Decision Support Systems 93 (2017) 51–61

Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

Evaluation of on-line trading systems: Markov-switching vs time-varying parameter models

Michele Fornaciari, Carlo Grillenzoni*

University IUAV of Venice, Venezia, Italy

ARTICLE INFO

Article history:
Received 4 March 2016
Received in revised form 28 July 2016
Accepted 10 September 2016
Available online 20 September 2016

Keywords:
Gain maximization
Maximum likelihood
Recursive least squares
Standard and Poor index
State probabilities
Sequential analysis

ABSTRACT

Automatic trading systems, to support the decisions of investors in financial markets, are increasingly used nowadays. Such systems process data on-line and provide signals of buy and sell in correspondence of pits and peaks of the market. Real-time detection of turning points in financial time series is a challenging issue and can only be performed with sequential methods. This paper considers non-linear and non-stationary dynamic models used in statistics and econometrics, and evaluates their performance. In particular, it compares Markov switching (MS) regression and time-varying parameter (TVP) methods; the latter extend moving-average (MA) techniques which are widely used by traders. The novel approach of this paper is to select the coefficients of the detection methods by optimizing the profit objective functions of the trading activity, using statistical estimates as initial values. The paper also develops a sequential approach, based on sliding windows, to cope with the time-variability of MS coefficients. An extensive application to the daily Standard & Poor 500 index (the world's leading indicator of stock values) in the period 1999–2015, provides evidence in favor of models with a few parameters. This seems a natural consequence of the complexity of the gain maximization problem, which usually admits multiple local solutions. Directions for further research are represented by multivariate detection methods and the development of recursive algorithms for gain optimization.

1. Introduction

In financial economics, *bull* and *bear* markets are common ways to describe periods where stock prices are persistently increasing or decreasing. In a bull period, the market shows confidence, high returns coupled with low variability, investors are positive on the economic perspectives. On the contrary, during a bear period the market is declining, usually with sharp drop and high volatility; investors show lack of confidence and traded volumes are stagnant. Following the decision rule to "*buy low* and *sell high*", traders should buy stocks as soon as a bull period starts, and values increase, whereas they should sell as soon as a bear market takes place, in order to avoid price drop. Reliable automatic methods are then necessary to forecast or to promptly detect turning points in financial time series [1,12].

Signaling methods used in finance, range from heuristic rules based on moving averages (MA [24]) up to complex *non-linear* filters, which have hidden Markov dynamics (HM [21,23]), threshold autoregression (TAR, [22]) and neural networks (NN [25]). In this paper we focus on Markov switching (MS) models developed by Hamilton [13,15] since they are used both in financial and real macroeconomics, e.g. [2,3,7,16,17]. In financial trading, MS models aim to capture the dynamics of stock values, by modeling time series as the result of two or more alternating regimes. The switching mechanism is governed by an unobserved state-variable that follows a first-order Markov chain, whose current value depends on its immediate past. This approach is capable to represent the dynamics and the distribution of stock returns, which have distinct patterns on different periods [2].

Many studies in finance deal with foreign exchange rates and aim to test the existence of non-linearity in the series and/or the efficacy of MS models to support trading decisions, see [6–8]. Useful empirical findings are that MS systems fit data better than linear models, but only slightly outperform linear and MA rules in terms of profitability [7, Table 3]. This issue has raised many interpretations as to whether the MS approach may represent agent expectations and behaviors; in fact, traders usually adopt MA rules to interact with the market volatility and the monetary policy [3,6]. However, the near equivalence of MS and MA in trading may arise from the fact that MS models are estimated with statistical criteria, rather than

* Corresponding author.
E-mail address: carlog@iuav.it (C. Grillenzoni).

http://dx.doi.org/10.1016/j.dss.2016.09.005

economic ones, e.g. [8]. This remark stimulates the development of technical solutions for improving the operational capabilities of the MS modeling.

Financial economists usually consider two states, e.g. [3,6,7,20], where the first represents the bull market and the second is the bear market. By monitoring MS state probabilities, they determine when a time series changes the regime, so as to indicate the investment decision. The parameters θ of the models are estimated with the maximum likelihood (ML) method, under the assumption of Gaussian innovations; with them one can generate the required state probabilities. Given a sample of size T, the *smoothed* (or backward) probabilities are usually considered to infer the dominant regime [3,4]; however, this solution fundamentally works *off-line*, on the data available at time $t < T$ (see [14, Chapter 22], [16]). In on-line conditions, where stock data flow continuously, one may only use *prediction* probabilities to identify the current state at time $t \geq T$.

When monitoring the state probabilities, one cannot immediately decide whether to buy or sell until a new trend is reliably assessed. In fact, one could observe *local* movements in stock prices even between *global* peaks and troughs of the series [17]. Since only persistent changes are of interest in trading, a threshold value $0 < \kappa < 1$ must be introduced to discriminate significant changes. This value could be selected on the basis of the standard errors of ML estimates; however, a general problem is that such errors satisfy a statistical criterion, not an economic one. In practice, ML estimates, and the implied prediction probabilities, may not be optimal in terms of profitability — just because the assumptions of Gaussian innovations (or other distributions) and the Markov chain dynamics of parameters may not be consistent with the data.

To avoid the constraints of MS models, one may consider linear models whose parameters change over time in a non-parametric way. The formal correspondence between non-linear and non-stationary systems has been discussed by Granger [9] and [22], who showed that parameter variability may arise from the approximation of non-linear processes (including the MS). Further, since adaptive models are only driven by a smoothing coefficients, they may be considered as members of the class of MA rules. It follows that, by sequentially estimating a linear model, one might detect bull and bear phases by using, as threshold κ, the off-line value of parameters. A common adaptive estimator is the recursive least square (RLS [11]) with discounted observations, which is tuned by a coefficient $0 < \lambda < 1$.

In this paper we develop a detection method for regime variations which correspond to global peaks and troughs, i.e. relevant *turning points* of financial time series. The approach consists of selecting MS parameters through a criterion that embeds the buy–sell return between turning points. Starting from classical ML estimates, one can re-optimize the parameters by maximizing a gain objective function with search algorithms. The new values of θ, κ yield, by definition, a grater gain, thus information useful for investments. This extends to MS the approach developed by [12] for adaptive regression models, to which the method is finally compared.

The content of the work is as follows: Section 2 presents the MS framework and develops its trading implementation. Section 3 describes the methodology for adaptive linear models. Section 4 performs extensive applications to the daily S&P index, and compares the results of the various solutions. Finally, we draw our conclusions in Section 5.

2. Markov switching parameters

Consider a financial time series $\{Y_t\}_{t=1}^T$ subject to trading, where t indicates minutes or days. Let $\{y_t\}$ be the detrended series and $\{e_t\}$ a Gaussian white noise; the models we consider in this paper are MS versions of schemes widely used in statistics and econometrics, such as mean level, auto-regression and linear trend, see [10,14]

$$\begin{aligned} & y_t = (Y_t - Y_{t-1}), \\ \text{mean level} \quad & y_t = \mu_{j_t} + \sigma_{j_t} e_t, \end{aligned} \tag{1}$$

$$\text{auto regression} \quad y_t = \phi_{j_t}\, y_{t-1} + \sigma_{j_t} e_t, \tag{2}$$

$$\text{unit-root model} \quad Y_t = \varphi_{j_t} Y_{t-1} + \sigma_{j_t} e_t, \tag{3}$$

$$\begin{aligned} \text{linear trend} \quad & Y_t = \alpha_{j_t} + \beta_{j_t} t + \sigma_{j_t} e_t, \\ & e_t \sim \mathrm{IN}(0,1), \\ & j_t = 1,2, \end{aligned} \tag{4}$$

where j_t is the parameter state. The representations (1)–(4) are motivated by the fact that stock values and exchange rates typically behave like random walks, with two possible states, which correspond to the bull and bear market respectively. For example, expanding the model (3) one has

$$\begin{aligned} t \in \text{Bull}\,(j=1): \quad & Y_t = \varphi_1 Y_{t-1} + \sigma_1 e_t, \quad \varphi_1 \geq \varphi_2, \\ t \in \text{Bear}\,(j=2): \quad & Y_t = \varphi_2 Y_{t-1} + \sigma_2 e_t, \quad \sigma_1 \leq \sigma_2, \end{aligned}$$

and similarly for the other models, where $\mu_1 \geq \mu_2$ and $\beta_1 \geq \beta_2$. Parameter inequalities are motivated by the fact that when the market grows one has $\varphi_1 > 1$ [11], and when the stock values plunge there is greater volatility. The switching between the two states is governed by the discrete process $\{j_t\}$ which is non-observable, but follows a stationary first-order Markov chain with transition probabilities $0 \leq \pi_{il} \leq 1$. For example, in correspondence of a market peak, the probability of passing from the bull to the bear state is given by $\mathrm{P}(j_t = 2 | j_{t-1} = 1) = \pi_{21}$, whereas the probability of staying in a bull market is given by $\pi_{11} = 1 - \pi_{21}$.

Models (1)–(4) can be written in a common vector regression form as

$$Y_t = \boldsymbol{\theta}_{j_t}' \mathbf{x}_t + \sigma_{j_t} e_t, \tag{5}$$

where, in the case of the Eq. (1), one has $\boldsymbol{\theta}_{j_t} = [1, \mu_{j_t}]'$ and $\mathbf{x}_t = [Y_{t-1}, 1]'$, and so forth for the other models. Given the two-state assumption, the whole set of parameters are given by $\boldsymbol{\delta} = [\boldsymbol{\theta}_1', \boldsymbol{\theta}_2', \sigma_1, \sigma_2, \pi_{11}, \pi_{22}]'$, and the log likelihood function of the model (5) can be written as

$$\ell_T(\boldsymbol{\delta}) = \sum_{t=2}^{T} \log \left[\sum_{i=1}^{2} f_\delta(Y_t | j_t = i)\, \mathrm{P}_\delta(j_t = i) \right], \tag{6}$$

where f_δ is the density function of Y_t. The above expression is motivated by the non-observable nature of the parameter state, and by the use of the rules of conditional and total probability applied to the joint density $f(Y_t, j_t)$.

Following [13,14,20], function (6) can be computed by means of an algorithm which also provides the estimate of the probability $P_{j,t}$ of the state j at time t. Under the assumption of Gaussian innovations, the algorithm can be written in *recursive* form as follows:

$$f_{j,t} = \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left[-\frac{\left(Y_t - \boldsymbol{\theta}_j' \mathbf{x}_t\right)^2}{2\sigma_j^2} \right], \quad j = 1,2, \tag{7a}$$