Partial Autocorrelation Function, PACF

Al Nosedal
University of Toronto

March 5, 2019

We have seen that the ACF is an excellent tool in identifying the order of an MA(q) process, because it is expected to "cut off" after lag q. However, we pointed out that the ACF is not as useful in the identification of the order of an AR(p) process for which it will most likely have a mixture of exponential decay and damped sinusoid expressions. Hence such behaviour, while indicating that the process might have an AR structure, fails to provide further information about the order of such structure. For that, we will define and employ the partial autocorrelation function (PACF) of the time series.

A reminder

Consider an AR(1) process, $x_t = \phi x_{t-1} + w_t$. Note that $x_{t-1} = \phi x_{t-2} + w_{t-1}$, substituting back

$$x_t = \phi^2 x_{t-2} + \phi w_{t-1} + w_t.$$

Again, $x_{t-2} = \phi x_{t-3} + w_{t-2}$, substituting back

$$x_t = \phi^3 x_{t-3} + \phi^2 w_{t-2} + \phi w_{t-1} + w_t.$$

A reminder

Continuing this process, we could rewrite x_t as

$$x_t = w_t + \phi w_{t-1} + \phi^2 w_{t-2} + \phi^3 w_{t-3} + ...$$

(note that x_t involves $\{w_t, w_{t-1}, w_{t-2}, w_{t-3}, ...\}$)

PACF

To formally define the PACF for mean-zero stationary time series, let $\hat{x}_{t+h}$, for $h \geq 2$, denote the regression of x_{t+h} on $\{x_{t+h-1}, x_{t+h-2}, ..., x_{t+1}\}$ which we write as

$$\hat{x}_{t+h} = \beta_1 x_{t+h-1} + \beta_2 x_{t+h-2} + ... + \beta_{h-1} x_{t+1}.$$

NO intercept is needed because the mean of x_t is zero.

PACF

In addition, let $\hat{x}_t$ denote the regression of x_t on $\{x_{t+1}, x_{t+2}, ..., x_{t+h-1}\}$, then

$$\hat{x}_t = \beta_1 x_{t+1} + \beta_2 x_{t+2} + ... + \beta_{h-1} x_{t+h-1}.$$

Definition

The partial autocorrelation function (PACF) of a stationary process, x_t, denoted ϕ_h^h (or ϕ_{hh}) , for $h = 1, 2, ...$ is

$$\phi_{11} = corr(x_{t+1}, x_t) = \rho(1)$$

and

$$\phi_{hh} = corr(x_{t+h} - \hat{x}_{t+h}, x_t - \hat{x}_t), \quad h \geq 2.$$

Example. PACF of an AR(1)

Consider the PACF of the AR(1) process given by $x_t = \phi x_{t-1} + w_t$, with $|\phi| < 1$.
By definition, $\phi_{11} = \rho(1) = \phi$ (Remember?)

Example. PACF of an AR(1)

To calculate ϕ_{22}, consider the regression of x_{t+2} on x_{t+1}, say $\hat{x}_{t+2} = \beta x_{t+1}$. We choose β to minimize

$$E[x_{t+2} - \hat{x}_{t+2}]^2 = E[x_{t+2} - \beta x_{t+1}]^2 = E[x_{t+2}^2] - 2\beta E[x_{t+1}x_{t+2}] + \beta^2 E[x_{t+1}^2]$$

equivalent to

$$E[x_{t+2} - \hat{x}_{t+2}]^2 = \gamma(0) - 2\beta\gamma(1) + \beta^2\gamma(0).$$

Example. PACF of an AR(1)

Now, we find the derivative w.r.t. β and set it equal to zero.

$$f(\beta) = \gamma(0) - 2\beta\gamma(1) + \beta^2\gamma(0).$$

$$f^{'}(\beta) = 2\gamma(1) + 2\beta\gamma(0)$$

(solving for β)

$$\beta = \rho(1) = \phi.$$

(Note that $f^{''}(\beta) > 0$, so $f(\beta)$ attains its minimum at ϕ).

Example. PACF of an AR(1)

Next, consider the regression of x_t on x_{t+1}, say $\hat{x}_t = \beta x_{t+1}$. We choose β to minimize

$$E[x_t - \hat{x}_t]^2 = E[x_t - \beta x_{t+1}]^2 = E[x_t^2] - 2\beta E[x_t x_{t+1}] + \beta^2 E[x_{t+1}^2]$$

equivalent to

$$E[x_t - \hat{x}_t]^2 = \gamma(0) - 2\beta\gamma(1) + \beta^2\gamma(0).$$

This is the same equation as before, so $\beta = \rho(1) = \phi$.

Example. PACF of an AR(1)

Hence,

$$\begin{aligned} cov(x_{t+2} - \hat{x}_{t+2}, x_t - \hat{x}_t) &= cov(x_{t+2} - \phi x_{t+1}, x_t - \phi x_t) \\ &= cov(x_{t+2} - \phi x_{t+1}, x_t - \phi x_t) \text{ (note that } w_{t+2} = x_{t+2} - \phi x_{t+1} \\ &= cov(w_{t+2}, x_t - \phi x_t) \text{ (check reminder).} \end{aligned}$$

Recall that x_t involves $\{w_t, w_{t-1}, w_{t-2}, ...\}$ and x_{t+1} involves $\{w_{t+1}, w_t, w_{t-1}, ...\}$ which are uncorrelated to w_{t+2}.

Thus, $cov(x_{t+2} - \hat{x}_{t+2}, x_t - \hat{x}_t) = 0$.

Example. PACF of an AR(1)

Therefore,

$$\phi_{22} = corr(w_{t+2}, x_t - \phi x_{t+1}) = 0.$$

It can be shown that $\phi_{hh} = 0$ for all $h > 1$.

What is the PACF

Suppose that we consider the "memory" in an AR(1) process. We know that its autocorrelation function is given by $\rho(k) = \phi^k$. Consider the dependency of observations one lag apart; they are correlated $\rho(1) = \phi$ for the AR(1) model. Now consider observations two lags apart. You will be prone to answer that they are correlated $\rho(2) = \phi^2$. Hence, observations x_t are correlated with observations x_{t+2} to the extent ϕ^2. But is x_{t+2} dependent on x_t after considering the intermediate link with x_{t+1}?

What is the PACF

The question can be answered by partial correlation. If the terms are denoted 1, 2, and 3 (for x_t, x_{t+1}, and x_{t+2}, respectively), we want to know if $\rho_{13.2}$ is zero, where $\rho_{13.2}$ is the correlation of x_t and x_{t+2} given (conditional on) x_{t+1}. The standard equation for partial correlation is

$$\rho_{13.2} = \frac{\rho_{13} - \rho_{12}\rho_{32}}{\sqrt{1-\rho_{12}^2}\sqrt{1-\rho_{32}^2}}$$

We know that for the AR(1), $\rho_{13} = \rho(2) = \phi^2$ and $\rho_{12} = \rho_{32} = \phi$. Hence, the numerator is $\phi^2 - \phi\phi = 0$. So the answer is: NO, there is no relationship between x_{t+2} and x_t after removing the intermediate association with x_{t+1}. All higher-order partials will also vanish. To summarize, if the process is AR(1), once we get to lag 2, all partial correlations are zero.

Example: Applying the Yule-Walker Equations

Suppose we suspect that $p = 2$; that is, we suspect that we are dealing with an AR(2) process. The Yule-Walker equations are:

$E[Y_t Y_{t-1}] = a_1 E[Y_{t-1}^2] + a_2 E[Y_{t-2} Y_{t-1}] + E[W_t Y_{t-1}]$

$\gamma(1) = a_1 \gamma(0) + a_2 \gamma(1)$

$\frac{\gamma(1)}{\gamma(0)} = a_1 + a_2 \frac{\gamma(1)}{\gamma(0)}$

$$\rho(1) = a_1 + a_2 \rho(1) \tag{1}$$

Example: Applying the Yule-Walker Equations (cont.)

$E[Y_t Y_{t-2}] = a_1 E[Y_{t-2}^2] + a_2 E[Y_{t-2} Y_{t-2}] + E[W_t Y_{t-2}]$
$\gamma(2) = a_1 \gamma(1) + a_2 \gamma(0)$
$\frac{\gamma(2)}{\gamma(0)} = a_1 \frac{\gamma(1)}{\gamma(0)} + a_2$

$$\rho(2) = a_1 \rho(1) + a_2 \quad (2)$$

Example: Applying the Yule-Walker Equations (cont.)

Or, in matrix form,

$$\begin{pmatrix} \rho(1) \\ \rho(2) \end{pmatrix} = \begin{pmatrix} 1 & \rho(1) \\ \rho(1) & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

Example: Applying the Yule-Walker Equations (cont.)

The second-order partial autocorrelation coefficient is a_2, written ϕ_{22}, which can be found using Cramer's rule,

$$a_2 = \phi_{22} = \frac{det \begin{pmatrix} 1 & \rho(1) \\ \rho(1) & \rho(2) \end{pmatrix}}{det \begin{pmatrix} 1 & \rho(1) \\ \rho(1) & 1 \end{pmatrix}} = \frac{\rho(2) - \rho^2(1)}{1 - \rho^2(1)}$$

Example

Consider the quarterly growth rate of U.S. real gross national product (GNP), seasonally adjusted, from the second quarter of 1947 to the first quarter of 1991. We shall try fitting an AR process to this series.
Data set is available at

```
gnp_url = "https://mcs.utm.utoronto.ca/~nosedal/data/q-gnp.txt
```

R Code

```
#Step 1. Reading data;
# url of gnp;
gnp_url =
"https://mcs.utm.utoronto.ca/~nosedal/data/q-gnp.txt"
# import data in R;
gnp= read.table(gnp_url, header = FALSE);

head(gnp);
```

R Code

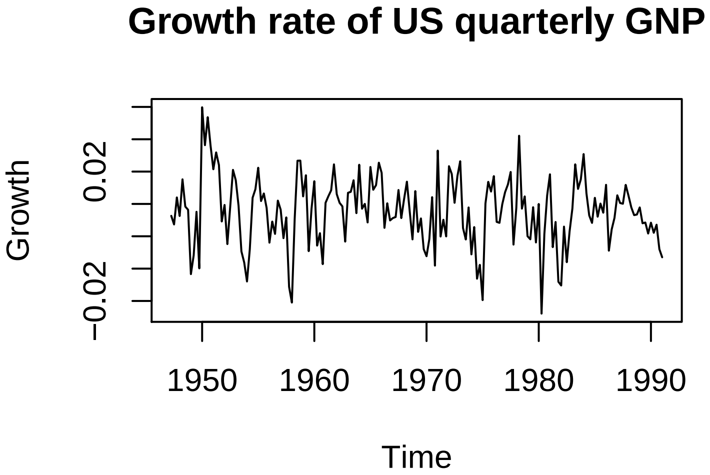
```
##          V1
## 1 0.00632
## 2 0.00366
## 3 0.01202
## 4 0.00627
## 5 0.01761
## 6 0.00918
```

Time Series Plot

```
# creating time series object;
gnp.ts=ts(gnp,frequency=4,start=c(1947,2));

plot(gnp.ts,main="Growth rate of US quarterly GNP",
ylab="Growth");
```

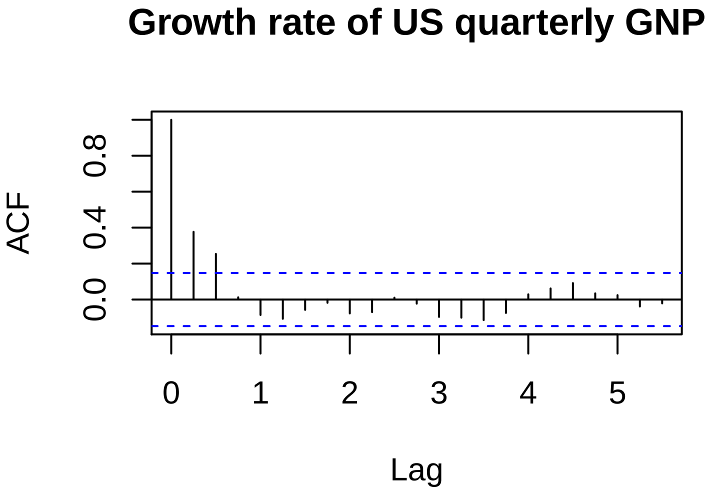
Time Series Plot



ACF

```
acf(gnp.ts,main="Growth rate of US quarterly GNP");
```

ACF



Autocorrelations (values)

```r
rhos=acf(gnp.ts,plot=FALSE)$acf;

rhos[1];

rhos[2];

rhos[3];

rhos[4];
```

Autocorrelations (values)

```
## [1] 1
## [1] 0.3768704
## [1] 0.253912
## [1] 0.01252511
```

Levinson-Durbin Method

For a given lag k, it can be shown that the ϕ_{kk} satisfy the Yule-Walker equations:

$$\rho_j = \phi_{k1}\rho_{j-1} + \phi_{k2}\rho_{j-2} + \phi_{k3}\rho_{j-3} + \cdots + \phi_{kk}\rho_{j-k}$$

for $j = 1, 2, \cdots, k$.
Here we are treating $\rho_1, \rho_2, \cdots, \rho_k$ as given and wish to solve for $\phi_{k1}, \phi_{k2}, \cdots, \phi_{kk}$.

Levinson-Durbin Method (cont.)

Levinson and Durbin gave an efficient method for obtaining the solutions to equations given on previous slide, for either theoretical or sample autocorrelations. They showed that these equations can be solved recursively as follows:

$$\phi_{kk} = \frac{\rho_k - \sum_{j=1}^{k-1} \phi_{k-1,j}\rho_{k-j}}{1 - \sum_{j=1}^{k-1} \phi_{k-1,j}\rho_j}$$

where

$$\phi_{k,j} = \phi_{k-1,j} - \phi_{kk}\phi_{k-1,k-j}$$

for $j = 1, 2, \cdots, k-1$

Partial Autocorrelations (values)

```
phis=pacf(gnp.ts,plot=FALSE);

phis;
```

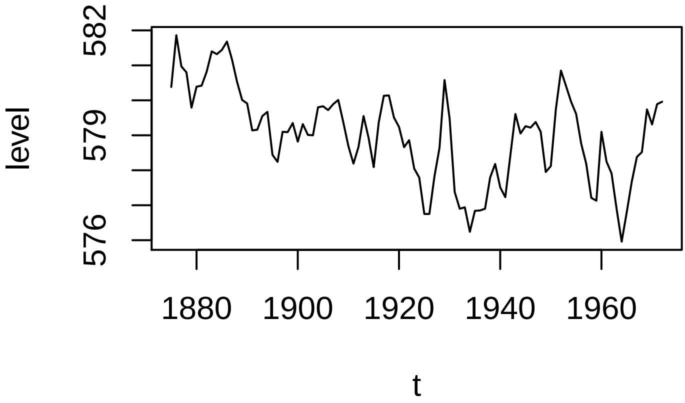
Partial Autocorrelations (values)

```
##
## Partial autocorrelations of series 'gnp.ts', by lag
##
##   0.25   0.50   0.75   1.00   1.25   1.50   1.75   2.00   2
##  0.377  0.130 -0.142 -0.099 -0.020  0.033  0.012 -0.111 -0.
##   2.75   3.00   3.25   3.50   3.75   4.00   4.25   4.50   4
## -0.037 -0.153 -0.051 -0.013  0.010  0.058 -0.011  0.032 -0.
##   5.25   5.50
## -0.057  0.018
```

PACF

```
pacf(gnp.ts,main="Growth rate of US quarterly GNP");
```

PACF



PACF (another way)

Another way to introduce PACF is to consider the following AR models in consecutive orders:

$$
\begin{aligned}
y_t &= \phi_{0,1} + \phi_{1,1} y_{t-1} + w_{1t}, \\
y_t &= \phi_{0,2} + \phi_{1,2} y_{t-1} + \phi_{2,2} y_{t-2} + w_{2t}, \\
y_t &= \phi_{0,3} + \phi_{1,3} y_{t-1} + \phi_{2,3} y_{t-2} + \phi_{3,3} y_{t-3} + w_{3t}, \\
y_t &= \phi_{0,4} + \phi_{1,4} y_{t-1} + \phi_{2,4} y_{t-2} + \phi_{3,4} y_{t-3} + \phi_{4,4} y_{t-3} + w_{4t}, \\
&\vdots
\end{aligned}
$$

PACF (another way)

where $\phi_{0,j}, \phi_{i,j}$, and $\{w_{jt}\}$ are, respectively, the constant term, the coefficient of y_{t-i}, and the error term of an AR(j) model. These models are in the form of a multiple linear regression and can be estimated by the least-squares method. The estimate $\hat{\phi}_{1,1}$ of the first equation is called the lag-1 sample PACF of y_t. The estimate $\hat{\phi}_{2,2}$ of the second equation is called the lag-2 sample PACF of y_t. The estimate $\hat{\phi}_{3,3}$ of the third equation is called the lag-3 sample PACF of y_t, an so on.

PACF (another way)

```
mod1=ar(gnp.ts,order.max=1);
mod2=ar(gnp.ts,order.max=2);
mod3=ar(gnp.ts,order.max=3);

mod1$ar;
mod2$ar;
mod3$ar;
```

PACF (another way)

```
## [1] 0.3768704
## [1] 0.3277258 0.1304018
## [1]  0.3462541  0.1769673 -0.1420867
```

For a stationary Gaussian AR(p) model, it can be shown that the sample PACF has the following properties:

- $\hat{\phi}_{p,p}$ converges to ϕ_p as the sample size T goes to infinity.
- $\hat{\phi}_{l,l}$ converges to zero for all $l > p$.
- The asymptotic variance of $\hat{\phi}_{l,l}$ is $\frac{1}{T}$ for $l > p$.

These results say that, for an AR(p) series, the sample PACF cuts off at lag p.