

Optimizing Sparse Mean Reverting Portfolios with AR-HMMs in the Presence of Secondary Effects

I. Róbert Sipos¹*, János Levendovszky¹

RESEARCH ARTICLE

Received 13 February 2014; accepted 06 February 2015

Abstract

In this paper we optimize mean reverting portfolios subject to cardinality constraints. First, the parameters of the corresponding Ornstein-Uhlenbeck (OU) process are estimated by auto-regressive Hidden Markov Models (AR-HMM) in order to capture the underlying characteristics of the financial time series. Portfolio optimization is then performed according to maximizing the mean return by the means of the introduced AR-HMM prediction algorithm. The optimization itself is carried out by stochastic search algorithms. The presented solutions satisfy the cardinality constraint thus providing a sparse portfolios which minimizes the transaction costs and maximizes the interpretability of the results.

The performance has been tested on historical data obtained from S&P 500 and FOREX. The results demonstrate that a good average return can be achieved by the proposed AR-HMM based trading algorithms in realistic scenarios. Furthermore, profitability can also be accomplished in the presence of secondary effects.

Keywords

mean reversion, Markov models, parameter estimation, financial time series, algorithmic trading

1 Introduction

Portfolio optimization was first investigated by Markowitz [22] in the context of diversification to minimize the associated risk and maximize predictability. Mean reversion is a good indicator of predictability, as a result, identifying mean reverting portfolios has become a key research area [4,9,30].

The novel contribution lies in the following facts: (i) in this paper we maximize the average return (which is optimized by the underlying marginal probability density functions (PDFs)) instead of the predictability parameter; (ii) we use the auto-regressive hidden Markov model (AR-HMM) approach, which is more general than the OU modeling. Furthermore, a multi state AR-HMM can be considered as a generalization of these PDFs, and also enables us to perform prediction based algorithmic trading on financial time series. HMMs are widely used for predictions [6,15,18,21].

Introducing cardinality constraints in order to minimize the transaction costs, which involves optimizing sparse portfolios, turns the original problem into NP hard [25]. In this paper, the sparse portfolio optimization was carried out by using stochastic search algorithms [10].

Finally, trading is perceived as a walk in the “buy/sell” action space, which is then tested numerically on S&P500 series and FOREX rates, and the results exhibit good average returns.

The structure of the paper is as follows:

- in section 2, the motivations and the previous results are outlined;
- in section 3, the model and the notations are introduced;
- in section 4, the concept of AR-HMM is briefly summarized, and its connection to the mean reverting processes is given;
- in section 5, we optimize the portfolio by maximizing the average prediction based return;
- in section 6, the computational model is mapped out;
- in section 7, a detailed performance analysis is given based on historical data;
- in section 8, some conclusions are drawn.

¹ Department of Networked Systems and Services
Budapest University of Technology and Economics
H-1521 Budapest, P.O.B. 91, Hungary

*Corresponding author, e-mail: siposr@hit.bme.hu

2 Motivation

While trading with mean reverting portfolios, we need to make a decision on the fact whether the observed process exhibits mean reverting properties, or following a Brownian motion or even driven by mean aversion. Since these properties can change over time, a more flexible time dependent distribution model is needed.

In the previous works [30], linear regression based methods proved to be sufficient for OU parameter estimation. However, the shortcomings of this method that even a higher accuracy would be desired for the purpose of algorithmic trading, and also they are unstable under certain circumstances (e.g. when the level of mean reversion is low).

The novel approach introduced in this paper manages to overcome the abovementioned problems and provides a more general model which can handle heterogeneous time series, and gives more accurate and stable predictions for the future prices.

3 Model

In this section we describe the model and the concept of mean reverting portfolios.

The time series describing the prices of assets is denoted by $\mathbf{s}_t^T = (s_{1,t}, \dots, s_{n,t})$ where $s_{i,t}$ is the price of asset i at time instant t . The portfolio vector is denoted by $\mathbf{x}^T = (x_1, \dots, x_n)$ where x_i gives the number of possessed quantity from asset i . The value of the portfolio at time t is denoted by $p(t)$ and defined as

$$p(t) = \mathbf{x}^T \mathbf{s}_t = \sum_{i=1}^n x_i s_{i,t} \quad (3.1)$$

Our objective is to find the optimal portfolio \mathbf{x}_{opt} which maximizes a pre-defined objective function, such as average return (described in Section 5), subject to cardinality constraint which specifies that the number of non-zero components in \mathbf{x}_{opt} must not exceed a given number l . The optimal portfolio is sought under the assumption that the portfolio value $p(t)$ exhibits mean reverting properties and follows an Ornstein-Uhlenbeck (OU) process [26]. This is a frequent assumption in trading [7,20,26,27] which follows from the VAR(1) nature of the underlying asset process.

The OU process is characterized by the following stochastic differential equation

$$dp(t) = \vartheta(\mu - p(t))dt + \sigma dW(t), \quad (3.2)$$

where $W(t)$ is a Wiener process and $\vartheta > 0$ (mean reversion coefficient), μ (long-term mean) and $\sigma > 0$ (volatility) are constants. By using the Itô-Doeblin formula [17], one can obtain the following solution:

$$p(t) = p(0)e^{-\vartheta t} + \mu(1 - e^{-\vartheta t}) + \int_0^t \sigma e^{-\vartheta(t-s)} dW(s) \quad (3.3)$$

which implies that

$$E(p(t)) = p(0)e^{-\vartheta t} + \mu(1 - e^{-\vartheta t}) \quad (3.4)$$

and asymptotically

$$\lim_{t \rightarrow \infty} p(t) \sim N\left(\mu, \sqrt{\frac{\sigma^2}{2\vartheta}}\right) \quad (3.5)$$

Parameter ϑ determines the convergence speed of the process towards the mean, and inversely indicating the level of uncertainty (via the standard deviation of the asymptotic Gaussian distribution (3.5)). Hence, for convergence trading, larger ϑ implies a better portfolio, as it quickly returns to the mean with a minimum amount of uncertainty.

4 Modeling OU processes with AR-HMMs

A Hidden Markov Model (HMM) [1] is a statistical model which is an extension of Markov chains. Such models are widely used in econometrics [13], especially for predictions [11,31]. In our case we use them to predict future values of financial time series. As a type of mixture models, having enough degree of freedom (number of hidden states) HMMs are capable of modeling a large class of distributions. By AR-HMM we can capture both the long and short range dependencies, as it combines a Markov chain on the hidden variables, and statistical dependencies on the observed variables [3].

In this model, the current state is no longer directly visible to the observer, but each state emits an observable output quantity denoted by $\mathbf{x} = \{p_1, p_2, \dots, p_T\}$. Unlike the standard HMM assumption, in the case of auto-regressive HMMs (AR-HMM) the emissions are conditionally not independent given the hidden state $\mathbf{Q} = \{q_1, q_2, \dots, q_T\}$ [24]. In our approach, the observable output was treated as a continuous value, described by a Gaussian probability density function. Then the probability of emitting a specific output is determined by the conditional probability

$$P(p(t)|p(t-1), q_t = j, \Theta) = N(p(t)|\varphi_j p(t-1) + \mu_j, \sigma_j) \quad (4.1)$$

In other words, the observation depends on the hidden state, and on the previous observation through an additive autoregressive component. (Note that the μ and σ symbols were already introduced in (3.2), however, in this case they are vector quantities.) The transition probabilities of the underlying Markov chain, describing the jumping probabilities from one state to another is given by the transition probability matrix $A_{ij} = P(q(t+1) = q_i | q(t) = q_j)$, while $\boldsymbol{\pi}_N$ denotes the initial distribution vector. The complete model then described by

$$\Theta = \{\boldsymbol{\pi}, \mathbf{A}, \boldsymbol{\varphi}, \boldsymbol{\sigma}\} \quad (4.2)$$

From the OU process, defined in (3.2), we can obtain [12]

$$E(p(t)) = e^{-\vartheta \Delta t} p(t-1) + (1 - e^{-\vartheta \Delta t}) \mu \quad (4.3)$$

Including the standard deviation, the probabilistic model becomes

$$P(p(t)|p(t-1), \vartheta, \mu, \sigma) = N\left(p(t) \left| e^{-\vartheta\Delta t} p(t-1) + (1 - e^{-\vartheta\Delta t}) \mu, \sigma \sqrt{\frac{1 - e^{-2\vartheta\Delta t}}{2\vartheta}} \right.\right) \quad (4.4)$$

This equation describes the OU process in the same form as (4.1), where the conditional observation probability of an AR-HMM is defined. By carrying out the following mappings

$$\varphi_0 = e^{-\vartheta\Delta t}, \quad \mu_0 = (1 - e^{-\vartheta\Delta t}) \mu \quad \text{and} \quad \sigma_0 = \sigma \sqrt{\frac{1 - e^{-2\vartheta\Delta t}}{2\vartheta}} \quad (4.5)$$

we can consider a single state AR-HMM being equivalent as an OU process. However, this can be either a mean reverting or a mean averting process. Setting $\vartheta = 0$ in (3.2) yields a Brownian motion (BM), more generally, an AR-HMM state can represent Brownian motion with a possible linear drift as well.

Based on this approach, a multi state AR-HMM can be considered as a generalization of the OU and BM SDEs. Furthermore, having multiple states brings even more flexibility into modeling the price distributions, and also captures if the process is driven by different models over different intervals in time (often called as regimes [14]).

5 Portfolio optimization

In this section we discuss the optimal portfolio selection subject to an objective function, namely maximizing the average return [30]. This objective function can be one of the most important attributes of any portfolio selection, however, its solution does not land itself to analytical tractability. We identify AR-HMM model parameters for an observed process, and then, as a novel approach, we optimize the prediction based average return. Starting from the observed initial value of $p(0)$, based on (3.4), the objective function can be expressed as follows:

$$\Psi(\mathbf{x}) = \max_{0 \leq t} E(p(t)) - p(0), \quad (5.1)$$

where

$$E(p(t)) = (E(p(t-1))\boldsymbol{\varphi} + \boldsymbol{\mu})^T \boldsymbol{\gamma}_t, \quad \text{and} \quad \boldsymbol{\gamma}_t = \mathbf{A}^T \boldsymbol{\gamma}_{t-1} \quad (5.2)$$

is described recursively, and $E(p(0)) = p(0) = \mathbf{x}^T \mathbf{s}_0$. Section 6.2 describes methods for estimating $\Theta = \{\boldsymbol{\pi}, \mathbf{A}, \boldsymbol{\varphi}, \boldsymbol{\mu}, \boldsymbol{\sigma}\}$ and $\boldsymbol{\gamma}_0$. If the predicted process is divergent, $E(p(\infty)) = \pm \infty$, then the mean reverting assumption does not hold, hence we set $\Psi(\mathbf{x}) = -\infty$.

In this way, portfolio optimization can be reduced to a constrained optimization problem:

$$\mathbf{x}_{\text{opt}} = \arg \max_{\mathbf{x}} \Psi(\mathbf{x}), \quad \text{card}(\mathbf{x}) \leq l. \quad (5.3)$$

Including the bid-ask spread into this model can be done in a straightforward manner, the portfolio prices for fitting the AR-HMM should be calculated as $p(t) = \mathbf{x}_{(long)}^T \mathbf{s}_t^{bid} + \mathbf{x}_{(short)}^T \mathbf{s}_t^{ask}$, while the current price in the objective function as $p(0) = \mathbf{x}_{(long)}^T \mathbf{s}_0^{ask} + \mathbf{x}_{(short)}^T \mathbf{s}_0^{bid}$, where $x_i^{(long)} = \begin{cases} x_i, & \text{if } x_i > 0 \\ 0, & \text{if } x_i \leq 0 \end{cases}$ and similarly $x_i^{(short)} = \begin{cases} x_i, & \text{if } x_i < 0 \\ 0, & \text{if } x_i \geq 0 \end{cases}$. Negative values in the portfolio vector

denote the ability to short sell assets.

6 Computational approach

The AR-HMM model parameter identification and the evaluation of prediction by calculating the expected return (5.1) is embedded into a stochastic search method, namely into simulated annealing (see 6.1) forming an optimal portfolio selection. Then, the identified portfolio is converted to a trading signal for taking the appropriate trading actions.

Our computational framework is shown on the following structural block diagram (Fig. 1) and detailed as follows:

- Generate a portfolio vector: driven by stochastic optimization (detailed in 6.1) a sparse portfolio candidate ($\mathbf{x}^T = (x_1, \dots, x_N)$ under the constraint $\text{card}(\mathbf{x}) \leq l$) is generated until the objective function is maximized;
- AR-HMM identification: model fitting (by using the learning algorithm described in section 6.2) based on a sliding window of observations;
- Evaluate objective function: using the identified model,

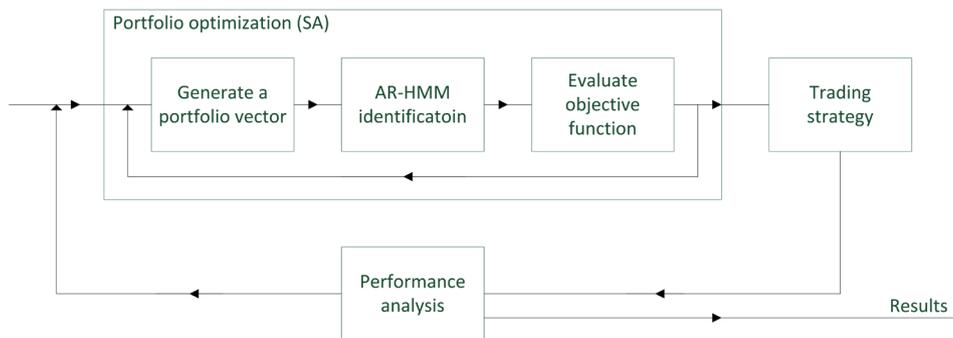


Fig. 1 Computational approach

a prediction for the future asset prices is calculated and evaluated by the objective function (see [chapter 5](#));

- Trading strategy forming the trading signal: based on the identified portfolio, the trading strategy ([section 6.3](#)) should decide on which trading action is to be launched;
- Performance analysis: a framework for trading and testing and evaluating various numerical indicators for the sake of comparison of the profitability of different methods ([chapter 7](#) contains further details).

6.1 Simulated annealing

This section focuses on the portfolio optimization. In the absence of proper analytical solutions for the constrained optimization problem posed in (5.3), we use simulated annealing for obtaining good quality heuristic solutions [10].

Simulated annealing [19] is a stochastic search method for finding the global optimum in a large search space. In this context the energy function $J(\mathbf{x})$ is the objective function maximizing the average return (5.1) for the selected portfolio:

$$J(\mathbf{x}) = \max_{0 \leq t} E(p(t)) - p(0). \quad (6.1)$$

With an appropriate neighbor function the cardinality constraint $card(\mathbf{x}) \leq l$ is automatically fulfilled at each step of the algorithm. The neighbor function on each iteration makes two steps: first, with a certain probability randomly changes the l dimensional subspace. In the second step, increase or decrease the volume of a likewise randomly chosen asset. The algorithm takes care not to exceed to available cash neither with the long positions, and if short selling is enabled, nor the short positions. Let \mathbf{x} be an arbitrary initialization vector, and then by calling a random number generation a vector \mathbf{x}' is generated subject to the abovementioned neighbor function. Accept the new vector if $J(\mathbf{x}') > J(\mathbf{x})$, or otherwise with $e^{-\frac{J(\mathbf{x})-J(\mathbf{x}')}{T}}$ probability. Continue the sampling while decreasing the T until zero. The last state vector is now the identified optimal sparse portfolio vector.

6.2 HMM parameter optimization

The model parameter estimation (learning) is a key aspect when we are using HMMs for prediction.

First, we need to sample the portfolio price with a sliding window of length T :

$$\mathbf{X} = \{p_{i-T+1}, \dots, p_i\} \quad (6.2)$$

During training, the likelihood (or in practice, due to the small order of magnitude of such probabilities, the log-likelihood) of the model is maximized based on the given observations [28]:

$$\Theta_{opt} = \arg \max_{\Theta} P(\mathbf{X} | \Theta) \quad (6.3)$$

The Baum-Welch expectation maximization (EM) algorithm [2] is a mechanism to iteratively update the model (4.2) starting from an arbitrary initial value and iterating until the likelihood of the model converges to a certain value. Since this is an iterative method, which can use the forward-backward algorithm, implemented in an efficient way by dynamic programming [29], this algorithm is relatively fast. On the other hand, it may get stuck in one of the local minima. Compared to the standard HMM, in the case of an AR-HMM, the forward-backward algorithm remains unchanged, while in the EM algorithm only a slight modification needed [5].

We can also obtain the posterior state probabilities from this procedure, the conditional probabilities of being at state j at time instance t given the observation sequence:

$$\gamma_t^{(j)} = P(q_t = j | \mathbf{X}) \quad (6.4)$$

6.2.1 Ornstein-Uhlenbeck parameter estimation with AR-HMM

As it was detailed in [section 4](#), the single state AR-HMM can be treated as a generalization of the OU process, hence it is suitable as a parameter estimation method. Estimating the long term mean (μ) of the process of portfolio valuations (3.1) is instrumental for mean reverting trading. For the sake of comparison with other estimation procedures [8,30], different methods were tested on artificially generated data.

Each observation was generated with the following parameters $\sigma = 3$, $T = 8$ and $\mu, \mu_0 \in [0;100]$, respectively. The comparison was done independently for mean reversion coefficients in $\mathcal{G} \in [0.5;1]$, and in each case the mean squared error was taken into account for 100 generated processes. The results are shown by [Fig. 2](#).

As one can see, the newly proposed way of OU parameter identification gives the most precise estimations, outperforming the traditional methods by an order of magnitude.

6.3 Trading strategy

In this section we describe the trading algorithm which is used to trade with the selected optimal portfolio based on the evaluation of the objective function [30]. In the proposed algorithm, trading is described as a walk in a binary state space in which either we already have a portfolio at hand or cash at hand, while the transitions are only affected by the evaluations of the potentially owned and the newly identified portfolios by the using the objective function (see [section 5](#)). The trading strategy is formalized by a state chart ([Fig. 3](#)).

As of (5.1), positive evaluation indicates a profitable portfolio, while negative evaluation indicates that the portfolio may produce a loss. Higher value in each objective function implies a better portfolio. Based on this, the agent buys a portfolio only if it has a positive evaluation. A new trading action is taken if a

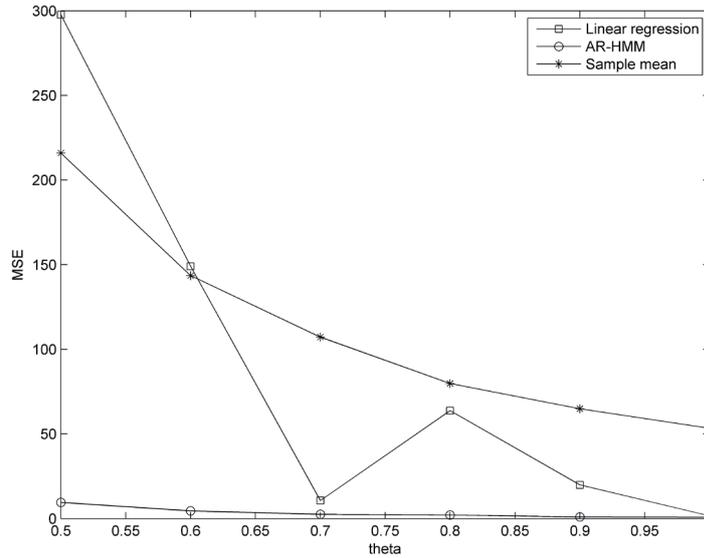


Fig. 2 Comparison of different mean estimation techniques

newly identified portfolio (x_{opt}) has higher and also positive evaluation than the present one (x). In this case one can sell the owned portfolio and buy the new one with higher expectations instead. This approach treats the present portfolio as a sunk cost, thus only the future expectations are taken into consideration. Hence, we do not have to give up the best available portfolio in favor of a presently unfavorable portfolio. In the case that neither the owned nor the currently identified portfolio has positive evaluation then the agent closes the positions (Fig. 3).

7 Performance analysis

An extensive back-testing framework was created to handle trading actions on various input data sets and provide numerical results for the sake of comparing different methods on different financial data series.

In this section, we show the numerical results obtained on the following data sets:

- Daily closing prices of 500 stocks from the S&P 500 (between July 2010 and July 2011) [33];
- FOREX bid and ask rates (EUR/USD, GBP/USD, AUD/USD, NZD/USD, USD/CHF, USD/CAD from the year of 2013 in daily resolution) [23].

For a detailed comparative analysis the following performance measures were calculated for each simulation: (i)

minimal value $G_{min} = \frac{1}{c_0} \min_{0 \leq t \leq T} c_t$; (ii) final value $G_{final} = \frac{c_T}{c_0}$;

(iii) maximal value $G_{max} = \frac{1}{c_0} \max_{0 \leq t \leq T} c_t$; (iv) average value

$G_{avg} = \frac{1}{c_0} \frac{1}{T} \sum_{t=1}^T c_t$, where c_t denotes the sum of owned cash

and the market value of the owned portfolio at time instance t , while c_0 denotes the initial cash (in each case the agent started with \$10,000).

Regarding the sparsity constraint, 3 assets were selected in each transaction.

7.1 Trading results on S&P500

Due to the large number of assets, $N=7$ hidden states were used. Unfortunately, there is no analytical method known to find the optimal value of this parameter [16]. Figure 4 compares the results achieved with different lengths of sliding window. As a benchmark, these results are compared to the performance achieved by the same objective function and trading strategy, but using linear regression for parameter estimation instead of AR-HMM (denoted by LR, further details in [30,32]).

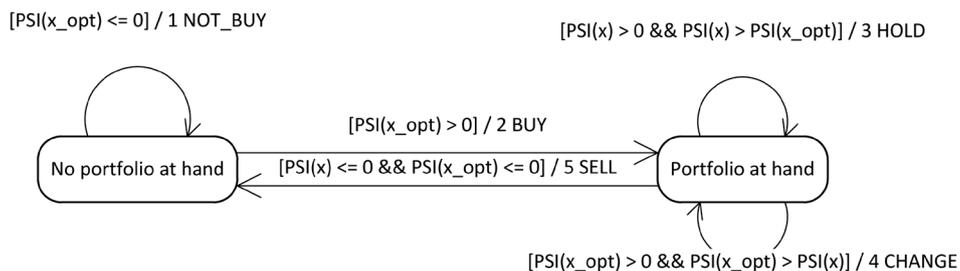


Fig. 3 Trading strategy

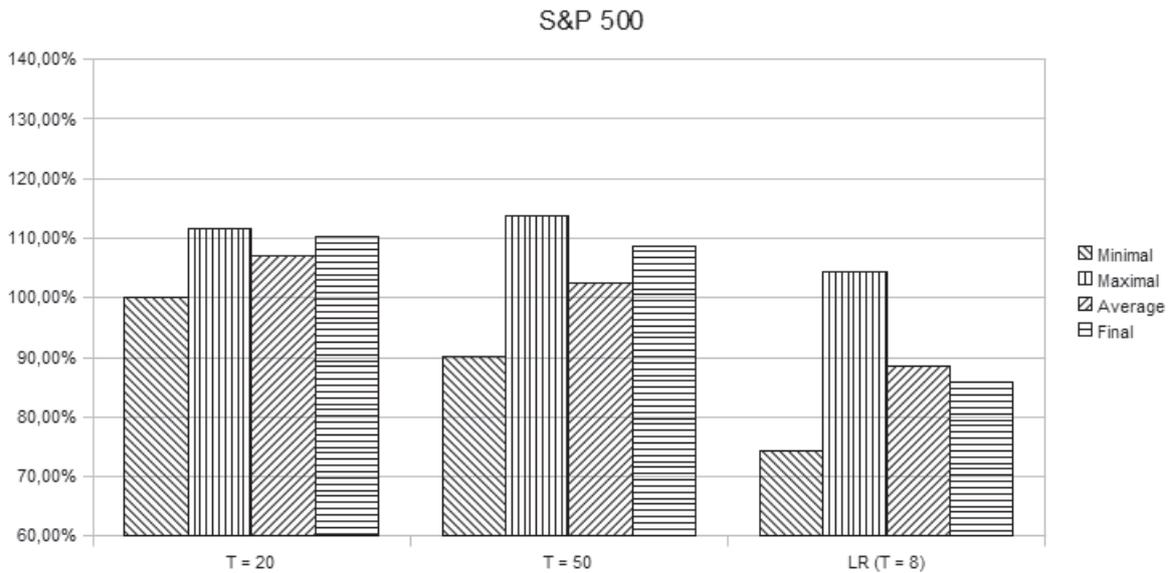


Fig. 4 Trading results on S&P 500 data

As one can see, the novel AR-HMM method was profitable and outperforms the linear regression.

7.2 Trading results on FOREX

For back-testing on FOREX time series, we used the *MetaTrader® 5* platform providing us a real environment with secondary effects including, like the bid-ask spread. The leverage was set to 10.

During the simulations $T=50$ days were used for model identification, while the effect of different number of hidden states in the trading results is shown in Fig. 5. Possibly because of the smaller number of available assets, less hidden states proved to be more favorable in this case. Also, the results are compared with the one achieved by linear regression parameter estimation (denoted by LR on the chart, see [30,32]).

In this period, the FOREX rates had a slightly decreasing tendency, an equally weighted portfolio would give us 97,76% of its initial price at the end of the year. The bar chart (Fig. 5) shows that the introduced methods with all settings beat this tendency, our methods achieved up to 110.12% yearly profit.

In the scenario when a longer ($T=100$) time window with $N=5$ hidden states were used, the trading was even more profitable with a 144.10% yearly profit including every secondary effects. However, considering longer sets of data, the training of the AR-HMM requires substantially longer computational time.

As it is shown in Fig. 6, using this strategy not only results in a favorable profit, without major drawdowns, the balance is almost monotonously increasing.

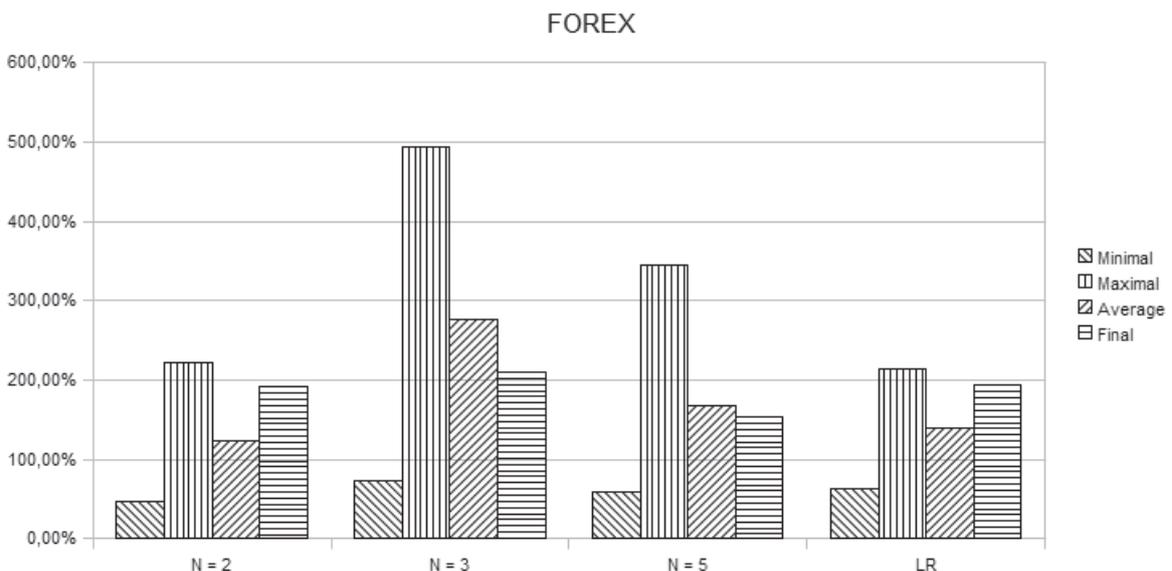


Fig. 5 Trading results on FOREX data

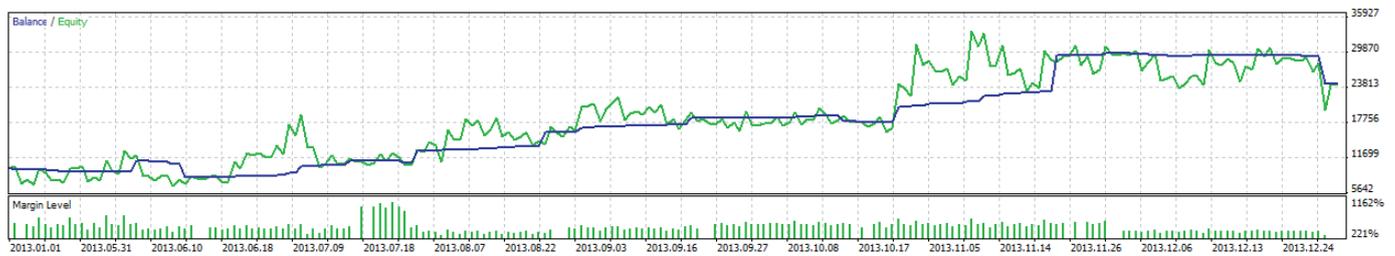


Fig. 6 Balance and equity through the one year period

8 Conclusions

In this paper we have proposed novel algorithms for OU parameter estimation and predicting future values of financial time series with AR-HMMs for portfolio optimization with cardinality constraints. The portfolio optimizations have been carried out by stochastic search with maximizing the mean return as an objective function.

The proposed trading algorithms has proven to be profitable on real financial time series taking into account the bid ask spread as well. The performance analysis demonstrated that the prediction based on the novel parameter identification could increase the trading efficiency and the profit compared to the traditional linear regression based method. However, there is a room for further improvement on the trading strategy in order to take other factors (e.g. to avoid early sells) into consideration, as well.

Acknowledgement

This work was partially supported by the European Union and the European Social Fund through project FuturICT.hu (grant no.: TAMOP-4.2.2.C-11/1/KONV-2012-0013).

References

- [1] Baum, L. E., Petrie, T. "Statistical inference for probabilistic functions of finite state Markov chains." *The annals of mathematical statistics*. 37 (6). pp. 1554-1563. 1966. DOI: [10.1214/aoms/1177699147](https://doi.org/10.1214/aoms/1177699147)
- [2] Baum, L. E., Petrie, T., Soules, G., Weiss, N. "A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains." *The Annals of Mathematical Statistics*. 41 (1). pp. 164-171. 1970. DOI: [10.1214/aoms/1177697196](https://doi.org/10.1214/aoms/1177697196)
- [3] Berchtold, A. "The double chain Markov model." *Communications in Statistics-Theory and Methods*. 28 (11). pp. 2569-2589. 1999.
- [4] D'Aspremont, A. "Identifying small mean-reverting portfolios." *Quantitative Finance*. 11 (3). pp. 351-364. 2011. DOI: [10.1080/14697688.2010.481634](https://doi.org/10.1080/14697688.2010.481634)
- [5] Dey, S., Krishnamurthy, V., Salmon-Legagneur, T. "Estimation of Markov-modulated time-series via EM algorithm." *IEEE Signal Processing Letters*. 1 (10). pp. 153-155. 1994. DOI: [10.1109/97.329841](https://doi.org/10.1109/97.329841)
- [6] Durbin, R. (ed.) "Biological sequence analysis: probabilistic models of proteins and nucleic acids." Cambridge University Press. 1998.
- [7] Fama, E., French, K. "Permanent and Temporary Components of Stock Prices." *Journal of Political Economy*. 96 (2). pp. 246-273. 1988. DOI: [10.1086/261535](https://doi.org/10.1086/261535)
- [8] Fogarasi, N., Levendovszky, J. "Improved parameter estimation and simple trading algorithm for sparse, mean reverting portfolios." *Annales Universitatis Scientiarum Budapestinensis de Ronaldo Eötvös Nominatae Sectio Computatorica*. pp. 121-144. 2012.
- [9] Fogarasi, N., Levendovszky, J. "A simplified approach to parameter estimation and selection of sparse, mean reverting portfolios." *Periodica Polytechnica Electrical Engineering and Computer Science*. 56 (1). pp. 21-28. 2013. DOI: [10.3311/PPee.7075](https://doi.org/10.3311/PPee.7075)
- [10] Fogarasi, N., Levendovszky, J. "Sparse, mean reverting portfolio selection using simulated annealing." *Algorithmic Finance*. 2 (3). pp. 197-211. 2013.
- [11] Fraser, A. M. "Hidden Markov models and dynamical systems." SIAM. 2008. DOI: [10.1137/1.9780898717747](https://doi.org/10.1137/1.9780898717747)
- [12] Gillespie, D. T. "Exact numerical simulation of the Ornstein-Uhlenbeck process and its integral." *Physical Review E*. 54 (2). pp. 2084-2091. 1996. DOI: [10.1103/physreve.54.2084](https://doi.org/10.1103/physreve.54.2084)
- [13] Hamilton, J. D. "Analysis of time series subject to changes in regime." *Journal of Econometrics*. 45 (1). pp. 39-70. 1990. DOI: [10.1016/0304-4076\(90\)90093-9](https://doi.org/10.1016/0304-4076(90)90093-9)
- [14] Hamilton, J. D., Susmel, R. "Autoregressive conditional heteroskedasticity and changes in regime." *Journal of Econometrics*. 64 (1). pp. 307-333. 1994. DOI: [10.1016/0304-4076\(94\)90067-1](https://doi.org/10.1016/0304-4076(94)90067-1)
- [15] Hassan, Md R., Nath, B. "Stock market forecasting using hidden Markov model: a new approach." In: *ISDA'05. Proceedings. 5th International Conference on*. pp. 192-196. 2005. DOI: [10.1109/isda.2005.85](https://doi.org/10.1109/isda.2005.85)
- [16] Hassan, Md R., Nath, B., Kirley, M. "A fusion model of HMM, ANN and GA for stock market forecasting." *Expert Systems with Applications*. 33 (1). pp. 171-180. 2007. DOI: [10.1016/j.eswa.2006.04.007](https://doi.org/10.1016/j.eswa.2006.04.007)
- [17] Ito, K. "Stochastic Integral." *Proceedings of the Imperial Academy*. 20. pp. 519-524. 1944.
- [18] Jurafsky, D., Martin, J. H. "Speech and Language Processing: An Introduction to Natural Language Processing. Computational Linguistics, and Speech Recognition." 2nd edition. London: Pearson Prentice Hall. 2006.
- [19] Kirkpatrick, S., Gelatt, C. D., Vecchi, M. P. "Optimization by Simulated Annealing." *Science*. 220. pp. 671-680. 1983.
- [20] Manzan, S. "Nonlinear Mean Reversion in Stock Prices." *Quantitative and Qualitative Analysis in Social Sciences*. 1 (3). pp. 1-20. 2007.
- [21] Mamon, R. S., Elliott, R. J. "Hidden Markov models in finance." Springer. 2007. DOI: [10.1007/978-1-4899-7442-6](https://doi.org/10.1007/978-1-4899-7442-6)
- [22] Markowitz, H. "Portfolio Selection." *The Journal of Finance*. 7 (1). pp. 77-91. 1952.
- [23] METATRADER FOREX time series [Online]. Available from: <http://www.metaquotes.net/en/metatrader5> [Accessed: February 2014].
- [24] Murphy, K. P. "Machine learning: a probabilistic perspective." MIT Press. 2012.
- [25] Natarjan, B. K. "Sparse approximate solutions to linear systems." *SIAM Journal on Computing*. 24 (2). pp. 227-234. 1995.

- [26] Ornstein, L. S., Uhlenbeck, G. E. "On the Theory of the Brownian Motion." *Physical Review*. 36 (5). p. 823. 1930.
- [27] Poterba, J. M., Summers, L. H. "Mean reversion in stock prices: Evidence and implications." *Journal of Financial Economics*. 22 (1). pp. 27-59. 1988. DOI: [10.3386/w2343](https://doi.org/10.3386/w2343)
- [28] Rabiner, L., Juang, B.-H. "An introduction to hidden Markov models." *ASSP Magazine*. 3 (1). pp. 4-16. 1986.
- [29] Rabiner, L. "A tutorial on hidden Markov models and selected applications in speech recognition." *Proceedings of the IEEE*. 77 (2). pp. 257-286. 1989.
- [30] Sipos, I. R., Levendovszky, J. "Optimizing sparse mean reverting portfolios." *Algorithmic Finance*. 2 (2). pp. 127-139. 2013.
- [31] Sipos, I. R., Levendovszky, J. "Trading with Hidden Markov Models." In: *Proceedings, 15th Applied Stochastic Models and Data Analysis (ASMDA2013)*. Barcelona. 2013.
- [32] Van Gelder, P. "Statistical modeling of financial markets." In: *TU Delft, UFS Workshop, Bloemfontein*. March 4-5 2009. 2009.
- [33] YAHOO. *S&P 500 asset series* [Online]. Available from: <http://finance.yahoo.com> [Accessed: July 2011].