ACF/PACF Estimation and AR/MA models

Maurice Stevenson Bartlett F.R.S.
(1910-2002)

President of the Royal Statistical Society 1966-1967

\[
\text{var}(\hat{\rho}_k) \approx \frac{1}{n} \sum_{i=1}^{\infty} \left( \rho_i^2 + \rho_{i+k} \rho_{i-k} - 4 \rho_k \rho_i \rho_{i-k} + 2 \rho_k^2 \rho_i^2 \right)
\]

\[
\approx \frac{1}{n} \left( 1 + 2 \rho_1^2 + 2 \rho_2^2 + \cdots + 2 \rho_m^2 \right)
\]
Outline

1 §2.5 (cont): ACF & PACF Estimation

2 §2.6 MA(∞) and AR(∞) Representations
Autocovariance Estimation

The sample autocovariance function.

\[
\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-h} (Z_{t+h} - \bar{Z})(Z_t - \bar{Z})
\]

And compare \( \hat{\gamma}(h) \) with \( \gamma(h) \).

\[
\gamma(h) = \text{E} \left[ (Z_{t+h} - \mu)(Z_t - \mu) \right]
\]

The following estimator is slightly less biased, but its variance is often larger.

\[
\hat{\gamma}(h) = \frac{1}{n-h} \sum_{t=1}^{n-h} (Z_{t+h} - \bar{Z})(Z_t - \bar{Z})
\]
Autocovariance Estimation

The sample autocovariance function.

\[ \hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-h} (Z_{t+h} - \bar{Z})(Z_t - \bar{Z}) \]

And compare \( \hat{\gamma}(h) \) with \( \gamma(h) \).

\[ \gamma(h) = E \left[ (Z_{t+h} - \mu)(Z_t - \mu) \right] \]

The following estimator is slightly less biased, but it’s variance is often larger.

\[ \hat{\gamma}(h) = \frac{1}{n-h} \sum_{t=1}^{n-h} (Z_{t+h} - \bar{Z})(Z_t - \bar{Z}) \]
Asymptotics of $\hat{\gamma}$

When the process $\{Z_t\}$ is Gaussian, Bartlett (1946) derived the following asymptotics of $\hat{\gamma}_k$:

$$\text{cov}(\hat{\gamma}_k, \hat{\gamma}_{k+j}) \approx \frac{1}{n} \sum_{i=1}^{\infty} (\gamma_i \gamma_{i+j} + \gamma_{i+k+j} \gamma_{i-k})$$

$$\text{var}(\hat{\gamma}_k) \approx \frac{1}{n} \sum_{i=-\infty}^{\infty} (\gamma_i^2 + \gamma_{i+k} \gamma_{i-k})$$

To guarantee $\hat{\gamma} \to 0$, the following condition is imposed:

$$\sum_{-\infty}^{\infty} |\gamma_i| < \infty$$
When the process \( \{ Z_t \} \) is Gaussian, Bartlett (1946) derived the following asymptotics of \( \hat{\gamma}_k \):

\[
\text{cov}(\hat{\gamma}_k, \hat{\gamma}_{k+j}) \approx \frac{1}{n} \sum_{i=1}^{\infty} (\gamma_i \gamma_{i+j} + \gamma_{i+k+j} \gamma_{i-k})
\]

\[
\text{var}(\hat{\gamma}_k) \approx \frac{1}{n} \sum_{i=-\infty}^{\infty} (\gamma_i^2 + \gamma_i \gamma_{i+k} \gamma_{i-k})
\]

To guarantee \( \hat{\gamma} \to 0 \), the following condition is imposed:

\[
\sum_{-\infty}^{\infty} |\gamma_i| < \infty
\]
Autocorrelation Estimation and Asymptotics

The sample autocorrelation function is defined by

\[
\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{\sum_{t=1}^{n-k} (Z_t - \bar{Z})(Z_{t+k} - \bar{Z})}{\sum_{t=1}^{n} (Z_t - \bar{Z})^2}
\]

with asymptotics given by

\[
\text{cov}(\hat{\rho}_k, \hat{\rho}_{k+j}) \approx \frac{1}{n} \sum_{i=-\infty}^{\infty} (\rho_i \rho_{i+j} + \rho_{i+k+j} \rho_{i-k} - 2 \rho_k \rho_i \rho_{i-k} - 2 \rho_{k+j} \rho_i \rho_{i-k} + 2 \rho_k \rho_{k+j} \rho_i^2)
\]

\[
\text{var}(\hat{\rho}_k) \approx \frac{1}{n} \sum_{i=1}^{\infty} (\rho_i^2 + \rho_{i+k} \rho_{i-k} - 4 \rho_k \rho_i \rho_{i-k} + 2 \rho_k^2 \rho_i^2)
\]

\[
\approx \frac{1}{n} \left( 1 + 2 \rho_1^2 + 2 \rho_2^2 + \cdots + 2 \rho_m^2 \right)
\]

The second variance approximation holds if \(\rho_k = 0\) for \(k > m\).
The sample partial autocorrelation function is computed via the Durbin-Levinson recursive algorithm (1960).

Start with \( \hat{\phi}_{11} = \hat{\rho}_1 \)
then recursively compute

\[
\hat{\phi}_{k+1,k+1} = \frac{\hat{\rho}_{k+1} - \sum_{j=1}^{k} \hat{\phi}_{kj} \hat{\rho}_{k+1-j}}{1 - \sum_{j=1}^{k} \hat{\phi}_{kj} \hat{\rho}_j}
\]

and

\[
\hat{\phi}_{k+1,j} = \hat{\phi}_{kj} - \hat{\phi}_{k+1,k+1} \hat{\phi}_{k,k+1-j} \quad j = 1, \ldots, k
\]

Quenouille (1949) showed that for a white noise process the following approximation holds

\[
\text{var}(\hat{\phi}_{kk}) \approx \frac{1}{n}
\]

Hence \( \pm 2/\sqrt{n} \) can be used as critical limits on \( \hat{\phi}_{kk} \) to test for white noise.
PACF Estimation

The sample partial autocorrelation function is computed via the Durbin-Levinson recursive algorithm (1960).

Start with $\hat{\phi}_{11} = \hat{\rho}_1$

then recursively compute

$$\hat{\phi}_{k+1,k+1} = \frac{\hat{\rho}_{k+1} - \sum_{j=1}^{k} \hat{\phi}_{kj} \hat{\rho}_{k+1-j}}{1 - \sum_{j=1}^{k} \hat{\phi}_{kj} \hat{\rho}_j}$$

and

$$\hat{\phi}_{k+1,j} = \hat{\phi}_{kj} - \hat{\phi}_{k+1,k+1} \hat{\phi}_{k,k+1-j} \quad j = 1, \ldots, k$$

Quenouille (1949) showed that for a white noise process the following approximation holds

$$\text{var}(\hat{\phi}_{kk}) \approx \frac{1}{n}$$

Hence $\pm 2/\sqrt{n}$ can be used as critical limits on $\phi_{kk}$ to test for white noise.
The sample partial autocorrelation function is computed via the Durbin-Levinson recursive algorithm (1960).

Start with $\hat{\phi}_{11} = \hat{\rho}_1$

then recursively compute

$$\hat{\phi}_{k+1,k+1} = \frac{\hat{\rho}_{k+1} - \sum_{j=1}^{k} \hat{\phi}_{kj} \hat{\rho}_{k+1-j}}{1 - \sum_{j=1}^{k} \hat{\phi}_{kj} \hat{\rho}_j}$$

and

$$\hat{\phi}_{k+1,j} = \hat{\phi}_{kj} - \hat{\phi}_{k+1,k+1} \hat{\phi}_{k,k+1-j} \quad j = 1, \ldots, k$$

Quenouille (1949) showed that for a white noise process the following approximation holds

$$\text{var}(\hat{\phi}_{kk}) \approx \frac{1}{n}$$

Hence $\pm 2 / \sqrt{n}$ can be used as critical limits on $\phi_{kk}$ to test for white noise.
PACF Estimation

The sample partial autocorrelation function is computed via the Durbin-Levinson recursive algorithm (1960).

Start with \( \hat{\phi}_{11} = \hat{\rho}_1 \)
then recursively compute

\[
\hat{\phi}_{k+1,k+1} = \frac{\hat{\rho}_{k+1} - \sum_{j=1}^{k} \hat{\phi}_{kj} \hat{\rho}_{k+1-j}}{1 - \sum_{j=1}^{k} \hat{\phi}_{kj} \hat{\rho}_j}
\]

and

\[
\hat{\phi}_{k+1,j} = \hat{\phi}_{kj} - \hat{\phi}_{k+1,k+1} \hat{\phi}_{k,k+1-j} \quad j = 1, \ldots, k
\]

Quenouille (1949) showed that for a white noise process the following approximation holds

\[
\text{var}(\hat{\phi}_{kk}) \approx \frac{1}{n}
\]

Hence \( \pm 2/\sqrt{n} \) can be used as critical limits on \( \phi_{kk} \) to test for white noise.
PACF Estimation

The sample partial autocorrelation function is computed via the Durbin-Levinson recursive algorithm (1960).

Start with $\hat{\phi}_{11} = \hat{\rho}_1$
then recursively compute

$$\hat{\phi}_{k+1,k+1} = \frac{\hat{\rho}_{k+1} - \sum_{j=1}^{k} \hat{\phi}_{kj} \hat{\rho}_{k+1-j}}{1 - \sum_{j=1}^{k} \hat{\phi}_{kj} \hat{\rho}_j}$$

and

$$\hat{\phi}_{k+1,j} = \hat{\phi}_{kj} - \hat{\phi}_{k+1,k+1} \hat{\phi}_{k,k+1-j} \quad j = 1, \ldots, k$$

Quenouille (1949) showed that for a white noise process the following approximation holds

$$\text{var}(\hat{\phi}_{kk}) \approx \frac{1}{n}$$

Hence $\pm 2/\sqrt{n}$ can be used as critical limits on $\phi_{kk}$ to test for white noise.
PACF Estimation

The sample partial autocorrelation function is computed via the Durbin-Levinson recursive algorithm (1960).

Start with $\hat{\phi}_{11} = \hat{\rho}_1$

then recursively compute

\[
\hat{\phi}_{k+1,k+1} = \frac{\hat{\rho}_{k+1} - \sum_{j=1}^{k} \hat{\phi}_{kj}\hat{\rho}_{k+1-j}}{1 - \sum_{j=1}^{k} \hat{\phi}_{kj}\hat{\rho}_j}
\]

and

\[
\hat{\phi}_{k+1,j} = \hat{\phi}_{kj} - \hat{\phi}_{k+1,k+1}\hat{\phi}_{k,k+1-j} \quad j = 1, \ldots, k
\]

Quenouille (1949) showed that for a white noise process the following approximation holds

\[
\text{var}(\hat{\phi}_{kk}) \approx \frac{1}{n}
\]

Hence $\pm 2/\sqrt{n}$ can be used as critical limits on $\hat{\phi}_{kk}$ to test for white noise.
Outline

1. §2.5 (cont): ACF & PACF Estimation
2. §2.6 MA(∞) and AR(∞) Representations
Theorem (Wold Decomposition)

A stationary time series $Z_t$ that is purely nondeterministic can always be expressed in the form

$$Z_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots = \mu + \sum_{j=0}^{\infty} \psi_j a_{t-j}$$

where $\psi_0 = 1$, $\{a_t\}$ is a zero mean white noise process, and $\sum_{j=0}^{\infty} \psi_j^2 < \infty$. Moreover, this decomposition is unique.
Backshift and Difference Operators

Definition (Backshift Operator)
The backshift operator, $B$, is defined by

$$BZ_t = Z_{t-1}$$

and repeated backshift operations are represented by

$$B^k Z_t = B B \cdots B Z_t = Z_{t-k} \quad (k \text{ times})$$

Definition (Difference Operator)
The difference operator, $\Delta$, is defined by

$$\Delta Z_t = (1 - B)Z_t = Z_t - Z_{t-1}$$

and repeated $d$ times is called differences of order $d$ which represented by

$$\Delta^d = (1 - B)(1 - B) \cdots (1 - B) Z_t \quad (d \text{ times})$$
Backshift and Difference Operators

Definition (Backshift Operator)
The backshift operator, $B$, is defined by

$$BZ_t = Z_{t-1}$$

and repeated backshift operations are represented by

$$B^k Z_t = BB \cdots B Z_t = Z_{t-k}$$

$k$ times

Definition (Difference Operator)
The difference operator, $\Delta$, is defined by

$$\Delta Z_t = (1 - B)Z_t = Z_t - Z_{t-1}$$

and repeated $d$ times is called differences of order $d$ which represented by

$$\Delta^d = (1 - B)(1 - B) \cdots (1 - B) Z_t$$

$d$ times
Moving Average Model — MA($q$)

**Definition (Moving average model — MA($q$))**

The moving average model of order $q$ is defined to be

$$Z_t = \mu + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \cdots + \theta_q a_{t-q}$$

where $\theta_1, \theta_2, \ldots, \theta_q$ are parameters in $\mathbb{R}$.

The above model can be compactly written as

$$Z_t = \mu + \theta(B)a_t$$

where $\theta(B)$ is the moving average operator.

**Definition (Moving Average Operator)**

The moving average operator is

$$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$$
Moving Average Model — MA($q$)

**Definition (Moving average model — MA($q$))**

The moving average model of order $q$ is defined to be

$$Z_t = \mu + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \cdots + \theta_q a_{t-q}$$

where $\theta_1, \theta_2, \ldots \theta_q$ are parameters in $\mathbb{R}$.

The above model can be compactly written as

$$Z_t = \mu + \theta(B) a_t$$

where $\theta(B)$ is the moving average operator.

**Definition (Moving Average Operator)**

The moving average operator is

$$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$$
Moving Average Model — MA($q$)

Definition (Moving average model — MA($q$))
The moving average model of order $q$ is defined to be

$$Z_t = \mu + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \cdots + \theta_q a_{t-q}$$

where $\theta_1, \theta_2, \ldots, \theta_q$ are parameters in $\mathbb{R}$.

The above model can be compactly written as

$$Z_t = \mu + \theta(B) a_t$$

where $\theta(B)$ is the moving average operator.

Definition (Moving Average Operator)
The moving average operator is

$$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$$
Consider the mean zero MA(1) process

\[ Z_t = a_t + \theta a_{t-1} \]

Then the autocovariance function is computed as

\[ \gamma(h) = \text{cov}(Z_{t+h}, Z_t) \]
MA(1)

Consider the mean zero MA(1) process

\[ Z_t = a_t + \theta a_{t-1} \]

Then the autocovariance function is computed as

\[ \gamma(h) = \text{cov}(Z_{t+h}, Z_t) \]
Consider the mean zero MA(1) process

\[ Z_t = a_t + \theta a_{t-1} \]

Then the autocovariance function is computed as

\[ \gamma(h) = \text{cov}(Z_{t+h}, Z_t) \]
\[ = E[(a_{t+h} + \theta a_{t+h-1})(a_t + \theta a_{t-1})] \]
Consider the mean zero MA(1) process

\[ Z_t = a_t + \theta a_{t-1} \]

Then the autocovariance function is computed as

\[ \gamma(h) = \text{cov}(Z_{t+h}, Z_t) \]

\[ = E [(a_{t+h} + \theta a_{t+h-1})(a_t + \theta a_{t-1})] \]

\[ = E(a_{t+h}a_t) + \theta E(a_{t+h-1}a_t) + \theta E(a_{t+h}a_{t-1}) + \theta^2 E(a_{t+h-1}a_{t-1}) \]

\[
\begin{array}{c}
h=0 & h=1 & h=-1 & h=0 \\
E(a_{t+h}a_t) & + \theta E(a_{t+h-1}a_t) & + \theta E(a_{t+h}a_{t-1}) & + \theta^2 E(a_{t+h-1}a_{t-1})
\end{array}
\]
Consider the mean zero MA(1) process

\[ Z_t = a_t + \theta a_{t-1} \]

Then the autocovariance function is computed as

\[ \gamma(h) = \text{cov}(Z_{t+h}, Z_t) = E[(a_{t+h} + \theta a_{t+h-1})(a_t + \theta a_{t-1})] \]

\[ = E(a_{t+h}a_t) + \theta E(a_{t+h-1}a_t) + \theta E(a_{t+h}a_{t-1}) + \theta^2 E(a_{t+h-1}a_{t-1}) \]

\[ = \begin{cases} 
(1 + \theta^2)\sigma^2, & |h| = 0 \\
\theta\sigma^2, & |h| = 1 \\
0, & |h| > 1
\end{cases} \]
MA(1) — ACF Computation

Therefore

\[ \rho(h) = \begin{cases} 
1, & h = 0 \\
\frac{\theta}{1 + \theta^2}, & |h| = 1 \\
0, & |h| > 1 
\end{cases} \]

Note that when \( \theta = \frac{1}{\theta'} \), we have

\[ \rho(h) = \begin{cases} 
1, & h = 0 \\
\frac{\theta'^{-1}}{1 + \theta'^{-2}} = \frac{\theta'}{\theta'^2 + 1}, & |h| = 1 \\
0, & |h| > 1 
\end{cases} \]
MA(1) — ACF Computation

Therefore

\[ \rho(h) = \begin{cases} 
1, & h = 0 \\
\frac{\theta}{1+\theta^2}, & |h| = 1 \\
0, & |h| > 1 
\end{cases} \]

Note that when \( \theta = \frac{1}{\theta'} \), we have

\[ \rho(h) = \begin{cases} 
1, & h = 0 \\
\frac{\theta' - 1}{1+\theta'^{-2}} = \frac{\theta'}{\theta'^2+1}, & |h| = 1 \\
0, & |h| > 1 
\end{cases} \]
Invertibility

Recall the MA($q$) process:

$$Z_t = \mu + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \cdots + \theta_q a_{t-q}$$

We wish MA($q$) processes to be uniquely determined by their ACFs. Invertibility guarantees this.

**Definition (Invertibility)**

An MA model process is called invertible if $Z_t$ has the AR($\infty$) representation

$$Z_t = \sum_{j=0}^{\infty} \pi_j Z_{t-j} + a_t$$

In the MA(1) case, this amounts to $|\theta| < 1$.

We later see an easy way of determining whether a given MA model is invertible based on the roots of the MA operator.
Recall the MA($q$) process:

$$Z_t = \mu + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \cdots + \theta_q a_{t-q}$$

We wish MA($q$) processes to be uniquely determined by their ACFs. Invertibility guarantees this.

**Definition (Invertibility)**

An MA model process is called invertible if $Z_t$ has the AR($\infty$) representation

$$Z_t = \sum_{j=0}^{\infty} \pi_j Z_{t-j} + a_t$$

In the MA(1) case, this amounts to $|\theta| < 1$.

We later see an easy way of determining whether a given MA model is invertible based on the roots of the MA operator.
Invertibility

Recall the MA($q$) process:

\[ Z_t = \mu + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \cdots + \theta_q a_{t-q} \]

We wish MA($q$) processes to be uniquely determined by their ACFs. Invertibility guarantees this.

**Definition (Invertibility)**

An MA model process is called invertible if $Z_t$ has the AR($\infty$) representation

\[ Z_t = \sum_{j=0}^{\infty} \pi_j Z_{t-j} + a_t \]

In the MA(1) case, this amounts to $|\theta| < 1$.

We later see an easy way of determining whether a given MA model is invertible based on the roots of the MA operator.
Invertibility

Recall the MA($q$) process:

$$Z_t = \mu + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \cdots + \theta_q a_{t-q}$$

We wish MA($q$) processes to be uniquely determined by their ACFs. Invertibility guarantees this.

**Definition (Invertibility)**

An MA model process is called invertible if $Z_t$ has the AR($\infty$) representation

$$Z_t = \sum_{j=0}^{\infty} \pi_j Z_{t-j} + a_t$$

In the MA(1) case, this amounts to $|\theta| < 1$.

We later see an easy way of determining whether a given MA model is invertible based on the roots of the MA operator.
Invertibility

Recall the MA($q$) process:

$$Z_t = \mu + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \cdots + \theta_q a_{t-q}$$

We wish MA($q$) processes to be uniquely determined by their ACFs. Invertibility guarantees this.

Definition (Invertibility)

An MA model process is called invertible if $Z_t$ has the AR($\infty$) representation

$$Z_t = \sum_{j=0}^{\infty} \pi_j Z_{t-j} + a_t$$

In the MA(1) case, this amounts to $|\theta| < 1$.

We later see an easy way of determining whether a given MA model is invertible based on the roots of the MA operator.