# Optimization of Trading Rules

## with a Penalty Term for Increased Risk-Adjusted Performance

**Thomas Hellström**
Department of Computing Science
Umeå University
Sweden
email: thomash@cs.umu.se
www.cs.umu.se/~thomash



## Overview

- Trading rules
- Performance Evaluation
- Benchmarks
- Constrained Optimization
- Empirical results



## Trading Rules

**are used for decision support:**

$$T(t) = \begin{cases} \text{Buy} & : \text{ if } g(t) = 1 \\ \text{Sell} & : \text{ if } g(t) = -1 \\ \text{Do nothing} & : \text{ if } g(t) = 0 \end{cases}$$

$g$ is a function of the previous stock prices $Close$:

$$g: \{Close(t), Close(t-1), ..., Close(t-k)\} \curvearrowright \{-1, 0, 1\}.$$

**g can often be parameterized as *g*[*X*], and optimized w.r.t *X***



## Example of a Trading Rule

**Example 1.** *Function g is defined as*

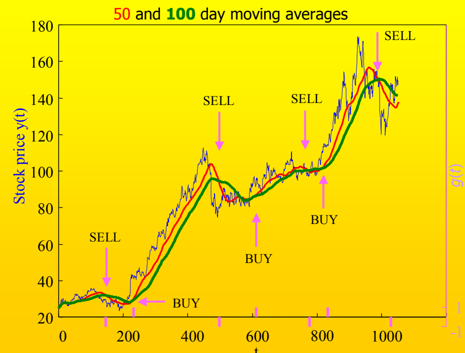
$$g(t) = \begin{cases} 1: & \text{if } mav_S(t) > mav_L(t) \ \wedge \ mav_S(t-1) \le mav_L(t-1) \\ -1: & \text{if } mav_S(t) < mav_L(t) \ \wedge \ mav_S(t-1) \ge mav_L(t-1) \\ 0: & \text{otherwise} \end{cases} \quad (3)$$

*where $mav_k(t)$ is a moving average of length k defined as*

$$mav_k(t) = \frac{1}{k} \sum_{m=0}^{k-1} Close(t-m). \quad (4)$$



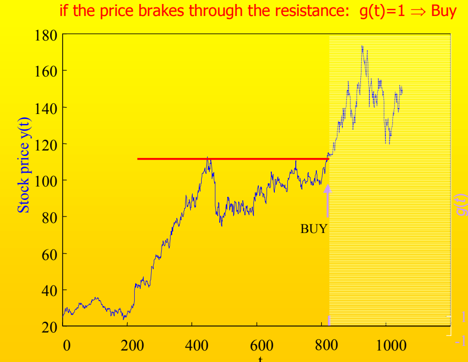
## Example of a Trading Rule





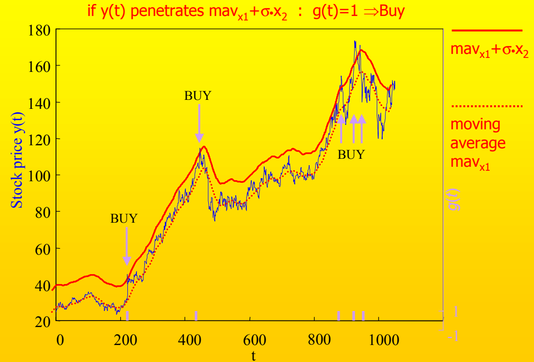

## Level of Resistance

if the price brakes through the resistance: g(t)=1 ⇒ Buy



## Trading Channel Breakout

if y(t) penetrates $mav_{x1}+\sigma \cdot x_2$ : g(t)=1 ⇒Buy






## Gaussian Volume

The Gaussian volume $V_n(t)$ is a transformation of the traded volume (number of stocks) $V(t)$ defined as

$$V_n(t) = (V(t) - m_V(t))/\sigma_V(t), \quad (5)$$

where $m_V(t)$ and $\sigma_V(t)$ are computed in a running window of length $n$ as

$$m_V(t) = \frac{1}{n}\sum_{i=1}^{n} V(t-i) \quad (6)$$

and

$$\sigma_V = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(V(t-i) - m_V(t))^2}. \quad (7)$$

$V_n$ expresses the number of standard deviations, by which the volume differs from its running mean. The normalization makes it possible to compare values of $V_n$ for different stocks and also for different times. In this paper the Gaussian volume $V_{10}$ is used and is denoted by $gvol10$, since this is the name of the ASTA [9] implementation of the function.



## Combining Moving Averages with Volume

The crossing moving averages is combined with the Gaussian volume into one compound trading rule *mav*:

$$mav(x_1, x_2, x_3) = Mavx(x_1, x_2) \wedge gvol10 > x_3,$$

where

$$Mavx(x_1, x_2) = mav_{x_1}(t) > mav_{x_2}(t) \wedge mav_{x_1}(t-1) \le mav_{x_2}(t-1)$$

and $mav_{x_1}$ and $mav_{x_2}$ are given by (4).



## Trading Channel Breakout

The main part of this trading rule is what is popularly known as *Bollinger Bands* (see e.g. page 91 in [15]). The complete trading rule is defined as

$$break(x_1, x_2, x_3) = breakout(x_1, x_2) \wedge gvol10 > x_3, \quad (10)$$

where the *breakout* function is defined as

$$breakout(x_1, x_2) = \begin{array}{l} Close(t) > (mav_{x_1}(t) + x_2 \cdot \sigma_{x_1}(t)) \quad \wedge \\ Close(t-1) \le (mav_{x_1}(t) + x_2 \cdot \sigma_{x_1}(t)) \end{array} \quad (11)$$

and $mav_{x_1}(t)$ is given by (4). Function $\sigma_{x_1}(t)$ computes the standard deviation of the *Close* as

$$\sigma_{x_1}(t) = \sqrt{\frac{1}{x_1 - 1}\sum_{i=0}^{x_1-1}(Close(t-i) - mav_{x_1}(t))^2}. \quad (12)$$

The idea is to define an upper boundary for a trading channel and generate a *Buy* signal when the *Close* penetrates this boundary from below. This upper boundary is defined as the sum of a moving average $mav_{x_1}$ and $x_2$ times an estimate of the standard deviation $\sigma_{x_1}$.



## Level of Resistance

The trading rule *Level of Resistance*, in this paper denoted *resist*, is based on a technique commonly executed by manual inspection of the stock charts. The general idea is to identify *peaks* in a window backwards, where the *Close* price is roughly the same. When such peaks are found, a *Buy* signal is generated if the *Close* price crosses from below the level for the found peaks. We define the trading rule *resist* as

$$resist(x_1, x_2, x_3, x_4) = xresist(x_1, x_2, x_3) \wedge gvol10 > x_4 \quad (13)$$

where

$$xresist(x_1, x_2, x_3) = Close(t) > plevel \wedge Close(t-1) \le plevel \quad (14)$$

and

$$plevel = \begin{cases} l: & \textit{if} \text{ at least } x_2 \text{ peaks in } Close \text{ that differs by less than } x_3\% \text{ can be identified at level } l \text{ in an } x_1\text{-day-long window backwards.} \\ 0: & \text{otherwise} \end{cases}$$



## Performance Evaluation

**Performance measures are needed twice:**

1) **In the learning phase when optimal parameters are determined**
2) **When the trading rules are tested out of sample**

**Two common ways to compute performance:**

- **Profit in Simulated Trading**
- **Hit rate at a fixed prediction horizon**

**In this paper the Hit rate is used.**

## Hit Rate

**The positive hit rate for a Buy rule is the fraction of buy signals, which are followed by an increase in stock price:**

For a time period $[1,...,T]$ and a set of stocks $S$, the h-day positive hit rate for a *Buy* rule $g$ is defined as

$$H_g^+ = \frac{card\{(t,s)|R_h^s(t+h) > 0, g_s(t) = 1, 1 \le t \le T-h, s \in S\}}{card\,\{(t,s)|R_h^s(t+h) \ne 0, g_s(t) = 1, 1 \le t \le T-h, s \in S\}} \quad (15)$$

where $g_s$ is the function specifying the trading rule as described in (1). The return $R_h^s$ is the relative change in price and is defined as

$$R_h^s(t) = 100 \cdot \frac{Close_s(t) - Close_s(t-h)}{Close_s(t-h)} \quad (16)$$

**The negative hit rate for a Sell rule is the fraction of sell signals, which are followed by a decrease in stock price**



## Benchmarks

**A benchmark should provide an alternative and standardized way to produce predictions.**

**The Naive Prediction of Return asserts today's return $R_h^s(t)$ as the best estimate of $R_h^s(t+h)$. For a time period $[1,...,T]$ and a set of stocks $S$, the $h$-day positive hit rate for the naive return predictor is computed as**

$$H_N^+ = \frac{card\,\{(t,s)|R_h^s(t+h) > 0, R_h^s(t) > 0, 1 \le t \le T-h, s \in S\}}{card\,\{(t,s)|R_h^s(t+h) \ne 0, R_h^s(t) > 0, 1 \le t \le T-h, s \in S\}}. \quad (17)$$

In this paper we use 1-day and 5-day returns ($h = 1$ and $h = 5$ in equation (17)) to form two benchmarks denoted *Naive-1$_+$* and *Naive-5$_+$* respectively.



## Benchmarks

**The Naive $\varepsilon$ prediction asserts today's price $Close_s(t)$ as the best estimate of $Close_s(t+h)$. For a time period $[1,\ldots,T]$ and a set of stocks $S$, the $h$-day positive hit rate for the Naive-$\varepsilon$ predictor is computed as**

$$H_\varepsilon = \frac{card\,\{(t,s)|R_h^s(t+h) > 0, 1 \le t \le T-h, s \in S\}}{card\,\{(t,s)|R_h^s(t+h) \ne 0, 1 \le t \le T-h, s \in S\}}.$$



## Optimizing the Trading Rules

**The function *g* is normally parameterized with a few parameters $X$ that can be determined to optimize performance on the training data.**
**The notation *g*[$X$] denotes this parameterization.**

**The trading rule normally issues Buy and Sell signals only for a minor part of the time steps. This is bad for two reasons:**

1) **Bad statistical significance for the performance**

2) **Risk for over optimization. I.e: bad generalization**



## A Constrained Optimization Problem

**We therefore formulate a constrained optimization problem for a Buy rule *g* :**

$$\begin{aligned}
&\arg\ \underset{x}{max} H^+_{g[x]} \\
&\text{s.t.} \\
&card\{(s,t)|g_s[x](t) = 1, t \le T-h,\ s \in S\} \ge N_0, \\
&x_L \le x \le x_H
\end{aligned}$$

**$X_L$ and $X_H$ are lower and upper bounds for the unknown parameters $X$ and the other constraint is the total number of generated Buy signals. Using a hard constraint leads to a non-smooth problem. Furthermore, it is hard to decide on a crisp value for $N_0$.**



## Reformulation to a Smooth Problem

$$\begin{aligned}
&\arg\ \underset{x}{max}\ H^+_{g[x]} \cdot\ support_{N_0}\,(card\{(s,t)|g_s[x](t) = 1, t \le T-h, s \in S\}) \\
&\text{s.t.} \qquad\qquad\qquad\qquad (20) \\
&x_L \le x \le x_H
\end{aligned}$$

where $support_{N_0}$ is given by the sigmoid function

$$support_{N_0}(n) = \frac{1}{1 + e^{-\alpha(n-\beta)}}. \quad (21)$$

The parameters $\alpha$ and $\beta$ are computed to fulfill the equations $support_{N_0}(N_0) = 0.99$ and $support_{N_0}(N_0 \cdot 0.5)\ =\ 0.01$

**The constraint acts as a regularizer since the search space for the function *g* is reduced by requiring a minimum number of trading signals.**

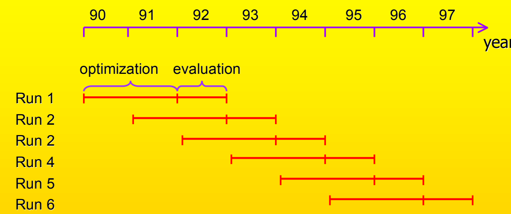
## Optimization

- The optimization problem is a box-bounded non-convex global optimization problem, where no derivates are available.
- In this paper we are using the DIRECT algorithm by Jones (1993). The algorithm estimates the Lipschitz constant and uses it to control the trade-off between global versus local search.



## Sliding Windows

Since we can't use cross validation




The performance is computed as the average performance for the 6 runs



## Results

**The 32 largest Swedish stocks have been used in the tests:**

Table 2: Hit rate and number of selected points for optimized trading rules. Totals from 6 1-year test periods (1992-1997) with the preceding 2 years for training. 5-day prediction horizon.

| | Method | $H_{tr}$ | $N_{tr}$ | $H_{te}$ | $N_{te}$ | 90%−low $H_{te}$ |
|---|---|---|---|---|---|---|
| Regularized trading rules | $resist_{100}$ | 63.61 | 753 | 58.72 | 516 | 55.03 |
| | $break_{100}$ | 68.36 | 708 | 64.22 | 450 | 60.33 |
| | $mav_{100}$ | 65.62 | 701 | 61.21 | 397 | 57.01 |
| Non regularized | $resist_1$ | 86.27 | 51 | 64.29 | 28 | 47.00 |
| | $break_1$ | 73.56 | 295 | 56.32 | 190 | 50.09 |
| | $mav_1$ | 85.11 | 94 | 61.70 | 47 | 48.67 |
| Non regularized | $Naive-e$ | 50.83 | 80456 | 52.35 | 42372 | 51.95 |
| | $Naive-1_+$ | 51.97 | 33437 | 52.77 | 18257 | 52.16 |
| | $Naive-5_+$ | 52.01 | 37519 | 52.17 | 20539 | 51.59 |

**The non-regularized optimization over-fits data and is no better than the bench marks out-of-sample. The regularized rules are all significantly better than the bench marks.**



## Example of Optimized Trading Rules for 1992

Table 4: Optimized trading rules for 1992. 5-day prediction horizon.

| Method | Optimized expression |
|---|---|
| $resist_{100}$ | $xresist(74,3,5.72) \wedge gvol10 > 0.67$ |
| $break_{100}$ | $breakout(38,1.39) > 0 \wedge gvol10 > 1.87$ |
| $mav_{100}$ | $Mavx(3,37) \wedge gvol10 > 0.33$ |
| $resist_1$ | $xresist(21,6,1.83) \wedge gvol10 > 0.67$ |
| $break_1$ | $breakout(124,2.5) > 0 \wedge gvol10 > 2.5$ |
| $mav_1$ | $Mavx(11,112) \wedge gvol10 > 1.44$ |

**Each year gets different optimal rules**



## Stability of the Found Optima

**The stability and the relevance of the found optima is also tested. The trading rules for 1992 are applied not only for 1992 but also for the following years up to 1997:**

Table 6: Hit rate for trading rules optimized with data from 1990-1991. 5-day prediction horizon.

| Method | 92 | 93 | 94 | 95 | 96 | 97 | Average | $H_{te}$ |
|---|---|---|---|---|---|---|---|---|
| $resist_{100}$ | 54.7 | 58.7 | 51.0 | 56.6 | 52.5 | 53.5 | 54.6 | 58.72 |
| $break_{100}$ | 59.3 | 62.9 | 59.4 | 52.2 | 57.0 | 53.9 | 57.1 | 64.22 |
| $mav_{100}$ | 56.7 | 70.3 | 53.9 | 51.2 | 61.5 | 59.7 | 58.7 | 61.21 |

**The average performance is lower than for the year-by-year optimized rules ($H_{te}$).**



## References


- **Optimization of Trading Rules with a Penalty Term for Increased Risk-Adjusted Performance.** Published in Advanced Modeling and Optimization, 2(3):135-149, 2000.
- **Global Optimization of Costly Nonconvex Functions, with Financial Applications.** (with Kenneth Holmström**.** Theory of Stochastic Processes 2000.
- **Parameter Tuning in Trading Algorithms Using ASTA.** (with Kenneth Holmström). Published in Y.S. Abu-Mostafa, B.LeBaron, A.W. Lo, and A.S. Weigend (Editors), Computational Finance 1999, p. 343-357.
- **ASTA - a Tool for Development of Stock Prediction Algorithms.** Published in Theory of Stochastic Processes, 5(21)(1-2):22-31, 1999.
- **ASTA - User's Reference Guide.** Technical Report UMINF-00.16 ISSN-0348-0542, Department of Computing Science Umeå University, Umeå Sweden, 2000.
- **D.R. Jones, C.D. Perttunen, and B.E. Stuckman.** Lipschitzian optimization without the Lipschitz constant. Journal of Optimization Theory and Applications, 79(1):157--181, October 1993.


email: thomash@cs.umu.se
www.cs.umu.se/~thomash