Automating Transition Functions: A Way To Improve Trading Profits with Recurrent Reinforcement Learning

Jin Zhang

Automating transition functions: a way to improve trading profits with recurrent reinforcement learning

Jin Zhang

Faculty of Economics and Business Administration, University of Basel, Peter Merian-Weg 6, Switzerland
jin.zhang@unibas.ch

Abstract. This paper presents an application of the logistic smooth transition function and recurrent reinforcement learning for designing financial trading systems. We propose a trading system which is an upgraded version of the regime-switching recurrent reinforcement learning (RS-RRL) trading system referred to in the literature. In our proposed system (RS-RRL 2.0), we use an automated transition function to model the regime switches in equity returns. Unlike the original RS-RRL trading system, the dynamic of the transition function in our trading system is driven by utility maximization, which is in line with the trading purpose. Volume, relative strength index, price-to-earnings ratio, moving average prices from technical analysis, and the conditional volatility from a GARCH model are considered as possible options for the transition variable in RS-RRL type trading systems. The significance of Sharpe ratios, the choice of transition variables, and the stability of the trading system are examined by using the daily data of 20 Swiss SPI stocks for the period April 2009 to September 2013. The results from our experiment show that our proposed trading system outperforms the original RS-RRL and RRL trading systems suggested in the literature in terms of better Sharpe ratios recorded in three consecutive out-of-sample periods.

Keywords: Daily equity trading, Recurrent reinforcement learning, Transition variable selection, Automated transition functions.

1 Introduction

Recent developments in algorithmic trading have shown that people never stop their effort of searching and developing trading strategies. In the computer science literature, artificial intelligence (AI) techniques have been increasingly applied in the field of technical analysis for equity trading, e.g. the application of a genetic algorithm (GA) in technical trading rules optimization [1], the adaptive reinforcement learning system which uses 14 commonly-used technical indicators as a part of the system inputs [5], and a Logitboost method to select combinations of technical trading rules for stocks and futures trading [4].

Although AI has been used widely in research on financial trading platform design, these studies have concentrated on the optimization of trading signals

in technical analysis, in isolation from financial fundamentals and other available tools. Various approaches based on technical analysis [13] and fundamental analysis [8] have been developed to forecast future trends in stock prices. Additionally, tools from financial engineering have been used to facilitate price movement analysis. In real world, financial news, quotes, company earnings, pre-market and after-hours data, market research, company analysis are commonly used for price movement analysis. For example, researchers suggest to use trading systems which use fundamental indicators to select equity and technical indicators to produce trading signals [14].

Recurrent reinforcement learning (RRL), an online learning technique which finds approximate solutions to stochastic dynamic programming problems, was used by researchers to tune financial trading systems for the purpose of utility maximization [12]. In the literature, most discussions of RRL trading have been in the context of high-frequency FX and equity indices trading (see [6, 5, 2, 7]). Due to the effectiveness of RRL in training financial trading systems, the technique has been adopted by researchers to build more sophisticated financial trading systems. For example, the multi-layer trading system integrated a risk management module with a RRL trading system for utility maximization [5].

Over the last few decades, models such as regime switching (RS) have become popular in econometrics. For example, the threshold autoregressive (TAR) model [15] and the smooth transition autoregressive (STAR) model allowing for smooth transition [3] draw people's attention. Although interest in RS models has grown, most papers on RS have focused on model development, with only a few applications of RS models concerned with artificial intelligence being found in the financial trading field.

In the literature, it has been found that transitions between economic regimes are often signaled by price patterns. Researchers proposed a regime-switching RRL trading system to cope with the price changes which is non-linear due to regime switches [9, 10]. The basic RS-RRL trading system consists of two independent RRL systems which produce two trading signals for two scenarios, i.e. for low and high volatility regimes. The trading signal is made on the basis of the weighted total value of the trading signals from the two RRL systems. The weight is an output from a logistic smooth transition autoregressive (LSTAR) function, in which the conditional variance from a GARCH model is used as the transition variable.

In this paper, we aim to add to the existing literature on RS-RRL. We do so by studying the transition variable selection problem in the RS-RRL trading system. Five indicator variables (i.e. volume, relative strength index, price to earnings ratio, conditional volatility and moving average prices) are considered as possible transition variables in the RS-RRL trading system. We also introduce an upgraded version of the original RS-RRL trading system for the purpose of enhancing trading profits. Our upgraded version, the RS-RRL 2.0 trading system uses an automated transition function which is different from the constant transition function in the original RS-RRL trading system.

The paper is organized as follows. Section 2 provides a brief overview of RRL-type trading systems and introduces the our proposed trading system: RS-RRL 2.0. Section 3 provides the results of the experiment. Section 4 concludes the paper.

2 RRL-type trading systems

2.1 Recurrent reinforcement learning

Recurrent reinforcement learning (RRL) was a technique to tune financial trading systems for the purpose of utility maximization [12]. The RRL technique is a stochastic gradient ascent algorithm which continuously optimizes a utility measure by using new market information. Although, in most discussions on RRL, the market information usually comprises a series of lagged price returns. The basic RRL trading system in [12] is designed to trade a single-asset with a two-position action (long/short), which is produced by using linear combinations of returns and a tanh function (see Figure 1). $\boldsymbol{w}$ denotes a series parameters of the input signals. F_{t-1} refers to the current holding position; and F_t denotes the asset's holding position on the following day.

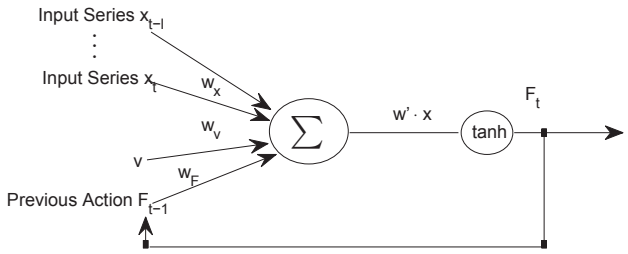

Fig. 1. Recurrent reinforcement learning

The goal of the RRL trading system is to maximize a wealth measure U_t by adjusting the parameter set $\boldsymbol{w}_t$ in a continuous manner:

$$\max U_t(R_t; \boldsymbol{w}_t), \tag{1}$$

$$R_t = \nu \cdot (\operatorname{sgn}(F_{t-1}) \cdot r_t - \delta \cdot |\operatorname{sgn}(F_t) - \operatorname{sgn}(F_{t-1})|), \tag{2}$$

and

$$F_t = \tanh(\boldsymbol{w}_t \times \mathbf{I}_t), \tag{3}$$

where ν is the number of shares and δ is the transactions cost rate. $\mathbf{I}_t$ denotes a series of inputs, including F_{t-1}, a constant v with a value of 1, and a set of lagged returns $r_t, r_{t-1}, r_{t-2}, \ldots, r_{t-l+1}$, $t = 1, \ldots, T$, and l is an integer number representing the length of the lags.

The signal parameters $\boldsymbol{w}_t$ are updated by RRL based on stochastic gradients. The gradients of U_t with respect to the signal parameter set $\boldsymbol{w}_t$ can be written as:

$$\frac{dU_t(\boldsymbol{w}_t)}{d\boldsymbol{w}_t} = \frac{dU_t}{dR_t}\left\{\frac{dR_t}{dF_t}\frac{dF_t}{d\boldsymbol{w}_t} + \frac{dR_t}{dF_{t-1}}\frac{dF_{t-1}}{d\boldsymbol{w}_{t-1}}\right\}, \tag{4}$$

$$\frac{dF_t}{d\boldsymbol{w}_t} \approx \frac{\partial F_t}{\partial \boldsymbol{w}_t} + \frac{\partial F_t}{\partial F_{t-1}}\frac{dF_{t-1}}{d\boldsymbol{w}_{t-1}}, \tag{5}$$

$$\frac{dR_t}{dF_{t-1}} = \nu \cdot (r_t + \delta \cdot \mathrm{sgn}(F_t - F_{t-1})), \tag{6}$$

$$\frac{dR_t}{dF_t} = -\nu \cdot \delta \cdot \mathrm{sgn}(F_t - F_{t-1}). \tag{7}$$

In the literature, the utility measure used in RRL trading systems is the differential Sharpe ratio (DSR), which is the first-order term after taking the Taylor series expansion of a performance measure, namely the exponential moving average Sharpe ratio (EMSR) at $\eta \to 0$:

$$EMSR_t = \frac{A_t}{K_\eta \cdot (B_t - A_t^2)^{1/2}}, \tag{8}$$

where $A_t = A_{t-1} + \eta \cdot (R_t - A_{t-1})$, $B_t = B_{t-1} + \eta \cdot (R_t^2 - B_{t-1})$, and $K_\eta = (\frac{1-\eta/2}{1-\eta})^{1/2}$. In other words, U_t can be written as:

$$DSR_t = \frac{B_{t-1} \cdot (R_t - A_{t-1}) - \frac{1}{2} \cdot A_{t-1} \cdot (R_t^2 - B_{t-1})}{(B_{t-1} - A_{t-1}^2)^{\frac{3}{2}}}, \tag{9}$$

and the derivative of U_t with respect to R_t in Eq. (4) can be written as:

$$\frac{dU_t}{dR_t} = \frac{B_{t-1} - A_{t-1} \cdot R_t}{(B_{t-1} - A_{t-1}^2)^{3/2}}. \tag{10}$$

At time t, the signal parameter set $\boldsymbol{w}_t$ is updated by using $\boldsymbol{w}_t = \boldsymbol{w}_{t-1} + \rho \cdot \frac{dU_t(\boldsymbol{w}_t)}{d\boldsymbol{w}_t}$, where ρ is the learning step.

2.2 Regime-switching RRL

Although the basic RRL trading system uses linear combinations to generate trading signals, nonlinear behavior in financial time series has been documented in the literature. To model the nonlinearities in price changes caused by regime switches, researchers proposed a RS-RRL trading system [9], in which the trading signal F_t is defined as:

$$F_t = F_t^A \cdot (1 - G_t) + F_t^B \cdot G_t, \tag{11}$$

$$F_t^A = \tanh\left(\boldsymbol{w}_{t-1}^A \times \mathbf{I}_{t-1}\right), \tag{12}$$

$$F_t^B = \tanh\left(\boldsymbol{w}_{t-1}^B \times \mathbf{I}_{t-1}\right). \tag{13}$$

In other words, the final output F_t of the system is a weighted sum of F_t^A and F_t^B from two individual RRL systems. The weight is a value generated by using a logistic function:

$$G_t(s_t;\gamma,c) = \frac{1}{1+\exp(-\gamma\cdot(s_t-c))}, \tag{14}$$

where s_t, γ and c refer to the transition variable, the transition rate and the threshold value respectively.

In the econometric literature, G_t refers to the transition function in a LSTAR model. The LSTAR models a univariate return series r_t as:

$$\begin{aligned} r_t = (\phi_{0,1}+\phi_{1,1}r_{t-1}+\ldots+\phi_{l,1}r_{t-l})\cdot(1-G(s_t;\gamma,c)) \\ +(\phi_{0,2}+\phi_{1,2}r_{t-1}+\ldots+\phi_{l,2}r_{t-l})\cdot G(s_t;\gamma,c)+\varepsilon_t. \end{aligned} \tag{15}$$

It is found that the conditional volatility variable from a GARCH model is a suitable transition variable for RS-RRL trading; and the transition rate and the threshold value in Eq. (14) are derived by using a quasi-maximum likelihood estimation (for further details, see [11]).

2.3 Regime-switching RRL 2.0

There are two main concerns with the original RS-RRL trading system. The first is that, if users do not manually update the parameters in G_t, the transition function will be constant. Constant transition models may not work well in the real world as the transition rate and the threshold value can be time-varying. The second concern is that, the likelihood based inference for the transition function does not serve the purpose of RRL in the trading context. In other words, the derivation of LSTAR parameters with quasi-maximum likelihood estimation (i.e. minimizing error terms), does not fit well with the purpose of trading, which is utility maximization.

We therefore propose a new RS-RRL trading system which we call RS-RRL 2.0 (see Figure 2), in order to improve the trading performance of the RS-RRL trading system. In our proposed trading system, in addition to the two RRL trading systems, we use a control unit which consists of a summation function and a sigmoid function to mimic the transition function G_t in the LSTAR model. Inputs of the control unit include the transition variable S_t, the threshold value c, and the current holding position F_{t-1}. To maximize the utility function, RRL updates the signal parameters in the trading units and the control unit. The design of RS-RRL 2.0 allows the transition function G_t and the utility measure to interact in an automated fashion in real time.

In the proposed RS-RRL 2.0 trading system, F_t is the weighted sum of the outputs F_t^A and F_t^B from the two individual RRL trading systems, which are given by Eq. (11), Eq. (12) and Eq. (13). G_t is a function of the transition variable, the threshold value, and the signal parameters $\boldsymbol{w}^G$:

$$G_t = \text{sigmoid}(\boldsymbol{w}_{t-1}^G\times\mathbf{I}_{t-1}) = \frac{1}{1+\exp(-\boldsymbol{w}_{t-1}^G\times\mathbf{I}_{t-1})}, \tag{16}$$

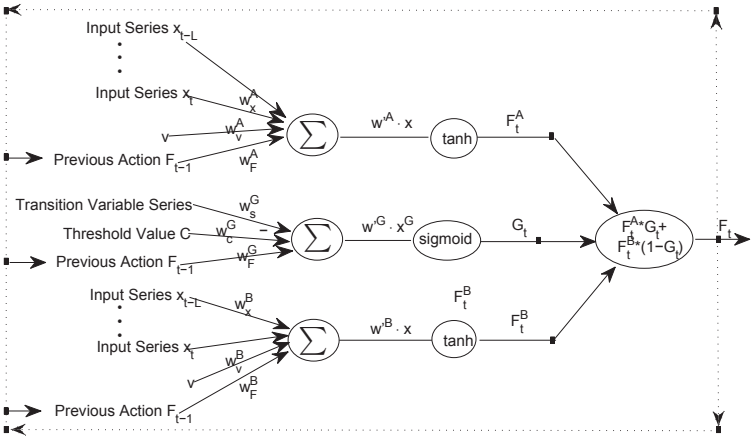

Fig. 2. The regime-switching recurrent reinforcement learning 2.0

which is similar to the logistic transition function in LSTAR models. We consider only one transition function in this paper; other transition functions from the econometric literature, such as exponential STAR (ESTAR) functions, also can be modified and used in the system.

Gradients of U_t with respect to the signal parameter sets $\boldsymbol{w}_t^A$, $\boldsymbol{w}_t^B$ and $\boldsymbol{w}_t^G$ can be written as:

$$\frac{dU_t(\boldsymbol{w}_t^A)}{d\boldsymbol{w}_t^A} = \frac{dU_t}{dR_t}\left\{\frac{dR_t}{dF_t}\frac{dF_t}{d\boldsymbol{w}_t^A} + \frac{dR_t}{dF_{t-1}}\frac{dF_{t-1}}{d\boldsymbol{w}_{t-1}^A}\right\}, \tag{17}$$

$$\frac{dU_t(\boldsymbol{w}_t^B)}{d\boldsymbol{w}_t^B} = \frac{dU_t}{dR_t}\left\{\frac{dR_t}{dF_t}\frac{dF_t}{d\boldsymbol{w}_t^B} + \frac{dR_t}{dF_{t-1}}\frac{dF_{t-1}}{d\boldsymbol{w}_{t-1}^B}\right\}, \tag{18}$$

and

$$\frac{dU_t(\boldsymbol{w}_t^G)}{d\boldsymbol{w}_t^G} = \frac{dU_t}{dR_t}\left\{\frac{dR_t}{dF_t}\frac{dF_t}{d\boldsymbol{w}_t^G} + \frac{dR_t}{dF_{t-1}}\frac{dF_{t-1}}{d\boldsymbol{w}_{t-1}^G}\right\}. \tag{19}$$

We use the reward measure and the utility function given in Eq. (2) and Eq. (9), so that the derivatives $\frac{dU_t}{dR_t}$, $\frac{dR_t}{dF_t}$ and $\frac{dR_t}{dF_{t-1}}$ are the same as Eq. (10), Eq. (6) and Eq. (7). The other components in Eq. (17), Eq. (18) and Eq. (19) can be found in the following:

$$\frac{dF_t}{d\boldsymbol{w}_t^A} \approx \frac{\partial F_t^A}{\partial \boldsymbol{w}_t^A} + \frac{\partial F_t}{\partial F_{t-1}}\frac{dF_{t-1}}{d\boldsymbol{w}_{t-1}^A}, \tag{20}$$

$$\frac{dF_t}{d\boldsymbol{w}_t^B} \approx \frac{\partial F_t^B}{\partial \boldsymbol{w}_t^B} + \frac{\partial F_t}{\partial F_{t-1}}\frac{dF_{t-1}}{d\boldsymbol{w}_{t-1}^B}, \tag{21}$$

and

$$\frac{dF_t}{d\boldsymbol{w}_t^G} \approx \frac{\partial G_t}{\partial \boldsymbol{w}_t^G} + \frac{\partial F_t}{\partial F_{t-1}}\frac{dF_{t-1}}{d\boldsymbol{w}_{t-1}^G}, \tag{22}$$

where

$$\frac{\partial F_t}{\partial F_{t-1}} = (G_t \frac{\partial F_t^A}{\partial F_{t-1}} + F_t^A \frac{\partial G_t}{\partial F_{t-1}}) - (G_t \frac{\partial F_t^B}{\partial F_{t-1}} + F_t^B \frac{\partial G_t}{\partial F_{t-1}}) + \frac{\partial F_t^B}{\partial F_{t-1}},$$

and all the other required components can be derived using the chain rule of derivatives.

3 Results of the experiment

3.1 The context of the experiment

The companies selected for this experiment are Swiss SPI stocks (1 April 2009 – 26 September 2013, each series containing 1126 observations). We consider volume, relative strength index (RSI), price-to-earnings (P/E) ratio and moving average price changes from technical analysis, and conditional volatility from a GARCH model as options for selecting the transition variable. Of over 100 SPI stocks, 20 have full historical information (downloaded from Bloomberg) of the above variables. The 1126 observations of each stock are partitioned into an initial training set consisting of the first 750 samples. We consider three consecutive out-of-sample periods, with each out-of-sample period consisting of 125 observations.

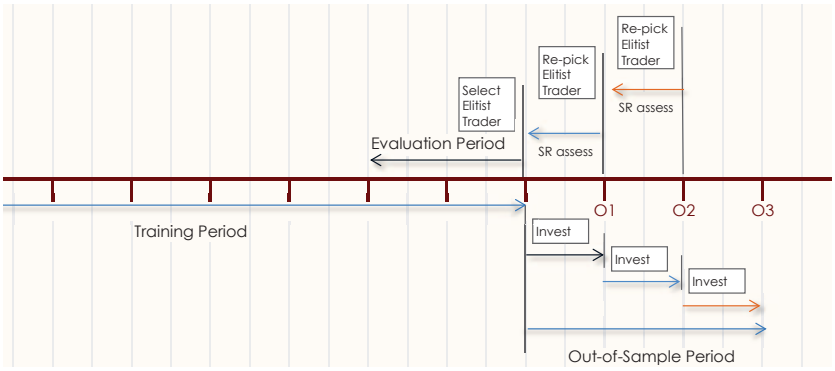

Fig. 3. Training and trading

The RS-RRL 2.0 trading system is designed to trade a single asset, there are 20 trading systems to trade the 20 SPI stocks. Every trading system consists of a group of 500 simulated traders (sim-traders). In the real world, as different traders may have different levels of information asymmetry, we initialize the sim-traders by using different numbers (i.e. random numbers from a Gaussian distribution with a mean of 0 and a standard deviation of 0.05). The sim-traders are then trained with the first 750 samples. At the end of the training period, we

select the best-performing trader (referred to as the elitist trader hereafter) from the 500 sim-traders according to their Sharpe ratio rankings in an evaluation period (i.e. the last 125 trades before out-of-sample trading). The elitist trader is reselected at the beginning of each out-of-sample period.

Based on preliminary tests, the following parameters were found to be suitable settings for the daily equity trading problem: the number of shares traded $\nu = 1$; the learning rate $\rho = 0.15$; the adaption rate $\eta = 0.05$. We expect that new information will be reflected in stock prices in a maximum period of two weeks, therefore we use a value of $l = 10$. The transaction cost δ has a value of 3 bps. The parameters of the transition variable and the threshold value have an initial value of 1. We use an optimization approach, differential evolution to estimate the threshold value (see [11]), other parameters of the system are initialized using random numbers.

3.2 Out-of-sample trading performance

Profitability and stability are two particularly important factors in a financial trading system. In this study, we use the Sharpe ratio to measure the profitability and we calculate the Sharpe ratio using daily returns. It should be noted that the trading performance of RRL-type trading systems relates directly to the initialization of signal parameters. Therefore, stability refers to the consistency of the Sharpe ratios recorded from independent restarts of the trading system.

We restart the RS-RRL 2.0 trading system 100 times, saving the daily Sharpe ratio of the elitist trader in each out-of-sample period from each trial. We collect a number of 100 Sharpe ratios for each stock. The mean value and the standard deviation (SD) of the 100 Sharpe ratios are considered as the measures for profitability and stability respectively. We check the daily Sharpe ratio statistics which are produced by using the basic RRL, the RS-RRL 1.0, and the RS-RRL 2.0 trading systems for each stock. We collect these Sharpe ratio means and SDs based on 100 restarts of the RRL and the RS-RRL 1.0 trading systems from the three out-of-sample periods. To make a fair comparison, the system parameters used in the RRL and RS-RRL 1.0 trading systems are the same as that used in RS-RRL 2.0. Both of the RRL and RS-RRL trading systems consist of a group of 500 sim-traders. We select an elitist trader from the 500 sim-traders for out-of-sample trading on the basis of their Sharpe ratio rankings from evaluation.

The means and standard deviations which are computed based the 100 restarts from the three out-of-sample periods are reported in Table 1 and Table 2. As these tables show that, the RS-RRL 2.0 trading system tends to perform at least as well as the RRL and the RS-RRL 1.0 trading system, and often better with respect to higher means and lower standard deviations. We perform the one-way analysis of variance (ANOVA) to compare the means of the Sharpe ratios produced by the RS-RRL 1.0 and the RS-RRL 2.0 trading systems. In Table 1, the Star symbols indicate the cases where the two Sharpe ratio means are significantly different from each other at a significance level of 5%, given a same transition variable. The results of the experiment show that most of the SR means from the RS-RRL 2.0 trading system are statistically greater than those

from the RS-RRL 1.0 trading system at a significance level of 5%. Although most Sharpe ratio means are not statistically greater than zero, we can still produce a positive return after netting the profits and losses from trading the 20 stocks by using the RS-RRL 2.0, which confirms the same findings in [11].

4 Conclusions

This paper presents an upgraded version of the RS-RRL trading system referred to in the literature. In the proposed RS-RRL 2.0 trading system, we use an automated transition function to model regime switches in equity returns. Unlike the original RS-RRL trading system, the dynamic of the transition function in our system is driven by utility maximization. Although the proposed system looks similar to an artificial neural network (ANN), we use a performance measure, i.e. the differential Sharpe ratio, which is different from the error function in traditional ANNs to implement the online learning for equity trading. We are able to reveal the time-varying feature of the threshold value by studying the parameter w_C^G (as shown in Figure 2), which provides a deeper insight into the impact of regime switches on stock returns comparing to the constant transition function suggested in the literature.

The results of our experiment show that regime-switching trading system brings value-added to trading performance. The RS-RRL 1.0 trading system referred to in the literature, produces higher Sharpe ratio means than the RRL trading system without impairing the stability of trading results. We also show that most the Sharpe ratios from the proposed RS-RRL 2.0 trading system are statistically greater than that from the RS-RRL 1.0 trading system at a significance level of 5%, given a same transition variable. With respect to transition variable selection, it seems that the price-to-earnings ratio is apt to produce higher Sharpe ratio means comparing the other four indicators; however, it is difficult to tell a 'one size fits all' indicator which improves the Sharpe ratio in general based on such a small sample of stocks.

Acknowledgements

This work has been partially supported by the Swiss National Science Foundation project SNSF-138095.

Table 1. Sharpe ratio means

		Volume		Con.Vol		RSI 9		PE		MV 30	
	RRL	RSRRL1	RSRRL2	RSRRL1	RSRRL2	RSRRL1	RSRRL2	RSRRL1	RSRRL2	RSRRL1	RSRRL2
1	0.033	0.004	0.098*	0.026	0.096*	0.028	0.094*	0.026	0.095*	0.040	0.101*
2	0.020	0.027	0.031	0.005	0.029*	0.016	0.029	0.024	0.025*	0.046	0.025
3	0.056	0.076	0.046*	0.026	0.025	0.060	0.055	0.033	0.027	0.013	0.036*
4	0.040	0.041	0.033	0.038	0.039	0.041	0.033	0.043	0.040	0.006	0.023*
5	0.008	-0.007	0.070*	0.007	0.070*	0.006	0.059*	0.018	0.072*	0.038	0.076*
6	0.057	0.023	0.113*	0.060	0.112*	0.061	0.107*	0.097	0.112*	0.009	0.110*
7	0.033	0.016	0.020	-0.003	0.042*	0.039	0.055*	0.009	0.021	0.009	0.014
8	0.014	0.020	0.079*	0.021	0.076*	0.011	0.059*	0.023	0.080*	0.044	0.077*
9	0.027	0.052	0.028*	0.035	0.031	0.024	0.023	0.067	0.041*	-0.006	0.019*
10	0.032	0.029	0.047*	0.062	0.074	0.025	0.019	0.016	0.034*	-0.002	0.054*
11	0.026	0.028	0.024	0.007	0.037*	0.017	0.031*	0.003	0.042*	0.023	0.036*
12	0.050	0.069	0.078	0.075	0.071	0.051	0.058	0.027	0.073*	0.048	0.080*
13	0.034	0.030	0.023	0.013	0.014	0.030	0.025	0.014	0.024*	0.029	0.015*
14	0.064	0.072	0.082	0.060	0.066	0.077	0.094*	0.056	0.073*	0.069	0.087*
15	0.032	0.043	0.068*	0.033	0.072*	0.036	0.061*	0.063	0.065	0.031	0.051*
16	0.020	0.014	-0.015*	-0.026	-0.022	0.023	0.001*	-0.003	0.045*	-0.011	0.008*
17	0.035	0.026	0.104*	0.029	0.087*	0.045	0.053	0.062	0.111*	0.004	0.106*
18	0.019	0.019	0.006*	0.001	0.011	0.025	0.002*	0.028	-0.009*	0.024	0.004*
19	0.035	0.042	0.067*	0.049	0.089*	0.043	0.087*	0.044	0.084*	0.036	0.081*
20	0.066	0.005	0.049*	0.048	0.055	0.062	0.090*	0.082	0.048*	0.040	0.055*

Table 2. Sharpe ratio standard deviations

		Volume		Con.Vol		RSI 9		PE		MV 30	
	RRL	RSRRL1	RSRRL2	RSRRL1	RSRRL2	RSRRL1	RSRRL2	RSRRL1	RSRRL2	RSRRL1	RSRRL2
1	0.081	0.077	0.044	0.077	0.043	0.078	0.052	0.089	0.037	0.075	0.039
2	0.070	0.072	0.060	0.069	0.053	0.067	0.059	0.065	0.047	0.081	0.049
3	0.072	0.075	0.051	0.075	0.050	0.068	0.065	0.063	0.050	0.083	0.052
4	0.071	0.070	0.042	0.069	0.043	0.069	0.051	0.060	0.043	0.059	0.055
5	0.078	0.077	0.052	0.069	0.049	0.077	0.078	0.065	0.054	0.083	0.054
6	0.073	0.072	0.037	0.077	0.039	0.075	0.050	0.064	0.036	0.095	0.038
7	0.091	0.085	0.068	0.083	0.077	0.095	0.087	0.092	0.075	0.077	0.062
8	0.099	0.092	0.105	0.094	0.112	0.099	0.107	0.086	0.111	0.104	0.112
9	0.079	0.071	0.057	0.075	0.063	0.076	0.070	0.070	0.061	0.077	0.069
10	0.081	0.081	0.126	0.099	0.154	0.085	0.110	0.076	0.131	0.094	0.145
11	0.079	0.091	0.056	0.078	0.065	0.080	0.055	0.065	0.073	0.077	0.068
12	0.068	0.081	0.060	0.072	0.063	0.074	0.054	0.080	0.076	0.075	0.080
13	0.071	0.076	0.047	0.072	0.055	0.073	0.044	0.056	0.045	0.076	0.057
14	0.072	0.073	0.064	0.078	0.067	0.076	0.058	0.069	0.069	0.076	0.079
15	0.080	0.073	0.050	0.070	0.051	0.074	0.055	0.065	0.045	0.084	0.056
16	0.082	0.095	0.085	0.077	0.080	0.077	0.082	0.086	0.080	0.079	0.096
17	0.075	0.087	0.047	0.076	0.059	0.078	0.066	0.085	0.048	0.077	0.055
18	0.078	0.073	0.083	0.086	0.072	0.072	0.090	0.071	0.083	0.090	0.068
19	0.074	0.068	0.077	0.074	0.078	0.076	0.093	0.083	0.078	0.067	0.077
20	0.095	0.085	0.092	0.082	0.073	0.094	0.097	0.093	0.079	0.076	0.087

References

1. Franklin Allen and Risto Karjalainen. Using genetic algorithms to find technical trading rules. *Journal of Financial Economics*, 51:245–271, 1999.
2. Francesco Bertoluzzo and Marco Corazza. Making financial trading by recurrent reinforcement learning. In B. Apolloni, R.J. Howlett, and L. Jain, editors, *Lecture Notes in Computer Science*, volume 4693, pages 619–626. Springer, KES 2, 2007.
3. K.S. Chan and Howell Tong. On estimating thresholds in autoregressive models. *Journal of time series analysis*, 7:179–190, 1986.
4. Germán Creamer. Model calibration and automated trading agent for euro futures. *Quantitative Finance*, 12:531–545, 2012.
5. M. A. H. Dempster and V. Leemans. An automated FX trading system using adaptive reinforcement learning. *Expert Systems with Applications*, 30:543–552, 2006.
6. Carl Gold. FX trading via recurrent reinforcement learning. In *Proceedings of IEEE international conference on computational intelligence in financial engineering*, pages 363–370, Los Alamitos, 2003. IEEE Computer Society Press.
7. Denise Gorse. Application of stochastic recurrent reinforcement learning to index trading. In *European Symposium on Artificial Neural Networks*, 2011.
8. Benjamin Graham and Jason Zweig. *The Intelligent Investor: The Definitive Book on Value Investing*. HarperCollins, 2003.
9. Dietmar Maringer and Tikesh Ramtohul. Regime-switching recurrent reinforcement learning for investment decision making. *Computational Management Science*, 9:89–107, 2012.
10. Dietmar Maringer and Tikesh Ramtohul. Regime-switching recurrent reinforcement learning in automated trading. In Anthony Brabazon, Michael O'Neill, and Dietmar G. Maringer, editors, *Natural Computing in Computational Finance*, volume 4, pages 93–121. Springer, 2012.
11. Dietmar Maringer and Jin Zhang. Transition variable selection for regime switching recurrent reinforcement learning. In *Proceedings of the 2014 IEEE Conference on Computational Intelligence for Financial Engineering and Economics*, pages 407–412, 2014.
12. John Moody, Lizhong Wu, Yuansong Liao, and Matthew Saffell. Performance functions and reinforcement learning for trading systems and portfolios. *Journal of Forecasting*, 17:441–470, 1998.
13. John Murphy. *Technical analysis of the financial markets : a comprehensive guide to trading methods and applications*. New York Institute of Finance, New York, 1999.
14. António Silva, Rui Neves, and Nuno Horta. Portfolio optimization using fundamental indicators based on multi-objective ea. In *Proceedings of the 2014 IEEE Conference on Computational Intelligence for Financial Engineering and Economics*, number 158–165, 2014.
15. Howell Tong. On a threshold model. In C Chen, editor, *Pattern recognition and signal processing*, NATO ASI Series E: Applied Sc. (29). Sijthoff & Noordhoff, Netherlands, 1978.