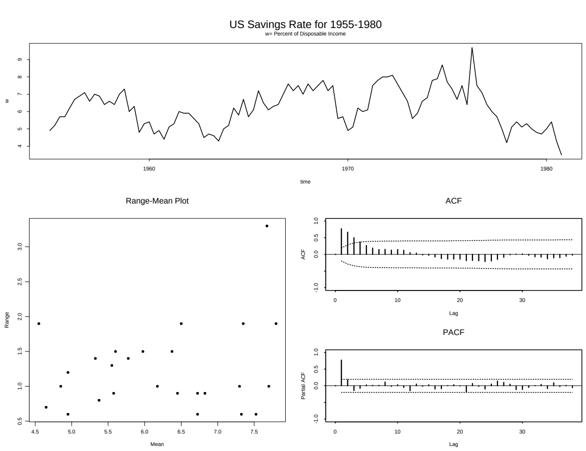
$$e_n(l) = Z_{n+l} - \widehat{Z}_n(l) = a_{n+l} + \psi_1 a_{n+l-1} + \cdots + \psi_{l-1}a_{n+1}$$

we can see that because

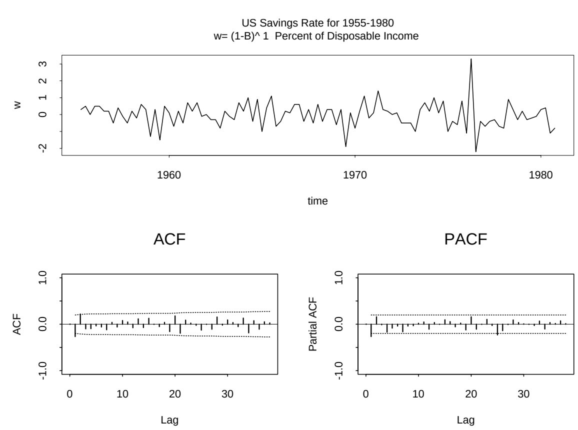
$$\begin{aligned} \text{Var}[e_n(l)] &= \sigma_a^2(1+\psi_1^2+\cdots+\psi_{l-1}^2), \\ \lim_{l\to\infty} \text{Var}[e_n(l)] &= \sigma_a^2(1+\psi_1^2+\psi_2^2+\cdots) = \text{Var}(Z_t) \end{aligned}$$

For nonstationary time series, things are more complicated, but the forecast-error variance grows without bound because the ψ weights do not die down.

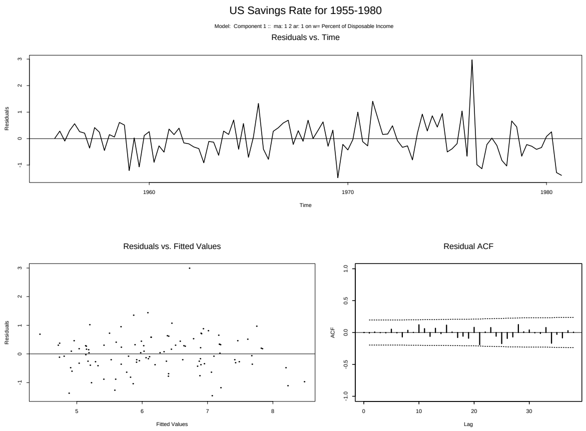
Function `iden` Output for the Savings Rate Data



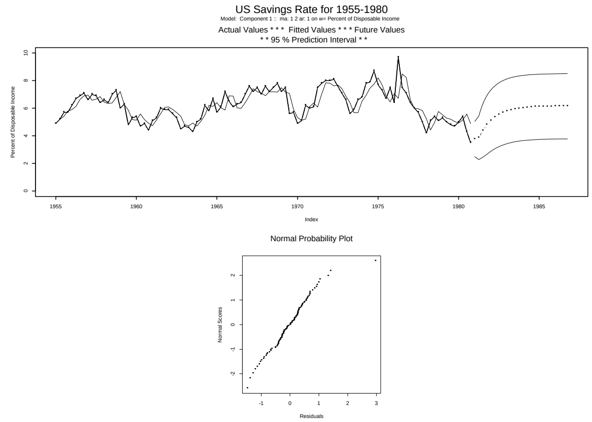
Function `iden` Output for the First Differences of the Savings Rate Data



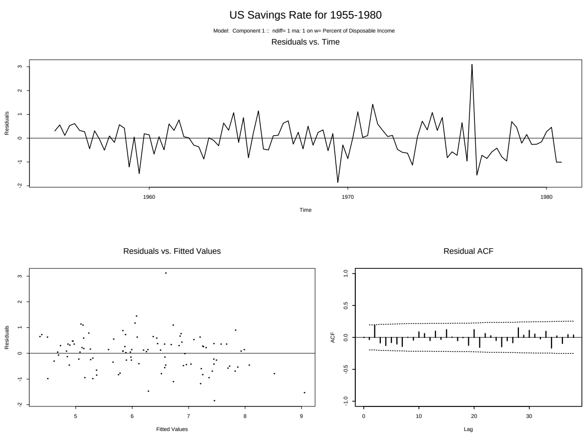
Function `esti` Output for the Savings Rate Data ARMA(1,2) Model—Part 1



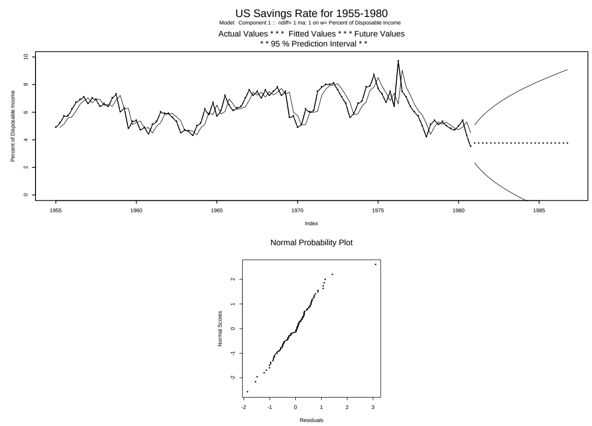
Function `esti` Output for the Savings Rate Data ARMA(1,2) Model—Part 2



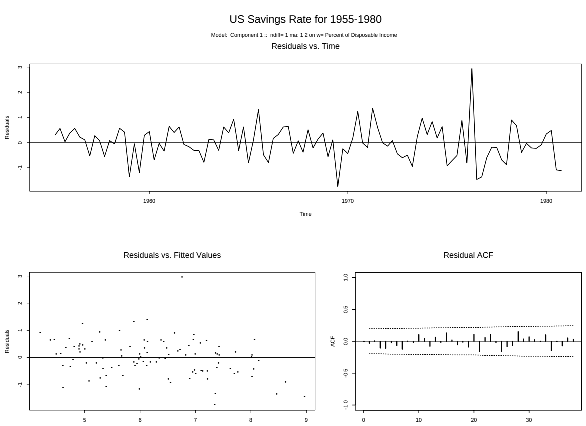
Function `esti` Output for the Savings Rate Data ARIMA(0,1,1) Model—Part 1



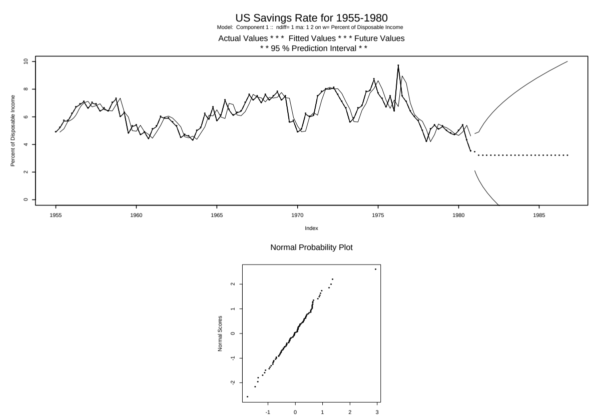
Function `esti` Output for the Savings Rate Data ARIMA(0,1,1) Model—Part 2



Function `esti` Output for the Savings Rate Data ARIMA(0,1,2) Model—Part 1



Function `esti` Output for the Savings Rate Data ARIMA(0,1,2) Model—Part 2



Reasons Needing a Long Realization

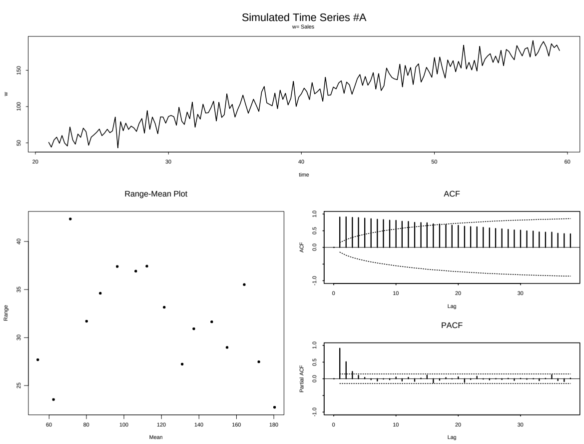
- Estimate correlation structure (i.e., the ACF and PACF) functions and get accurate standard errors.
- Estimate seasonal pattern (need at least 4 or 5 seasonal periods).
- Approximate prediction intervals assume that parameters are known (good approximation if realization is large).
- Fewer estimation problems (likelihood function better behaved).
- Possible to check forecasts by withholding recent data .
- Can check model stability by dividing data and analyzing both sides.

Reasons For Using a Parsimonious Model

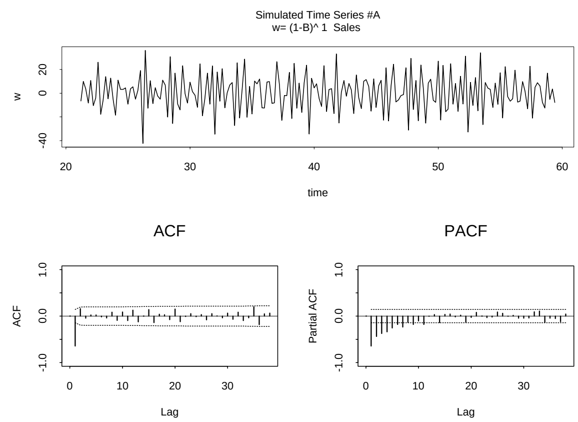
- Fewer numerical problems in estimation.
- Easier to understand the model.
- With fewer parameters, forecasts less sensitive to deviations between parameters and estimates.
- Model may applied more generally to similar processes.
- Rapid real-time computations for control or other action.

Having a parsimonious model is less important if the realization is large.

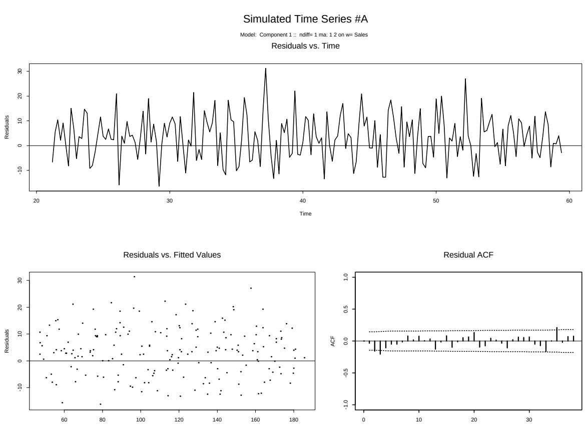
Function `iden` Output for Simulated Series A



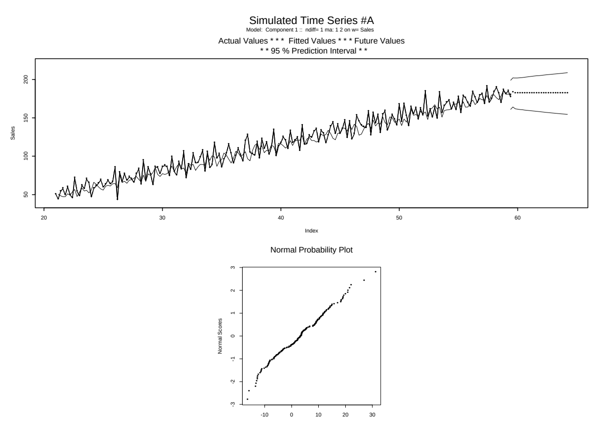
Function `iden` Output for the First Differences of Simulated Series A



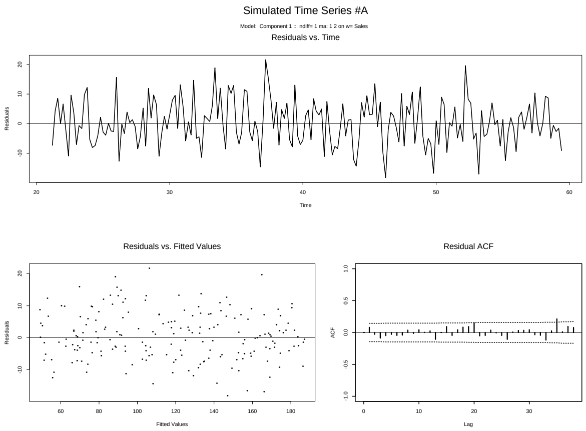
Function `esti` Output for Simulated Series A IMA(1,2) Model—Part 1



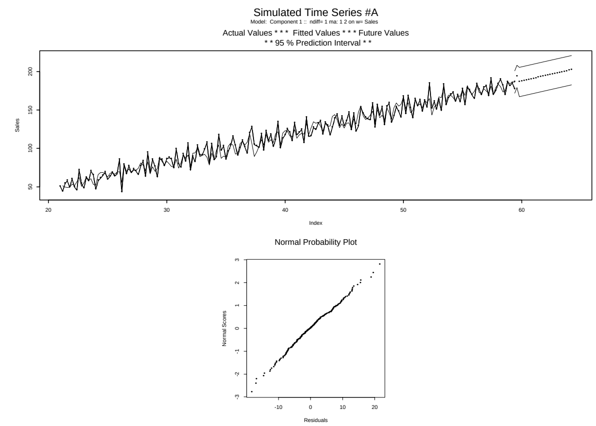
Function `esti` Output for for Simulated Series A IMA(1,2) Model—Part 2



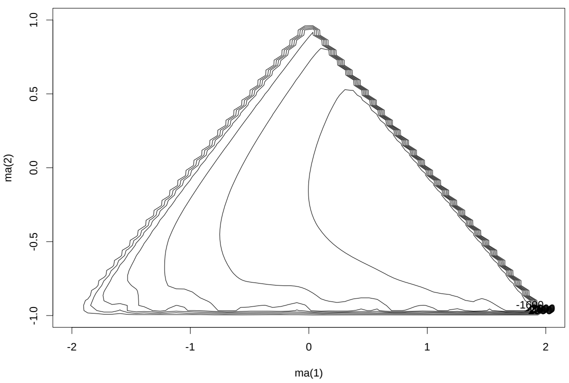
Function `esti` Output for Simulated Series A IMA(1,2) Model with Deterministic Trend—Part 1



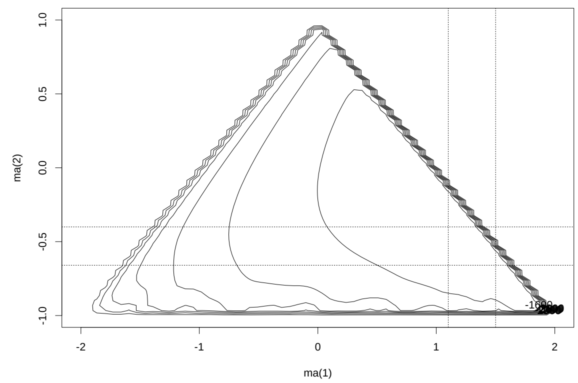
Function `esti` Output for for Simulated Series A IMA(1,2) Model with Deterministic Trend—Part 2



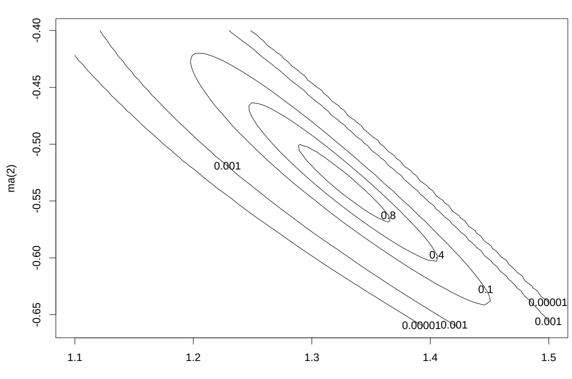
Plot of IMA(1,2) Model Log-likelihood Surface for Simulated Series A



Plot of IMA(1,2) Model Log-likelihood Surface for Simulated Series A



Plot of IMA(1,2) Model Relative Likelihood Surface for Simulated Series A—Close-up View



Perspective Plot of IMA(1,2) Model Relative Likelihood Surface for Simulated Series A—Close-up View