Comparison of trading algorithms*

Jan Zeman

4th year of PGS, email: janzeman3@seznam.cz
Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, CTU
advisor: T. V. Guy, Institute of Information Theory and Automation, ASCR

Abstract. The paper continues the previous research aimed at design the automatic trading system. The paper concerns rating the quality of designed approaches. It reviews both general methods and methods specialized to trading. The proposed method is a combination of them.

Abstrakt. Článek navazuje na předchozí výzkum týkající se obchodování s futures. Téma je zaměřeno na hodnocení dříve navržených algoritmů. Článek reviduje hodnotící metody jak obecné tak zaměřené na problematiku obchodování. Výsledkem je kombinovaná metoda, která je testována a hodnocena v závěrečné části.

1 Introduction

The paper towards automatic trading system for the futures contracts. The previous research concerns the task definition and basic solution [3, 4]. The previous work proposed many approaches and we have to compare them in order to select the most suitable one. Two subtask are considered: First is how to recognize the good approach standalone, and second deals with comparison of two approaches and selecting the better one.

To recognize a good approach, a final profit can be used as the measure of a success. However in trading applications, the continuous development of the cumulative profit has higher impact than the final profit. The analyzing the cumulative is more complex due to working with the whole sequence, but can bring better insight to approach quality.

The comparison of two approaches seems to be easy, when the approaches are tested on common data set. When even more data sets are available, the comparison becomes complex, because each data set produces one dimension in results, then the comparison of multidimensional results is needed. The typical problem is: Approach A makes a total profit at five data sets $ 100000 USD, but profit was positive at only two data sets. Approach B makes a total profit only $ 50000 USD, but it makes positive profit at four of five data sets. Which approach is better? Both approaches can win, but the best should be chosen according to the preference of trader.

The paper proposes a small review of the comparison methods and applies the methods to one of the solved problems.

The paper contains two main parts. Section 2 introduces the problematics and defines the task (Sec. 2.1), defines a coefficient characterizing the quality of approach using the cumulative gain (Sec. 2.2) and introduces methods for multi-dimension comparing (Sec. 2.3). Section 3 introduces futures trading (Sec. 3.1) and coefficients used in trading (Sec.

*This work has been supported by the grant MŠMT 1M0572.

3.2), defines algorithm of approaches rating (Sec. 3.3). The algorithm is applied and commented in Sec. 3.4.

2 Comparing methods

The section deals with definition of the solved task and given assumptions.

2.1 Task of interest

We assume a decision maker and system. Decision maker is human or machine with aims related to the system. The decision maker obtains a data y_t at the system, and design the decision u_t to reach his aims. The process is repeated each discrete time instant $t \in \{1, \ldots, T\}$. The aims of decision maker are characterized by a gain function G, which maps the system output and decisions to a real number. Higher value indicates higher success. The decision maker tries to maximize the gain function.

We focus on quality evaluation of designed decisions, hence we assume the knowledge of a whole data $y_1, \ldots, y_T$ and decision sequence $u_1, \ldots, u_T$. Moreover, we assume the knowledge of the gain function:

$$G : (y_1, \ldots, y_T, u_1, \ldots, u_T) \rightarrow \mathbb{R} \tag{1}$$

and its additive shape

$$G = \sum_{i=1}^{T} g_i, \qquad \text{where} \qquad g_i : (y_1, \ldots, y_i, u_1, \ldots, u_i) \rightarrow \mathbb{R}, \tag{2}$$

and g_i is called a one-step gain.

Let us define *cumulative gain* via:

$$G^t = \sum_{i=1}^{t} g_i. \tag{3}$$

The gain is a sum over all time instants $\{1, \ldots, T\}$, whereas cumulative gain is sum over the first t time steps $\{1, \ldots, t\}$, $t \leq T$. Hence, we use the term *final gain* for the gain from here onward. Moreover, the cumulative gain can be viewed as a sequence $G^1, \ldots, G^T$ and characterizes the approach behavior.

We assume that there are M different approaches trying to maximize the gain (2) and N testing data sets or experiment data available to compare the success of the approaches. In summary, we have $M \times N$ final gains to decide, which approach is the best. Moreover, we can obtain $M \times N \times T$ values, in order to analyze the approaches using the cumulative gains.

2.2 Cumulative gain comparison

It is disputable, whether the final gain is a good criterion for rating of the approaches. In some tasks, the good final gain can be reached only by a few last steps, hence the analysis of the cumulative gain is required. But working with a whole sequence of cumulative gain containing T values is difficult. Hence, it is needed to characterize the quality of cumulative gain by one coefficient, and this section defines such a coefficient.

The ideal cumulative gain increases, therefore the knowledge of a trend is important. To reach this knowledge, the sequence can be fitted by a linear function $y(t) = at + b$, where a, b are parameters. We assume a sequence of values $G^1, G^2, \ldots, G^T$, and we search the best values of coefficients a, b to minimize squared error $\min_{a,b} \sum_{t=1}^{T} (G^t - y(t))^2$. The obtained coefficients a_{min}, b_{min} characterize the nearest linear approximation of the original sequence. Hence, the values of a_{min}, b_{min} can be used to evaluate the success of the approach.

The coefficient a_{min} reflects a trend of cumulative gain. The positive value characterizes an increase, the negative one a decrease. The value of coefficient a is related to strength of the increase, higher value means sharper increase. Thus, it can be used as a relatively good criterion of the approach quality.

On the other hand, the linear approximation is not suitable, when the difference between original sequence and approximation $(G^t - a_{min}t - b_{min})$ is not normal distributed. This property cannot be warranted by any cumulative gain. Hence, the credibility of the coefficient a_{min} is lowered. The credibility of coefficient a_{min} is given by value of error squares $s = \sum_{t=1}^{T} (G^t - a_{min}t - b_{min})^2$, the less value of s brings better credibility of a_{min}. To obtain one characteristic coefficient, let us define *increase coefficient* c_I as follows:

$$c_I = \frac{a}{\log_{10}(s)}, \quad \text{with} \quad s = \sum_{t=1}^{T} (G^t - at - b)^2, \tag{4}$$

where a_{min}, b_{min} are coefficients of the best linear approximation of the cumulative gain sequence. The logarithm is used due to big differences in values of s for the trading task.

The higher value of c_I is rated as better result of an approach. The positive value of coefficient c_I characterizes the increase of cumulative gain, the weighting by difference s lowers the value of coefficient for bad fitted sequences. The coefficient c_I covers our requirements for working with cumulative gain, hence the further sections deals with comparing results obtained on more data sets.

2.3 Multi-dimension comparing

As was introduced, the comparison of two approach is simply, when they are tested at one data set, but when more data set is available, the decision become complex. The complexity originates from fact that the comparison has nature of multidimensional task, where each data set forms one dimension of compared vectors. Following two subsections deals with this task. Section 2.3.1 try to transform the multidimensional task to one-dimensional by weighted summing. Whereas, the Section 2.3.2 let the task multidimensional and defines comparison of vectors.

Analogical with Sec. 2.1, we assume M approaches and N testing data sets. The aim is select the best approach, hence we form M vectors $R^1, \ldots, R^M$ containing the results, which are quality measures related to each data sets. The quality measures can be final gain, increase coefficient, or other variable characterizing the approach quality. Thus, each vector contains N values $R^i = (r_1^i, \ldots, r_N^i)$. Our aim is to chose the best approach using only this vectors.

2.3.1 Weighted sum

The first simply solution is to summarize the results and evaluate

$$\mathcal{S}^m = \sum_{n=1}^{N} r_n^m$$

for each approach $m \in \{1, \ldots, M\}$. Then each approach is characterized by one real number and it is simple to compare them.

Summing the results is simply and effective, but has a lot of disadvantages. When one of data sets produces outstanding results, the total sum is influenced by this outlayer and the results are not correct. Moreover, the maximal obtainable results must be comparable for all data sets, because the higher potential gives higher weight to given data set. The maximal and minimal possible value of results can be calculated for some special tasks and using them the following coefficient can be defined:

$$FP_n^m = \frac{r_n^m - G_n^{min}}{G_n^{max} - G_n^{min}} \times 100\%, \tag{5}$$

where G_n^{min} and G_n^{max} are minimal and maximal result values obtainable at nth data set. Let the coefficient is called *final percentage*. The final percentage express the percentage of success reached by approach according to maximal and minimal potential results reachable on the given data set. Summing FP_n^m over $n \in \{1, \ldots, N\}$ brings the equivalent results, where each experiment has the same weight independent on its potential. Instead of summing, it is better to calculate the mean value:

$$MFP^m = \frac{1}{N} \sum_{n=1}^{N} FP_n^m \tag{6}$$

the results can be interpreted as mean potential percentage of the approach m. Let coefficient MFP^m is called *mean final percentage*. The coefficient (6) is generalized weighted sum. When the minimal results potential equals zero ($G_n^{min} = 0$), then it is equivalent to weighted sum with weights: $w_n = 1/G_n^{max}$.

The coefficient MFP assigns each approach one number and the searching the best approach is transformed to sorting the number.

2.3.2 Efficient solution

Another way to compare the vectors $R^1, \ldots, R^M$ is by defining dominating and efficient solution.

The vector $R^i = (r_1^i, \ldots, r_N^i)$ is *dominated* by vector $R^j = (r_1^j, \ldots, r_N^j)$ even if following inequalities are valid:

$$\forall n \in \{1, \ldots, N\} \quad r_n^i \leq r_n^j,$$

and

$$\exists n \in \{1, \dots, N\} \quad r_n^i < r_n^j.$$

Efficient solution is such a vector from the set $\{R^1, \dots, R^M\}$, which is not dominated by any other vector. The term of efficient solution is taken from multiobjective optimization [1].

Taking only efficient solutions, the set of outstanding solutions can be found. The efficiency does not mix results reached on different data sets, i.e. the outstanding results on one data set cannot help the approach rating such as in poor summing the gains.

On the other hand, the efficient solutions typically forms a subset of $\{R^1, \dots, R^M\}$. Hence, the method does not lead to one best approach, but it excludes a small set of outstanding approaches. The method cannot prefer one of efficient solutions, until the additional information about preferences is not added.

3 Example: commodity futures trading

The commodity futures trading is challenging task related to trading on stock exchanges and prices speculation. The commodity futures means an contract for delivering the commodity to given date in future. The price of contract is often object of speculation.

The speculator can speculate for following situations:

Price increase, the speculator buys the contract, it is said to open the *long* position. Then, he waits, until the price increases, and sells the contract (it is said to close the long position).

The profit is the difference of buy/sell contract price. The difference, whether speculator makes profit or loss, depends, whether the price follows his expectation. Hence, the profit from the long position is made, when the price increases, whereas the speculator loses the same value, when the price decreases.

Price decrease, the speculator sells the contact, it is said to open the *short* position. The fact, that he can sell not-owned contract, is related to principles of given exchange, the speculator can lend the contract for this operation. Then, he will buy the contract back, it is said to close the short position.

Indefinite, the speculator has no opened position. He is in so called *flat* position, or *out of market*. Speculator neither profits nor loses by this operation.

A transaction cost must be paid for each contract, which changes the position.

The period from entering the non-flat position at market to leaving the position is called *trade*. The trade is very important, because the profit in cumulative gain is only hypothetical. But at the end of the trade, the cumulative gain corresponds with the real realized profit.

3.1 Task definition

Let denote the price in time t by y_t and position held in time t by u_t. The structure of u_t is following: the absolute value $|u_t|$ sets the number of contracts in an open position;

and the signum of u_t sets the kind of position, minus for short and plus for long position. The flat position is characterized by $u_t = 0$.

For this notation the gain function is defined as:

$$G = \sum_{t=1}^{T} g_t = \sum_{t=1}^{T} \underbrace{(y_t - y_{t-1})u_{t-1} - C|u_t - u_{t-1}|}_{g_t}, \tag{7}$$

where C is the normalized transaction cost. For offline experiments, the transaction cost is artificially increased by so-called *slippages*. Slippages are required due to delay between prompting the market command and its realization, during this short time period the price can change. Second reason for slippages is that the action on market changes the price itself and this is often not included in off-line experiments. Both reasons causes that the price in real trading could be different from the value stored in data sets. To avoid this difference, the transaction cost has two parts $C = c + s$ for our task, where c is transaction cost payed to exchange provider for each contract in position, and s are slippages, which artificially make the transaction cost higher.

The slippages are estimated by an economic specialist. We use values obtained from Colosseum a.s. due our cooperation. Although the slippages makes the task more difficult, the trading system profitable at off-line data with slippages has big chance to be profitable in real trading.

3.2 Requirements to applicability

The economist have designed a lot of additional criteria to rate, whether the approach is good or bad. This criteria are closely related to the trading task. Moreover, the economist will decide, whether the approach will be applied in practice, hence is important to take this coefficients and criteria into a consideration. This section overview the main coefficients and introduces the criteria required to application of the approaches.

3.2.1 Main coefficients

Net profit is the same variable as the final gain (7).

Gross profit is the net profit calculated only over the profitable trades. The profitable trade is trade which starts with lower value of cumulative gain than finishes.

Gross loss is analogy with gross profit, but for non-profitable trades. The Gross profit is positive number, gross loss is negative number and net profit is sum of them.

Total cost is total amount of transaction cost c payed for realization of decision as was introduced in Sec. 3.1. The total cost is calculated via: $(-1)\sum_{t=1}^{T} c|u_t - u_{t-1}|$.

Total slippages is total amount of slippages s, calculated in analogy with transaction cost $(-1)\sum_{t=1}^{T} s|u_t - u_{t-1}|$. The slippages can be used for analyzing the results, because in the trading task is typical that slippages make the result negative (see [2]).

Trades is count of trades done during the experiment.

Winning/Losing trades is count of trades with positive/negative profit.

Days long/short/flat is count of time instants, when a contract was held in long/short/flat position. (The word 'days' is related to fact that we work with a day-data.)

Maximal drawdown is the biggest negative difference in cumulative gain sequence. This variable characterizes the risk related to given approach. The drawdown of bad approach is relatively same value as the final gain.

Length of drawdown characterizes the length of the maximal drawdown, i.e. how many time instants was the drawdown realized. Again, the bad approach has drawdown with comparable length as the data sequence.

3.2.2 Combinations of coefficients

The previous coefficient are raw coefficient obtainable from result. Following coefficients can be computed from the raw coefficients and give us criteria for identifying the good approach.

Percent profit gives percentage of winning trades:

$$\text{Percent profit} = \frac{\text{Winning trades}}{\text{Winning trades} + \text{Losing trades}}.$$

Profit factor is ratio of earned and lost money:

$$\text{Profit factor} = -\frac{\text{Gross profit}}{\text{Loss}}.$$

Profit per trade is average profit obtained in trade

$$\text{Profit per trade} = \frac{\text{Net profit}}{\text{Trades}}.$$

3.2.3 Criteria on good approach

There is a difference between theoretical design of approaches and its applicability in practice. Whereas, the theoretical success is each small bettering of an approach, the practical application demands significantly good results. The criteria to application of the tested approach for futures trading were designed by economic specialist from Colosseum a.s. The criteria are presented in Table 1.

3.3 Algorithm of rating

The decision, which approach is best, should be done using following rules:

1. The non-efficient approaches are excluded, the final gain is taken as measure of approach quality. This step chooses a subset of the original approaches.

Coefficient	Relation	Value
Net profit	greater than	0
Maximal drawdown	less than	1/10 net profit
Length of drawdown	less than	250 days
Percent profit	greater than	0.4
Profit factor	greater than	1.5
Profit per trade	greater than	$100 USD

Table 1: Requirements on approach to applicability in practice.

Ticker	Commodity	Exchange
CC	Cocoa	CSCE
CL	Petroleum-Crude Oil Light	NMX
FV2	5-Year U.S. Treasury Note	CBT
JY	Japanese Yen	CME
W	Wheat	CBT

Table 2: Reference markets, their tickers and exchanges.

2. The non-efficient approaches are excluded, the coefficient c_I is taken as measure of approach quality. This step chooses a subset of the original approaches.

3. The approaches are sorted by their MFP - the highest value as first.

4. The approaches are tested consequently, whether suffice the requirements on applicable approach. The proving is done over all data sets, hence each approach must satisfy $6 \times N$ conditions. The first, sufficient is rated as the best approach, because is efficient and has highest MFP.

3.4 Tuning the parameters

We have available price history from five market (see Tab. 2) and approach presented in [4], where are 2 parameters the length of regressor $l \in \{1, 2, \ldots, 10\}$ and the forgetting factor $\lambda \in \{1, 0.999, 0.99, 0.9\}$. (The explanation of the parameters is not important.) Thus, we have 40 couples of parameters and our aim is to estimate, which couple is the best. Due to availability of five data sets, the count of experiments is 200.

Table 3 reviews the results obtained by presented method (see Sec. 3.3). The values in the table were constructed by ordering the MFP coefficients (see Sec. 2.3.1), where the highest value of MFP was denoted by 1, second highest by 2 etc. And the highlighted approaches were marked as efficient in both steps 1 and 2 of algorithm from Sec. 3.3.

For last step of the algorithm, there is no approach satisfying all requirements for applicability. The nearest is the approach with the parameters $l = 1$ and $\lambda = 1$, where are satisfied 20 conditions from 30.

For the further research, the parameters couple $l = 1$ and $\lambda = 1$ will be used, although the non-applicability. The reason for this choice is that the given approach is the most successful and moreover the analysis with respect to c_I coefficient define in Sec. 2.2 reaches also the best results (see Tab. 4).

The testing of c_I coefficient showed that approaches with value $c_I > 1.5$ have increasing cumulative gain without big drawdowns. Hence, the coefficient c_I can be used for rating the best approach in further research.

4 Conclusion

The paper concerns with the criteria of comparing approaches testing on data sets. The algorithm of the best approach choosing is designed. The algorithm is applied on the results obtained in tuning approach for futures trading task, and it chooses the best approach.

The main advantage of the designed algorithm lies in possibility to compare the approaches tested on more data sets. The algorithm combines the simply method of weighted sum with efficients solutions and applicability of approach. This combination is also great advantage.

The disadvantage of given algorithm is that the algorithm can exclude all approaches due to applicability conditions. And opposite, the efficient solution often selects big subset.

The algorithm will be tested in further research, but it make the ground idea for further algorithms in rating the approaches.

References

[1] M. Ehrgott. *Multicriteria optimization.* Lecture Notes in Economics and Mathematical Systems. Springer-Verlag, 2000.

[2] M. Kárný, J. Šindelář, Š. Pírko, and J. Zeman. Adaptively optimized trading with futures. Technical report, 2010.

[3] J. Zeman. Futures trading: Design of a strategy. In *Proceedings of the International Conference on Operations Research and Financial Engineering 2009.* WASET, 2009.

[4] J. Zeman. A new approach to estimating the bellman function. In Pavelková Lenka Hofman Radek, Šmídl Václav, editor, *Proceedings of the 10th International PhD Workshop on Systems and Control.* ÚTIA, AV ČR, 2009.

l	$\lambda = 1$	$\lambda = 0.999$	$\lambda = 0.99$	$\lambda = 0.9$
1	[redacted]	[redacted]	26	28
2	[redacted]	17	23	30
3	[redacted]	15	25	27
4	[redacted]	19	18	33
5	[redacted]	[redacted]	20	35
6	[redacted]	12	22	36
7	5	14	31	37
8	[redacted]	16	29	38
9	[redacted]	24	34	39
10	[redacted]	7	32	40

Table 3: Comparison of 40 approaches for Bellman function estimation, each approach is defined by couple l and λ, the efficient solutions are highlighted and the numbers in table are order of approaches by MFP.

l	$\lambda = 1$	$\lambda = 0.999$	$\lambda = 0.99$	$\lambda = 0.9$
1	1.0551	-0.3355	-1.2594	-1.7522
2	0.2444	-0.1365	-0.3234	-1.6807
3	0.6385	0.3818	-0.5466	-1.7164
4	0.4861	-0.0215	0.1084	-1.8719
5	0.6014	0.2383	-0.2796	-2.3504
6	0.6046	0.0992	-0.4663	-2.8133
7	0.5481	0.1318	-1.6669	-3.4924
8	0.5002	-0.0869	-1.2192	-3.7682
9	0.3632	-0.7274	-1.9563	-4.3346
10	0.2865	0.4667	-1.7901	-5.0938

Table 4: The mean value increase coefficient c_I calculated over available data sets.