Introduction to the `bbdetection` package

Valeriy Zakamulin*

This revision: March 26, 2021

Contents

1 Introduction **2**

2 Data **2**

3 Dating of Turning Points **2**

3.1 The Algorithm of Lunde and Timmermann (2004) 3

3.2 The Algorithm of Bry and Boschan (1971) 4

3.2.1 The Algorithm 4

3.2.2 Our Modification of the Algorithm of Bry and Boschan (1971) 4

3.2.3 Known Problems 5

4 Other Functions **6**

5 Implementation of the Algorithms **6**

5.1 Implementation of the Filtering Algorithm 6

5.2 Implementation of the Dating Algorithm 7

6 Examples of Using the Package **8**

References **12**

*School of Business and Law, University of Agder, Service Box 422, 4604 Kristiansand, Norway, E-mail: `Valeri.Zakamouline@uia.no`, Website: `http://vzakamulin.weebly.com/`

1 Introduction

The bbdetection package aims to implement two algorithms of detecting Bull and Bear markets in stock prices. The package also contains functions for printing out the dating the Bull and Bear states of the market, the descriptive statistics of the states, and functions for plotting the results. For the sake of convenience, the package includes the monthly and daily data on the prices (not adjusted for dividends) of the S&P 500 stock market index.

The package is provided AS IS, without any implied warranty as to its accuracy or suitability.

2 Data

The daily and monthly closing prices of the S&P 500 stock market index are downloaded from the Yahoo Finance website[1]. The data start from January 3rd, 1950.

Both the data series are stored as zoo time series objects. The names of the objects are sp500d and sp500m; the former one is for daily data and the latter one is for monthly data.

3 Dating of Turning Points

The bbdetection package implements two well-known algorithms for detecting turnings points between the Bull and Bear phases of the stock market: the algorithm of Lunde and Timmermann (2004) and the algorithm of Bry and Boschan (1971). The algorithms are different because there is no generally accepted formal definition of Bull and Bear markets in the finance literature. There is a common consensus among financial analysts that a Bull (Bear) market denotes a period of generally rising (falling) prices. However, when it comes to the dating of Bull and Bear markets, financial analysts are broken up into two distinct groups. One group insists that in order to qualify for a Bull (Bear) market phase, the stock market price should increase (decrease) substantially. For example, the rise (fall) in the stock market price should be greater than 20% from the previous local trough (peak) in order to qualify for being a distinct Bull (Bear) market. The other group believes that in order to qualify for a Bull (Bear) name, the stock market price should increase (decrease) over a substantial period

[1] http://finance.yahoo.com/

of time. For instance, the stock market price should rise (fall) over a period of greater than 5 months in order to qualify for being a distinct Bull (Bear) market.

3.1 The Algorithm of Lunde and Timmermann (2004)

The algorithm of Lunde and Timmermann (2004) is based on imposing a minimum on the price change since the last peak or trough. This dating rule is implemented in the following manner. Let λ_1 be a scalar defining the threshold of the movement in stock prices that triggers a switch from a Bear state to a Bull state, and let λ_2 be the threshold for shifts from a Bull state to a Bear state. Denote by P_t the stock market price at time t and suppose that a trough in P has been detected at time $t_0 < t$. Therefore the algorithm knows that a Bull state begins from time $t_0 + 1$. The algorithm first finds the maximum value of P on the time interval $[t_0, t]$

$$P_{t_0,t}^{\max} = \max\{P_{t_0}, P_{t_0+1}, \dots, P_t\}$$

and then computes the (inverse of the) relative change in P where the maximum value serves as the reference value

$$\delta_t = \frac{P_{t_0,t}^{\max} - P_t}{P_{t_0,t}^{\max}}.$$

If $\delta_t > \lambda_2$, then a new peak is detected at time t_{peak} at which P attains maximum on $[t_0, t]$. The period $[t_0 + 1, t_{\text{peak}}]$ is labeled as a Bull state. A Bear state begins from $t_{\text{peak}} + 1$.

If, on the other hand, a peak in P has been detected at time $t_0 < t$, then the algorithm finds the minimum value of P on the time interval $[t_0, t]$

$$P_{t_0,t}^{\min} = \min\{P_{t_0}, P_{t_0+1}, \dots, P_t\}$$

and computes the relative change in P from the minimum value

$$\delta_t = \frac{P_t - P_{t_0,t}^{\min}}{P_{t_0,t}^{\min}}.$$

If $\delta_t > \lambda_1$, then a new trough is detected at time t_{trough} at which P attains minimum on $[t_0, t]$. The period $[t_0 + 1, t_{\text{trough}}]$ is labeled as the Bear state. A Bull state begins from $t_{\text{trough}} + 1$.

3.2 The Algorithm of Bry and Boschan (1971)

3.2.1 The Algorithm

The algorithm of Bry and Boschan (1971) was designed to detect turning points in the business cycle. This algorithm was successfully employed for dating Bull and Bear markets in the studies by Pagan and Sossounov (2003) and Gonzalez, Powell, Shi, and Wilson (2005). This dating algorithm consists of two major stages: determination of initial turning points in P and censoring operations. In order to determine the initial turning points, first of all one uses a window of length τ_{window} months on either side of the date and identifies a peak (trough) in P as a point higher (lower) than other points in the window. Second, one enforces the alternation of turning points by selecting highest of multiple peaks and lowest of multiple troughs. Censoring operations require: eliminating peaks and troughs in the first and last τ_{censor} months; eliminating the phases that last less than τ_{phase} months (unless the relative change in P over a month exceeds the threshold θ); and eliminating cycles[2] that last less than τ_{cycle} months. The censoring operations are usually repeated many times before the sequence of turning points satisfy all constraints.

In the study by Pagan and Sossounov (2003) the following parameters are used:

$$\tau_{\text{window}} = 8, \tau_{\text{censor}} = 6, \tau_{\text{phase}} = 4, \tau_{\text{cycle}} = 16, \theta = 20\%$$

Gonzalez et al. (2005) used a slightly different set of parameters:

$$\tau_{\text{window}} = 6, \tau_{\text{censor}} = 6, \tau_{\text{phase}} = 5, \tau_{\text{cycle}} = 15, \theta = 20\%$$

3.2.2 Our Modification of the Algorithm of Bry and Boschan (1971)

It is worth emphasizing that in all previous applications of the algorithm of Bry and Boschan (1971), the researchers used exclusively data at the monthly frequency. Whereas the results of the algorithm of Lunde and Timmermann (2004) are rather stable when the data frequency is changed to either weekly or daily (because the parameters of the algorithm are two relative changes in the stock price), the results of the algorithm of Bry and Boschan (1971) are very

[2] A cycle denotes two subsequent phases, either upswing and consequent downswing, or downswing and consequent upswing.

sensitive to the data frequency. It is relatively straightforward to change the sizes of the four windows in the algorithm ($\tau_{\text{window}}, \tau_{\text{censor}}, \tau_{\text{phase}}, \tau_{\text{cycle}}$). For example, if the data frequency is changed from monthly to daily, one can re-scale the sizes of the four windows using the rule $\tau_i^{\text{daily}} = \tau_i^{\text{monthly}} \times 21$ (assuming 21 trading days in a single month). However, the value of the parameter θ is related to a relative change in the stock price over a single month. If one re-scales the value of this parameter, one can get the states of the market that last one single day only.

To make the dating algorithm to be easily accommodated to any data frequency, in our implementation the value of θ defines the relative change in P, **over the whole phase of the market**, that invalidates the elimination of the phases that last less than τ_{phase} observations. It is worth noting that our definition of θ is more logical in the light of the view that "the rise (fall) in the stock market price should be greater than θ% from the previous local trough (peak) in order to qualify for being a distinct Bull (Bear) market". The logic behind this view is that the value of θ is related to the whole phase of the market, not a single month of the phase.

3.2.3 Known Problems

The algorithm of Lunde and Timmermann (2004) is very robust; the only difficulty in this algorithm is the correct identification of the very first state of the market. Unfortunately, the algorithm of Bry and Boschan (1971) makes use of pattern recognition techniques and the correct outcome is, strictly speaking, not guaranteed. Sometimes the algorithm does not correctly identify the bull-bear states of a financial market. The problem seems to lie in the arbitrary choice of parameters. Specifically, the problem seems to occur because usually one defines $\tau_{\text{window}} > \tau_{\text{phase}}$ (see the sets of parameters used by Pagan and Sossounov (2003) and Gonzalez et al. (2005)). That is, usually the size of the window used to determine the initial turning points is larger than the minimum phase length; this does not seem to be logical. As a rule, reducing the value of τ_{window} mitigates the problem. Therefore, the users should be aware of possible problems and **are advised to always visually check the correctness of the result** (during Bull states the prices should generally increase, during the Bear states decrease).

4 Other Functions

The package includes two functions that print out (in the console window) the dating of Bull-Bear states (`bb.dating.states`) and the descriptive statistics of the Bull-Bear states (`bb.summary.stat`). The outcome of each function is a table in LaTeX format that can be incorporated in a TeX (LaTeX) document.

Both functions report the duration of the states and the amplitude of the states. The duration is measured in the number of observations. For example, if the data are at a monthly frequency, the duration is measured in the number of months. The amplitude is measured in percentages.

The package also includes function `bb.plot` that plots the log of prices and highlights the bear states of the market.

5 Implementation of the Algorithms

It is worth noting that the algorithm of Lunde and Timmermann (2004) is motivated by the idea that, in order to qualify for a distinct (bear or bull) phase, the stock market price should change substantially from the previous peak or trough. On the other hand, the algorithm of Bry and Boschan (1971) exploits the idea that, in order to qualify for a distinct phase, the trend in the stock market price should continue over a substantial period from the previous peak or trough. In the subsequent exposition, in order to shorten the terminology, we denote the algorithms of Lunde and Timmermann (2004) and Bry and Boschan (1971) as the *filtering algorithm* and the *dating algorithm* respectively.

5.1 Implementation of the Filtering Algorithm

The application of the filtering rule requires making an arbitrary choice of two parameters $\{\lambda_1, \lambda_2\}$. It is unclear how to make an appropriate choice in each case. Lunde and Timmermann (2004) report the empirical results for many alternative sets of parameters where the value of λ_i, $i \in \{1, 2\}$ is varied between 5% and 30% with a step of 5%.

Before running the algorithm, one needs to define the parameters $\{\lambda_1, \lambda_2\}$. This is done via a call to the `setpar_filtering_alg(tr_bull, tr_bear)` function where

- `tr_bull` is the value of λ_1
- `tr_bear` is the value of λ_2

By default, the values of these parameters are $\{\lambda_1 = 15, \lambda_2 = 15\}$.

The algorithm is performed via a call to the `run_filtering_alg(index)` function where

- `index` is a vector that contains the stock market prices

This function returns a logical vector (of the same size as the vector `index`) that contains **TRUE** for Bull states of the market and **FALSE** for Bear states of the market.

A typical usage of these two functions

```
setpar_filtering_alg(tr_bull, tr_bear)
bull <- run_filtering_alg(index)
```

Figure 1 provides an illustration of the content of the returning vector bull and how to interpret the results of the filtering algorithm. Vector `bull` is of length 10; this length corresponds to the length of vector `index`. Each element of vector `bull` is either `FALSE` (for Bear states) or `TRUE` (for Bull states). The number of Bull/Bear states equals the number of disjunctive sequences of `TRUE/FALSE` values. In this example, the first state of the market is Bull, which lasts 3 periods corresponding to elements 1-3 of vector `bull`. The next state of the market is Bear that lasts 4 periods corresponding to elements 4-7 of vector `bull`. The final state of the market is again Bull that lasts 3 periods. Totally, there are 2 Bull states and 1 Bear state. A local peak in the `index` is located in the last period of a Bull phase. Similarly, a local trough in the `index` is located in the last period of a Bear phase. In our example in Figure 1, there is 1 peak (element 3 of vector `bull`) and 1 trough (element 7 of vector `bull`). Even though the last state of the market is Bull, there is uncertainty about whether to consider the element 10 of vector `bull` as a local peak because we do not know the subsequent dynamics of the market. To be on the safe side, we do not label element 10 as a local peak.

5.2 Implementation of the Dating Algorithm

Before running the algorithm, one needs to define the parameters $\{\tau_{\text{window}}, \tau_{\text{censor}}, \tau_{\text{phase}}, \tau_{\text{cycle}}, \theta\}$. This is done via a call to the `setpar_dating_alg(t_window, t_censor, t_phase, t_cycle, max_chng)` function where

Turning point	-	-	Peak	-	-	-	Trough	-	-	-
Index of the element	1	2	3	4	5	6	7	8	9	10
Value of the element	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE
State of the market	**Bull**			**Bear**				**Bull**		

Figure 1: Illustration of the content of the returning vector `bull` and how to interpret this content.

- `t_window` is the value of τ_{window}
- `t_censor` is the value of τ_{censor}
- `t_phase` is the value of τ_{phase}
- `t_cycle` is the value of τ_{cycle}
- `max_chng` is the value of θ

By default, the values of these parameters are: $\{\tau_{\text{window}} = 8, \tau_{\text{censor}} = 6, \tau_{\text{phase}} = 4, \tau_{\text{cycle}} = 16, \theta = 20\}$ (as in Pagan and Sossounov (2003)).

The algorithm is performed via a call to the `run_dating_alg(index)` function where

- `index` is a vector that contains the stock market prices

Similarly to `run_filtering_alg` function, this function returns a logical vector (of the same size as the vector `index`) that contains **TRUE** for Bull states of the market and **FALSE** for Bear states of the market.

A typical usage of these two functions

```
setpar_dating_alg(t_window, t_censor, t_phase, t_cycle, max_chng)
bull <- run_dating_alg(index)
```

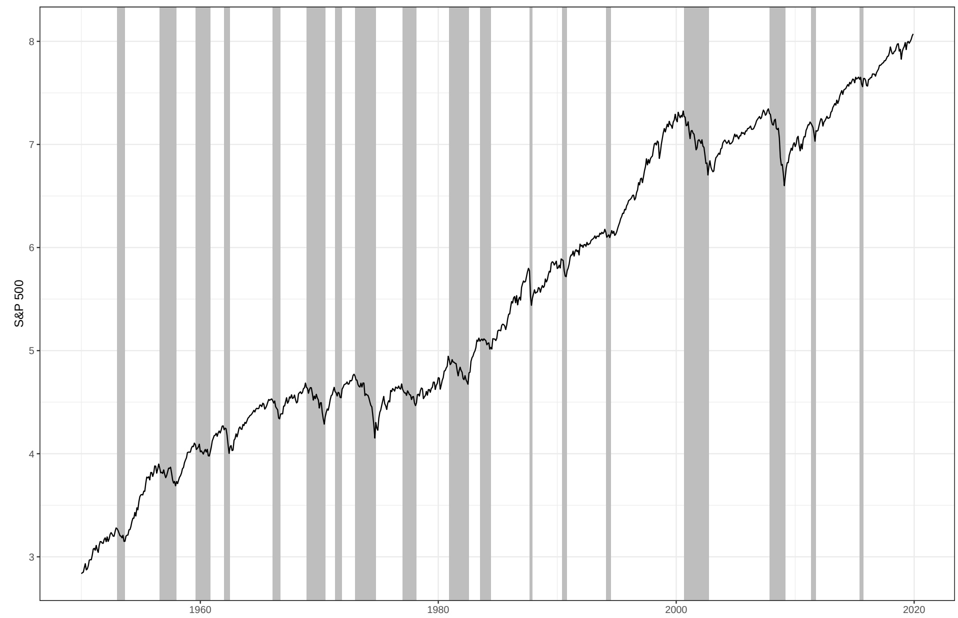
6 Examples of Using the Package

The program below uses the monthly prices of S&P 500 index, detects the bull-bear states using the dating algorithm, plots the graph, prints out the dating of bull-bear states, and finally prints out the summary statistics of bull-bear states:

```
> rm(list=ls(all=TRUE))
> library(bbdetection)
```

```
> library(zoo)
> library(xtable)
> library(ggplot2)
> prices <- as.vector(coredata(sp500m)) # retrieve prices
> dates <- index(sp500m) # retrieve dates from zoo-object
> setpar_dating_alg(8, 6, 4, 16, 20) # parameters for monthly data
> bull <- run_dating_alg(prices) # detect bull-bear states

> # plot the result
> bb.plot(prices, bull, dates, "S&P 500")
```




```
> dates <- as.yearmon(dates) # convert to "yearmon" format if monthly data
> Sys.setlocale("LC_TIME", "English") # Use English names for months

> # prints out the dating of bull-bear states
> df <- bb.dating.states(prices, bull, dates)

> # print out the summary statistics of bull-bear states
> df <- bb.summary.stat(prices, bull)
```

The next program uses the daily prices of S&P 500 index, detects the bull-bear states using the filtering algorithm, and plots the graph:

Bull markets			Bear markets		
Dates	Duration	Amplitude	Dates	Duration	Amplitude
Jan 1950 to Dec 1952	36	56	Jan 1953 to Aug 1953	8	-12
Sep 1953 to Jul 1956	35	112	Aug 1956 to Dec 1957	17	-16
Jan 1958 to Jul 1959	19	45	Aug 1959 to Oct 1960	15	-10
Nov 1960 to Dec 1961	14	29	Jan 1962 to Jun 1962	6	-20
Jul 1962 to Jan 1966	43	60	Feb 1966 to Sep 1966	8	-16
Oct 1966 to Nov 1968	26	35	Dec 1968 to Jun 1970	19	-30
Jul 1970 to Apr 1971	10	33	May 1971 to Nov 1971	7	-6
Dec 1971 to Dec 1972	13	16	Jan 1973 to Sep 1974	21	-45
Oct 1974 to Dec 1976	27	45	Jan 1977 to Feb 1978	14	-15
Mar 1978 to Nov 1980	33	58	Dec 1980 to Jul 1982	20	-21
Aug 1982 to Jun 1983	11	40	Jul 1983 to May 1984	11	-7
Jun 1984 to Aug 1987	39	115	Sep 1987 to Nov 1987	3	-28
Dec 1987 to May 1990	30	46	Jun 1990 to Oct 1990	5	-15
Nov 1990 to Jan 1994	39	49	Feb 1994 to Jun 1994	5	-5
Jul 1994 to Aug 2000	74	231	Sep 2000 to Sep 2002	25	-43
Oct 2002 to Oct 2007	61	75	Nov 2007 to Feb 2009	16	-50
Mar 2009 to Apr 2011	26	71	May 2011 to Sep 2011	5	-16
Oct 2011 to May 2015	44	68	Jun 2015 to Sep 2015	4	-7
Oct 2015 to Dec 2019	51	54			

```
> rm(list=ls(all=TRUE))
> library(bbdetection)
> library(zoo)
> library(xtable)
> library(ggplot2)
> prices <- as.vector(coredata(sp500d)) # retrieve prices
> dates <- index(sp500d) # retrieve dates from zoo-object
> setpar_filtering_alg(15, 15) # same parameters for daily and monthly data
> bull <- run_filtering_alg(prices) # detect bull-bear states

> # plot the result
> bb.plot(prices, bull, dates, "S&P 500")
```

	Bull	Bear
Number of phases	19	18
Minimum duration	10	3
Average duration	32	12
Median duration	30	12
Maximum duration	74	25
Minimum amplitude	26	-8
Average amplitude	77	-24
Median amplitude	61	-19
Maximum amplitude	242	-53

References

Bry, G. and Boschan, C. (1971). *Cyclical Analysis of Time Series: Selected Procedures and Computer Programs*. NBER.

Gonzalez, L., Powell, J. G., Shi, J., and Wilson, A. (2005). “Two Centuries of Bull and Bear Market Cycles”, *International Review of Economics and Finance*, *14*(4), 469 – 486.

Lunde, A. and Timmermann, A. (2004). “Duration Dependence in Stock Prices: An Analysis of Bull and Bear Markets”, *Journal of Business and Economic Statistics*, *22*(3), 253–273.

Pagan, A. R. and Sossounov, K. A. (2003). “A Simple Framework for Analysing Bull and Bear Markets”, *Journal of Applied Econometrics*, *18*(1), 23–46.