The `partialAR` package for modeling time series with both permanent and transient components

Matthew Clegg
matthewcleggphd@gmail.com
Stationary Autoregressive Processes

Key Features:
- Well-defined mean
- Finite variance
- Mean-reverting

When they can be found in real life, profitable trading opportunities exist.

But: It assumes all shocks are transient. This is unrealistic except in special cases.

\[ M_t = 0.98 M_{t-1} + \varepsilon_{M, t} \]
Unit Root Processes

Key Features:
- Undefined variance
- Random walk

Often this seems to be a more realistic model of price movements, but it is not useful for trading.

In this model, all shocks are permanent.

\[ R_t = R_{t-1} + \varepsilon_{R_t} \]
Partially Autoregressive Processes

Key Features:

• Globally it looks like a random walk
• Locally, it is mean reverting

Shocks contain both permanent and transient components.

Possibly useful when a mean-reverting financial time series is contaminated with a small random walk.

Previously considered by Summers [1986] and Poterba and Summers [1988].
Model Parameters

\( \rho \) \hspace{1cm} \text{Coefficient of mean reversion}

\( \sigma_{M}^{2} \) \hspace{1cm} \text{Variance of innovations in mean-reverting component}

\( \sigma_{R}^{2} \) \hspace{1cm} \text{Variance of innovations in random walk component}

Estimation can be performed using a Kalman filter.
The partialAR package

fit.par(X)
Finds the maximum likelihood fit of a partially autoregressive model to the zoo series X.

test.par(X)
Tests the alternative hypothesis of partial autoregression against the null hypothesis that X is either a random walk or stationary AR(1).
Proportion of Variance Attributable to Mean Reversion

\[ R_{MR}^2 = \frac{2\sigma_M^2}{2\sigma_M^2 + (1 + \rho)\sigma_R^2} \]

Measures whether a particular series is closer to a random walk or a mean-reverting series:

- For a pure random walk, it will be 0
- For a partially autoregressive series, it will be between 0 and 1
- For a pure autoregressive series, it will be 1

When two series are cointegrated, this value will be one when computed for the residual series.
Example: Testing Cointegration of Price of Coca-Cola vs. Pepsi, 2013

Sample R Code:

```r
> library(quantmod)
> library(tseries)
> library(egcm)
> library(partialAR)
> getSymbols("KO")
> getSymbols("PEP")
> KO.PEP <- merge(KO, PEP)
> KO.PEP2013 <- window(KO.PEP, start=as.Date("2013-01-01"), end=as.Date("2013-12-31"))
> KO.PEP2013 <- KO.PEP2013[, c("KO.Adjusted", "PEP.Adjusted")]
> colnames(KO.PEP2013) <- c("KO", "PEP")
> egcm(KO.PEP2013)
> adf.test(lm(KO~PEP, KO.PEP2013)$residuals)
> pp.test(lm(KO~PEP, KO.PEP2013)$residuals)
> fit.par(egcm(KO.PEP2013)$residuals)
> test.par(egcm(KO.PEP2013)$residuals)
```
Example: Testing Cointegration of price of Coca-Cola vs. Pepsi, 2013
Coca Cola vs. Pepsi 2013: Finding a partially autoregressive fit

> fit.par(egcm(KO.PEP2013)$residuals)

Fitted model:
\[ X[t] = M[t] + R[t] \]
\[ M[t] = 0.6396 \times M[t-1] + \epsilon_{M,t}, \quad \epsilon_{M,t} \sim N(0, 0.4623^2) \]
\[ (0.2201) \]
\[ R[t] = R[t-1] + \epsilon_{R,t}, \quad \epsilon_{R,t} \sim N(0, 0.4569^2) \]
\[ (0.1088) \]

\[ M_0 = 0.0000, \quad R_0 = -6.0535 \]
\[ (NA) \quad (0.6536) \]

Proportion of variance attributable to mean reversion (pvmr) = 0.5553

Negative log likelihood = 249.01
Coca Cola vs. Pepsi 2013: Testing the Goodness of Fit

```r
> test.par(egcm(KO.PEP2013)$residuals)
```

Test of [Random Walk or AR(1)] vs Almost AR(1) [LR test for AR1]

data: egcm(KO.PEP2013)$residuals

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random Walk</td>
<td>-2.83</td>
<td>0.058</td>
</tr>
<tr>
<td>AR(1)</td>
<td>-2.67</td>
<td>0.010</td>
</tr>
<tr>
<td>Combined</td>
<td></td>
<td>0.040</td>
</tr>
</tbody>
</table>
Limitations

• It works best when the half-life of mean reversion is short.

• Test does not have great power, especially when $\rho$ is close to 1.

• If used on a price series with significant bid-ask bounce, the model may focus on the bid-ask spread.

• Results of fit can be biased by GARCH effects.

• See paper for further details.
Conclusion

Partial autoregression is a possibly useful model when confronted with a mean-reverting time series that has been contaminated with a random walk component.

If you think your data may fit this model, the partialAR package can be used to find the best fit.

Feel free to contact me for further details and assistance:

matthewcleggphd@gmail.com
