

Introduction to pair trading -Based on cointegration-

- Shinichi Takayanagi
- Kohta Ishikawa

Topics

1. What is pair trading?
2. What is cointegration?
3. Idea of pair trading based on cointegration
4. Simulation by R language
5. Summary & concluding remarks

1. What is pair trading?

Pair trading was pioneered by ...

- **Gerry Bamberger and Nunzio Tartaglia**
- **Quantitative group at Morgan Stanley**
- **Around 1980s**
- **D.E. Shaw & Co. is famous for this strategy**

Pair trading is ...

Market neutral trading strategy

Pair trading belongs to ...

Physics, Information theory

Statistical Arbitrage

PCA, ICA, Autoregression

Neural Net

Pattern Recognition

Basic idea of pair trading ...

**Select two stocks
which move **similarly****

Basic idea of pair trading ...

Sell high priced stock

Buy low priced stock

Basic idea of pair trading ...

Stock A : -----

Stock B : -----



Basic idea of pair trading ...

Usually, monitor
the difference
between two stock prices

Basic idea of pair trading ...

the difference between two stock prices



2. What is cointegration?

Cointegration is ...

- **Pioneered by Engle and Granger**
- **Statistical property of time series**
- **Around 1990s**

Cointegration is ...

Not correlation

Cointegration and correlation

- **Correlation**

- Specify co-movement of **return**
- **Short term** relationship

- **Cointegration**

- Specify co-movement of **price**
- **Long term** relationship

(weak) Stationary time series

Not depend on time

- $E(X_t) = \mu$
- $\text{var}(X_t) = \sigma^2$
- $\text{cov}(X_t, X_{t-s}) = \gamma(s)$

Example of stationary time series

White noise

- $E(\varepsilon_t) = 0$
- $\text{var}(\varepsilon_t) = \sigma^2$
- $\text{cov}(\varepsilon_t, \varepsilon_s) = 0, t \neq s$

Non stationary time series

Depend on time

- $E(X_t) = \mu_t$
- $\text{var}(X_t) = \sigma_t^2$
- $\text{cov}(X_t, X_{t-s}) = \gamma(t, s)$

Example of non stationary time series

Brownian motion

- $E(W_t) = 0$
- $\text{var}(W_t) = t$
- $\text{cov}(W_t, W_{t-s}) = t - s$

Lag operator L

- $LX_t = X_{t-1}$
- $(1 - L)X_t = X_t - X_{t-1} = \Delta X_t$

Integrated of order P

X_t : non stationary

$(1-L)^p X_t$: stationary



$$X_t \sim I(p)$$

Example of “integrate”

$Z_t = Z_{t-1} + \varepsilon_t$: Random walk

ε_t : White noise



Calculate difference

$\Delta Z_t = Z_t - Z_{t-1} = \varepsilon_t$: Stationary

$\therefore Z_t \sim I(1)$

X_t and Y_t are cointegrated if ...

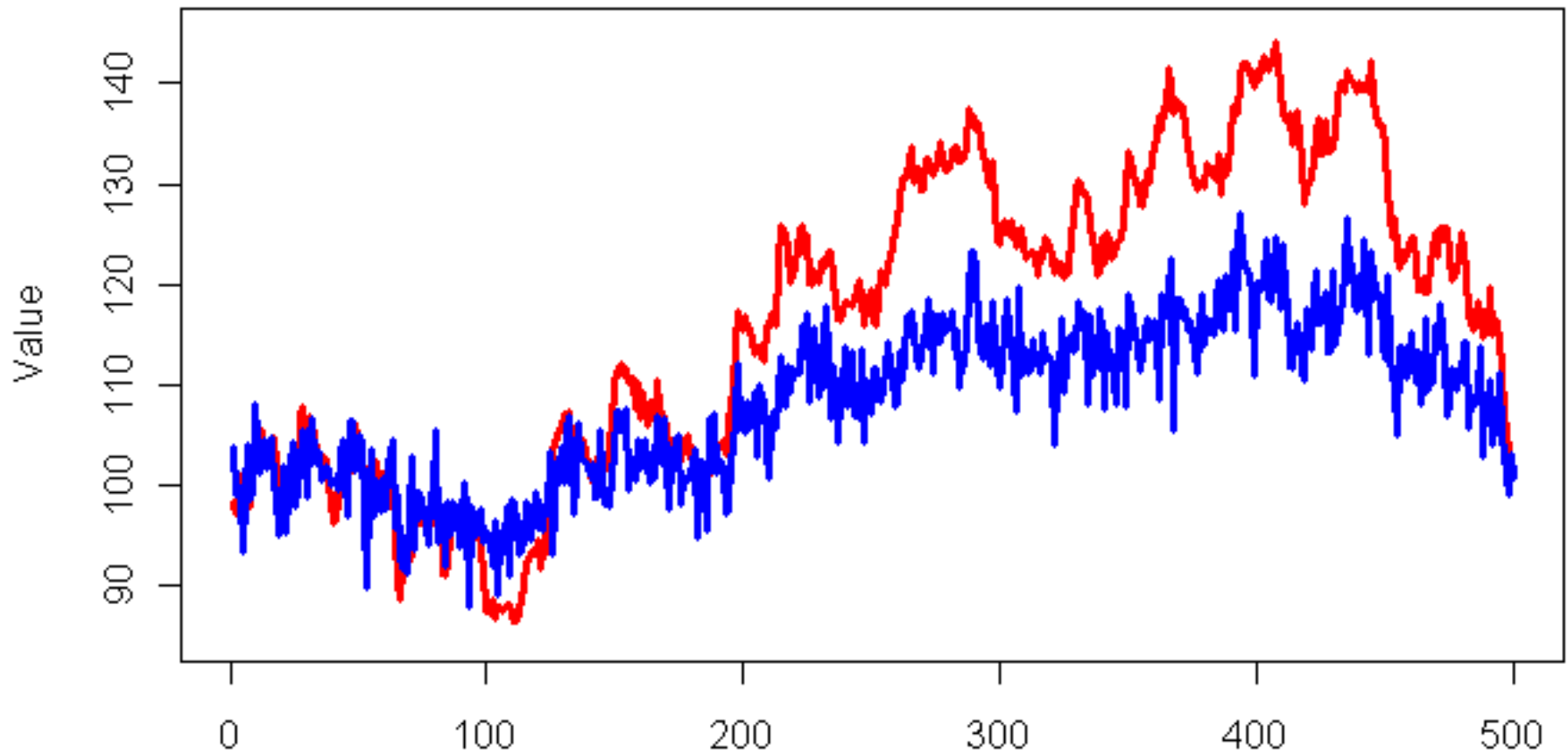
$$u_t = Y_t - (\alpha + \beta X_t)$$

$$u_t : \sim I(0), \text{ stationary process}$$

$$X_t, Y_t : \sim I(1)$$

*This is a special version of general cointegration for $I(1)$

Example of cointegrated time series



X_t : -----

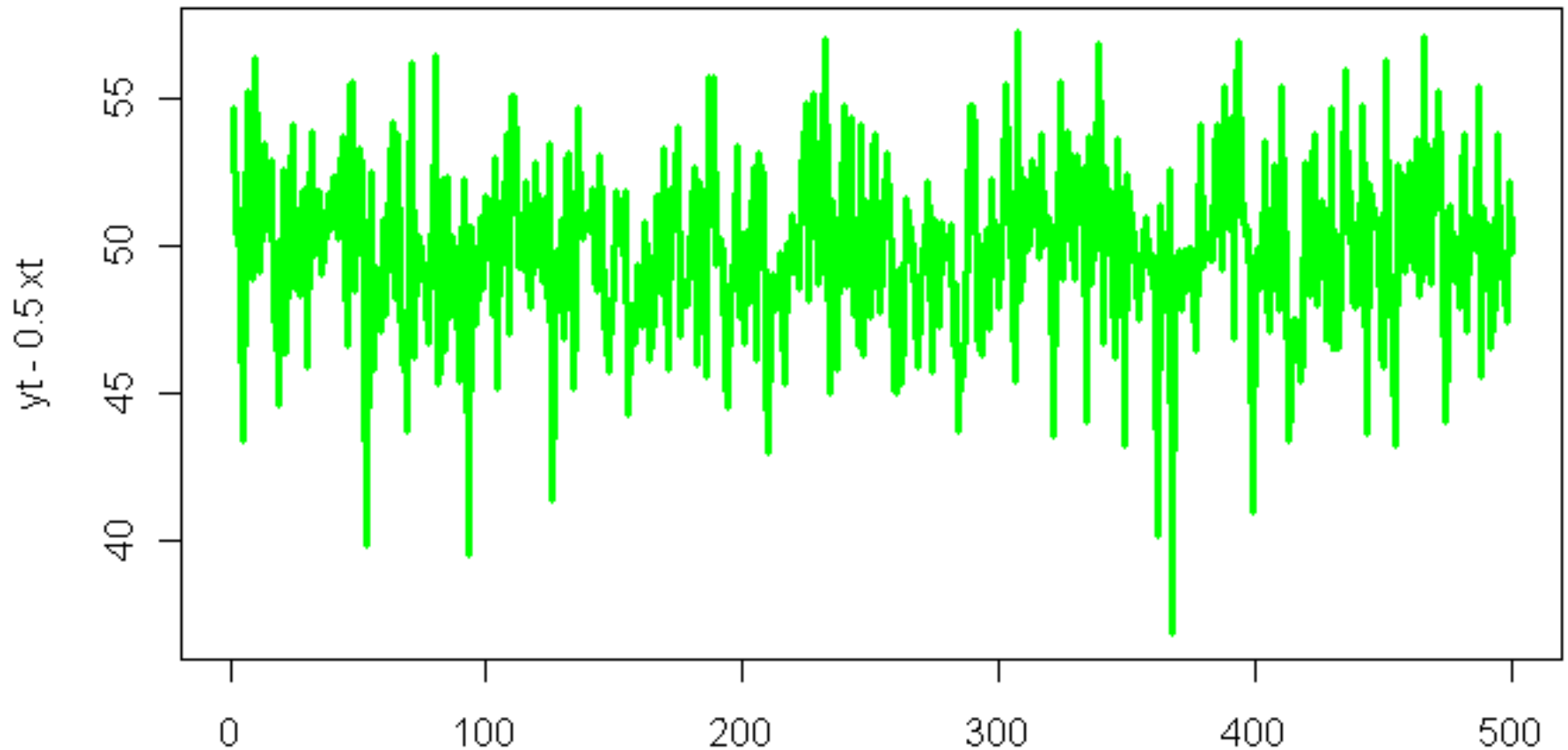
Y_t : -----

$$Y_t = 50 + 0.5X_t + u_t$$

X_t : $100 + 2 \times \text{Normal brownian motion}$

u_t : $3 \times \text{Gaussian noise}$

Example of cointegrated time series



Plot : $u_t = Y_t - 0.5 X_t$

ut seems to be...

Stationary

&

Mean reversion

Question

**Can we apply this idea
to trading strategy?**

Then, we can apply

**Cointegration idea
to log stock price**

Log price spread(*) is...

Stationary

&

Mean reversion

※ $Spread_t := \log(Y_t) - (\alpha + \beta \log(X_t))$, X_t, Y_t : stock price

Simple trading idea

if $Spread_t > \text{very high}$: Buy X_t , Sell Y_t

if $Spread_t < \text{very low}$: Buy Y_t , Sell X_t

$$Spread_t = \log(Y_t) - (\alpha + \beta \log(X_t))$$

X_t, Y_t : stock price

4. Simulation by R language

Process

1. Find two likely cointegrated stocks
2. Estimate spreads
3. Check stationarity
4. Create trading signal
5. Run back-test

1. Find two likely cointegrated stocks

```
> library(PairTrading)
> #load sample stock price data
> data(stock.price)
> #select 2 stocks
> price.pair <- stock.price[,1:2]["2008-12-31::"]
> head(price.pair)
```

```
      7201 7203
```

```
2009-01-05  333 3010
```

```
2009-01-06  341 3050
```

```
2009-01-07  374 3200
```

```
2009-01-08  361 3140
```

* Just load sample data in this case....

2. Estimate spreads

```
> reg <- EstimateParameters(price.pair, method = lm)
```

```
> str(reg)
```

List of 3

\$ spread :An 'xts' object from 2008-12-30 to 2011-08-05 containing:

Data: num [1:635, 1] -0.08544 -0.0539 -0.04306 -0.00426 -0.01966 ...

- attr(*, "dimnames")=List of 2

..\$: NULL

..\$: chr "B"

Indexed by objects of class: [Date] TZ:

xts Attributes:

NULL

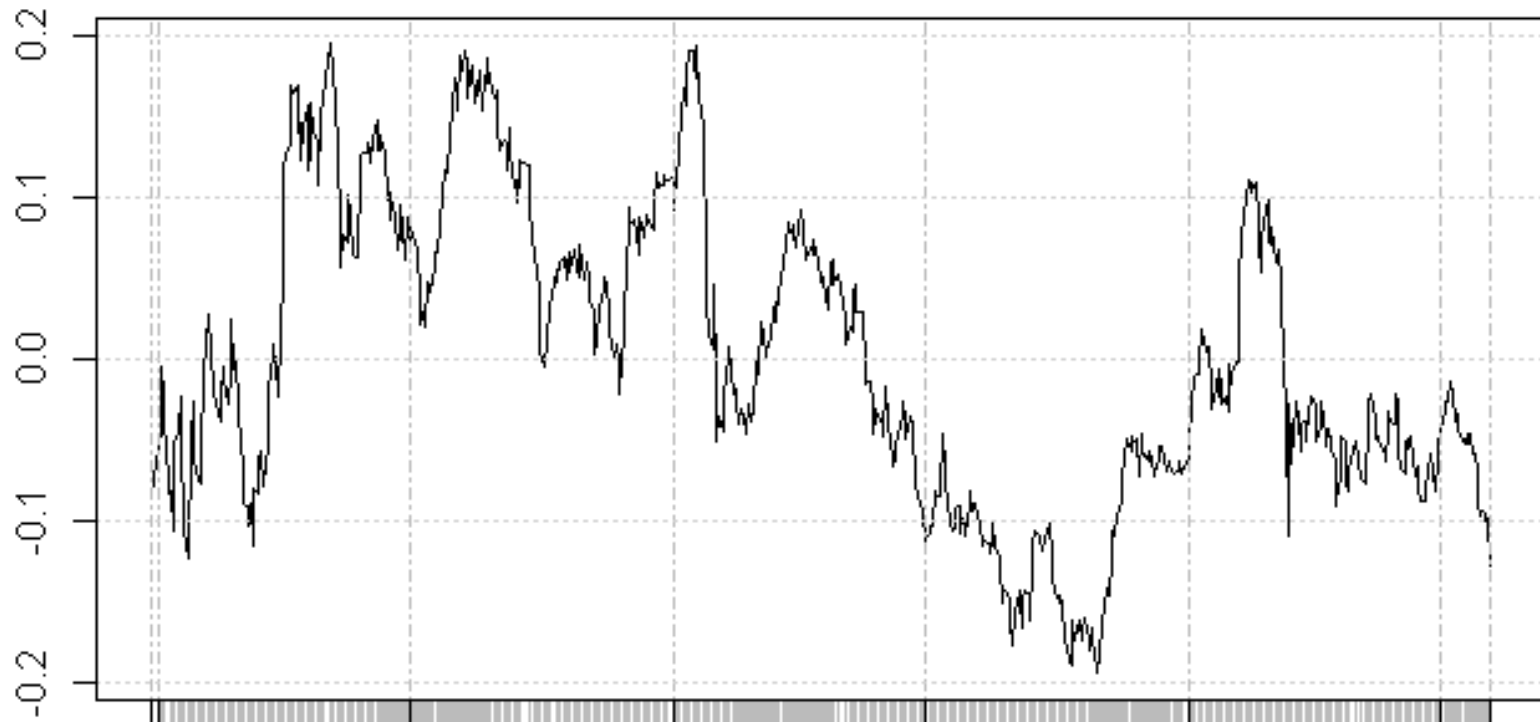
\$ hedge.ratio: num 0.0997

\$ premium : num 7.48

2. Estimate spreads

```
> plot(reg$spread, main = "Spread")
```

Spread



12 30 2008 7 01 2009 1 04 2010 7 01 2010 1 04 2011 7 01 2011

$$Spread_t = \log(Y_t) - (\alpha + \beta \log(X_t)), X_t, Y_t : \text{stock price}$$

3. Check stationarity

```
> PP.test(as.numeric(reg$spread))
```

Phillips-Perron Unit Root Test

```
data: as.numeric(reg$spread)
```

Dickey-Fuller = -3.2299, Truncation lag parameter = 6, p-value
= 0.08278

```
> adf.test(as.numeric(reg$spread))
```

Augmented Dickey-Fuller Test

```
data: as.numeric(reg$spread)
```

Dickey-Fuller = -3.6462, Lag order = 8, p-value = 0.02825

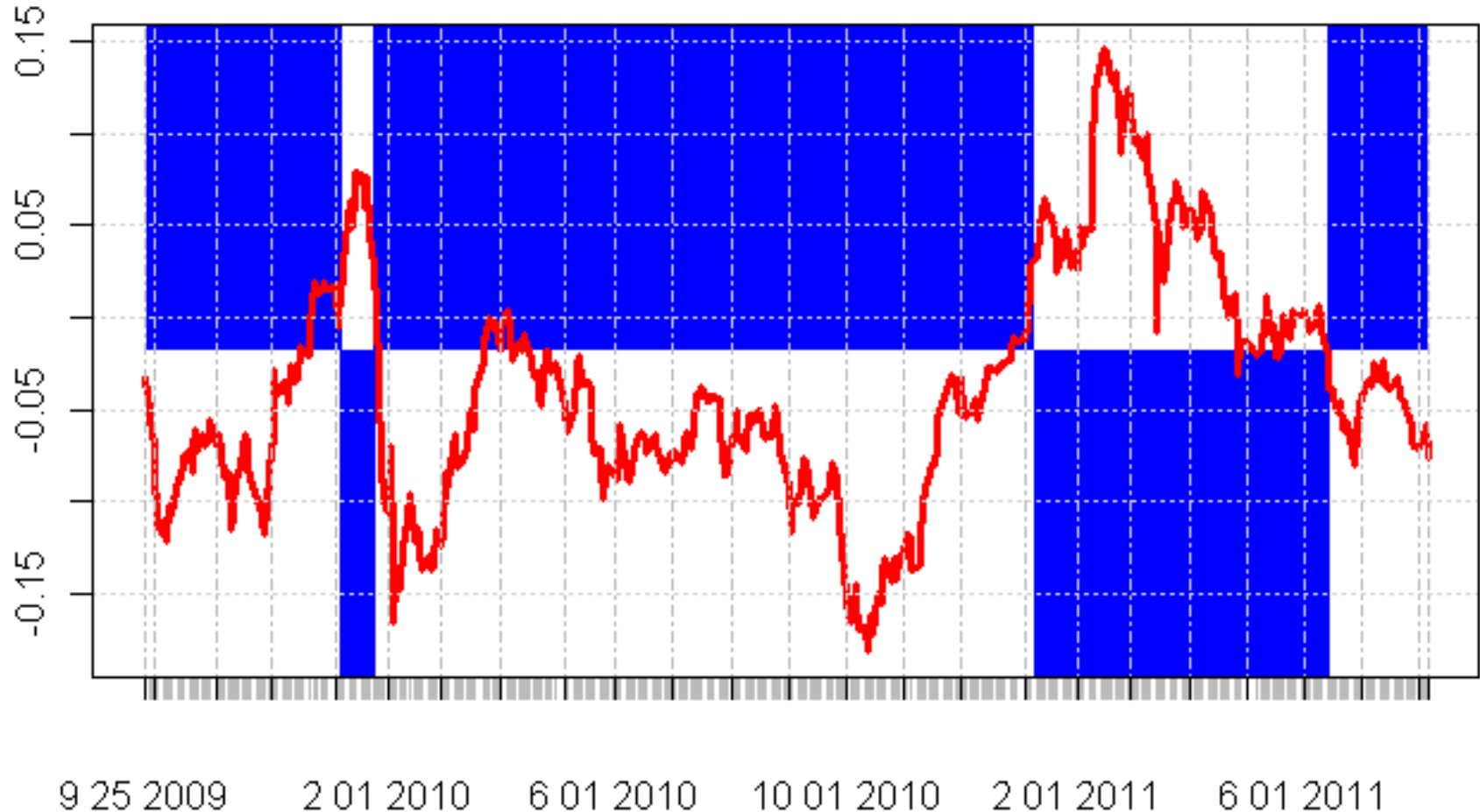
alternative hypothesis: stationary

4. Create trading signal

```
> params <-  
  EstimateParametersHistorically(price.pair,  
    period = 180)  
  
> signal <- Simple(params$spread, 0.05)  
  
> barplot(signal,col="blue",space = 0, border =  
  "blue",xaxt="n",yaxt="n",xlab="",ylab="")  
  
> par(new=TRUE)  
  
> plot(params$spread, type="l", col = "red",  
  lwd = 3, main = "Spread & Signal")
```

4. Create trading signal

Spread & Signal

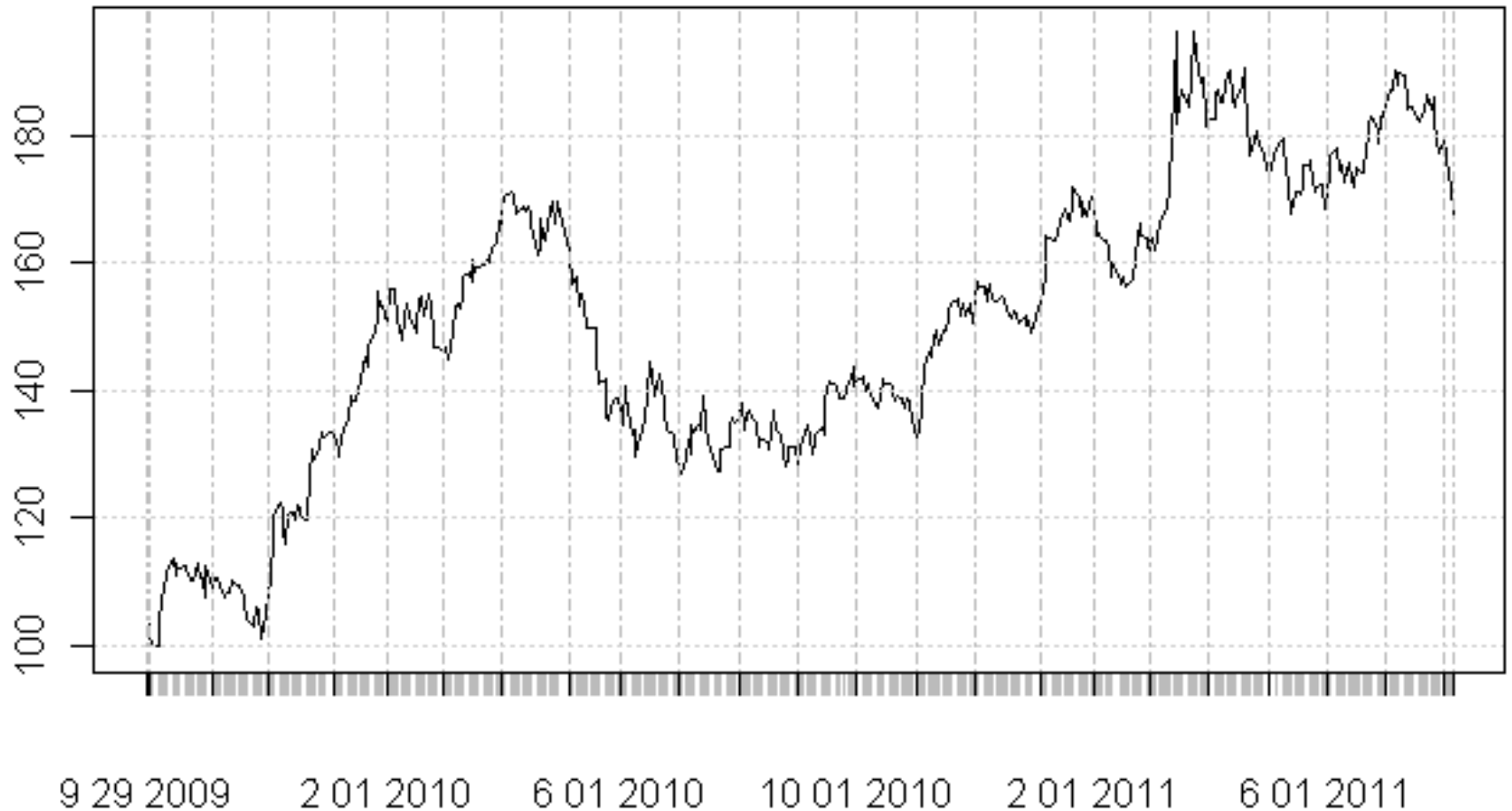


5. Run back-test

```
> return.pairtrading <-  
  Return(price.pair, lag(signal),  
    lag(params$hedge.ratio))  
  
> plot(100 * cumprod(1 +  
  return.pairtrading), main =  
  "Performance of pair trading")
```

5. Run back-test

Performance of pair trading



5. Summary & concluding remarks

Summary & concluding remarks

- Pair trading is simple quantitative trading strategy
- Cointegration is long term relation ship of time series
- Idea of cointegration may give a chance to make a profit from financial market by pair trading
- Next step
 - Sophisticate parameter estimation & trading rule
 - Make a simulation close to real

Reference

- Pairs trade(http://en.wikipedia.org/wiki/Pairs_trade)
- Cointegration(<http://en.wikipedia.org/wiki/Cointegration>)
- Andrew Neil Burgess, “A Computational Methodology for Modeling the Dynamics of Statistical Arbitrage”
- Russell Wojcik, “Pairs Trading: A Professional Approach”
- Daniel Herlemont, “Pairs trading, convergence trading, cointegration”
- Paul Teetor, “Using R to Test Pairs of Securities for Cointegration”(<http://quanttrader.info/public/testForCoint.html>)
- Ganapathy Vidyamurthy, “Pairs Trading: Quantitative Methods and Analysis “