Learning Volatility of Discrete Time Series Using Prediction with Expert Advice

Vladimir V. V'yugin

Institute for Information Transmission Problems, Russian Academy of Sciences,
Bol'shoi Karetnyi per. 19, Moscow GSP-4, 127994, Russia
vyugin@iitp.ru

Abstract. In this paper the method of prediction with expert advice is applied for learning volatility of discrete time series. We construct arbitrage strategies (or experts) which suffer gain when "micro" and "macro" volatilities of a time series differ. For merging different expert strategies in a strategy of the learner, we use some modification of Kalai and Vempala algorithm of following the perturbed leader where weights depend on current gains of the experts. We consider the case when experts one-step gains can be unbounded. New notion of a volume of a game v_t is introduced. We show that our algorithm has optimal performance in the case when the one-step increments $\Delta v_t = v_t - v_{t-1}$ of the volume satisfy $\Delta v_t = o(v_t)$ as $t \to \infty$.

1 Introduction

In this paper we construct arbitrage strategies which suffer gain when "micro" and "macro" volatilities of a time series differ. For merging different expert strategies in a strategy of the learner, we use methods of the theory of prediction with expert advice.

Using Cheredito [2] results in mathematical finance, Vovk [14] considered two strategies which suffer gain when prices of a stock follow fractional Brownian motion: the first strategy shows a large return in the case when volatility of the time series is high, the second one shows a large return in the opposite case.

The main peculiarity of these arbitrage strategies is that their one-step gains and losses can be unrestrictedly large. Also, there is no appropriate type of a loss function for these strategies, and we are forced to consider general gains and losses. We construct a learning algorithm merging online these strategies into a one *Learner's* strategy protected from these unbounded fluctuations as much as possible.

Prediction with Expert Advice considered in this paper proceeds as follows. A "pool" of expert strategies (or simply, experts) $i = 1, \dots N$ is given. We are asked to perform sequential actions at times $t = 1, 2, \dots, T$. At each time step t, experts $i = 1, \dots N$ receive results of their actions in form of their gains or losses s^i_t. In what follows we refer to s^i_t as to gains.

At the beginning of the step t *Learner*, observing cumulated gains $s^i_{1:t-1} = s^i_1 + \dots + s^i_{t-1}$ of all experts $i = 1, \dots N$, assigns non-negative weights w^i_t (summing

O. Watanabe and T. Zeugmann (Eds.): SAGA 2009, LNCS 5792, pp. 16–30, 2009.

to 1) to each expert i and suffers a gain equal to the weighted sum of experts gains $\tilde{s}_t = \sum_{i=1}^{N} w_t^i s_t^i$.

The cumulative gain of the learner on first T steps is equal to $\tilde{s}_{1:T} = \sum_{t=1}^{T} \tilde{s}_t$.

This can be interpreted in probabilistic terms as follows. On each time step t, *Learner* choose to follow an expert i according to the internal distribution $P\{I_t = i\} = w_t^i$, $i = 1, \ldots N$; at the end of step t *Learner* receives the same gain s_t^i as the ith expert and suffers *Learner's* cumulative gain $s_{1:t} = s_{1:t-1} + s_t^i$.

Let $E(s_t)$ denote the Learner's expected one-step gain according to this randomization; it coincides with the weighted sum $\tilde{s}_t$ of the experts gains. The cumulative expected gain of our learning algorithm on first T steps is equal to $E(s_{1:T}) = \sum_{t=1}^{T} E(s_t)$. Evidently, $\tilde{s}_{1:T} = E(s_{1:T})$ for all T.

The goal of the learner's algorithm is to perform in terms of $\tilde{s}_{1:T} = E(s_{1:T})$ almost as well as the best expert in hindsight in the long run.

In the traditional framework of prediction with expert advice, it is supposed that one-step gains of experts are bounded, for example, $0 \le s_t^i \le 1$ for all i and t. We allow gains to be unbounded. We consider also the case when the notions of loss and gain functions are not used.

In this paper we use the method of following the perturbed leader. This method was discovered by Hannan [5]. Kalai and Vempala [7] rediscovered this method and published a simple proof of the main result of Hannan. They called the algorithm of this type FPL (Following the Perturbed Leader). Hutter and Poland [6] presented a further developments of the FPL algorithm for countable class of experts, arbitrary weights and adaptive learning rate. Also, FPL algorithm is usually considered for bounded one-step losses: $0 \le s_t^i \le 1$ for all i and t.

The similar results can be achieved by other aggregate strategies, like Weighted Majority (WM) algorithm of Littlestone and Warmuth [8] or algorithm "hedge" of Freund and Schapire [4]. The FPL algorithm has the same performance as the WM-type algorithms up to a factor $\sqrt{2}$. A major advantage of the FPL algorithm is that its analysis remains easy for an adaptive learning rate, in contrast to WM-derivatives (see remark in [6]).

Most papers on prediction with expert advice either consider bounded losses or assume the existence of a specific loss function. The setting allowing unbounded one-step losses (or gains) do not have wide coverage in literature; we can only refer reader to [1], [10], where polynomial bounds on one-step losses were considered.

In this paper, we present some modification of Kalai and Vempala [7] algorithm of following the perturbed leader (FPL) for the case of unrestrictedly large one-step expert gains $s_t^i \in (-\infty, +\infty)$ not bounded in advance. This algorithm uses adaptive weights depending on past cumulative gains of the experts.

We introduce new notions of *volume of a game* $v_t = \sum\limits_{j=1}^{t} \max_i |s_j^i|$ and *scaled fluctuation* of the game $\mathrm{fluc}(t) = \Delta v_t / v_t$, where $\Delta v_t = v_t - v_{t-1}$ for $t \ge 1$.

In Section 2 we present a game with two zero-sum experts which suffer gain or loss when micro and macro volatilities of stock prices differ. These gains and losses cannot be bounded in advance. The notion of the volume of a game is natural from this financial point of view.

In Section 3 we consider a more general problem - we consider a game with N experts, where these experts suffer unbounded gains $s_t^i \in (-\infty, +\infty)$. We present some probabilistic learning algorithm merging the experts decisions and prove that this algorithm is performed well under very broad assumptions.

In Section 4 we discuss a derandomized version this algorithm for the case of two zero-sum experts with arbitrary one-step gains from Section 2.

We show in Theorem 1 (Section 3) that if $\text{fluc}(t) \le \gamma(t)$ for all t, where $\gamma(t)$ is a non-increasing computable real function such that $0 < \gamma(t) \le 1$ for all t, then the algorithm of following the perturbed leader with adaptive weights constructed in Section 3 has the performance

$$E(s_{1:T}) \ge \max_{i=1,\dots N} s^i_{1:T} - 2\sqrt{6(1+\ln N)} \sum_{t=1}^{T} (\gamma(t))^{1/2} \Delta v_t.$$

If $\gamma(t) \to 0$ as $t \to \infty$ this algorithm is asymptotically consistent in a modified sense

$$\liminf_{T\to\infty} \frac{1}{v_T} E(s_{1:T} - \max_{i=1,\dots N} s^i_{1:T}) \ge 0, \tag{1}$$

where $s_{1:T}$ is the total gain of our algorithm on steps $\le T$.

Proposition 1 of Section 3 shows that if the condition $\Delta v_t = o(v_t)$ is violated the cumulative gain of any probabilistic prediction algorithm can be much less than the gain of some expert of the pool.

2 Arbitrage Strategies

We consider a time series as a series of some stock prices. Then each expert from the pool will represent a method of playing on the stock market.

Let K, M and T be positive integer numbers and let the time interval $[0, KT]$ be divided on a large number KM of subintervals. Let also $S = S(t)$ be a function representing a stock price at time t. Define a discrete time series of stock prices

$$S_0 = S(0), S_1 = S(T/(KM)), S_2 = S(2T/(KM)) \dots, S_{KM} = S(T). \tag{2}$$

In this paper, volatility is an informal notion. We say that the sum

$$\sum_{i=0}^{K-1} (S_{(i+1)T} - S_{iT})^2$$

represents the macro volatility and the sum

$$\sum_{i=0}^{KT-1} (\Delta S_i)^2,$$

where $\Delta S_i = S_{i+1} - S_i$, $i = 1, \dots KT$, represent the micro volatility of the time series (2). In this paper for simplicity we consider the case $K = 1$.

Informally speaking, the first strategy will show a large return if $(S_T - S_0)^2 \gg \sum\limits_{i=0}^{T-1} (\Delta S_i)^2$; the second one will show a large return when $(S_T - S_0)^2 \ll \sum\limits_{i=0}^{T-1} (\Delta S_i)^2$. There is an uncertainty domain for these strategies, i.e., the case when both $\gg$ and $\ll$ do not hold.[1]

We consider the game between an investor and the market. The investor can use the long and short selling. At beginning of time step t Investor purchases the number C_t of shares of the stock by S_{t-1} each. At the end of trading period the market discloses the price S_{t+1} of the stock, and the investor incur his current income or loss $s_t = C_t \Delta S_t$ at the period t. We have the following equality

$$(S_T - S_0)^2 = (\sum_{t=0}^{T-1} \Delta S_t)^2 =$$
$$\sum_{t=0}^{T-1} 2(S_t - S_0)\Delta S_t + \sum_{t=0}^{T-1} (\Delta S_t)^2. \quad (3)$$

The equality (3) leads to the two strategies for investor which are represented by two experts. At the beginning of step t Experts 1 and 2 hold the number of shares

$$C_t^1 = 2C(S_t - S_0), \quad (4)$$
$$C_t^2 = -C_t^1, \quad (5)$$

where C is an arbitrary positive constant.

These strategies at step t earn the incomes $s_t^1 = 2C(S_t - S_0)\Delta S_t$ and $s_t^2 = -s_t^1$. The strategy (4) earns in first T steps of the game the income

$$s_{1:T}^1 = \sum_{t=1}^{T} s_t^1 = 2C((S_T - S_0)^2 - \sum_{t=1}^{T-1} (\Delta S_t)^2).$$

The strategy (5) earns in first T steps the income $s_{1:T}^2 = -s_{1:T}^1$.

The number of shares C_t^1 in the strategy (4) or number of shares $C_t^2 = -C_t^1$ in the strategy (5) can be positive or negative. The one-step gains s_t^1 and $s_t^2 = -s_t^1$ are unbounded and can be positive or negative: $s_t^i \in (-\infty, +\infty)$.

[1] The idea of these strategies is based on the paper of Cheredito [2] (see also Rogers [11], Delbaen and Schachermayer [3]) who have constructed arbitrage strategies for a financial market that consists of money market account and a stock whose price follows a fractional Brownian motion with drift or an exponential fractional Brownian motion with drift. Vovk [14] has reformulated these strategies for discrete time. We use these strategies to define a mixed strategy which incur gain when macro and micro volatilities of time series differ. There is no uncertainty domain for continuous time.

We analyze this game in the follow leader framework. We introduce *Learner* that can choose between two strategies (4) and (5).

The following simple example of zero sum experts shows that even in the case where one-step gains of experts are bounded Learner can perform much worse than each expert: let the current gains of two experts on steps $t = 0, 1, \ldots, 6$ be $s^1_{0,1,2,3,4,5,6} = (1/2, -1, 1, -1, 1, -1, 1)$ and $s^2_{0,1,2,3,4,5,6} = (0, 1, -1, 1, -1, 1, -1)$. Suppose that $s^1_0 = s^2_0 = 0$. The "Follow Leader" algorithm always chooses the wrong prediction; its income is $s_{1:6} = -5.5$.

We solve this problem in Section 3 using randomization of the experts cumulative gains.

3 The Follow Perturbed Leader Algorithm with Adaptive Weights

We consider a game of prediction with expert advice. Let at each step t of the game all N experts receive arbitrary one-step gains $s^i_t \in (-\infty, +\infty)$, $i = 1, \ldots N$, and the cumulative gain of the ith expert after step t is equal to $s^i_{1:t} = s^i_{1:t-1} + s^i_t$. A probabilistic learning algorithm of choosing an expert presents for any i the probabilities $P\{I_t = i\}$ of following the ith expert given the cumulative gains $s^i_{1:t-1}$ of the experts $i = 1, \ldots N$ in hindsight.

Probabilistic algorithm of choosing an expert.

FOR $t = 1, \ldots T$

Given past cumulative gains of the experts $s^i_{1:t-1}$ choose the expert i, where $i = 1, \ldots N$, with probability $P\{I_t = i\}$.

Receive the one-step gains at step t of the expert s^i_t and suffer one-step gain $s_t = s^i_t$ of the master algorithm.

ENDFOR

The performance of this probabilistic algorithm is measured in its *expected regret*

$$E(\max_{i=1,\ldots N} s^i_{1:T} - s_{1:T}),$$

where the random variable $s_{1:T}$ is the cumulative gain of the master algorithm, $s^i_{1:T}$, $i = 1, \ldots N$, are the cumulative gains of the expert algorithms, E is the mathematical expectation.

In this section we explore asymptotic consistency of probabilistic learning algorithms in case of unbounded one-step gains. A probabilistic algorithm is called *asymptotically consistent* if

$$\liminf_{T\to\infty} \frac{1}{T} E(s_{1:T} - \max_{i=1,\ldots N} s^i_{1:T}) \geq 0. \tag{6}$$

Notice that then $0 \leq s^i_t \leq 1$ all expert algorithms have total gain $\leq T$ on first T steps. This is not true for the unbounded case, and there are no reasons to divide the expected regret (6) on T. We modify the definition (6) of the normalized expected regret as follows. Define *the volume* of a game at step t

$$v_t = \sum_{j=1}^{t} \max_i |s_j^i|.$$

Evidently, $v_{t-1} \le v_t$ for all $t \ge 1$. Put $v_0 = 0$.

A probabilistic learning algorithm is called *asymptotically consistent* (in the modified sense) in a game with N experts if

$$\liminf_{T\to\infty} \frac{1}{v_T} E(s_{1:T} - \max_{i=1,\dots N} s_{1:T}^i) \ge 0. \tag{7}$$

A game is called *non-degenerate* if $v_t \to \infty$ as $t \to \infty$.

Denote $\Delta v_t = v_t - v_{t-1}$ for $t \ge 1$. The number

$$\text{fluc}(t) = \frac{\Delta v_t}{v_t} = \frac{\max_i |s_t^i|}{v_t}, \tag{8}$$

is called *scaled fluctuation* of the game at the step t (put $0/0 = 0$).

By definition $0 \le \text{fluc}(t) \le 1$ for all $t \ge 1$.

The following simple proposition shows that any probabilistic learning algorithm cannot be asymptotically optimal for some game such that $\text{fluc}(t) \not\to 0$ as $t \to \infty$. For simplicity, we consider the case of two experts.

Proposition 1. *For any probabilistic algorithm of choosing an expert and for any $\epsilon > 0$ two experts exist such that*

$$\text{fluc}(t) \ge 1 - \epsilon,$$
$$\frac{1}{v_t} E(\max_{i=1,2} s_{1:t}^i - s_{1:t}) \ge \frac{1}{2}(1 - \epsilon)$$

for all t, where $s_{1:t}$ is the cumulative gain of this algorithm.

Proof. Given a probabilistic algorithm of choosing an expert and ϵ such that $0 < \epsilon < 1$, define recursively one-step gains s_t^1 and s_t^2 of expert 1 and expert 2 at any step $t = 1, 2, \dots$ as follows. By $s_{1:t-1}^1$ and $s_{1:t-1}^2$ denote the cumulative gains of these experts incurred at steps $\le t - 1$; let v_{t-1} be the corresponding volume, where $t = 1, 2, \dots$.

Define $v_0 = 1$ and $M_t = 4v_{t-1}/\epsilon$ for $t \ge 1$.

For $t \ge 1$, define $s_t^1 = 0$ and $s_t^2 = M_t$ if $P\{I_t = 1\} > \frac{1}{2}$, and define $s_t^1 = M_t$ and $s_t^2 = 0$ otherwise. Let for $t \ge 1$, v_t be the volume of the game.

Let s_t be one-step gain of the master algorithm and $s_{1:t}$ be its cumulative gain at step $t \ge 1$. By definition for all $t \ge 1$,

$$E(s_t) = s_t^1 P\{I_t = 1\} + s_t^2 P\{I_t = 2\} \le \frac{1}{2} M_t.$$

Evidently, $E(s_{1,1}) = E(s_1)$ and $E(s_{1:t}) = E(s_{1:t-1}) + E(s_t)$ for all $t \ge 2$. Then $E(s_{1:t}) \le \frac{1}{2}(1 + \epsilon/2) M_t$ for all $t \ge 1$. Evidently, $v_t = v_{t-1} + M_t = M_t(1 + \epsilon/4)$ and $M_t \le \max_i s_{1:t}^i$ for all $t \ge 1$.

Therefore, the normalized expected regret of the master algorithm is bounded from below

$$\frac{1}{v_t}E(\max_i s^i_{1:t} - s_{1:t}) \ge \frac{\frac{1}{2}(1-\epsilon/2)M_t}{M_t(1+\epsilon/4)} \ge \frac{1}{2}(1-\epsilon).$$

for all t. $\triangle$

Let $\xi^1,\dots\xi^N$ be a sequence of i.i.d random variables distributed according to the exponential law with the density $p(x)=\exp\{-x\}$.

Let $\gamma(t)$ be a non-increasing real function such that $0<\gamma(t)\le 1$ for all t; for example, $\gamma(t)=t^{-\delta}$, where $\delta>0$ and $t\ge 1$. Define

$$\alpha_t = \frac{1}{2}\left(1-\frac{\ln\frac{1+\ln N}{6}}{\ln\gamma(t)}\right) \text{ and} \tag{9}$$

$$\mu_t = (\gamma(t))^{\alpha_t} = \sqrt{\frac{6}{1+\ln N}}(\gamma(t))^{1/2}. \tag{10}$$

for all t.[2]

We consider an FPL algorithm with a variable learning rate

$$\epsilon_t = \frac{1}{\mu_t v_{t-1}}, \tag{11}$$

where μ_t is defined by (10) and the volume v_{t-1} depends on experts actions on steps $<t$. By definition $v_t \ge v_{t-1}$ and $\mu_t \le \mu_{t-1}$ for $t=1,2,\dots$. Also, if $\gamma(t)\to 0$ as $t\to\infty$ then $\mu_t\to 0$ as $t\to\infty$.

We suppose without loss of generality that $s^i_0 = v_0 = 0$ for all i and $\epsilon_0=\infty$.

The FPL algorithm is defined as follows:

FPL algorithm.

FOR $t=1,\dots T$

Define $I_t = \mathrm{argmax}_i\{s^i_{1:t-1} + \frac{1}{\epsilon_t}\xi^i\}$, where ϵ_t is defined by (11) and $i = 1,2,\dots N$.[3]

Receive one-step gains s^i_t for experts $i=1,\dots,N$, and receve one-step gain $s^{I_t}_t$ of the FPL algorithm.

ENDFOR

Let $s_{1:T} = \sum\limits_{t=1}^{T} s^{I_t}_t$ be the cumulative gain of the FPL algorithm.

The following theorem gives a bound for regret of the FPL algorithm. It shows also that if the game is non-degenerate and $\Delta v_t = o(v_t)$ as $t\to\infty$ with algorithmic bound then the FPL-algorithm with variable parameters μ_t is asymptotically consistent.

---

[2] The choice of the optimal value of α_t will be explained later. It will be obtained by minimization of the corresponding member of the sum (40) below.
The definition (9) is invalid for $\gamma(t)=1$. In that follows for $\gamma(t)=1$, we will use the values $(\gamma(t))^{\alpha_t}$ and $(\gamma(t))^{1-\alpha_t}$ defined by (10).

[3] If the maximum is achieved for more then one different i choose the minimal such i.

Theorem 1. *Let $\gamma(t)$ be a non-increasing computable real function such that $0 < \gamma(t) \le 1$ and*

$$\mathrm{fluc}(t) \le \gamma(t) \tag{12}$$

for all t. Then the expected gain of the FPL *algorithm with variable learning rate (11) is satisfied*

$$E(s_{1:T}) \ge \max_i s^i_{1:T} - 2\sqrt{6(1+\ln N)}\sum_{t=1}^{T}(\gamma(t))^{1/2}\Delta v_t. \tag{13}$$

for all T.

If the game is non-degenerate and $\gamma(t) \to 0$ as $t \to \infty$ this algorithm is asymptotically consistent

$$\liminf_{T\to\infty}\frac{1}{v_T}E(s_{1:T} - \max_{i=1,\dots N} s^i_{1:T}) \ge 0. \tag{14}$$

Proof. The analysis of optimality of the FPL algorithm is based on an intermediate predictor IFPL (Infeasible FPL).

IFPL algorithm.

FOR $t = 1, \dots T$

Define

$$\epsilon'_t = \frac{1}{\mu_t v_t}, \tag{15}$$

where v_t is the volume of the game at step t and μ_t is defined by (10).

Also, define $J_t = \mathrm{argmax}_i\{s^i_{1:t} + \frac{1}{\epsilon'_t}\xi^i\}$.

Receive the one step gain $s^{J_t}_t$ of the IFPL algorithm.

ENDFOR

The IFPL algorithm predicts under the knowledge of ϵ'_t and $s^i_{1:t}$, $i = 1, \dots N$, which may not be available at beginning of step t. Using unknown value of ϵ'_t and $s^i_{1:t}$ is the main distinctive feature of our version of IFPL.

For any t we have

$$I_t = \mathrm{argmax}_i\{s^i_{1:t-1} + \frac{1}{\epsilon_t}\xi^i\},$$

$$J_t = \mathrm{argmax}_i\{s^i_{1:t} + \frac{1}{\epsilon'_t}\xi^i\} = \mathrm{argmax}_i\{s^i_{1:t-1} + s^i_t + \frac{1}{\epsilon'_t}\xi^i\}.$$

The expected one-step and cumulated gains of the FPL and IFPL algorithms at step t are denoted

$$l_t = E(s^{I_t}_t) \text{ and } r_t = E(s^{J_t}_t),$$

$$l_{1:T} = \sum_{t=1}^{T} l_t \text{ and } r_{1:T} = \sum_{t=1}^{T} r_t,$$

respectively, where $s_t^{I_t}$ is the one-step gain of the FPL algorithm at step t and $s_t^{J_t}$ is the one-step gain of the IFPL algorithm, and E denotes the mathematical expectation.

Lemma 1. *The total expected gains of the* FPL *and* IFPL *algorithms satisfy the inequality*

$$l_{1:T} \geq r_{1:T} - 6\sum_{t=1}^{T}(\gamma(t))^{1-\alpha_t}\Delta v_t \tag{16}$$

for all T.

Proof. For any $j = 1, \dots n$ and fixed reals $c_1, \dots c_N$ define

$$m_j = \max_{i\neq j}\{s^i_{1:t-1} + \frac{1}{\epsilon_t}c_i\},$$
$$m'_j = \max_{i\neq j}\{s^i_{1:t} + \frac{1}{\epsilon'_t}c_i\} = \max_{i\neq j}\{s^i_{1:t-1} + s^i_t + \frac{1}{\epsilon'_t}c_i\}.$$

Let $m_j = s^{j_1}_{1:t-1} + \frac{1}{\epsilon_t}c_{j^1}$ and $m'_j = s^{j_2}_{1:t-1} + s^{j_2}_t + \frac{1}{\epsilon'_t}c_{j_2}$. By definition of m'_j and since $j_1 \neq j$ we have

$$m'_j = s^{j_2}_{1:t-1} + s^{j_2}_t + \frac{1}{\epsilon'_t}c_{j_2} \geq s^{j_1}_{1:t-1} + \frac{1}{\epsilon'_t}c_{j^1} =$$
$$s^{j_1}_{1:t-1} + \frac{1}{\epsilon_t}c_{j_1} + \left(\frac{1}{\epsilon'_t} - \frac{1}{\epsilon_t}\right)c_{j_1} =$$
$$m_j + \left(\frac{1}{\epsilon'_t} - \frac{1}{\epsilon_t}\right)c_{j_1}. \tag{17}$$

We compare conditional probabilities $P\{I_t = j|\xi^i = c_i, i \neq j\}$ and $P\{J_t = j|\xi^i = c_i, i \neq j\}$.

The following chain of equalities and inequalities is valid:

$$P\{I_t = j|\xi^i = c_i, i \neq j\} =$$
$$P\{s^j_{1:t-1} + \frac{1}{\epsilon_t}\xi^j \geq m_j|\xi^i = c_i, i \neq j\} =$$
$$P\{\xi^j \geq \epsilon_t(m_j - s^j_{1:t-1})|\xi^i = c_i, i \neq j\} =$$
$$P\{\xi^j \geq \epsilon'_t(m_j - s^j_{1:t-1}) + (\epsilon_t - \epsilon'_t)(m_j - s^j_{1:t-1})|\xi^i = c_i, i \neq j\} =$$
$$P\{\xi^j \geq \epsilon'_t(m_j - s^j_{1:t-1}) + (\epsilon_t - \epsilon'_t)(s^{j_1}_{1:t-1} - s^j_{1:t-1} + \frac{1}{\epsilon_t}c_{j_1})|\xi^i = c_i, i \neq j\} = \tag{18}$$
$$\exp\{-(\epsilon_t - \epsilon'_t)(s^{j_1}_{1:t-1} - s^j_{1:t-1})\} \times \tag{19}$$
$$P\{\xi^j \geq \epsilon'_t(m_j - s^j_{1:t-1}) + (\epsilon_t - \epsilon'_t)\frac{1}{\epsilon_t}c_{j_1}|\xi^i = c_i, i \neq j\} \geq \tag{20}$$
$$\exp\{-(\epsilon_t - \epsilon'_t)(s^{j_1}_{1:t-1} - s^j_{1:t-1})\} \times$$
$$P\{\xi^j \geq \epsilon'_t(m'_j - s^j_{1:t} + s^j_t - \left(\frac{1}{\epsilon'_t} - \frac{1}{\epsilon_t}\right)c_{j_1}) + (\epsilon_t - \epsilon'_t)\frac{1}{\epsilon_t}c_{j_1}|\xi^i = c_i, i \neq j\} = \tag{21}$$

$$\exp\{-(\epsilon_t - \epsilon'_t)(s^{j_1}_{1:t-1} - s^j_{1:t-1}) - \epsilon'_t s^j_t\} \times \quad (22)$$

$$P\{\xi^j \ge \epsilon'_t(m'_j - s^j_{1:t}) | \xi^i = c_i, i \ne j\} = \quad (23)$$

$$\exp\left\{-\left(\frac{1}{\mu_t v_{t-1}} - \frac{1}{\mu_t v_t}\right)(s^{j_1}_{1:t-1} - s^j_{1:t-1}) - \frac{s^j_t}{\mu_t v_t}\right\} \times$$

$$P\{\xi^j > \frac{1}{\mu_t v_t}(s^j_{1:t}) - m'_j | \xi^i = c_i, i \ne j\} \ge$$

$$\exp\left\{-\frac{\Delta v_t}{\mu_t v_t}\frac{(s^j_{1:t-1} - s^{j_1}_{1:t-1})}{v_{t-1}} - \frac{\Delta v_t}{\mu_t v_t}\right\} \times \quad (24)$$

$$P\{\xi^j > \frac{1}{\mu_t v_t}(m'_j - s^j_{1:t}) | \xi^i = c_i, i \ne j\} =$$

$$\exp\left\{-\frac{\Delta v_t}{\mu_t v_t}\left(1 + \frac{s^{j_1}_{1:t-1} - s^j_{1:t-1}}{v_{t-1}}\right)\right\} P\{J_t = 1 | \xi^i = c_i, i \ne j\}. \quad (25)$$

Here we have used twice, in (18)-(19) and in (21)-(22), the equality $P\{\xi > a + b\} = e^{-b}P\{\xi > a\}$ for any random variable ξ distributed according to the exponential law. The equality (20)-(21) follows from (17) and $\epsilon_t \ge \epsilon'_t$ for all t. We have used in (24) the equality $v_t - v_{t-1} = \max_i |s^i_t|$.

The expression in the exponent (25) is bounded

$$\left|\frac{s^{j_1}_{1:t-1} - s^j_{1:t-1}}{v_{t-1}}\right| \le 2, \quad (26)$$

since $\left|\frac{s^i_{1:t-1}}{v_{t-1}}\right| \le 1$ for all t and i.

Therefore, we obtain

$$P\{I_t = j | \xi^i = c_i, i \ne j\} \ge$$

$$\exp\left\{-\frac{3}{\mu_t}\frac{\Delta v_t}{v_t}\right\} P\{J_t = j | \xi^i = c_i, i \ne j\} \ge \quad (27)$$

$$\exp\{-3(\gamma(t))^{1-\alpha_t}\} P\{J_t = j | \xi^i = c_i, i \ne j\}. \quad (28)$$

Since, the inequality (28) holds for all c_i, it also holds unconditionally

$$P\{I_t = j\} \ge \exp\{-3(\gamma(t))^{1-\alpha_t}\} P\{J_t = j\}. \quad (29)$$

for all $t = 1, 2, \ldots$ and $j = 1, \ldots N$.

Since $s^j_t + \Delta v_t \ge 0$ for all j and t, we obtain from (29)

$$l_t + \Delta v_t = E(s^{I_t}_t + \Delta v_t) = \sum_{j=1}^N (s^j_t + \Delta v_t) P(I_t = j) \ge$$

$$\exp\{-3(\gamma(t))^{1-\alpha_t}\} \sum_{j=1}^N (s^j_t + \Delta v_t) P(J_t = j) =$$

$$\exp\{-3(\gamma(t))^{1-\alpha_t}\}(E(s^{J_t}_t) + \Delta v_t) =$$

$$\exp\{-3(\gamma(t))^{1-\alpha_t}\}(r_t+\Delta v_t)\geq$$
$$(1-3(\gamma(t))^{1-\alpha_t})(r_t+\Delta v_t)=$$
$$r_t+\Delta v_t-3(\gamma(t))^{1-\alpha_t}(r_t+\Delta v_t)\geq$$
$$r_t+\Delta v_t-6(\gamma(t))^{1-\alpha_t}\Delta v_t. \quad (30)$$

In the last line of (30) we have used the inequality $|r_t|\leq\Delta v_t$ for all t and the inequality $\exp\{-3r\}\geq 1-3r$ for all r.

Subtracting Δv_t from both sides of the inequality (30) and summing it by $t=1,\dots T$, we obtain

$$l_{1:T}\geq r_{1:T}-6\sum_{t=1}^{T}(\gamma(t))^{1-\alpha_t}\Delta v_t$$

for all T. Lemma 1 is proved. △

The following lemma gives a lower bound for the gain of the IFPL algorithm.

Lemma 2. *The expected cumulative gain of the* IFPL *algorithm with the learning rate (15) is bounded by*

$$r_{1:T}\geq\max_i s^i_{1:T}-(1+\ln N)\sum_{t=1}^{T}(\gamma(t))^{\alpha_t}\Delta v_t \quad (31)$$

for all T.

Proof. The proof is along the line of the proof from Hutter and Poland [6] with an exception that now the sequence ϵ'_t is not monotonic.

Let in this proof, $s_t=(s^1_t,\dots s^N_t)$ be a vector of one-step gains and $s_{1:t}=(s^1_{1:t},\dots s^N_{1:t})$ be a vector of cumulative gains of the experts algorithms. Also, let $\xi=(\xi^1,\dots\xi^N)'$ be a vector whose coordinates are i.i.d according to the exponential law random variables.

Recall that $\epsilon'_t=1/(\mu_t v_t)$ and $v_0=0$, $\epsilon_0=\infty$.

Define $\tilde{s}_{1:t}=s_{1:t}+\frac{1}{\epsilon'_t}\xi$ for $t=1,2,\dots$. Consider the one-step gains $\tilde{s}_t=s_t+\xi\left(\frac{1}{\epsilon'_t}-\frac{1}{\epsilon'_{t-1}}\right)$ for the moment. For any vector s and a unit vector d denote

$$M(s)=\text{argmax}_{d\in D}\{d\cdot s\},$$

where $D=\{(0,\dots 1),\dots(1,\dots 0)\}$ is the set of N unit vectors of dimension N and "·" is the inner product of two vectors. We first show that

$$\sum_{t=1}^{T}M(\tilde{s}_{1:t})\cdot\tilde{s}_t\geq M(\tilde{s}_{1:T})\cdot\tilde{s}_{1:T}. \quad (32)$$

For $T=1$ this is obvious. For the induction step from $T-1$ to T we need to show that

$$M(\tilde{s}_{1:T})\cdot\tilde{s}_T\geq M(\tilde{s}_{1:T})\cdot\tilde{s}_{1:T}-M(\tilde{s}_{1:T-1})\cdot\tilde{s}_{1:T-1}.$$

This follows from $\tilde{s}_{1:T} = \tilde{s}_{1:T-1} + \tilde{s}_T$ and

$$M(\tilde{s}_{1:T}) \cdot \tilde{s}_{1:T-1} \le M(\tilde{s}_{1:T-1}) \cdot \tilde{s}_{1:T-1}.$$

We rewrite (32) as follows

$$\sum_{t=1}^{T} M(\tilde{s}_{1:t}) \cdot s_t \ge M(\tilde{s}_{1:T}) \cdot \tilde{s}_{1:T} - \sum_{t=1}^{T} M(\tilde{s}_{1:t}) \cdot \xi \left(\frac{1}{\epsilon'_t} - \frac{1}{\epsilon'_{t-1}} \right). \quad (33)$$

By the definition of M we have

$$M(\tilde{s}_{1:T}) \cdot \tilde{s}_{1:T} \ge M(s_{1:T}) \cdot \left(s_{1:T} + \frac{\xi}{\epsilon'_T} \right) = \max_{d \in D} \{ d \cdot s_{1:T} \} + M(s_{1:T}) \cdot \frac{\xi}{\epsilon'_T}. \quad (34)$$

We have

$$\sum_{t=1}^{T} \left(\frac{1}{\epsilon'_t} - \frac{1}{\epsilon'_{t-1}} \right) M(\tilde{s}_{1:t}) \cdot \xi = \sum_{t=1}^{T} (\mu_t v_t - \mu_{t-1} v_{t-1}) M(\tilde{s}_{1:t}) \cdot \xi. \quad (35)$$

We will use the inequality

$$0 \le E(M(\tilde{s}_{1:t}) \cdot \xi) \le E(M(\xi) \cdot \xi) = E(\max_i \xi^i) \le 1 + \ln N. \quad (36)$$

The proof of this inequality uses an idea of Lemma 1 from [6]. We have for the exponentially distributed random variables ξ^i, $i = 1, \dots N$,

$$P\{\max_i \xi^i \ge a\} = P\{\exists i (\xi^i \ge a)\} \le \sum_{i=1}^{N} P\{\xi^i \ge a\} = N \exp\{-a\}. \quad (37)$$

Since for any non-negative random variable η, $E(\eta) = \int\limits_0^\infty P\{\eta \ge y\} dy$, by (37) we have $E(\max_i \xi^i - \ln N) = \int\limits_0^\infty P\{\max_i \xi^i - \ln N \ge y\} dy \le$ $\int\limits_0^\infty N \exp\{-y - \ln N\} dy = 1$. Therefore, $E(\max_i \xi^i) \le 1 + \ln N$.

By (36) the expectation of (35) has the upper bound

$$\sum_{t=1}^{T} E(M(\tilde{s}_{1:t}) \cdot \xi)(\mu_t v_t - \mu_{t-1} v_{t-1}) \le (1 + \ln N) \sum_{t=1}^{T} \mu_t \Delta v_t.$$

Here we have used the inequality $\mu_t \le \mu_{t-1}$ for all t,

Since $E(\xi^i) = 1$ for all i, the expectation of the last term in (34) is equal to

$$E \left(M(s_{1:T}) \cdot \frac{\xi}{\epsilon'_T} \right) = \frac{1}{\epsilon'_T} = \mu_T v_T. \quad (38)$$

Combining bounds (33)-(35) and (38), we obtain

$$r_{1:T} = E\left(\sum_{t=1}^{T} M(\tilde{s}_{1:t}) \cdot s_t\right) \geq$$

$$\max_i s^i_{1:T} + \mu_T v_T - (1 + \ln N) \sum_{t=1}^{T} \mu_t \Delta v_t \geq$$

$$\max_i s^i_{1:T} - (1 + \ln N) \sum_{t=1}^{T} \mu_t \Delta v_t. \tag{39}$$

Lemma is proved. △.

We finish now the proof of the theorem.

The inequality (16) of Lemma 1 and the inequality (31) of Lemma 2 imply the inequality

$$E(s_{1:T}) \geq \max_i s^i_{1:T} - \sum_{t=1}^{T} (6(\gamma(t))^{1-\alpha_t} + (1 + \ln N)(\gamma(t))^{\alpha_t}) \Delta v_t. \tag{40}$$

for all T.

The optimal value (9) of α_t can be easily obtained by minimization of each member of the sum (40) by α_t. In this case μ_t is equal to (10) and (40) is equivalent to (13).

Let $\gamma(T) \to 0$ as $T \to \infty$ and the game is non-degenerate, i.e., $v_T \to \infty$ as $T \to \infty$. We have $\sum_{t=1}^{T} \Delta v_t = v_T$ for all T. Then by Toeplitz lemma [13]

$$\frac{1}{v_T}\left(2\sqrt{6(1 + \ln N)} \sum_{t=1}^{T} (\gamma(t))^{1/2} \Delta v_t\right) \to 0$$

as $T \to \infty$. This limit and the inequality (13) imply the expected asymptotic consistency (14). Theorem is proved. △

In [1] and [10] polynomial bounds on one-step losses were considered. We also present a bound of this type.

Corollary 1. *Suppose that* $\max_i |s^i_t| \leq t^a$ *for all* t *and and* $\liminf_{t\to\infty} \frac{v_t}{t^{a+\delta}} > 0$, *where* $i = 1, \ldots N$ *and* a, δ *are positive real numbers. Then*

$$E(s_{1:T}) \geq \max_i s^i_{1:T} - O(\sqrt{(1 + \ln N)}) T^{1-\frac{1}{2}\delta + a}$$

as $T \to \infty$, *where* $\gamma(t) = t^{-\delta}$ *and* μ_t *is defined by (10).*

4 Zero-Sum Game

In this section we derandomize the zero-sum game from Section 2. We interpret the expected one-step gain $E(s_t)$ gain as the weighted average of one-step gains of experts strategies. In more detail, at each step t, *Learner* divide his investment in proportion to the probabilities of expert strategies (4) and (5) computed by the FPL algorithm and suffers the gain

$$G_t = 2C(S_t - S_0)(P\{I_t = 1\} - P\{I_t = 2\})\Delta S_t$$

at any step t, where C is an arbitrary positive constant; $G_{1:T} = \sum_{t=1}^{T} G_t = E(s_{1:T})$ is the *Learner's* cumulative gain.

If the game satisfies $|s_t^1|/\sum_{i=1}^{t}|s_i^1| \le \gamma(t)$ for all t then by (13) we have the lower bound

$$G_{1:T} \ge |s_{1:T}^1| - 8\sum_{t=1}^{T}(\gamma(t))^{1/2}|s_t^1|$$

for all T.

Assume that $\gamma(t) = \mu$ for all t. Then $G_{1:T} \ge |s_{1:T}^1| - 8\mu^{1/2}v_T$ for all T.

5 Conclusion

In this paper we study two different problems: the first of them is how use the fractional Brownian motion of prices to suffer gain on financial market; the second one consists in extending methods of the theory of prediction with expert advice for the case when experts one-step gains are unbounded. Though these problems look independent, the first of them serves as a motivating example to study averaging strategies in case of unbounded gains and losses.

The FPL algorithm with variable learning rates is simple to implement and it is bringing satisfactory experimental results when prices follow fractional Brownian motion.

There are some open problems for further research. How to construct a defensive strategy for *Learner* in sense of Shafer and Vovk's book [12]? This means that *Learner* starting with some initial capital never goes to debt and suffer a gain when macro and micro volatilities differ. Also, a general problem is to develop another strategies which suffer gain when prices follow fractional Brownian motion.

In the theory of prediction with expert advice, it is useful to analyze the performance of the well known algorithms (like WM) for the case of unbounded gains and losses in terms of the volume of a game.

There is a gap between Proposition 1 and Theorem 1, since we assume in this theorem that the game satisfies $\text{fluc}(t) \le \gamma(t) \to 0$, where $\gamma(t)$ is computable. Does there exists an asymptotically consistent learning algorithm in case where $\text{fluc}(t) \to 0$ as $t \to \infty$ with no computable upper bound?

Acknowledgments

This research was partially supported by Russian foundation for fundamental research: 09-07-00180-a and 09-01-00709-a. Author thanks Volodya Vovk and Yuri Kalnishkan for useful comments.

References

[1] Allenberg, C., Auer, P., Gyorfi, L., Ottucsak, G.: Hannan consistency in on-Line learning in case of unbounded losses under partial monitoring. In: Balcázar, J.L., Long, P.M., Stephan, F. (eds.) ALT 2006. LNCS (LNAI), vol. 4264, pp. 229–243. Springer, Heidelberg (2006)

[2] Cheredito, P.: Arbitrage in fractional Brownian motion. Finance and Statistics 7, 533 (2003)

[3] Delbaen, F., Schachermayer, W.: A general version of the fundamental theorem of asset pricing. Mathematische Annalen 300, 463–520 (1994)

[4] Freund, Y., Schapire, R.E.: A decision-theoretic generalization of on-line learning and an application to boosting. Journal of Computer and System Sciences 55, 119–139 (1997)

[5] Hannan, J.: Approximation to Bayes risk in repeated plays. In: Dresher, M., Tucker, A.W., Wolfe, P. (eds.) Contributions to the Theory of Games 3, pp. 97–139. Princeton University Press, Princeton (1957)

[6] Hutter, M., Poland, J.: Prediction with expert advice by following the perturbed leader for general weights. In: Ben-David, S., Case, J., Maruoka, A. (eds.) ALT 2004. LNCS (LNAI), vol. 3244, pp. 279–293. Springer, Heidelberg (2004)

[7] Kalai, A., Vempala, S.: Efficient algorithms for online decisions. In: Schölkopf, B., Warmuth, M.K. (eds.) COLT/Kernel 2003. LNCS (LNAI), vol. 2777, pp. 26–40. Springer, Heidelberg (2003); Extended version in Journal of Computer and System Sciences 71, 291–307 (2005)

[8] Littlestone, N., Warmuth, M.K.: The weighted majority algorithm. Information and Computation 108, 212–261 (1994)

[9] Lugosi, G., Cesa-Bianchi, N.: Prediction, Learning and Games. Cambridge University Press, New York (2006)

[10] Poland, J., Hutter, M.: Defensive universal learning with experts for general weight. In: Jain, S., Simon, H.U., Tomita, E. (eds.) ALT 2005. LNCS (LNAI), vol. 3734, pp. 356–370. Springer, Heidelberg (2005)

[11] Rogers, C.: Arbitrage with fractional Brownian motion. Mathematical Finance 7, 95–105 (1997)

[12] Shafer, G., Vovk, V.: Probability and Finance. It's Only a Game! Wiley, New York (2001)

[13] Shiryaev, A.N.: Probability. Springer, Berlin (1980)

[14] Vovk, V.: A game-theoretic explanation of the $\sqrt{dt}$ effect, Working paper #5 (2003), `http://www.probabilityandfinance.com`