Identifying Market Price Levels using Differential Evolution

Michael Mayo

University of Waikato, Hamilton, New Zealand
mmayo@waikato.ac.nz
WWW home page: http://www.cs.waikato.ac.nz/~mmayo/

Abstract. Evolutionary data mining is used in this paper to investigate the concept of support and resistance levels in financial markets. Specifically, Differential Evolution is used to learn support/resistance levels from price data. The presence of these levels is then tested in out-of-sample data. Our results from a set of experiments covering five years worth of daily data across nine different US markets show that there is statistical evidence for price levels in certain markets, and that Differential Evolution can uncover them.

Keywords: differential evolution, finance, support and resistance

1 Introduction

Do price levels exist in market series? The idea of price levels – often referred to as "support" and "resistance" – has been prevalent anecdotally for nearly as long as financial markets have existed. These phenomenon are a staple feature in most trading and finance textbooks (e.g. [3]), which typically teach people to buy assets at support lines (because prices are likely to rise from them) and conversely sell assets at resistance lines (because prices are likely to fall). But is there statistical evidence that support and resistance lines exist?

Answering this question is the focus of this paper. We adopt a machine learning-based methodology, utilizing Differential Evolution (DE) [4], in an attempt to learn price levels from market data. We then compare these optimized levels to randomly selected levels in order to determine appreciable differences. Specifically, if a set of levels where price reverses are found in the training/in-sample data, then we want to know if these levels continue to persist in chronologically subsequent testing/out-of-sample data. If the levels do persist, then it can be said that the concepts of support and resistance have foundation. On the other hand, if the best levels found in-sample cannot be used to predict reversals out-of-sample, then we can conclude confidently that the dual concepts of support and resistance have no foundation.

To date, there has been little consideration of this question in the applied machine learning/finance research literature. Most other approaches deal with pattern-based turning point prediction and often ignore absolute prices. This research, on the other hand, focuses on turning points based on absolute price levels. The question is whether or not the market "remembers" these old levels and therefore whether history may repeat predictably. Note that the levels used here are basically horizontal lines; we leave the generalization of the method to angled trend lines and channels for future work.

Our results described in this paper show that in some markets, especially indices such as the Dow Jones Industrial Average, the presence of price levels can be detected with statistical significance.

2 Background

In this section, we cover the concept of a "price level" in a little more detail and briefly explain the variant of DE used here.

2.1 Price Levels in Markets

A support or resistance line is, by definition, a point in the market where price has an increased probability of reversal. For the purposes of this paper, a good price level is therefore any price that the market often reaches but infrequently penetrates. To illustrate this concept, consider Figures 1 and 2.

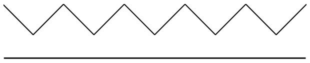

Fig. 1. Example of a poor level with excessive distance from price.

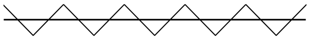

Fig. 2. Example of a poor level with excessive penetrations by price.

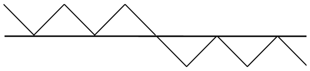

Fig. 3. Example of a strong level with a high ratio of failed to successful penetrations.

These figures both depict poor levels. In the first case, price never reaches the level and therefore there is no opportunity for price to reverse at the level.

In the second case, price frequently penetrates the level, so the level is clearly not a support or resistance for price.

Figure 3, on the other hand, depicts conceptually a good level. In this case, price frequently reverses at the level, and when it does eventually break through, the level that was formally support now becomes resistance. The same concept is depicted again in Figure 4, this time somewhat more realistically using candlesticks.

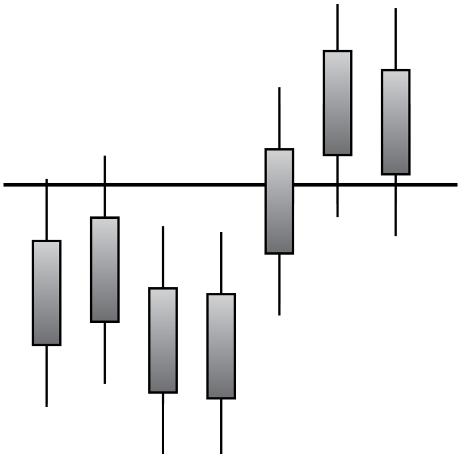

Fig. 4. Figure similar to Fig. 3 but depicted more realistically using candlesticks.

In this paper, we define a level as any price point with a high probability of a "failed penetration" on the daily charts. A failed penetration occurs when a candlestick's wick touches the level, but the open and close prices are both on the same side of the level.

Table 1. Key parameters in differential evolution and the values used in this paper for $VS = 10$ as suggested by [2].

Parameter	Value	Definition
VS	10	Individual size (i.e. dimension)
NP	18	Population size
F	0.67	Amplification factor
CR	0.50	Probability of crossover
E	20,000	Max. number of function evals

For example, in Figure 4, the first and last pair of candlesticks depict failed penetrations; however the fifth candlestick is a successful penetration because the open and close prices straddle the level. We can therefore estimate the probability of a failed penetration at the level as $\frac{4}{5}$. This is the essence of how we measure the value of a price level in this paper.

2.2 Differential Evolution

Lack of space prevents a complete description of the DE algorithm here. Essentially, DE is an evolutionary optimization algorithm for continuous spaces first proposed in [4]. The algorithm has been used in financial applications previously [1].

There are two main reasons for using DE in this research. Firstly, DE's proponents claim that the algorithm's global convergence properties are very good, and this has been demonstrated in past studies. Secondly, DE requires very few user-specified parameters. Furthermore, a recent study [2] has elucidated the best combinations of parameters for different problem sizes. In the experiments described here, the problem size (i.e. the number of dimensions) VS is set to 10, and the corresponding set of parameter values (according to [2]) that we use is given in Table 1.

In all of our experiments, we use the standard *DE/rand/1/bin* variant of DE described in [4].

3 Differential Evolution for Price Level Identification

In this section, the basic adaptation of DE for price level identification is described. We also describe the randomized algorithm used as a control case in the experiments in the next section.

3.1 Individual Representation

Recall that the aim of this research is to use DE to optimize a set of levels, such that the probability of penetration failures at those levels is maximized. The set of levels found would thus correspond to a set of support/resistance levels. Ideally, the size of this set of levels should be variable because, for example, in one market at a particular time there may be many active levels, but in another market (or in the same market but at a different time) there may only be a few active levels.

However, varying the number of levels like this poses difficulties when DE is applied to the problem, because DE expects vectors to be fixed-length. We therefore modify the problem slightly and require that the cardinality of the set of levels be fixed to VS. In other words, a set S of price levels can be defined according to Equation 1. The levels l_0, l_1, etc, can therefore be encoded directly as vector elements for DE to optimize.

$$S = \{l_0, l_1, ..., l_{VS-1}\}, l_i \in \mathbf{R} \tag{1}$$

An issue is that the individual levels, if unconstrained, may be "optimized" to the same single best value. This was in fact the case in early testing of the algorithm: if a single level proved to have high value, then it was usually the case that DE would set l_0, l_1, etc, all to that price, with the net result being that only a single level was discovered.

The solution to this problem is to limit the values of each variable $l_i \in S$ to non-overlapping ranges. Let us assume that the price series being used is divided into a training series T_0 that is followed chronologically by a testing series T_1. T_0 is used for optimization, and T_1 is set aside (it will be used later to calculate the out-of-sample values of the best individual from the T_0 optimization phase).

From the training series T_0, a price range can be straightforwardly calculated, and this range can be divided by the number of levels VS to give a range size per level, as Equation 2 shows.

$$rangeSize = \frac{high(T_0) - low(T_0)}{VS} \tag{2}$$

Once the range is known, then upper and lower bounds on the value of each level can be calculated, Equations 3 and 4 show.

$$lower_i = low(T_0) + (i \times rangeSize) \tag{3}$$

$$upper_i = lower_i + rangeSize \tag{4}$$

We therefore modify the standard DE algorithm and require that:

$$lower_i \leq l_i \leq upper_i, \forall l_i \in S. \tag{5}$$

This is achieved by (i) initializing all new individuals with values for l_i selected uniformly and randomly in the appropriate range and (ii) scoring all offspring individuals with levels outside of the appropriate range with maximum negative value, i.e. such individuals are allowed to be generated by the search process but are immediately “aborted”.

3.2 Value Function

Evolutionary search algorithms require a value function that is either maximized or minimized by the search process. In our case, the value function is an estimate of the probability of the failure of price to penetrate the levels in the current set of levels S being considered. Value should therefore be maximized.

Recall that we have a training set T_0 used to estimate value for each set S – and more generally, let T be any series of daily Open-High-Low-Close (OHLC) data.

Given some T and a particular level $l \in S$, let us define two useful functions: $f(l, T)$, specifically the number of failed penetration attempts of price against level l in the series T (i.e. the number of wicks that intersect l where the open and close of the day are both on the same side of the level); and $t(l, T)$, the total number of intersections between price and l in T (i.e. the number of bars in T that overlap l regardless of whether it is only a wick or a full candlestick body).

Obviously, then, it follows that $t(l, T) \geq f(l, T)$ will always hold.

To illustrate these functions, consider Figure 4. If the figure depicts the only touches to the level l in T, then $f(l, T) = 4$ and $t(l, T) = 5$.

Clearly, then, for each level l we will want to maximize the ratio of the first function to the second function in order to find good levels that are likely to "repel" price. Mathematically, this is expressed in the function defined by Equation 6.

$$V(S,T) = \frac{\sum_{l \in S} f(l,T)}{(\sum_{l \in S} t(l,T)) + K} \tag{6}$$

A constant K is added to the denominator in order to reduce the value of levels with only a small number of touches. For example, suppose $VS = 1$, $S = \{l\}$, $f(l,T) = 1$ and $t(l,T) = 1$. Then the ratio of f to g expressed as a percentage is 100%, but the sample size is very small and this result is therefore unreliable.

With K fixed to 10 (which is K's value in all experiments reported here), the value of this level is reduced appropriately to $\frac{1}{11}$. For the level depicted in Figure 4, the value would therefore be $\frac{4}{5+10} = \frac{4}{15}$, which is an underestimate of the true probability of a failed penetration but certainly reasonably greater than $\frac{1}{11}$.

In summary, the optimization problem that we will use DE for can be expressed as the problem of finding a set S_{DE} such that $V(S_{DE}, T_0)$ is maximized, with the out-of-sample value of interest denoted by $V(S_{DE}, T_1)$.

3.3 Control Case

In order to evaluate the efficacy of DE for price level identification, we need something to compare it to. In this research, we consider the case of simply randomly generating the levels subject to the constraints expressed by Equations 1-5. This is performed using the training data T_0 and is used as the control for our evaluation. We refer to such a randomly generated set of levels by S_{RN} and therefore the out-of-sample value (of the control algorithm for a single trial) is $V(S_{RN}, T_1)$.

Note that we also use the exact same method to randomly initialize new individuals at the start of each DE run – thus the key difference being tested is the ability of DE to optimize levels within each level's range.

4 Evaluation

In this section, the details of the evaluation and the results are covered.

4.1 Datasets

Three US stock indices and six US company stock markets were selected for the evaluation. The indices (S & P 500, Dow Jones Industrial Average and the Nasdaq Composite) were chosen because they are representative of the market at large. The six companies were chosen randomly from those that make up the

Table 2. Markets analyzed in this study. Market data is daily OHLC bars for five years to 3rd August 2012. Three month volume data from [5] as at 14th August 2012 is also given for the company stock markets, so as to provide an indication of the liquidity of each market.

Symbol	Market	Volume (3m)
SPY	S & P 500	na
DJIA	Dow Jones Industrial Average	na
COMP	Nasdaq Composite Index	na
AAPL	Apple	14,696,400
BA	Boeing Airlines	4,163,520
CELG	Celgene	3,622,660
JEC	Jacobs Engineering Group	955,114
JNJ	Johnson & Johnson	15,793,800
KMB	Kimberly-Clark	2,845,560

respective indices, the only criteria being that (i) they needed to have at least five year's worth of price data available and (ii) no two companies from the same industry were chosen. Table 2 lists details of the selected markets.

For each market, daily OHLC data for the five years to 3rd August 2012 was downloaded from Yahoo! Finance [5].

4.2 Method

For each of the market data sets listed in Table 2, we conducted 50 trials.

Each trial consisted of (i) selecting a random 1 year subsample from the five year's worth of data, of which the first 6 months was used for training (T_0) and the second six months for testing (T_1);(ii) generating a random set of levels S_{RN} given T_0 which are subject to the constraints specified by Equations 1-4; and (iii), generating an optimized set S_{DE} using differential evolution to maximize the value function on T_0.

The out-of-sample values $Val(S_{RN}, T_1)$ and $Val(S_{DE}, T_1)$ were then computed and recorded.

Thus, a total of nine markets × 50 trials or 450 experiments were conducted.

4.3 Results

We present the results first graphically using standard box-plots in Figures 5-7 . Each figure depicts the results of applying each algorithm – randomized control and DE – fifty times to a random 1 year sub-sample of each five year market price series. We will then we discuss the statistical significance of the results.

Examining firstly the results on the index data in Figure 5, we can see that the DE algorithm does indeed find levels in two out of the three indices that lead to out-of-sample improvements. For example, in the Dow Jones Industrial Average data, median out-of-sample value improves from approximately 0.47 to

0.51. Similarly, the S & P 500 index also appears to exhibit price levels, but the mean out-of-sample value is less: the improvement is from a median of 0.34 to 0.36 instead. (Remember that value is a significant underestimate of the true probability of a failed penetration due to the smoothing factor K in the value function.)

The NASDAQ Composite Index, on the other hand, exhibits an out-of-sample experimental decrease in median value from 0.36 to 0.34. However, it should be noted that out that the variance in the level values for this index is quite high in the control case, but when DE is applied, the optimization process decreases variance considerably. In other words, DE makes the level quality more predictable, albeit at the cost of a slight decrease in median value.

Figures 6-7 depict the results of applying the control and DE algorithms to the six stocks. Again, DE results in improvements in some cases (namely, Apple, Johnson & Johnson and Boeing Airlines), but in other cases it fails to improve median value. The probable explanation for this may be market volume: according to Table 2, these are the three stocks with largest volume. The remaining three markets have less volume, therefore less liquidity, and therefore the presence of price levels appears to be more difficult to detect.

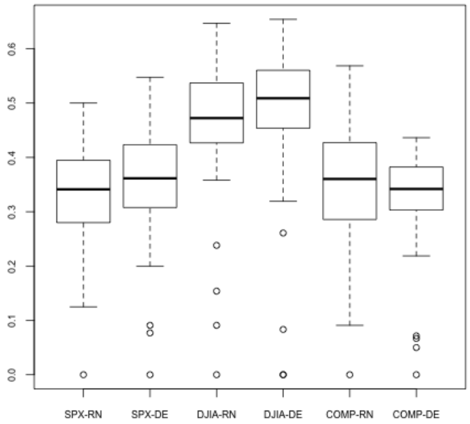

Fig. 5. Market-algorithm (x axis) vs out-of-sample value for best individual (y axis), for indices.

Finally, we performed a series of statistical significance tests to verify the results inferred visually from the box plots. For each market, two statistical tests were carried out: a standard T-Test comparing the mean out-of-sample performance of the control algorithm vs DE, and a non-parametric Wilcoxon signed rank test comparing the medians. Both tests are those implemented in the widely used statistical computing package R [6].

The results of these tests are shown in Table 3, where lower p-values indicate increased likelihood that the out-of-sample means/medians over fifty trials are not the same. A number of the tests, particularly those on the Dow Jones Industrial Average data and Boeing Airlines, are significant with 95% confidence (in fact, the DJIA tests are also significant at 99% level). Conversely, the tests show no significant differences for some of the other markets such as Apple and Celgene – although several of the test are *close* to significant.

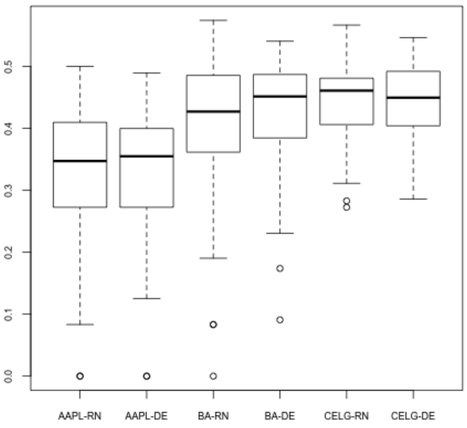

Fig. 6. Market-algorithm (x axis) vs out-of-sample value for best individual (y axis), for stocks 1-3.

5 Conclusion

To conclude, this paper has investigated the concept of price levels – anecdotally "support" and "resistance" – in markets. We have used Differential Evolution to learn these levels in markets. The learned levels were then compared to levels selected randomly. The results indicate that in some markets (especially those with higher liquidity) these levels do exist, and their presence can be detected with statistical significant using our approach.

References

1. Brabazon, A. & O'Neill M. *Biologically Inspired Algorithms for Financial Modelling.* Natural Computing Series, Springer, 2006.
2. Pederson M. *Good Parameters for Differential Evolution.* Hvass Laboratories, Technical Report HL1002, 2010.
3. Pring M. *Technical Analysis Explained.* McGraw-Hill, 2002.
4. Storn, R. & Price, K. Differential Evolution – A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces. *Journal of Global Optimization* 11: 341-359, 1997.
5. URL: http://nz.finance.yahoo.com/
6. R Development Core Team, *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. URL: http://www.R-project.org.

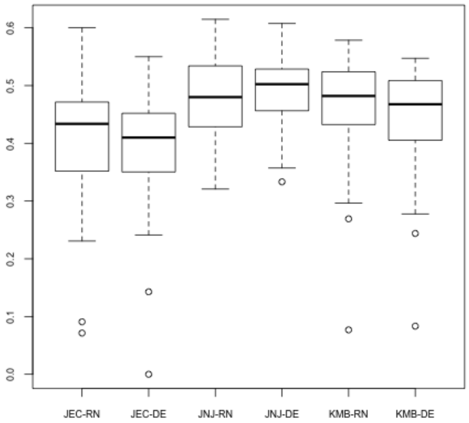

Fig. 7. Market-algorithm (x axis) vs out-of-sample value for best individual (y axis), for stocks 4-6.

Table 3. P-values from paired statistical significance tests comparing the mean (T-Test) and median (Wilcoxon) out-of-sample values of the best individuals obtained using DE compared to randomized level selection. Circles denote significant differences in the means/medians at 95% confidence.

	T-Test	Wilcoxon
SPY	0.062	0.013 ●
DJIA	0.003 ●	0.003 ●
COMP	0.069	0.121
AAPL	0.677	0.478
BA	0.023 ●	0.043 ●
CELG	0.861	0.905
JEC	0.293	0.255
JNJ	0.071	0.135
KMB	0.104	0.107