

Tools and Strategies for Development of Stock Prediction Algorithms

The Second International School on Actuarial and Financial Mathematics
Kyiv Ukraine
10th of June 1999

Thomas Hellström
University of Umeå
Sweden

<http://www.cs.umu.se/~thomash/>
email: thomash@cs.umu.se

© Thomas Hellström 2011

1

Contents:

- ASTA - introduction
- Optimization of trading rules
- Predicting Ranks instead of Returns
- Time series prediction with ASTA

© Thomas Hellström 2011

2

Wanted:

- A test bench for old and new trading strategies
- An interactive development tool for trading rules
- A non-interactive development tool for trading rules
- Data generation for post processing

© Thomas Hellström 2011

3

Two Formulations of the Prediction Task

Methods with a fixed prediction horizon

- "The Time Series Approach"
- "The Trading Rule Approach"

Problem:

We don't necessarily intend to sell the stocks h days after we bought them

Methods with no fixed prediction horizon

- Simulated trading with buy- and sell-rules

Problem:

Fewer points give lower statistical significance

© Thomas Hellström 2011

4

The Trading Rule Approach:

$$T(t) = \begin{cases} \text{Buy} & : g(X(t)) > 0 \\ \text{Sell} & : g(X(t)) < 0 \\ \text{Do nothing} & : \text{otherwise} \end{cases}$$

$$X(t) = (R_1(t), \dots, R_N(t))$$

X can be: Past values of C, H, L, V
or derived entities: Volatility, Trend, ...

Learning Task:

Find a function g that produces the best performance at a fixed prediction horizon OR when applying the trading rule T

Drawback: Statistical significance;
The Buy and Sell signals are $\ll N$

© Thomas Hellström 2011

5

Design Idea for an Artificial Trader

The trading algorithm is expressed as a Buy Rule and a Sell Rule:

Buy Rule :

$\text{Close}(T) > \text{Close}(T-1) \ \& \ \text{Volume}(T) > \text{mean}(\text{Volume}(T-10:T))$

Sell Rule :

$\text{Loss} > 10\% \mid \text{Profit} > 20\% \mid \text{Trend} < -2\%$

© Thomas Hellström 2011

6

What Is Good Performance?

- Which performance measures are relevant
 - Total profit
 - Average profit per day
 - Profit per year
 - Number of trades
- What is the best way to measure the performance
- To what do we compare when we say it's good ?
 - Benchmarks: Official Stock Indexes

© Thomson Intelligence 2000

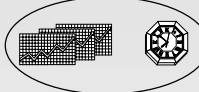
7

Basic Architecture of ASTA

- Operations**
- Initialize
 - Step in time

- Attributes**
- Stock data
 - Index
 - Current time T

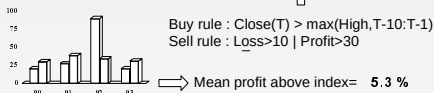
Market



Trader

- Operations**
- Initialize
 - Buy recommendations
 - Sell recommendations
 - Buy
 - Sell

- Attributes**
- Portfolio
 - Cash
 - Wealth
 - Buy rule
 - Sell rule
 - Other parameters



© Thomson Intelligence 2000

8

Simulation of Simple Buy and Sell Rules

ASTA

Stocks: 320 From date: 87 To date: 97 Courtage (%): 0.15 Min Courtage: 90 Min buy (%) per trade: 5 Max buy (%) per trade: 20 Initial Cash: 100000

Buy rule: $\text{Close}(T) > \max(\text{High}, 10; T-1)$

Sell rule: $\text{Loss} > 10 \mid \text{Profit} > 30$

Run Save graph Optimize

Market: 32 stocks Dates: 820104-980409 (4071 days)

Load Generate Save

Dump Trades: ☐ To window ☐ To file ☐ Diagram

Buy price: ☐ Today's ☒ Tomorrow's

Sell price: ☐ Today's ☒ Tomorrow's

Performance:

Annual profits:	87	88	89	90	91	92	93	94	95	96	97	Mean	Total
Strategy profit	-3.5	48.4	30.2	-23.7	12.2	-4.7	61.0	10.3	41.6	32.5	23.5	21.6	583.3
Index profit	-7.9	51.9	22.9	-29.7	5.4	-0.0	52.1	6.4	18.3	38.2	23.8	15.2	310.3
Difference profit	4.3	-3.5	7.3	7.0	6.8	-4.7	9.7	13.7	23.3	-6.7	-0.3	6.3	273.0
Number of trades	79	34	46	64	51	39	51	54	27	26	25	45	496

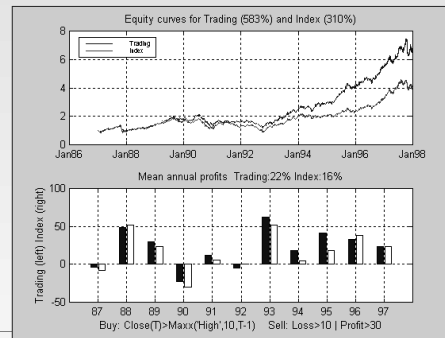
Switch to graph window for performance plots

Help End

© Thomson Intelligence 2000

9

Performance of Simple Buy and Sell Rules



© Thomson Intelligence 2000

10

Trading Simulation of the Stochastics Indicator

ASTA

Stocks: 320 From date: 87 To date: 97 Courtage (%): 0.15 Min Courtage: 90 Min buy (%) per trade: 5 Max buy (%) per trade: 20 Initial Cash: 100000

Buy rule: $\text{Stoch}(30, 3, 20, 80) > 0$

Sell rule: $\text{Stoch}(30, 3, 20, 80) < 0$

Run Save graph Optimize

Market: 32 stocks Dates: 820104-980409 (4071 days)

Load Generate Save

Dump Trades: ☐ To window ☐ To file ☐ Diagram

Buy price: ☐ Today's ☒ Tomorrow's

Sell price: ☐ Today's ☒ Tomorrow's

Performance:

Annual profits:	87	88	89	90	91	92	93	94	95	96	97	Mean	Total
Strategy profit	19.7	57.5	75.2	-11.8	18.7	34.5	279.4	26.0	24.5	33.0	2.9	51.1	1758.8
Index profit	-7.9	51.9	22.9	-29.7	5.4	-0.0	52.1	6.4	18.3	38.2	23.8	15.2	310.3
Difference profit	27.6	5.6	52.3	17.9	13.3	34.5	227.3	23.5	6.2	-5.1	-20.9	35.9	1448.5
Number of trades	23	36	38	61	64	85	45	69	49	50	64	54	686

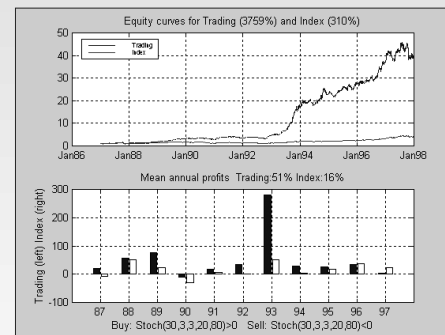
Switch to graph window for performance plots

Help End

© Thomson Intelligence 2000

11

Performance of the Stochastics Indicator



© Thomson Intelligence 2000

12

Predefined Functions in ASTA

- **Feature function:**
 - Close, High, Low, Volume, Rank, Trend, Volatility,...
- **Indicator functions:**
 - Stochastics, MACD, RSI, Keyrev, ClosePos, Runs, Days, Months, Daysin, Profit, Loss, Potential...
- **Operator functions:**
 - Min, Max, Mav, Crossings, Repeats,...

© 1999-2004 by ASTA Software AG

13

Using the Trader as an Objective Function

p = fasta(BuyRule, SellRule)

Example:

p = fasta('Trend5>0.5 & Gvol5>1.5', 'Loss>10 | Profit>20')

Strategies:

- **Parameterize the Buy Rule and Sell Rule:**
 - Buy Rule: 'Trend5>a & Gvol5>b '
 - Sell Rule: ' Loss>c | Profit>d '
 - Optimize P w.r.t a, b, c and d
- **The Buy Rule and Sell Rule are symbolic expressions**
 - Create new rules by combining functions and operators with genetic methods for example

© 2009-2016 H. E. K. S. S.

The Stochastics Indicator

Definition:

$$\text{Stoch}[K, Ks, D, \text{Buylevel}, \text{Sellevel}](t) = \frac{(\text{Close}(T) - L)}{(H - L)}$$

where

$$L = \text{mav}(\min(\text{Low}(t-K:t)), Ks)$$
$$H = \text{mav}(\max(\text{High}(t-K:t)), Ks)$$

Decision rule:

Buy if $\text{Stoch}(t) > \text{Buylevel}$

Sell if $\text{Stoch}(t) < \text{Sellevel}$

© Thomas W. Priddy.com 2015 15

The Stochastics Indicator

The top chart, titled "Electrolyse B", displays the price of the asset from January 1996 to January 1997. The y-axis ranges from 200 to 450. The price starts around 300, fluctuates, and then shows a significant upward trend starting in late 1996, peaking near 400 in early 1997.

The bottom chart, titled "Stochastics indicator", shows the Stochastics indicator for the same period. The y-axis ranges from 0 to 150. The indicator fluctuates between 50 and 100, with a notable peak in late 1996 followed by a sharp decline and a subsequent recovery. Vertical dashed lines indicate buy and sell signals: "Buy: Stoch(30,3,3,20,80,PL0T)>0" in late 1996 and "Sell: Stoch(30,3,3,20,80)<0" in early 1997.

P(Sellevel) for 11 Years

The graph displays the mean excess profit percentage as a function of the selection level (Sellevel) over an 11-year period. The y-axis, labeled 'Mean excess profit %', ranges from -5 to 45 in increments of 5. The x-axis, labeled 'Sellevel', ranges from 0 to 80 in increments of 10. The data is represented by a jagged line that fluctuates between approximately 20% and 40% profit for Sellevel values between 0 and 30. At Sellevel 35, there is a sharp vertical dashed line, and an arrow points to it with the text 'Maximum?'. Following this point, the profit percentage drops significantly, fluctuating between 10% and 20% for Sellevel values between 40 and 60, and then declining further to around 5% for Sellevel values between 65 and 80.

Sellevel	Mean excess profit %
0	22
5	20
10	25
15	30
20	35
25	32
30	35
35	40
40	25
45	20
50	15
55	10
60	15
65	10
70	5
75	2
80	0

One Curve per Year

Where is the Maximum?

© T. J. O'Connell 2011

18

Risk

The function $P(\theta)$ is stochastic and has no well-defined "maximum"

Entities to maximize:

- The mean value: $E[P(\theta)]$
- The lower limit of a confidence interval:

$$P_{\text{low}} = E[P(\theta)] - \sqrt{V[P(\theta)]}$$
- The Sharpe Ratio:
$$\frac{E[P(\theta)]}{\sqrt{V[P(\theta)]}}$$

© Thomas H. Edison 2011

19

Modified Sharpe Ratio: $\frac{E[P(\theta)]}{\sqrt{V[P(\theta)]}}$



© Thomas H. Edison 2011

20

No Trading, Just Short Interest Rate

ASTA

Stocks	From date	To date	Transaction cost (%)	Min transaction cost	Min buy (%) per trade	Max buy (%) per trade	Initial Cash
ALL	88	97	0.15	90	5	20	100000

Buy rule:
 Sell rule:
 Predict:
 Predictor:

Market: 30 stocks, Dates: 02/01/04-8/04/09 (4141 days)
 Load: auto0
 Generate:
 Save:

Dump Trades: ☐ To window ☐ To file ☐ Diagram
 Buy price: Today's Tomorrow's
 Sell price: Today's Tomorrow's

Performance:

Annual profits:	88	89	90	91	92	93	94	95	96	97	Mean	Total
Strategy profit	10.3	11.9	13.3	12.4	13.9	8.8	7.7	9.1	6.1	4.2	9.8	192.5
Index profit	31.9	32.9	29.7	5.4	-0.0	82.1	4.6	10.3	38.2	23.8	18.7	245.3
Difference profit	-41.5	-11.0	43.0	6.9	13.9	-43.3	3.1	-29.1	-32.1	-19.6	-9.0	-199.8
Number of trades	0	0	0	0	0	0	0	0	0	0	0	0

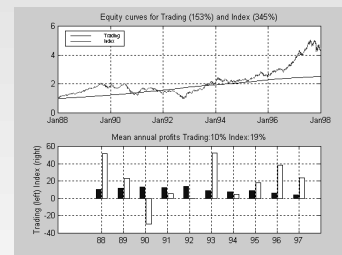
Switch to graph window for performance plots

Multiple runs: Random: Parameter: Values:
 Run: Sweep: Save graph: Help:
 Menu: End:

© Thomas H. Edison 2011

21

Surprisingly Enough.... Stocks Are NOT Clearly the Best Alternative Historically!



- * Stocks have been worse than a bank account 4 out of 11 years...
- * Equally good between Jan88 and Jan95

© Thomas H. Edison 2011

22

Dow Jones University

DJU

TEST YOUR KNOWLEDGE

3. Chartists often interpret trading volume as indicating bullish or bearish conditions. When a market is falling in price and the volume of trading activity is decreasing, it usually signals which of the following:

A. Market strength (bullish)
 B. Market weakness (bearish)
 C. Can't tell

The correct answer is

A. Market strength (bullish)
 A decline in volume often indicates that a trend is running out of steam. In this situation, in which prices are falling, a technical analyst could conclude that the decline in volume is a bullish sign because the downtrend may be ending. In our course, you'll learn to use technical analysis as a leading, directional indicator for individual stocks or whole markets.

© Thomas H. Edison 2011

23

Buy in Downtrend Combined with Decline in Traded Volume

ASTA

Stocks	From date	To date	Transaction cost (%)	Min transaction cost	Min buy (%) per trade	Max buy (%) per trade	Initial Cash
ALL	87	97	0.15	90	5	20	100000

Buy rule: Trend(20) < 0 & Gvol10 < 0
 Sell rule: Loss=10 | Profit=20
 Predict:
 Predictor:

Market: auto0
 Load: auto0
 Generate:
 Save:

Dump Trades: ☐ To window ☐ To file ☐ Diagram
 Buy price: Today's Tomorrow's
 Sell price: Today's Tomorrow's

Performance:

Mean Annual profits for 10 trading simulations (random=50):	87	88	89	90	91	92	93	94	95	96	97	Mean	Total
Strategy profit	-4.1	33.9	18.3	-28.4	15.1	10.2	73.0	7.9	12.4	38.2	42.7	22.4	385.0
Index profit	-7.9	31.9	22.9	-29.7	5.4	-0.0	82.1	4.6	10.3	38.2	23.8	18.7	245.3
Difference profit	1.7	4.0	-6.7	1.1	9.7	20.2	18.8	3.0	-5.8	0.0	18.9	5.3	274.7
St. dev. diff. profit:	7	6	11	7	8	8	9	9	5	7	9		
Number of trades	79	46	60	95	74	87	61	48	45	50	63	64	707

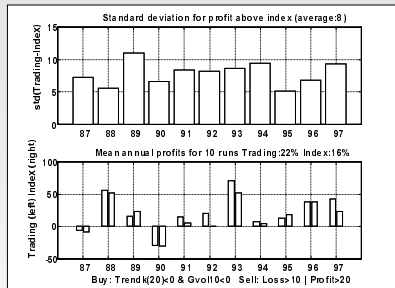
Switch to graph window for performance plots

Multiple runs: Random: Parameter: Values:
 Run: Sweep: Save graph: Help:
 Menu: End:

© Thomas H. Edison 2011

24

Buy in Downtrend combined with Decline in traded Volume



Better than index 9 out of 11 years.
The mean excess profit 5.9% is however less than the standard deviation 8%

© Thomas H. Edrington, 2011

25

Ranking Stock Returns $R(t)$

The k - day rank for stock m is defined as :

$$\text{Rank}_k^m(t) = \frac{\text{order}(R_k^m(t), \{R_k^1(t), \dots, R_k^N(t)\}) - 1}{N - 1} - 0.5$$

The stock with highest R gets rank 0.5

The stock with lowest R gets rank -0.5

The median stock gets Rank=0

Nice things about ranks:

- ✓ Predicting the rank is as good as predicting the returns
- ✓ Clear benchmark: hit rate for sign > 50%
- ✓ Uniform distribution
- ✓ The effect of global events gets automatically incorporated

© Thomas H. Edrington, 2011

26

One-Day Memory in the Ranks

$\text{Rank}_1(t+1)$ tabulated as a function of $\text{Rank}_1(t)$
207 Swedish stocks 1987-1997

Rank(t)	-0.45	-0.35	-0.25	-0.15	-0.05	0.05	0.15	0.25	0.35	0.45
Fraction% > 0	59.4	52.9	49.1	47.3	48	49.6	49.5	48.2	47.8	46.4
Mean rank	0.067	0.017	-0.005	-0.011	-0.011	-0.004	-0.005	-0.01	-0.014	-0.033
no. of obs.	30878	30866	31685	30837	30434	31009	31258	30539	30951	31550

With 59.4% probability:

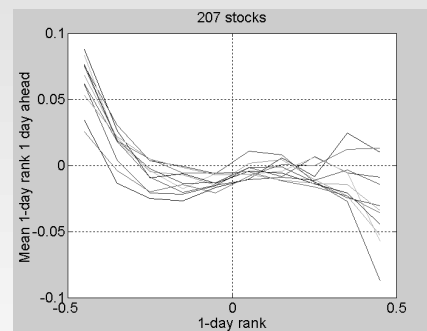
- ✓ The worst performing 10% of the stocks will be in the upper half next day.
- ✓ This prediction can be done EVERY day (since there is always a worst performing 10%).

© Thomas H. Edrington, 2011

27

One-Day Memory in the Ranks

one curve per year 1987-1997



© Thomas H. Edrington, 2011

28

Five-Day Memory in the Ranks

$\text{Rank}_5(t+5)$ tabulated as a function of $\text{Rank}_5(t)$
207 Swedish stocks 1987-1997

Rank(t)	-0.45	-0.35	-0.25	-0.15	-0.05	0.05	0.15	0.25	0.35	0.45
Fraction% > 0	66.6	51.9	50.1	50.3	50	50.2	49.4	48.4	48.4	44.3
Mean rank	0.04	0.01	0	0	0	0	0	-0.01	-0.01	-0.04
no. of obs.	30692	30689	31485	30652	30286	30772	30955	30349	30644	31026

With 56.5% probability:

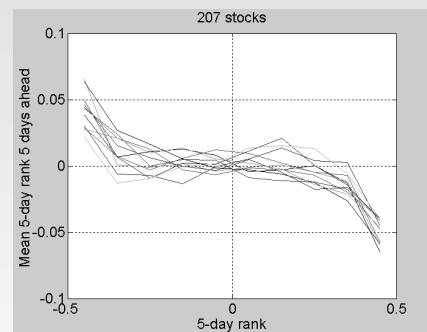
- ✓ The worst performing 10% of the stocks will be in the upper half next day.
- ✓ This prediction can be done EVERY day (since there is always a worst performing 10%).

© Thomas H. Edrington, 2011

29

Five-Day Memory in the Ranks

one curve per year 1987-1997



© Thomas H. Edrington, 2011

30

How About Predicting Ranks?

Ideas:

- ✓ Use the rank variables as inputs and/or outputs in a time series prediction. E.g:
 $\text{Rank}_i(t+1) = g(\text{Rank}_i(t), \text{Rank}_i(t-1), \dots, \text{Rank}_i(t-k))$
or

$$\text{Rank}_i(t+1) = g(R_1(t), R_5(t), \dots, R_k(t))$$

the function g may be a neural network, fuzzy rule base etc.

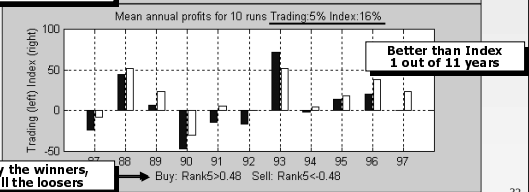
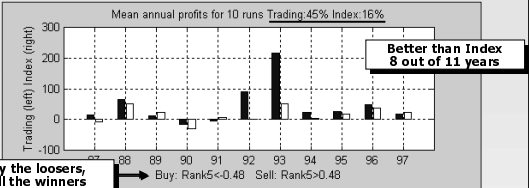
- ✓ Formulate Trading rules based on rank variables. E.g:

Buy if $\text{Rank}_1(t) < -0.40$ AND $\text{Rank}_5(t) < 0.3$

© Thomas H. Anderson 2011

31

Trading the 5-day rank in ASTA. 207 stocks



© Thomas H. Anderson 2011

32

Refined trading rules with 5-day rank.

ASTA

Stocks	From date	To date	Transaction cost (%)	Min transaction cost	Min buy (%) per trade	Max buy (%) per trade	Initial Cash
ALL	87	97	0.15	90	5	20	100000

Buy rule: $\text{Rank}_5 < -0.48 \& \text{Nstock} = 0$

Sell rule: $\text{Rank}_5 > 0.48 \mid \text{Loss} > 10 \mid \text{Profit} > 30$

Predict:

Predictor:

Market: 207 stocks, Dates: 82/104-98/04/09 (4156 days)

Load: Generate: Save:

Dump Trades: ☐ To window ☐ To file ☐ Diagram

Buy price: ☐ Today's ☐ Tomorrow's

Sell price: ☐ Today's ☐ Tomorrow's

Performance:

	87	88	89	90	91	92	93	94	95	96	97	Mean	Total
Strategy profit	155.5	59.8	13.8	3.3	32.7	226.1	318.9	49.4	12.0	53.7	17.0	76.6	18976.9
Index profit	-7.9	81.9	22.9	-29.7	9.4	-0.0	32.1	4.6	18.3	38.2	23.8	14.3	320.3
Difference profit	63.4	-2.1	-9.2	33.0	27.3	226.1	286.7	44.8	-6.3	15.6	-6.8	62.3	18656.6
St.dev. Diff.profit	21	15	11	18	33	104	130	12	9	12	15	40.2	1491
Number of trades	117	67	90	111	157	232	372	344	119	82	241	136	1491

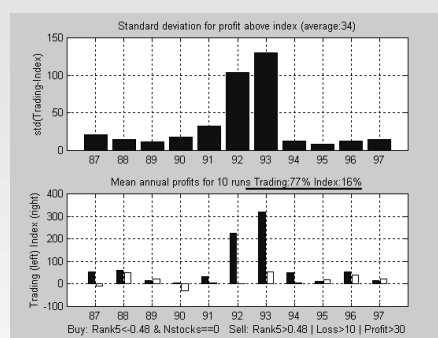
Graph window has been saved!

Multiple runs:

© Thomas H. Anderson 2011

33

Refined trading rules with 5-day rank.



© Thomas H. Anderson 2011

34

The Time Series Approach

Find a function g :

$$g(y(t), y(t-1), \dots, y(t-k)) \approx y(t+h)$$

where y is the returns for the stock prices

Minimize the RMSE:

$$P_h(t) = \sqrt{\frac{1}{N} \sum_{t=1}^N (g(t) - y(t+h))^2}$$

OR Maximize the hit rate for the sign predictions:

$$H = \frac{|\{t | y(t)g(t) > 0, t = 1, N\}|}{|\{t | y(t)g(t) \neq 0, t = 1, N\}|}$$

© Thomas H. Anderson 2011

35

ASTA for Time Series Analysis Prediction of Generalindex(T) from Generalindex(T-1)

ASTA

Stocks	From date	To date	Transaction cost (%)	Min transaction cost	Min buy (%) per trade	Max buy (%) per trade	Initial Cash
320828	89	97	0.15	90	0	20	100000

Buy rule:

Sell rule:

Predict: $\text{Trend}(1, T)$

Predictor: $\text{Trend}(1, T-1)$

Market: 32 stocks, Dates: 82/104-98/04/09 (4141 days)

Load: Generate: Save:

Dump Trades: ☐ To window ☐ To file ☐ Diagram

Buy price: ☐ Today's ☐ Tomorrow's

Sell price: ☐ Today's ☐ Tomorrow's

Performance:

	87	88	89	90	91	92	93	94	95	96	97	Mean	Total
Strategy profit	1.43	35.17	2152	57.42	2102	57.42	2102	57.42	2102	57.42	2102	57.42	2102
Index profit	1.43	35.17	2152	57.42	2102	57.42	2102	57.42	2102	57.42	2102	57.42	2102
Difference profit	0	0	0	0	0	0	0	0	0	0	0	0	0
St.dev. Diff.profit	0	0	0	0	0	0	0	0	0	0	0	0	0
Number of trades	0	0	0	0	0	0	0	0	0	0	0	0	0

Switch to graph window for performance plots

Multiple runs:

© Thomas H. Anderson 2011

36

