MATHEMATICAL ENGINEERING
TECHNICAL REPORTS

Masayuki KUMON, Akimichi TAKEMURA and Kei
TAKEUCHI

Sequential optimizing strategy in multi-dimensional bounded forecasting games

Masayuki Kumon[*]

Akimichi Takemura [†]

Kei Takeuchi [‡]

November 2009

Abstract

We propose a sequential optimizing betting strategy in the multi-dimensional bounded forecasting game in the framework of game-theoretic probability of Shafer and Vovk (2001). By studying the asymptotic behavior of its capital process, we prove a generalization of the strong law of large numbers, where the convergence rate of the sample mean vector depends on the growth rate of the quadratic variation process. The growth rate of the quadratic variation process may be slower than the number of rounds or may even be zero. We also introduce an information criterion for selecting efficient betting items. These results are then applied to multiple asset trading strategies in discrete-time and continuous-time games. In the case of continuous-time game we present a measure of the jaggedness of a vector-valued continuous process. Our results are examined by several numerical examples.

Keywords and phrases: game-theoretic probability, Hölder exponent, information criterion, Kullback-Leibler divergence, quadratic variation, strong law of large numbers, universal portfolio.

1 Introduction

Since the advent of the game-theoretic probability and finance by Shafer and Vovk [10], the field has been expanding rapidly. The present authors have been contributing to this emerging field by showing the essential role of the Kullback-Leibler divergence for

[*]Risk Analysis Research Center, The Institute of Statistical Mathematics, Research Organization of Information and Systems, masayuki_kumon@smile.odn.ne.jp

[†]Graduate School of Information Science and Technology, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, JAPAN, takemura@stat.t.u-tokyo.ac.jp

[‡]Emeritus, Graduate School of Economics, University of Tokyo

the strong law of large numbers (SLLN) [8, 9] and by proposing a new approach to continuous-time games [12, 13]. Our approach to continuous-time games has been further developed by V. Vovk [14, 15, 16].

In this paper we propose a sequential optimizing betting strategy for the multi-dimensional bounded forecasting game in discrete time and apply it as a high-frequency limit order type betting strategy for vector-valued continuous price processes.

Our strategy is very flexible and the analysis of its asymptotic behavior allows us to generalize game-theoretic statements of SLLN to a wide variety of cases. SLLN for the bounded forecasting game is already established in Chapter 3 of [10]. In [8] we gave a simple strategy forcing SLLN with the rate of $O(\sqrt{\log n/n})$, where n is the number of rounds. However the convergence rate of SLLN should depend on the growth rate of the quadratic variation process. For example, in view of Kolmogorov's three series theorem (e.g. Section IV.2 of [11]), the sum $s_n = x_1 + \cdots + x_n \in \mathbb{R}^1$ of centered independent measure-theoretic random variables converges a.s. if the sum of their variances converges (i.e. $\sum_n \mathrm{Var}(x_n) < \infty$). Therefore in this case the sample average $\bar{x}_n = s_n/n$ is of order $O(1/n)$. By our sequential optimizing betting strategy, we can give a unified game-theoretic treatment on the asymptotic behavior of s_n, which depends on the asymptotic behavior of $\sum_{i=1}^n x_i^2$ as $n \to \infty$.

The strength of our results can be seen when we interpret our results in the standard measure-theoretic framework. Let $s_n = x_1 + \cdots + x_n$ be a one-dimensional measure-theoretic martingale w.r.t. a filtration $\{\mathcal{F}_n\}$ with uniformly bounded differences $\|x_n\| \leq 1$, a.e. Let $V_n = x_1^2 + \cdots + x_n^2$. Then with probability one the sequence $\|s_n\|/\sqrt{\max(0, V_n \log V_n)}$, $n = 1, 2, \ldots$, is bounded. See Proposition 2.1 below. Note that in this statement no assumption is made on the growth rate of V_n. The rate itself may be random.

From more practical viewpoint, our sequential betting strategy is very simple to implement even for high dimensions and shows a very competitive performance when applied to various price processes. In Section 6 we compare the performance of our strategy with the well-known universal portfolio strategy developed by Thomas Cover and collaborators ([3, 4, 5, 6]). The performance of our sequential betting strategy is competitive against the universal portfolio. Note that the numerical integration needed for implementing universal portfolio is computationally heavy for high dimensions.

When we can bet on a large number of price processes, it is not always best to form a portfolio comprising all price processes, because estimating the best weight vector for the price processes might take a long time. By approximating the capital process of our sequential optimizing strategy, we will introduce an information criterion for selecting price processes in a portfolio.

The organization of this paper is as follows. In Section 2 we formulate the multi-dimensional bounded forecasting game, introduce our sequential optimizing strategy and state our main theoretical result. In Section 3 we give a proof of our result by analyzing asymptotic behavior of its capital process. We also introduce an information criterion for selecting efficient betting items. These results are then applied to multiple asset trading games in Section 4. In Section 4 we formulate the multiple asset trading game in continuous time and based on high-frequency limit order type betting strategies we present

2

a measure for the jaggedness of a path of a vector-valued continuous process. In Section 5, as indicating the generality of our results, we provide a multiple type of Girsanov's theorem for geometric Brownian motion and an argument concerning the mutual information quantity between betting games. In Section 6 we give numerical results for several Japanese stock price processes. We conclude the paper with some remarks in Section 7.

2 A sequential optimizing strategy and its implication to strong law of large numbers

We treat the following type of discrete time bounded forecasting game between Skeptic and Reality. $\mathcal{K}_0$ is the initial capital of Skeptic, D is a compact region in $\mathbb{R}^d$ such that its convex hull $\operatorname{co} D$ contains the origin in its interior, and $\cdot$ denotes the standard inner product of $\mathbb{R}^d$.

DISCRETE TIME BOUNDED FORECASTING GAME
Protocol:
    $\mathcal{K}_0 := 1$.
    FOR $n = 1, 2, \ldots$:
        Skeptic announces $M_n \in \mathbb{R}^d$.
        Reality announces $x_n \in D$.
        $\mathcal{K}_n = \mathcal{K}_{n-1} + M_n \cdot x_n$.
    END FOR

In this paper we regard d-dimensional vectors such as $x_n = (x_n^1, \ldots, x_n^d)^{\mathrm{t}}$ as column vectors with $^{\mathrm{t}}$ denoting the transpose. $\|x\| = \sqrt{x^{\mathrm{t}}x} = \sqrt{x \cdot x}$ denotes the Euclidean norm of x. Letting $\alpha_n = M_n/\mathcal{K}_{n-1}$, we can rewrite Skeptic's capital as $\mathcal{K}_n = \mathcal{K}_{n-1}(1 + \alpha_n \cdot x_n)$, $\alpha_n \in \mathbb{R}^d$. In the protocol, we require that Skeptic observes his collateral duty, in the sense that $\mathcal{K}_n \geq 0$ for all n irrespective of Reality's moves $x_1, x_2, \ldots$.

For constructing a strategy of Skeptic, consider that Skeptic himself generates 'training data' $\{x_{-n_0+1}, x_{-n_0+2}, \ldots, x_0\}$ of size $n_0 \geq d+1$. This operation is similar to a construction of a prior distribution in Bayesian statistics, where a prior distribution can be specified by a set of prior observations. Throughout this paper we fix an arbitrarily $\epsilon_0 \in (0,1)$ and choose the training data $\{x_{-n_0+1}, \ldots, x_0\}$ in such a way that

$$1 + \alpha \cdot x_n \geq 0, \ n = -n_0 + 1, \ldots, 0, \quad \Rightarrow \quad 1 + \alpha \cdot x \geq \epsilon_0, \ \forall x \in D. \tag{1}$$

Let $P_{n_0, \epsilon_0}^d = \{\alpha \mid 1 + \alpha \cdot x_n \geq 0, \ n = -n_0 + 1, \ldots, 0\}$. Then (1) is equivalent to

$$P_{n_0, \epsilon_0}^d \subset -(1 - \epsilon_0)(\operatorname{co} D)^{\perp},$$

where $(\operatorname{co} D)^{\perp}$ denotes the convex dual of $\operatorname{co} D$. For example, for $d = 1$ and $D = [-1, 1]$, we can take $x_{-1} = 1/(1 - \epsilon_0)$ and $x_0 = -1/(1 - \epsilon_0)$. Then α has to satisfy $|\alpha| \leq 1 - \epsilon_0$

and the right-hand side of (1) holds. For general $D \subset \mathbb{R}^d$ let $\bar{\delta} = \max_{x \in D} \|x\|$ and let $c = \bar{\delta}\sqrt{d}/(1 - \epsilon_0)$. Then we can take $n_0 = 2d$ training vectors as

$$(0, \ldots, 0, \pm c, 0, \ldots, 0)^{\mathrm{t}},$$

where c is in the i-th coordinate ($1 \leq i \leq d$). Then each element α^i, $1 \leq i \leq d$, of $\alpha = (\alpha^1, \ldots, \alpha^d)^{\mathrm{t}}$ has to satisfy $|\alpha^i| \leq 1/c$ and $\|\alpha\| \leq (1 - \epsilon_0)/\bar{\delta}$. Hence the right-hand side of (1) holds by Cauchy-Schwarz inequality. It should also be noted that (1) implies that the training vectors span the whole $\mathbb{R}^d$.

The strategy with a constant vector $\alpha_n \equiv \alpha \in \mathbb{R}^d$ is called a constant proportional betting strategy. For $N \geq 0$ we define

$$\Phi_{0,N}(\alpha) = \sum_{n=-n_0+1}^{N} \log(1 + \alpha \cdot x_n), \tag{2}$$

which is the log capital at round N under the constant proportional betting strategy, including the training data. We add '0' to the subscript to indicate that the training data are included in a summation. Since the game starts at time 1, actually the log capital of the constant proportional betting strategy is $\Phi_{0,N}(\alpha) - \Phi_{0,0}(\alpha) = \sum_{n=1}^{N} \log(1 + \alpha \cdot x_n)$.

Let us consider the maximization of $\Phi_{0,N}(\alpha)$ with respect to $\alpha \in \mathbb{R}^d$. The maximum corresponds to the log capital at time N of a 'hindsight' constant proportional betting strategy. Note that $\Phi_{0,N}(\alpha)$ is a strictly concave function of α. The condition (1) ensures that the maximum of $\Phi_{0,N}(\alpha)$ is attained at the unique point $\alpha = \alpha_N^*$ in the interior of P_{n_0,ϵ_0}^d so that

$$\frac{\partial \Phi_{0,N}}{\partial \alpha} \bigg|_{\alpha=\alpha_N^*} = \sum_{n=-n_0+1}^{N} \frac{x_n}{1 + \alpha_N^* \cdot x_n} = 0. \tag{3}$$

From numerical viewpoint we note that the numerical maximization of $\Phi_{0,N}(\alpha)$ is straightforward even in high dimensions.

We now define *sequential optimizing strategy* (SOS) of Skeptic, which is a realizable strategy unlike the hindsight strategy. It is given by

$$\alpha_n = \alpha_{n-1}^*, \quad n \geq 1. \tag{4}$$

The idea of SOS is very simple. We employ the empirically best constant proportion until the previous round for betting at the current round. Note that SOS depends on the choice of the training data. Skeptic's log capital $\log \mathcal{K}_{1,N}^*$ at round N under SOS is written as

$$\log \mathcal{K}_{1,N}^* = \sum_{n=1}^{N} \log(1 + \alpha_{n-1}^* \cdot x_n).$$

Let $\xi = x_1 x_2 \cdots \in D^\infty$ denote a *path*, which is an infinite sequence of Reality's moves. The set $\Omega = D^\infty$ of paths is called the sample space and any subset E of Ω is called an

event. $\xi^n = x_1 \ldots x_n$ denotes a partial path of Reality until the round n. A strategy $\mathcal{P}$ specifies α_n in terms of ξ^{n-1}, i.e. $\alpha_n = \mathcal{P}(\xi^{n-1})$. The capital process under $\mathcal{P}$ is given as $\mathcal{K}_{1,N}^{\mathcal{P}} = \prod_{n=1}^{N}(1 + \mathcal{P}(\xi^{n-1}) \cdot x_n)$. $\mathcal{P}$ is called prudent, if Skeptic observes his collateral duty by $\mathcal{P}$, i.e. $\mathcal{K}_{1,N}^{\mathcal{P}} \geq 0$, $\forall N \geq 0$, irrespective of Reality's moves $x_1, x_2, \ldots$. In this paper we only consider prudent strategies of Skeptic. We say that Skeptic can weakly force an event $E \subset \Omega$ by a strategy $\mathcal{P}$ if $\limsup_N \mathcal{K}_{1,N}^{\mathcal{P}} = \infty$ for every $\xi \notin E$. As in Section 1 we write

$$s_N = x_1 + \cdots + x_N \in \mathbb{R}^d, \quad V_N = x_1 x_1^{\mathrm{t}} + \cdots + x_N x_N^{\mathrm{t}} \quad (: d \times d). \tag{5}$$

Then $\operatorname{tr} V_N = \sum_{n=1}^{N} \|x_n\|^2$. We are now ready to state our main theorem.

Theorem 2.1. *By the sequential optimizing strategy Skeptic can weakly force*

$$E : \limsup_N \frac{\|s_N\|}{\sqrt{\max(1, \operatorname{tr} V_N \log(\operatorname{tr} V_N))}} < \infty. \tag{6}$$

The maximum in the denominator is needed only for the case that $\sup_N \operatorname{tr} V_N \leq 1$, such as $0 \in D$ and Reality always chooses $x_n \equiv 0$. It is important to emphasize that E in (6) is weakly forced irrespective of the rate of growth of $\operatorname{tr} V_N$, including the zero-growth case, i.e. the case that $\operatorname{tr} V_N$ converges to a finite value. A measure-theoretic interpretation of our result shows the flexibility of our result. When x_n's are measure-theoretic martingale differences, then the capital process under SOS is a non-negative measure-theoretic martingale, which converges to a finite value almost surely. Therefore as in Chapter 8 of [10] we have the following proposition. We use the same notation as above.

Proposition 2.1. *Let $s_n = x_1 + \cdots + x_n$ be a d-dimensional measure-theoretic martingale w.r.t. a filtration $\{\mathcal{F}_n\}$. Assume that the differences $x_n \in D$ are uniformly bounded a.e.. Then with probability one the sequence $\|s_n\|/\sqrt{\max(1, \operatorname{tr} V_n \log(\operatorname{tr} V_n))}$, $n = 1, 2, \ldots$, is bounded.*

Let $\lambda_{\max,N}$ and $\lambda_{\min,N}$ denote the maximum and the minimum eigenvalues of V_N. Consider the event

$$E' : \lim_N \frac{\log \lambda_{\max,N}}{\lambda_{\min,N}} = 0. \tag{7}$$

Theorem 2.1 gives only the order of s_N. If we condition the paths on the event E', then we can derive a more accurate numerical bound as follows.

Theorem 2.2. *By the sequential optimizing strategy Skeptic can weakly force*

$$E' \implies \limsup_N \frac{s_N^{\mathrm{t}} V_N^{-1} s_N}{\log |V_N|} \leq 1.$$

This theorem follows from the fact that on E' Skeptic can weakly force $\alpha_N^* \to 0$, as shown in the proof of this theorem in Section 3.4.

5

Note that $\lambda_{\min,N} \to \infty$ on E'. Note also that E in (6) holds if and only if $\limsup_N \|s_N\|/\sqrt{\max(1, \lambda_{\max,N} \log \lambda_{\max,N})} < \infty$, because $\lambda_{\max,N} \le \operatorname{tr} V_N \le d\lambda_{\max,N}$. Hence on E' we have

$$1 \ge \limsup_N \frac{s_N^{\mathrm{t}} V_N^{-1} s_N}{\log |V_N|} \ge \limsup_N \frac{\|s_N\|^2}{d\lambda_{\max,N} \log \lambda_{\max,N}}.$$

Therefore, although we only have a conditional statement in Theorem 2.2, it gives a more accurate numerical bound than Theorem 2.1.

3 Proof of the theorem and some other results on sequential optimizing strategy

In this section we provide proofs of the above theorems and present other results on the sequential optimizing strategy. For readability, we divide the section into several subsections.

3.1 Properties of α_N^* and the empirical risk neutral distribution

Let δ_x denote a unit point mass at $x \in \mathbb{R}^d$ and let $g_N = \sum_{n=-n_0+1}^{N} \delta_{x_n}/(N+n_0)$ denote the empirical distribution of the training data and Reality's moves $x_1, \ldots, x_N$ up to round N. In view of (3) we define the *empirical risk neutral distribution* g_N^* up to round N by

$$g_N^* = \frac{1}{N+n_0} \sum_{n=-n_0+1}^{N} \frac{\delta_{x_n}}{1 + \alpha_N^* \cdot x_n}.$$

For notational simplicity we omit '0' from the subscript of g_N and g_N^*, although they involve the training data. g_N^* is indeed a probability measure, because by (3) we have

$$\sum_{x_n} g_N(\{x_n\}) = \frac{1}{N+n_0} \sum_{n=-n_0+1}^{N} \frac{1}{1 + \alpha_N^* \cdot x_n} = \frac{1}{N+n_0} \sum_{n=-n_0+1}^{N} \frac{1 + \alpha_N^* \cdot x_n}{1 + \alpha_N^* \cdot x_n} = 1,$$

where the summation on the left-hand side is over distinct values of x_n, $n = -n_0+1, \ldots, N$. By $E_{g_N^*}[\cdot]$ we denote the expected value under g_N^*. Then (3) is written as $E_{g_N^*}[x] = 0$.

The log capital $\log \bar{\mathcal{K}}_{0,N}^* = \Phi_{0,N}(\alpha_N^*)$ of the constant hindsight strategy α_N^* up to round N including the training data is expressed as

$$\log \bar{\mathcal{K}}_{0,N}^* = \Phi_{0,N}(\alpha_N^*) = (N+n_0) \sum_{x_n} g_N(\{x_n\}) \log \frac{g_N(\{x_n\})}{g_N^*(\{x_n\})} = (N+n_0) D(g_N \| g_N^*), \quad (8)$$

where $D(g_N \| g_N^*)$ denotes the Kullback-Leibler divergence between two probability distributions g_N and g_N^*.

Now note that $\log(1+\alpha_{n-1}^* \cdot x_n) = \Phi_n(\alpha_{n-1}^*) - \Phi_{n-1}(\alpha_{n-1}^*)$. By summation by parts, the difference $\log \bar{\mathcal{K}}_{0,N}^* - \log \mathcal{K}_{1,N}^*$ between the hindsight strategy and SOS can be expressed as

$$\log \bar{\mathcal{K}}_{0,N}^* - \log \mathcal{K}_{1,N}^* = \sum_{n=1}^{N} \Delta\Phi_n + \Phi_{0,0}(\alpha_0^*), \tag{9}$$

where $\Delta\Phi_n = \Phi_n(\alpha_n^*) - \Phi_n(\alpha_{n-1}^*) \geq 0$ and $\Phi_{0,0}(\alpha_0^*)$ is a constant depending only on the training data. We will analyze the behavior the log capital $\log \mathcal{K}_{1,N}^*$ of SOS by analyzing $\log \bar{\mathcal{K}}_{0,N}^*$ and $\sum_{n=0}^{N} \Delta\Phi_n$.

We call

$$\bar{V}_N^* = \frac{1}{N+n_0} V_{0,N}^* = E_{g_N^*}[xx^{\mathrm{t}}] = \frac{1}{N+n_0} \sum_{n=-n_0+1}^{N} \frac{x_n x_n^{\mathrm{t}}}{1 + \alpha_N^* \cdot x_n}$$

the empirical risk neutral covariance matrix for Reality's moves up to round N. Write

$$s_{0,N} = \sum_{n=-n_0+1}^{N} x_n, \quad \bar{x}_{0,N} = \frac{1}{N+n_0} s_{0,N}.$$

Noting $g_N(\{x_n\}) = (1 + \alpha_N^* \cdot x_n) g_N^*(\{x_n\})$, we have

$$\bar{x}_{0,N} = E_{g_N}[x] = E_{g_N^*}[(1 + \alpha_N^* \cdot x)x] = E_{g_N^*}[x] + E_{g_N^*}[xx^{\mathrm{t}}]\alpha_N^* = E_{g_N^*}[xx^{\mathrm{t}}]\alpha_N^*.$$

Therefore α_N^* is expressed as

$$\alpha_N^* = \bar{V}_N^{*-1} \bar{x}_{0,N} = V_{0,N}^{*-1} s_{0,N}. \tag{10}$$

Since $\bar{V}_N^*$ itself contains α_N^*, (10) does not give an explicit expression of α_N^*. However it is a very useful exact relation for our analysis.

We now consider $\Delta\alpha_n^* = \alpha_n^* - \alpha_{n-1}^*$. In the following we use the notation

$$x_n(\alpha) = \frac{x_n}{1 + \alpha \cdot x_n}.$$

Taking the difference of the following two equalities

$$0 = \sum_{i=-n_0+1}^{n} x_i(\alpha_n^*), \qquad 0 = \sum_{i=-n_0+1}^{n-1} x_i(\alpha_{n-1}^*)$$

we obtain

$$0 = \sum_{i=-n_0+1}^{n-1} x_i \frac{(\alpha_{n-1}^* - \alpha_n^*) \cdot x_i}{(1 + \alpha_{n-1}^* \cdot x_i)(1 + \alpha_n^* \cdot x_i)} + x_n(\alpha_n^*).$$

Therefore

$$\Big(\sum_{i=-n_0+1}^{n-1} \frac{x_i x_i^{\mathrm{t}}}{(1+\alpha_{n-1}^* \cdot x_i)(1+\alpha_n^* \cdot x_i)} \Big)(\alpha_n^* - \alpha_{n-1}^*) = x_n(\alpha_n^*).$$

Note that the denominator on the left-hand side is a scalar and the matrix on the left-hand side is positive definite. Then

$$\Delta\alpha_n^* = V_{0,n-1}(\alpha_{n-1}^*, \alpha_n^*)^{-1} x_n(\alpha_n^*), \tag{11}$$

where $V_{0,n-1}(\alpha, \beta) = \sum_{i=-n_0+1}^{n-1} x_i(\alpha) x_i(\beta)^{\mathrm{t}}$.

Concerning the behavior of $\Delta\alpha_n^*$ we state the following lemma, which will be used in Section 3.4.

Lemma 3.1. $\lim_n \Delta\alpha_n^* = 0$ *for every* $\xi \in D^\infty$.

We give a proof of this lemma in Appendix.

3.2 Bounding the difference between the hindsight strategy and SOS from above

We now give a detailed analysis of $\sum_{n=1}^N \Delta\Phi_n$ on the right-hand side of (9) and bound it from above. We note the following simple fact on $\Delta\Phi_n$:

$$\Delta\Phi_n = \Phi_n(\alpha_n^*) - \Phi_n(\alpha_{n-1}^*) = \Phi_{n-1}(\alpha_n^*) - \Phi_{n-1}(\alpha_{n-1}^*) + \log \frac{1 + \alpha_n^* \cdot x_n}{1 + \alpha_{n-1}^* \cdot x_n}$$

$$\leq \log \frac{1 + \alpha_n^* \cdot x_n}{1 + \alpha_{n-1}^* \cdot x_n} = \log \Big(1 + \frac{\Delta\alpha_n^* \cdot x_n}{1 + \alpha_{n-1}^* \cdot x_n}\Big), \tag{12}$$

where the inequality holds since α_{n-1}^* maximizes $\Phi_{n-1}(\alpha)$. Substituting (11) into the right-hand side we obtain

$$\Delta\Phi_n \leq \log \big(1 + x_n(\alpha_n^*)^{\mathrm{t}} V_{0,n-1}(\alpha_{n-1}^*, \alpha_n^*)^{-1} x_n(\alpha_{n-1}^*)\big). \tag{13}$$

Note that we can also rewrite

$$1 + x_n(\alpha_n^*)^{\mathrm{t}} V_{0,n-1}(\alpha_{n-1}^*, \alpha_n^*)^{-1} x_n(\alpha_{n-1}^*) = \frac{|V_{0,n}(\alpha_{n-1}^*, \alpha_n^*)|}{|V_{0,n-1}(\alpha_{n-1}^*, \alpha_n^*)|}, \tag{14}$$

where we used a well-known relation between determinants (e.g. Corollary A.3.1 of [1]).

Let

$$C_1 = \max \Big(\sup_{\alpha \in -(\mathrm{co}\, D)^\perp,\, x \in D} \big(1 + \alpha \cdot x\big), \sup_{\substack{-n_0+1 \leq n \leq 0 \\ \alpha \in P_{n_0, \epsilon_0}^d}} \big(1 + \alpha \cdot x_n\big)\Big), \tag{15}$$

which is a constant depending only on the training data. The first argument $C_{1,0} = \sup_{\alpha \in -(\mathrm{co}\, D)^\perp,\, x \in D}(1 + \alpha \cdot x)$ on the right-hand side of (15) corresponds to the maximum one-step growth rate of Skeptic's capital under the collateral duty and $C_{1,0}$ equals 2 if

8

D is symmetric w.r.t. the origin. $C_{1,0}$ may be large if D is highly asymmetric w.r.t. the origin. For example, for $d = 1$ and $D = [-0.1, 1]$, we have $C_{1,0} = 11$.

For two symmetric matrices A, B, let $A \geq B$ mean that $A - B$ is non-negative definite. Then

$$V_{0,n-1}(\alpha_{n-1}^*, \alpha_n^*) \geq \frac{1}{C_1^2} V_{0,n-1},$$

where $V_{0,n-1} = V_{0,n-1}(0,0) = \sum_{i=-n_0+1}^n x_i x_i^{\mathrm{t}}$. Note that $V_{0,n-1}$ is positive definite because of the training data, although V_{n-1} in (5) may be singular. Note also that $1 + \alpha_m^* \cdot x_n \geq \epsilon_0$ for $m, n \geq 1$. Therefore

$$x_n(\alpha_n^*)^{\mathrm{t}} V_{0,n-1}(\alpha_{n-1}^*, \alpha_n^*)^{-1} x_n(\alpha_{n-1}^*) \leq C_2 x_n^{\mathrm{t}} V_{0,n-1}^{-1} x_n, \qquad C_2 = \frac{C_1^2}{\epsilon_0^2}.$$

Hence we can bound

$$\sum_{n=1}^N \Delta\Phi_n \leq \sum_{n=1}^N \log(1 + C_2 x_n^{\mathrm{t}} V_{0,n-1}^{-1} x_n).$$

Write $a_n = x_n^{\mathrm{t}} V_{0,n-1}^{-1} x_n \geq 0$. Note that $1 + a_n = |V_{0,n}|/|V_{0,n-1}|$. Also for $c \geq 1$ and $a \geq 0$ we have $1 + ac \leq (1 + a)^c$ and hence

$$\log(1 + ac) \leq c \log(1 + a).$$

Therefore

$$\sum_{n=1}^N \log(1 + C_2 a_n) \leq C_2 \sum_{n=1}^N \log(1 + a_n) = C_2 \log \frac{|V_{0,N}|}{|V_{0,0}|}.$$

Now we have proved the following lemma.

Lemma 3.2. *The difference $\sum_{n=1}^N \Delta\Phi_n$ on the right-hand side of (9) is bounded from above as*

$$\sum_{n=1}^N \Delta\Phi_n \leq C_2(\log |V_{0,N}| - \log |V_{0,0}|).$$

Since $|V_{0,N}|$ involves the training data, for simplicity in our statement we further bound it as follows. By the inequality between the geometric mean and arithmetic mean we have $|V_{0,N}|^{1/d} \leq \operatorname{tr} V_{0,N}/d$. Hence

$$\log |V_{0,N}| \leq d \log \operatorname{tr} V_{0,N} - d \log d = d \log(\operatorname{tr} V_N + \operatorname{tr} V_{0,0}) - d \log d$$
$$\leq d \log(\operatorname{tr} V_N + \operatorname{tr} V_{0,0}).$$

If $\operatorname{tr} V_N \leq 1$, then $\log(\operatorname{tr} V_N + \operatorname{tr} V_{0,0}) \leq \log(1 + \operatorname{tr} V_{0,0}) \leq \operatorname{tr} V_{0,0}$. On the other hand if $\operatorname{tr} V_N > 1$, then

$$\log(\operatorname{tr} V_N + \operatorname{tr} V_{0,0}) = \log \operatorname{tr} V_N + \log(1 + \frac{\operatorname{tr} V_{0,0}}{\operatorname{tr} V_N}) \leq \log \operatorname{tr} V_N + \frac{\operatorname{tr} V_{0,0}}{\operatorname{tr} V_N}$$
$$\leq \log \operatorname{tr} V_N + \operatorname{tr} V_{0,0}.$$

Therefore for both cases $\log |V_{0,N}| \leq d \max(0, \log \operatorname{tr} V_N) + d \operatorname{tr} V_{0,0}$. Let $C_3 = C_2(d \operatorname{tr} V_{0,0} - \log |V_{0,0}|)$. In summary we have the following bound.

Lemma 3.3. *The difference $\sum_{n=1}^{N} \Delta\Phi_n$ on the right-hand side of (9) is bounded from above as*

$$\sum_{n=1}^{N} \Delta\Phi_n \leq dC_2 \max(0, \log \operatorname{tr} V_N) + C_3, \tag{16}$$

where C_2, C_3 depend only on the training data.

Note that (16) is true even for the case that $\lim_N \operatorname{tr} V_N < \infty$. Note also that $\log \operatorname{tr} V_N$ is of order $O(\log N)$ even for $V_N = O(N^\gamma)$, $0 < \gamma < 1$. However for $\operatorname{tr} V_N = \log N$ we have $\log \operatorname{tr} V_N = \log \log N$.

3.3 Bounding the hindsight strategy from below

In this subsection we bound the hindsight strategy from below and thus finish the proof of Theorem 2.1.

Consider the function $(1 + t)\log(1 + t)$, $t > -1$. By Taylor expansion we have

$$(1 + t)\log(1 + t) = t + \frac{1}{2}\frac{t^2}{1 + \theta t}, \qquad 0 < \theta < 1.$$

By the definition of C_1 in (15) we have

$$(1 + \alpha_N^* \cdot x_n)\log(1 + \alpha_N^* \cdot x_n) \geq \alpha_N^* \cdot x_n + \frac{(\alpha_N^* \cdot x_n)^2}{2C_1}, \quad \forall N \geq 1, \ -n_0 + 1 \leq \forall n \leq N,$$

and

$$\begin{aligned}
\Phi_{0,N}(\alpha_N^*) &= \sum_{n=-n_0+1}^{N} (1 + \alpha_N^* \cdot x_n)\log(1 + \alpha_N^* \cdot x_n)\frac{1}{1 + \alpha_N^* \cdot x_n} \\
&\geq \sum_{n=-n_0+1}^{N} \frac{\alpha_N^* \cdot x_n}{1 + \alpha_N^* \cdot x_n} + \frac{1}{2C_1}\sum_{n=-n_0+1}^{N} \frac{(\alpha_N^* \cdot x_n)^2}{1 + \alpha_N^* \cdot x_n} \\
&= \frac{1}{2C_1}\sum_{n=-n_0+1}^{N} \frac{(\alpha_N^* \cdot x_n)^2}{1 + \alpha_N^* \cdot x_n}, \tag{17}
\end{aligned}$$

where we have used the fact $E_{g_N^*}[x] = 0$. By (10) the summation on the right-hand side can be written as

$$\sum_{n=-n_0+1}^{N} \frac{(\alpha_N^* \cdot x_n)^2}{1 + \alpha_N^* \cdot x_n} = \alpha_N^{*\,\mathrm{t}} V_{0,N}^* \alpha_N^* = \alpha_N^{*\,\mathrm{t}} s_{0,N} = s_{0,N}^{\mathrm{t}} V_{0,N}^{*-1} s_{0,N}. \tag{18}$$

In analyzing the behavior of (18) we need to be careful about the following fact: $1 + \alpha_N^* \cdot x_n$, $n \leq 0$, may be arbitrarily close to zero for the training data $x_{-n_0+1}, \ldots, x_0$.

In particular we might have different behavior between eigenvalues of $V_{0,N}^*$ and and those of V_N. To assess the effect of training data let

$$A_N = \sum_{n=-n_0+1}^{0} \frac{1}{1 + \alpha_N^* \cdot x_n}$$

and define the following event

$$E_1 : \limsup_N \frac{A_N}{\max(0, \log \operatorname{tr} V_N)} < \infty.$$

Again max is needed only for the case that $\operatorname{tr} V_N \leq 1$ for all N. We now show that Skeptic can weakly force E_1. Fix an arbitrary $\xi \in E_1^c$, where E_1^c denotes the complement of E_1. Then $\limsup_N A_N / \max(0, \log \operatorname{tr} V_N) = \infty$ and hence there exists some $n_1 \leq 0$ such that

$$\limsup_N \frac{1/(1 + \alpha_N^* \cdot x_{n_1})}{\max(0, \log \operatorname{tr} V_N)} = \infty.$$

Then there exits a subsequence of rounds $N_1 < N_2 < \cdots$ such that

$$\lim_k \frac{1/(1 + \alpha_{N_k}^* \cdot x_{n_1})}{\max(0, \log \operatorname{tr} V_{N_k})} = \infty.$$

Because $\alpha_{N_k}^* \cdot x_{n_1} \to -1$ we have

$$\limsup_N \frac{\frac{(\alpha_N^* \cdot x_{n_1})^2}{1 + \alpha_N^* \cdot x_{n_1}}}{\max(0, \log \operatorname{tr} V_N)} = \infty.$$

If we compare this with the left-hand side of (18), we see that a single term x_{n_1} of the training data contributes arbitrary large gain to Skeptic in comparison to the right-hand side of (16). This implies that $\limsup_N \log \mathcal{K}_{1,N}^* = \infty$. We have proved that by SOS Skeptic can weakly force E_1. Therefore from now we only consider $\xi \in E_1$.

At this point we distinguish two cases 1) $E_2 : \lim_N \operatorname{tr} V_N < \infty$ or 2) $E_2^c : \lim_N \operatorname{tr} V_N = \infty$. Consider the first case and fix an arbitrary $\xi \in E_2 \cap E_1$. For such a ξ there exists $\delta(\xi) > 0$ such that $\liminf_N (1 + \alpha_N^* \cdot x_n) \geq \delta(\xi)$ for all $n < 0$. Then

$$V_{0,N}^* \leq \frac{1}{\max(\epsilon_0, \delta(\xi))} \sum_{n=-n_0+1}^{N} x_n x_n^{\mathrm{t}}$$

and hence the maximum eigenvalue $\lambda_{\max,0,N}$ of $V_{0,N}^*$ is bounded. Then

$$s_{0,N}^{\mathrm{t}} V_{0,N}^{*-1} s_{0,N} \geq \frac{\|s_{0,N}\|^2}{\lambda_{\max,0,N}}$$

and $\limsup_N \log \mathcal{K}_{1,N}^* = \infty$ if $\limsup_N \|s_{0,N}\|^2 = \infty$. Noting that $\limsup_N \|s_{0,N}\|^2 = \infty$ if and only if $\limsup_N \|s_N\|^2 = \infty$, we have shown that by SOS skeptic can weakly force

$$\lim_N \operatorname{tr} V_N < \infty \implies \limsup_N \|s_N\|^2 < \infty.$$

11

Now consider the second case $E_2^c \cap E_1$. On $E_2^c \cap E_1$

$$\lim_N \frac{\log \operatorname{tr} V_N}{\operatorname{tr} V_N} = 0$$

always holds. Also on $E_2^c \cap E_1$

$$\limsup_N \frac{\operatorname{tr} V_{0,0}(\alpha_N^*)}{\log \operatorname{tr} V_N} < \infty \quad \text{where} \quad V_{0,0}(\alpha_N^*) = \sum_{n=-n_0+1}^{0} \frac{x_n x_n^{\mathrm{t}}}{1 + \alpha_N^* \cdot x_n}.$$

Therefore on $E_2^c \cap E_1$

$$\lim_N \frac{\operatorname{tr} V_{0,0}(\alpha_N^*)}{\operatorname{tr} V_N} = 0.$$

Also

$$\operatorname{tr}(V_{0,N}^* - V_{0,0}(\alpha_N^*)) \le \frac{1}{\epsilon_0} \operatorname{tr} V_N,$$

and hence on $E_2^c \cap E_1$

$$\limsup_N \frac{\operatorname{tr} V_{0,N}^*}{\operatorname{tr} V_N} = \limsup_N \frac{\operatorname{tr} V_{0,0}(\alpha_N^*) + \operatorname{tr}(V_{0,N}^* - V_{0,0}(\alpha_N^*))}{\operatorname{tr} V_N} \le \frac{1}{\epsilon_0}.$$

Now on the right-hand side of (18), for every $\xi \in E_2^c \cap E_1$ there exists $N_0 = N_0(\xi)$ such that for all $n \ge N_0$

$$\Phi_{0,N}(\alpha_N^*) \ge \frac{1}{2C_1} \frac{\|s_{0,N}\|^2}{\operatorname{tr} V_{0,N}^*} \ge \frac{\epsilon_0}{4C_1} \frac{\|s_{0,N}\|^2}{\operatorname{tr} V_N}.$$

Hence if for this ξ

$$\limsup_N \frac{\|s_{0,N}\|^2}{\operatorname{tr} V_N \log \operatorname{tr} V_N} = \infty$$

then $\limsup \log \mathcal{K}_{1,N}^* = \infty$ in view of Lemma 3.3. However on E_2^c the following two events are equivalent:

$$\limsup_N \frac{\|s_{0,N}\|^2}{\operatorname{tr} V_N \log \operatorname{tr} V_N} = \infty \quad \Leftrightarrow \quad \limsup_N \frac{\|s_N\|^2}{\operatorname{tr} V_N \log \operatorname{tr} V_N} = \infty.$$

This completes the proof of Theorem 2.1.

3.4 Better approximation to the capital process of SOS

Note that (13) is convenient because it gives an upper bound which always holds. However bounding by $\Phi_n(\alpha_n^*) - \Phi_n(\alpha_{n-1}^*) \le 0$ in (12) is not very accurate. By expanding $\Phi_n(\alpha_{n-1}^*)$ at $\alpha = \alpha_n^*$ and by noting $\partial \Phi_n(\alpha_n^*) = 0$, we have

$$\Delta \Phi_n = \frac{1}{2} \Delta \alpha_n^{*\mathrm{t}} I_n(\bar{\alpha}_n^*) \Delta \alpha_n^*, \quad \bar{\alpha}_n^* = \theta \alpha_{n-1}^* + (1-\theta)\alpha_n^*, \ 0 < \theta < 1, \tag{19}$$

where $I_n(\alpha) = V_{0,n}(\alpha, \alpha)$ is a $d \times d$ positive-definite matrix given by

$$I_n(\alpha) = -\partial \partial^{\mathrm{t}} \Phi_n(\alpha) = \sum_{i=-n_0+1}^{n} x_i(\alpha) x_i(\alpha)^{\mathrm{t}}.$$

Comparing (19) with the right-hand side of (12), we see that the upper bound in (12) is about twice the actual value of $\Delta\Phi_n$. Now by Lemma 3.1 and (14), we can approximate $\Delta\Phi_n$ as

$$\Delta\Phi_n \sim \frac{1}{2} \log \frac{|I_n(\alpha_n^*)|}{|I_{n-1}(\alpha_n^*)|}.$$

Then accumulated these sum is approximated as

$$\sum_{n=1}^{N} \Delta\Phi_n \sim \frac{1}{2} \sum_{n=1}^{N} \log \frac{|I_n(\alpha_n^*)|}{|I_{n-1}(\alpha_n^*)|} = \frac{1}{2} \log [I_N], \quad [I_N] = \prod_{n=1}^{N} \frac{|I_n(\alpha_n^*)|}{|I_{n-1}(\alpha_n^*)|}. \tag{20}$$

Hence from (8), (9) and (20) we obtain

$$\log \mathcal{K}_{1,N}^* = \log \bar{\mathcal{K}}_N^* - \sum_{n=1}^{N} \Delta\Phi_n - \Phi_{0,0}(\alpha_0^*) \sim ND(g_N \| g_N^*) - \frac{1}{2} \log [I_N].$$

The above result is summarized in the following theorem.

Theorem 3.1. *The log capital of the sequential optimizing strategy* $\log \mathcal{K}_{1,N}^*$ *is approximated as*

$$\log \mathcal{K}_{1,N}^* \sim ND(g_N \| g_N^*) - \frac{1}{2} \log [I_N], \quad [I_N] = \prod_{n=1}^{N} \frac{|I_n(\alpha_n^*)|}{|I_{n-1}(\alpha_n^*)|}. \tag{21}$$

Here we note that the quantity $|I_n(\alpha_n^*)|$ also appeared in the evaluation of Cover's universal portfolio [3] in the name of sensitivity (curvature, volatility) index. Differently from the form $[I_N]$ in SOS, only the last term $|I_N(\alpha_N^*)|$ enters in the sensitivity index. This difference reflects the fact that SOS depends on the intermediate moves of Reality's path $\xi^N = x_1 \cdots x_N$, whereas the universal portfolio is independent of them.

We found that the approximation (21) is extremely accurate in practice (cf. Section 6). Thus we propose to use this approximation as an information criterion for selecting betting items. Let us denote the betting game with d items by $Game(d)$, and suppose that there is a sequence of nested betting games such that

$$Game(1) \subset Game(2) \subset \cdots \subset Game(\bar{d}).$$

We also write the main terms of (21) in $Game(d)$ as

$$\log \mathcal{K}_{1,N}^*(d) \sim ND_d(g_N \| g_N^*) - \frac{1}{2} \log [I_N]_d.$$

As functions of d, $D_d(g_N\|g_N^*)$ increases monotonically and $\log [I_N]_d$ is also expected to increase monotonically (cf. Section 6). Hence due to the trade-off between $D_d(g_N\|g_N^*)$ and $\log [I_N]_d$ with respect to d, we can expect that $\max_{1\le d\le \bar{d}} \log \mathcal{K}_{1,N}^*(d)$ provides the optimal number d^* of betting items. Including this subject, we will examine the obtained results by numerical examples in Section 6.

Finally we give a brief proof of Theorem 2.2. The point of the proof is to show that $\alpha_N^* \to 0$ on E'.

Proof of Theorem 2.2. E' in (7) holds only if $\lambda_{\min,N} \to \infty$. Then by (17) and (18) we have
$$\Phi_{0,N}(\alpha_N^*) \ge \frac{1}{C_1^2}\|\alpha_N^*\|^2 \lambda_{\min,N}.$$

Note that $\log |V_N| \le d \log \lambda_{\max,N}$. Therefore if $\limsup \|\alpha_N^*\| > 0$ then $\limsup_N \mathcal{K}_{1,N}^* = \infty$. This shows that conditional on E' Skeptic can weakly force the event $\alpha_N^* \to 0$.

However when $\alpha_N^* \to 0$, for all sufficiently large N we can approximate
$$\Phi(\alpha_N^*) \sim \frac{1}{2} s_N^{\mathrm{t}} V_N^{-1} s_N, \qquad \sum_{n=1}^N \Delta \Phi_n \sim \frac{1}{2} \log |V_N|.$$

Since $\log |V_N| \to \infty$ on E', if $\limsup_N s_N^{\mathrm{t}} V_N^{-1} s_N / \log |V_N| > 1$ then $\limsup_N \log K_{1,N}^* = \infty$. Therefore conditional on E', by SOS Skeptic can weakly force $\limsup_N s_N^{\mathrm{t}} V_N^{-1} s_N / \log |V_N| \le 1$. $\qquad\square$

4 High frequency limit order SOS in multiple asset trading games in continuous time

In this section we generalize the results of [12] to the multi-dimensional case and apply SOS as a high-frequency limit order type investing strategy to multiple asset trading games in continuous time. We follow the notation and the definitions in [12]. For simplicity of statements we make convenient assumptions and only present salient aspects of SOS.

Let Ω^d denote the set of d-dimensional (component-wise) positive continuous functions on $[0,\infty)$. Market (Reality) chooses an element $S(\cdot) \in \Omega^d$. Investor (Skeptic) enters the market at time $t = t_0 = 0$ with the initial capital of $\mathcal{K}(0) = 1$ and he will buy or sell any amount of the assets $S(t) = (S^1(t),\ldots,S^d(t))^{\mathrm{t}}$ at discrete time points $0 = t_0 < t_1 < t_2 < \cdots$, provided that his capital always remains non-negative. His repeated tradings up to time t_i determine $M_i = (M_i^1,\ldots,M_i^d)^{\mathrm{t}} \in \mathbb{R}^d$, where M_i^j denotes the amount of the asset $S^j(t)$ he holds for the time interval $[t_i, t_{i+1})$. Let $\mathcal{K}(t)$ denote the capital of Investor at time t, which is written as
$$\mathcal{K}(t) = \mathcal{K}(t_i) + M_i \cdot (S(t) - S(t_i)) \quad \text{for} \ \ t_i \le t < t_{i+1}, \tag{22}$$
with $\mathcal{K}(0) = 1$. By defining
$$\alpha_i = (\alpha_i^1,\ldots,\alpha_i^d)^{\mathrm{t}}, \quad \alpha_i^j = \frac{M_i^j S^j(t_i)}{\mathcal{K}(t_i)},$$

we rewrite (22) as

$$\mathcal{K}(t) = \mathcal{K}(t_i)\left(1 + \alpha_i \cdot \frac{S(t) - S(t_i)}{S(t_i)}\right) \quad \text{for} \quad t_i \le t < t_{i+1}$$

in terms of the returns of the assets given by

$$\frac{S(t) - S(t_i)}{S(t_i)} = \left(\frac{S^1(t) - S^1(t_i)}{S^1(t_i)}, \dots, \frac{S^d(t) - S^d(t_i)}{S^d(t_i)}\right)^{\text{t}}.$$

Investor takes some constant $\delta > 0$ and decides the trading times $t_1, t_2, \dots$ by the "limit order" type strategy as follows. After t_i is determined, let t_{i+1} be the first time after t_i when

$$\left\|\frac{S(t_{i+1}) - S(t_i)}{S(t_i)}\right\| = \delta \tag{23}$$

happens. This process leads to a discrete time bounded forecasting game embedded into the asset trading game in the following manner. Let

$$x_n = (x_n^1, \dots, x_n^d)^{\text{t}} \in C_\delta, \quad x_n^j = \frac{S^j(t_{n+1}) - S^j(t_n)}{S^j(t_n)},$$

where C_δ denotes the sphere of radius δ in $\mathbb{R}^d$ given by (23), and also write $\mathcal{K}_n = \mathcal{K}(t_{n+1})$. Then we have the protocol of an embedded discrete time bounded forecasting game.

Embedded Discrete Time Bounded Forecasting Game
Protocol:
$\quad \mathcal{K}_0 := 1, \delta > 0$.
$\quad$ FOR $n = 1, 2, \dots$:
$\quad\quad$ Investor announces $\alpha_n \in \mathbb{R}^d$.
$\quad\quad$ Market announces $x_n \in C_\delta$.
$\quad\quad \mathcal{K}_n = \mathcal{K}_{n-1}(1 + \alpha_n \cdot x_n)$.
$\quad$ END FOR

We now fix $T > 0$, and Investor trades in the time interval $[0, T]$ by SOS in (4). For $A > 0$ let

$$E_{A,0,T} = \{S \in \Omega^d \mid |\log S^j(x) - \log S^j(y)| \le A, \ \exists j \in \{1, \dots, d\}, \ 0 \le \forall x < \forall y \le T\}.$$

Market is assumed to choose $S(\cdot) \in E_{A,0,T}^c$, which means that all d items are active in some time interval in $[0, T]$. We define $N = N(T, \delta, S(\cdot))$ by $t_N < T \le t_{N+1}$. Note that by taking δ sufficiently small,

$$N(T, \delta, S(\cdot)) \ge \frac{A}{\delta}$$

15

for every $S(\cdot) \in E^c_{A,0,T}$, so that $N \to \infty$ as $\delta \to 0$. Investor's capital $\mathcal{K}_\delta(T)$ at $t = T$ is written as

$$\mathcal{K}_\delta(T) = \mathcal{K}^*_{1,N}\left(1 + \alpha^*_{N-1} \cdot \frac{S(T) - S(t_N)}{S(t_N)}\right).$$

Since $\left\|\frac{S(T)-S(t_N)}{S(t_N)}\right\| \le \delta$, we have from (21)

$$\log \mathcal{K}_\delta(T) = \log \mathcal{K}^*_{1,N} + O(1) \sim ND(g_N\|g^*_N) - \frac{1}{2}\log[I_N]. \tag{24}$$

The strategy (10) is written as $\alpha^*_N = \alpha^*_N(T, \delta, S(\cdot)) = V^{*-1}_{0,N}s_{0,N}$. We now assume (cf. Theorem 2.2) that $\delta\alpha^*_N \to 0$ as $\delta \to 0$, i.e., Market chooses a path $S(\cdot) \in E'_T$, where

$$E'_T = \{S(\cdot) \in \Omega^d \mid \lim_{\delta \to 0} \delta\alpha^*_N(T, \delta, S(\cdot)) = 0\}.$$

Then

$$\alpha^*_N = \Big(\sum_{n=-n_0+1}^{N} x_n x^{\mathrm{t}}_n\Big)^{-1} \sum_{n=-n_0+1}^{N} x_n\,(1 + O(\delta)) = V^{-1}_{0,N}\Big(L(T) + \frac{1}{2}v_{0,N}\Big)(1 + O(\delta)),$$

where

$$V_{0,N} = \sum_{n=-n_0+1}^{N} x_n x^{\mathrm{t}}_n, \quad v_{0,N} = \Big(\sum_{n=-n_0+1}^{N} (x^1_n)^2, \ldots, \sum_{n=-n_0+1}^{N} (x^d_n)^2\Big)^{\mathrm{t}},$$
$$L(T) = \log S(T) - \log S(0).$$

We consider the first term $ND(g_N\|g^*_N)$ in (24). As was indicated by (18),

$$ND(g_N\|g^*_N) = \frac{1}{2}\alpha^{*\mathrm{t}}_N V^*_{0,N}\alpha^*_N(1 + O(\delta)) = \frac{1}{2}\alpha^{*\mathrm{t}}_N V_{0,N}\alpha^*_N(1 + O(\delta))$$
$$= \frac{1}{2}\Big[L(T)^{\mathrm{t}}V^{-1}_{0,N}L(T) + \frac{1}{2}\big(L(T)^{\mathrm{t}}V^{-1}_{0,N}v_{0,N} + v^{\mathrm{t}}_{0,N}V^{-1}_{0,N}L(T)\big)$$
$$+ \frac{1}{4}v^{\mathrm{t}}_{0,N}V^{-1}_{0,N}v_{0,N}\Big](1 + O(\delta)). \tag{25}$$

The middle term is dominated by the first term and the third term by Cauchy-Schwarz:

$$|L(T)^{\mathrm{t}}V^{-1}_{0,N}v_{0,N} + v^{\mathrm{t}}_{0,N}V^{-1}_{0,N}L(T)| \le 2\sqrt{L(T)^{\mathrm{t}}V^{-1}_{0,N}L(T)}\sqrt{v^{\mathrm{t}}_{0,N}V^{-1}_{0,N}v_{0,N}}\ .$$

Thus we consider the behavior of the first term and the third term. Because C_δ is the sphere of radius δ we have

$$\mathrm{tr}V_N = \mathrm{tr}D_N = N\delta^2,$$

where

$$V_N = \sum_{n=1}^{N} x_n x^{\mathrm{t}}_n, \quad D_N = \mathrm{diag}\Big(\sum_{n=1}^{N}(x^1_n)^2, \ldots, \sum_{n=1}^{N}(x^d_n)^2\Big).$$

Also the training data are of order δ. Hence $\mathrm{tr}V_{0,N} - \mathrm{tr}V_N = \mathrm{tr}D_{0,N} - \mathrm{tr}D_N = O(\delta^2)$.

Let us decompose $V_{0,N}$ and $v_{0,N}$ as

$$V_{0,N} = D_{0,N}^{1/2} R_{0,N} D_{0,N}^{1/2}, \quad v_{0,N} = D_{0,N} 1_d, \quad 1_d = (1, \ldots, 1)^{\mathrm{t}},$$

$$D_{0,N} = \mathrm{diag}\Big(\sum_{n=-n_0+1}^{N} (x_n^1)^2, \ldots, \sum_{n=-n_0+1}^{N} (x_n^d)^2 \Big),$$

where $R_{0,N}$ is the correlation matrix in $\{x_n^1, \ldots, x_n^d\}$, $n = -n_0 + 1, \ldots, N$. Then

$$v_{0,N}^{\mathrm{t}} V_{0,N}^{-1} v_{0,N} = 1_d^{\mathrm{t}} D_{0,N}^{1/2} R_{0,N}^{-1} D_{0,N}^{1/2} 1_d \geq \frac{1}{d} \mathrm{tr}D_{0,N},$$

because the maximum eigenvalue of $R_{0,N}$ is less than or equal to d.

Suppose that the Hölder exponent of $S(\cdot)$ is $0 < H < 1$ in the sense that

$$S(\cdot) \in E_{H,T} = \{S(\cdot) \mid 0 < \liminf_{\delta \to 0} \frac{\mathrm{tr}V_N}{\delta^{(2-\frac{1}{H})}} \leq \limsup_{\delta \to 0} \frac{\mathrm{tr}V_N}{\delta^{(2-\frac{1}{H})}} < \infty\}.$$

By combining the arguments so far, if $S(\cdot) \in E_{A,0,T}^c \cap E_T' \cap E_{H,T}$ then the following implications hold:

$$H > 0.5 \implies \mathrm{tr}D_N \to 0 \implies L(T)^{\mathrm{t}} V_{0,N}^{-1} L(T) \to \infty,$$
$$H < 0.5 \implies \mathrm{tr}D_N \to \infty \implies v_{0,N}^{\mathrm{t}} V_{0,N}^{-1} v_{0,N} \to \infty.$$

Also it is easily shown that the second term $\frac{1}{2} \log[I_N]$ in (24) is of smaller order than $ND(g_N \| g_N^*)$. We summarize our result as a theorem, which is a multi-dimensional generalization of Theorem 3.1 in [12].

Theorem 4.1. *By a high frequency $(\delta \to 0)$ limit order type sequential optimizing strategy in multiple asset trading games in continuous time, Investor can essentially force $H = 0.5$ for $S(\cdot) \in E_{A,0,T}^c$ in the sense*

$$S(\cdot) \in E_{A,0,T}^c \cap E_T' \cap E_{H,T} \text{ and } H \neq 0.5 \implies \mathcal{K}_\delta(T) \to \infty \text{ as } \delta \to 0.$$

5 Generality of high frequency limit order SOS

In this section we show a generality of the high-frequency limit order SOS developed in the previous section, which implies that when the asset price $S(t)$ follows the vector-valued geometric Brownian motion, our strategy automatically incorporates the well-known constant proportional betting strategy originated with Kelly ([7]) and yields the likelihood ratio in the Girsanov's theorem for geometric Brownian motion. The convergence results in this section are of measure-theoretic almost everywhere convergence.

When $S(t)$ is subject to the d-dimensional geometric Brownian motion with drift vector μ and non-singular volatility matrix σ,

$$L(T) = \left(\mu - \frac{1}{2}\sigma^2\right)T + \sigma W(T),$$

where $W(\cdot)$ denotes the d-dimensional standard Brownian motion, and σ^2 denotes the d-dimensional vector with the diagonal elements of $\sigma\sigma^{\mathrm{t}}$. In this section we let $T \to \infty$ and also let $\delta = \delta_T \to 0$ in such a way that $|\log \delta_T| = o(\sqrt{T})$. We have

$$V_{0,N} = (\sigma\sigma^{\mathrm{t}})T(1 + O(\delta_T)),$$

and hence we can evaluate

$$\alpha_N^* = \left[(\sigma\sigma^{\mathrm{t}})^{-1}\mu + \frac{(\sigma^{-1})^{\mathrm{t}}W(T)}{T}\right](1 + O(\delta_T)). \tag{26}$$

The first term in the right-hand side of (26) is the constant vector, which is derived also from the so-called Kelly criterion of maximizing $E[\log \mathcal{K}(T)]$.

Next consider $ND(g_N \| g_N^*)$ in (24), which was also indicated by (25),

$$\begin{aligned}
ND(g_N \| g_N^*) &= \frac{1}{2}\alpha_N^{*\mathrm{t}}V_{0,N}\alpha_N^*(1 + O(\delta_T)) \\
&= \left[\frac{T}{2}\mu^{\mathrm{t}}(\sigma\sigma^{\mathrm{t}})^{-1}\mu + \frac{1}{2}\left((\sigma^{-1}\mu)^{\mathrm{t}}W(T) + W(T)^{\mathrm{t}}(\sigma^{-1}\mu)\right)\right](1 + O(\delta_T)).
\end{aligned}$$

The log capital (24) is then expressed as

$$\begin{aligned}
\log \mathcal{K}_{\delta_T}(T) = \Big[\frac{1}{2}&\left((\sigma^{-1}\mu)^{\mathrm{t}}W(T) + W(T)^{\mathrm{t}}(\sigma^{-1}\mu)\right) + \frac{T}{2}\mu^{\mathrm{t}}(\sigma\sigma^{\mathrm{t}})^{-1}\mu \\
&- \frac{1}{2}\log T + \log \delta_T\Big](1 + O(\delta_T)) + O(1).
\end{aligned}$$

Hence the main terms on the right-hand side

$$-\log \mathcal{K}(T) = -\frac{1}{2}\left((\sigma^{-1}\mu)^{\mathrm{t}}W(T) + W(T)^{\mathrm{t}}(\sigma^{-1}\mu)\right) - \frac{T}{2}\mu^{\mathrm{t}}(\sigma\sigma^{\mathrm{t}})^{-1}\mu + o(\sqrt{T})$$

provide the likelihood ratio of the unique martingale measure known as the Girsanov's theorem in multiple assets case, and we obtain

$$\lim_{T \to \infty} \frac{\log \mathcal{K}(T)}{T} = \frac{1}{2}\mu^{\mathrm{t}}(\sigma\sigma^{\mathrm{t}})^{-1}\mu. \tag{27}$$

Finally we discuss mutual information quantities among subgames of the multi-dimensional bounded forecasting game. Let us denote the quadratic form in the right-hand side of (27) by

$$Q(S) = Q(S^1, \ldots, S^d) = \frac{1}{2}\mu^{\mathrm{t}}(\sigma\sigma^{\mathrm{t}})^{-1}\mu, \tag{28}$$

which designates the optimal exponential growth rate of Investor's capital process with d joint betting items $S = (S^1, \ldots, S^d)$. We partition S into the following form

$$S_{[1]} = (S^{j_1}, \ldots, S^{j_{k_1}}), \ S_{[2]} = (S^{j_{k_1+1}}, \ldots, S^{j_{k_2}}), \ \ldots, \ S_{[m]} = (S^{j_{k_{m-1}+1}}, \ldots, S^{j_{k_m}}),$$

and assume that Investor is allowed to trade the above m groups of joint sub-betting items successively during the one period of the d joint trading. Then the corresponding optimal exponential growth rate of Investor's capital process becomes

$$Q(S_{[1]}) + Q(S_{[2]}) + \cdots + Q(S_{[m]}). \tag{29}$$

Note that among (28) and (29) for all possible partitions there is no general dominance relations and this argument leads to the notion of mutual information quantity between betting games, which will be treated in a forthcoming paper.

6 Numerical examples

In this section we give some numerical examples on the stock price data from the Tokyo Stock Exchange. The data are daily closing prices from January 4th in 2000 to March 31st in 2006 for several Japanese companies listed on the first section of the TSE. There are $T = 1536$ daily closing prices.

From this data we construct the bounded forecasting game in the following manner. At first the daily returns $s_t^j = (S_{t+1}^j - S_t^j)/S_t^j, \ t = 1, \ldots, T-1, \ j = 1, \ldots, d$ of d items are transformed to $[-1, 1]$ by

$$z_t^j = \frac{2s_t^j - \bar{s}_t^j - \underline{s}_t^j}{\bar{s}_t^j - \underline{s}_t^j} \in [-1, 1], \quad \bar{s}_t^j = \max_{1 \leq t \leq T-1} s_t^j, \quad \underline{s}_t^j = \min_{1 \leq t \leq T-1} s_t^j.$$

Next 2^d training data $\tilde{z}_t = (\pm 1, \ldots, \pm 1)^t, \ t = 1, \ldots 2^d$, and a forecasting time $F = cT, \ 0 < c < 1$ are prepared, and forecasting value for the j-th component is

$$\rho^j = \frac{1}{2^d + F} \Big(\sum_{t=1}^{2^d} \tilde{z}_t^j + \sum_{t=1}^{F} z_t^j \Big), \quad j = 1, \ldots, d.$$

Then the bounded variables $x_n = (x_n^1, \ldots, x_n^d)^t$ in the protocol are introduced as

$$x_n^j = \begin{cases} \tilde{z}_n^j - \rho^j, & 1 \leq n \leq 2^d \\ z_{n-2^d+F}^j - \rho^j, & 2^d < n \leq N = 2^d + T - 1 - F. \end{cases}$$

Figures 1-5 and Figures 6-10 exhibit the cases of three items Takeda, Toyota, Kirin with $F = 0.17T$ and $F = 0.25T$, respectively. The notations in the figures are as follows

and their final values at the end of round N are indicated in the figures.

$$K_n^0 = \bar{\mathcal{K}}_n^* = \exp(nD(g_n\|g_n^*)), \quad K_n^1 = \mathcal{K}_n^*, \quad K_n^2 = \frac{\bar{\mathcal{K}}_n^*}{\sqrt{[I_n]}},$$

$$LK_n^0 = \log\bar{\mathcal{K}}_n^* = nD(g_n\|g_n^*), \quad LK_n^1 = \log\mathcal{K}_n^*, \quad LK_n^2 = nD(g_n\|g_n^*) - \frac{1}{2}\log[I_n],$$

$$LD_n^1 = \log\bar{\mathcal{K}}_n^* - \log\mathcal{K}_n^*, \quad LD_n^2 = \frac{1}{2}\log[I_n], \quad LD_n^3 = \frac{3}{2}\log n,$$

$$GR_n = D(g_n\|g_n^*), \quad QR_n = \frac{1}{2}\bar{x}_n^{\mathrm{t}}\bar{V}_n^{*-1}\bar{x}_n, \quad DR_n = \frac{\log[I_n]}{2n}.$$

As suggested in Section 3.4, K_n^1 and K_n^2, LK_n^1 and LK_n^2, LD_n^1 and LD_n^2 are almost overlapped in the figures. We can also see that the actual log deficiency LD_n^1 or LD_n^2 is far less than LD_3 which is the typical log deficiency in the case of finite items such as in the horse race game. Furthermore Figures 5,10 show that the deficiency rate process DR_n gives the precise convergence border rate for the growth rate process GR_n or its approximated quadratic rate process QR_n.

Figures 11-16 illustrate the cases of composite games

$$Game(1) \subset Game(2) \subset Game(3) \subset Game(4) \subset Game(5)$$

with five items 1. Takeda, 2. Toyota, 3. Kirin, 4. Tepco, 5. NNK in this order. As expected the following trade-off can be seen in the figures.

$$\begin{aligned}
LK_n^0 &: \; G(1) \; < G(2) \; < G(3) \; < G(4) \; < G(5) \;, \\
LD_n^2 &: \; G(1) \; < G(2) \; < G(3) \; < G(4) \; < G(5) \;, \\
LK_n^1 &: \; G(1) \; < G(5) \; < G(2) \; < G(4) \; < G(3) \;.
\end{aligned}$$

Hence the choice of the three items 1. Takeda, 2. Toyota, 3. Kirin is the most profitable one in the above composite games.

Figures 17-20 compare the sequential optimizing strategy with the universal portfolio for one item Takeda, Toyota, Kirin, an imaginary data, respectively. The universal portfolio in its simplest form with one item can be performed in the following way.

Divide the closed interval $A = \{\alpha \in \mathbb{R} \mid 1+\alpha x \geq 0, \forall x \in D\}$ of prudent strategies into disjoint subintervals $A_1, \ldots, A_M$. Then for the m-th account A_m with the initial capital $\mathcal{K}_0^{(m)} = 1/M$, Skeptic continues the game with constant betting ratio $\alpha_m \in A_m$, $m = 1, \ldots, M$. His capital at the end of round n is expressed as $\mathcal{K}_n^U = \sum_{m=1}^M \mathcal{K}_n^{(m)}$. The figures are the cases with $M = 100$ and the notations are

$$\begin{aligned}
K_n^{U0} &= \mathcal{K}_n^U \text{ without the training data } \{-1,\ 1\}, \\
K_n^{U1} &= \mathcal{K}_n^U \text{ with the training data } \{-1,\ 1\}.
\end{aligned}$$

Figures 20-22 show the case of an imaginary data given by

$$x_{-1} = -1, \quad x_0 = 1, \quad x_n = \frac{1}{n+1}, \quad n = 1, \ldots, 2000.$$

In this case $LK_n^1 \sim a \log n - c$, $0 < a < 1$, $c > 0$, which contrasts with the case of coin-tossing game $LK_n^1 \sim nD(\bar{x}_n \| \rho) - \frac{1}{2} \log n$.

Figures 17-20 suggest that there is no general superiority between the sequential optimizing strategy and the universal portfolio.

7 Some discussions

In this paper we proposed a sequential optimizing strategy in multi-dimensional bounded forecasting game and showed that it is a very flexible strategy. From a theoretical viewpoint it allowed us to prove a generalized form of the strong law of large numbers. From a practical viewpoint the strategy is easy to implement even in high dimensions and its performance is competitive against universal portfolio.

Theoretical comparison of our strategy with universal portfolio needs more detailed asymptotic investigation of the capital processes of these strategies. This is left to our future research.

In Section 4 as a limit order type strategy we considered successive stopping times defined by a sphere of radius δ for the vector of returns (cf. (23)), which is based on the standard Euclidean norm in $\mathbb{R}^d$. We note that other boundaries based on other norms which are equivalent to the standard one provide the same result stated in Theorem 4.1.

Theorem 2.1 for the case of $\sup_N V_N < \infty$ does not provide a game-theoretic version of Kolmogorov's three series theorem. It only implies that S_N, $N = 1, 2, \ldots$, are bounded. However we expect that a game-theoretic version of Kolmogorov's three series theorem can be established by appropriate modification of our strategy. This topic is also left to our future research.

A A convergence lemma

Let $u_1, u_2, \ldots$ be a sequence of points in $\mathbb{R}^d$. We assume that they are bounded: $\|u_n\| \leq 1$, $\forall n$, and that $u_1, \ldots, u_d$ are linearly independent. Define

$$y_n = (u_1 u_1^{\mathrm{t}} + u_2 u_2^{\mathrm{t}} + \cdots + u_{n-1} u_{n-1}^{\mathrm{t}})^{-1} u_n \in \mathbb{R}^d.$$

Then we have the following lemma. It is trivial for $d = 1$, but for $d > 1$ we need a careful argument.

Lemma A.1.
$$y_n \to 0, \qquad (n \to \infty).$$

Proof. We first show that y_n is bounded. Let $\lambda_{\min,d} > 0$ denote the minimum eigenvalue of $u_1 u_1^{\mathrm{t}} + \cdots + u_d u_d^{\mathrm{t}}$. Then all the eigenvalues of $u_1 u_1^{\mathrm{t}} + \cdots + u_n u_n^{\mathrm{t}}$, $n \geq d$, are greater then equal to $\lambda_{\min,d}$. Then all the eigenvalues of $(u_1 u_1^{\mathrm{t}} + \cdots + u_n u_n^{\mathrm{t}})^{-2}$ are less than or equal to $\lambda_{\min,d}^{-2}$. Hence

$$\|y_n\|^2 \leq \lambda_{\min,d}^{-2} \|u_n\|^2 \tag{30}$$

Figure 1 : Closing prices of Takeda, Toyota,
Kirin $F = 0.17T$
Figure 2 : Capital processes
$$K_n^0, \ K_n^1, \ K_n^2$$
Figure 3 : Log capital processes
$$LK_n^0, \ LK_n^1, \ LK_n^2$$
Figure 4 : Log deficiency processes
$$LD_n^1, \ LD_n^2, \ LD_n^3$$
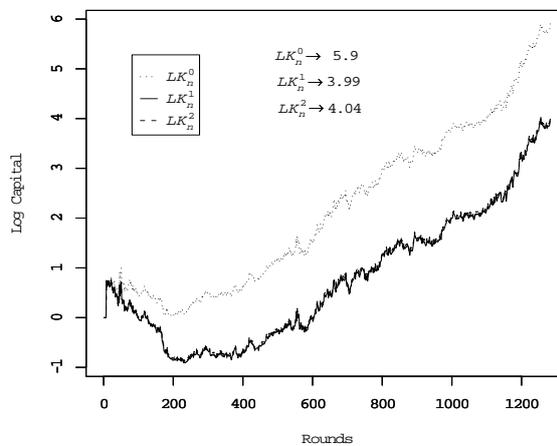Figure 5 : Rate processes
$$GR_n, \ QR_n, \ DR_n$$
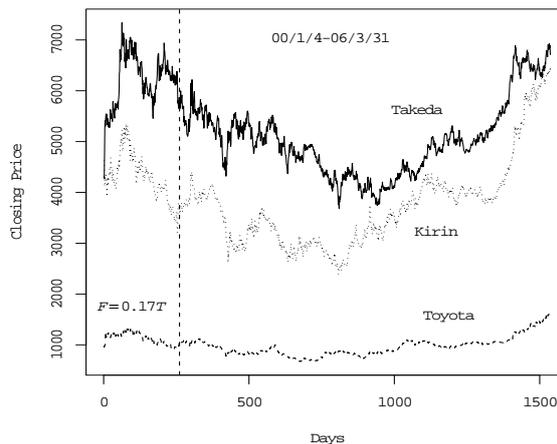
Figure 1: Closing prices
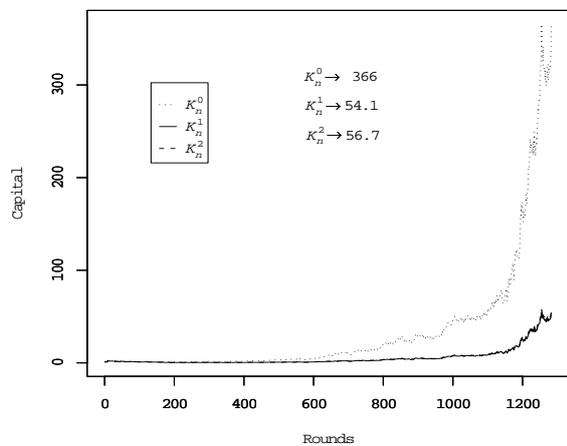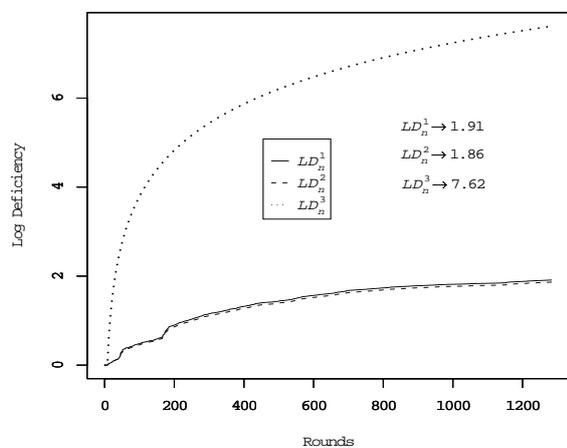
Figure 2: Capital processes

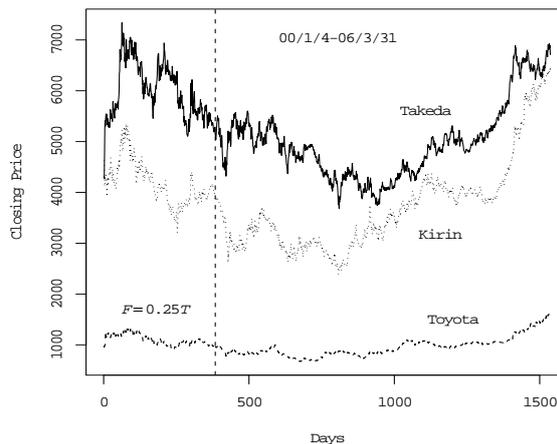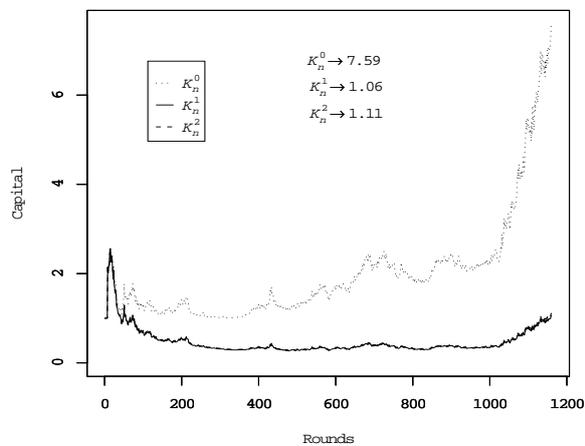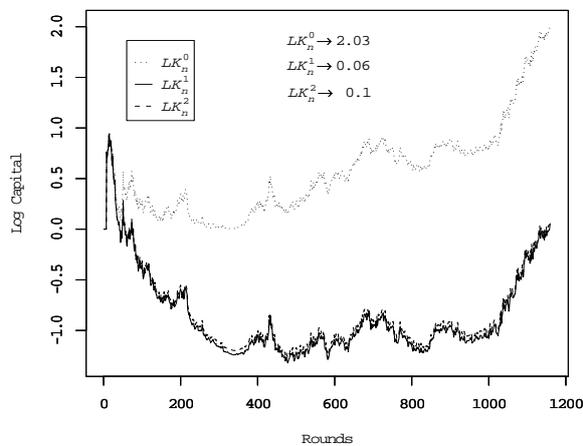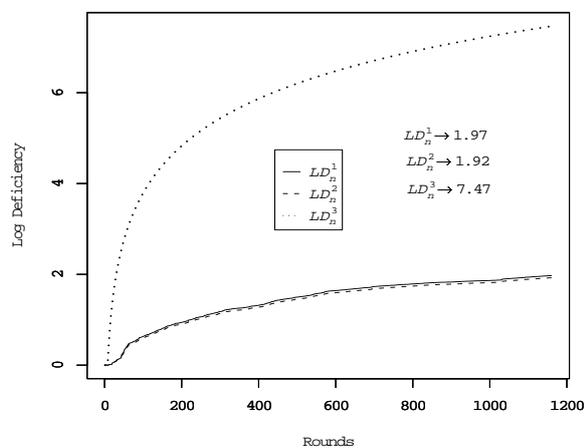Figure 3: Log capital processes

Figure 4: Log deficiency processes

Figure 5: Rate processes

22

Figure 6: Closing prices

Figure 7: Capital processes

Figure 8: Log capital processes

Figure 9: Log deficiency processes
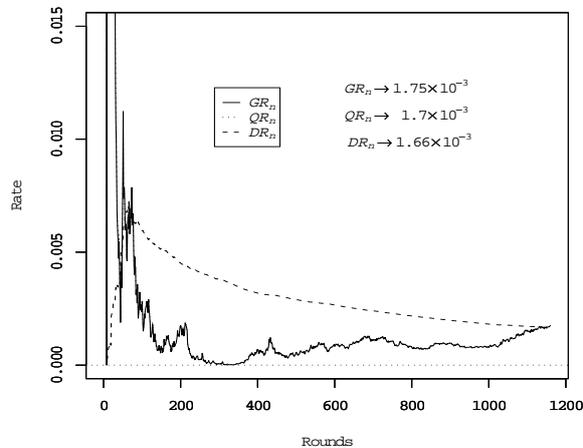
Figure 10: Rate processes

23
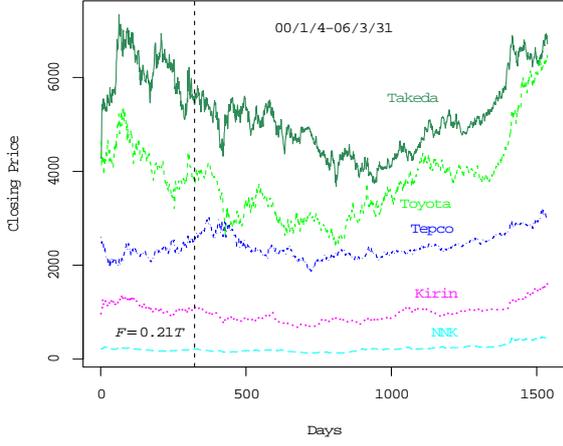
Figure 11: Closing prices
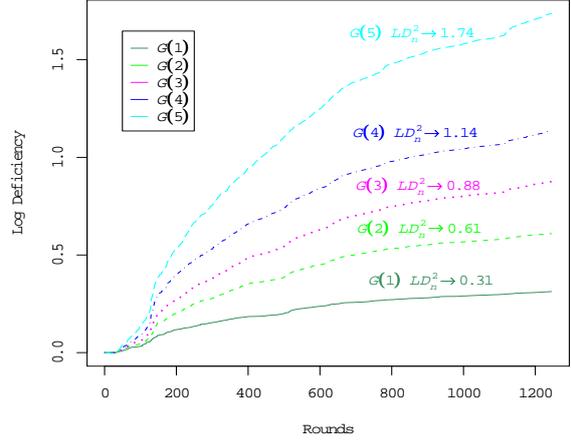
Figure 12: Log deficiency processes LD_n^2
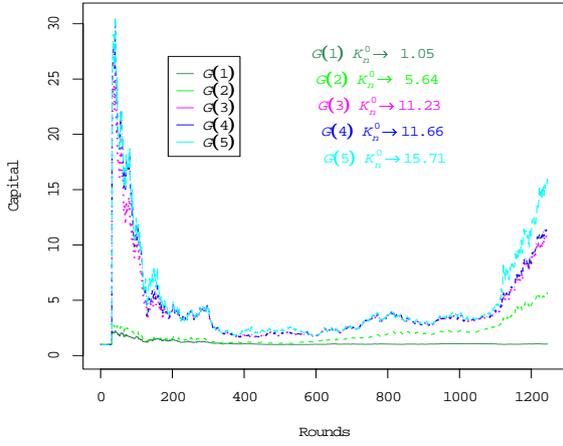
Figure 13: Capital processes K_n^0
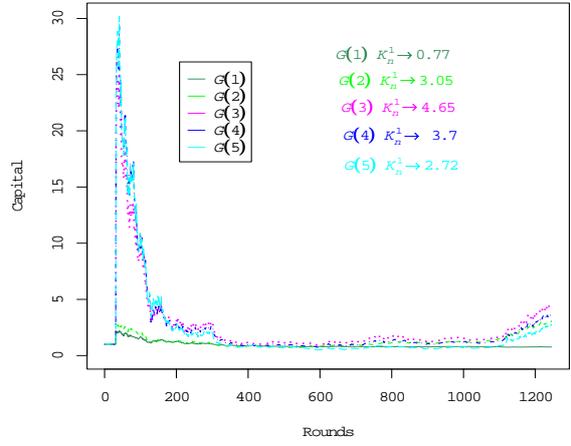
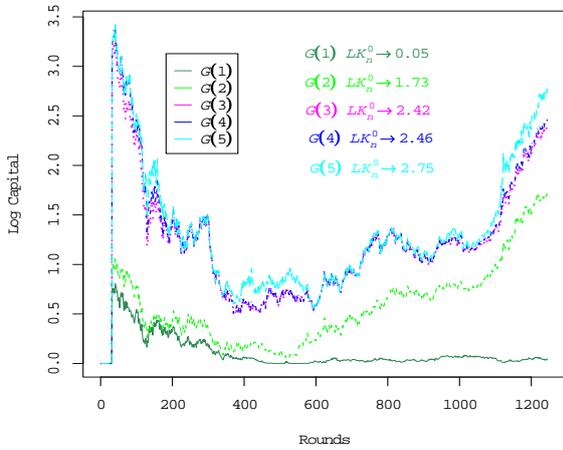Figure 14: Capital processes K_n^1
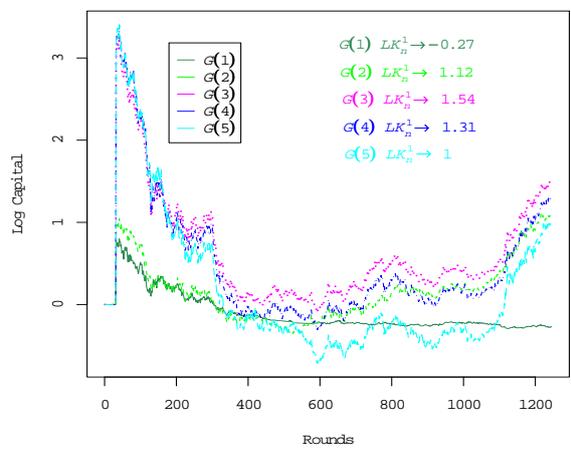

Figure 15: Log capital processes LK_n^0
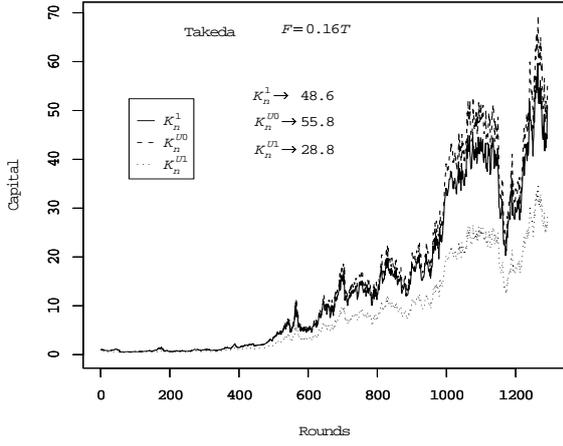

Figure 16: Log capital processes LK_n^1

24

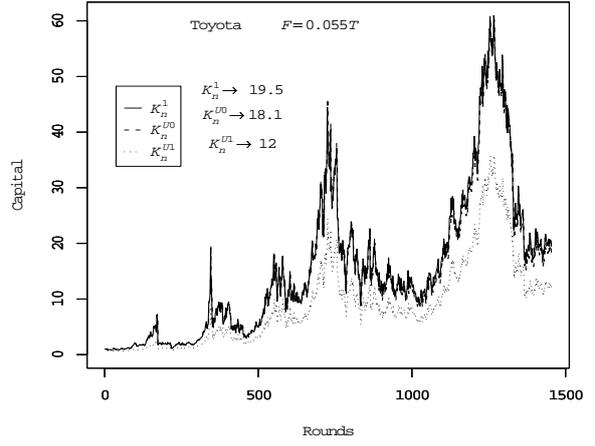Figure 17: $K_n^1,\ K_n^{U0},\ K_n^{U1}$

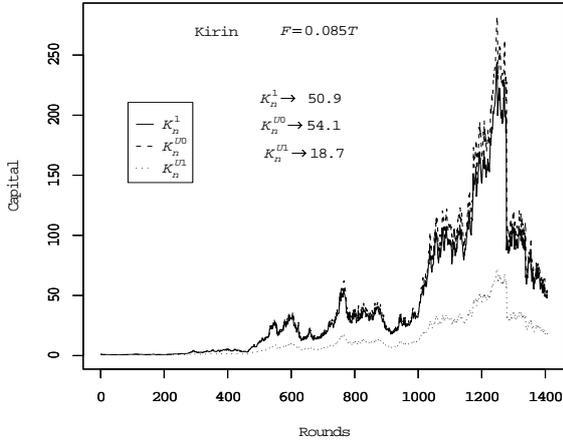

Figure 18: $K_n^1,\ K_n^{U0},\ K_n^{U1}$
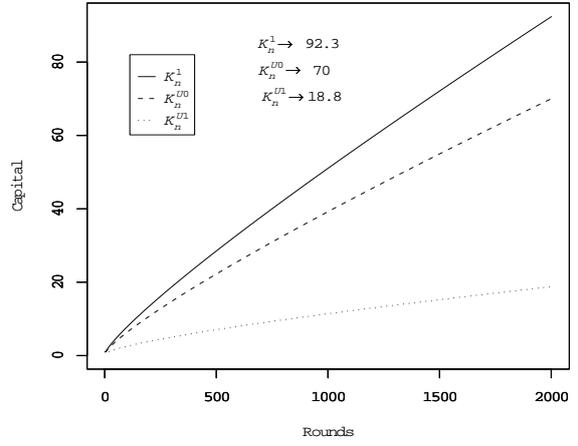

Figure 19: $K_n^1,\ K_n^{U0},\ K_n^{U1}$

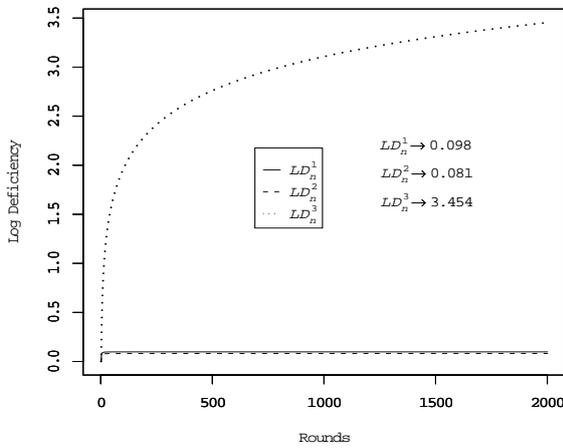Figure 20: $K_n^1,\ K_n^{U0},\ K_n^{U1}$

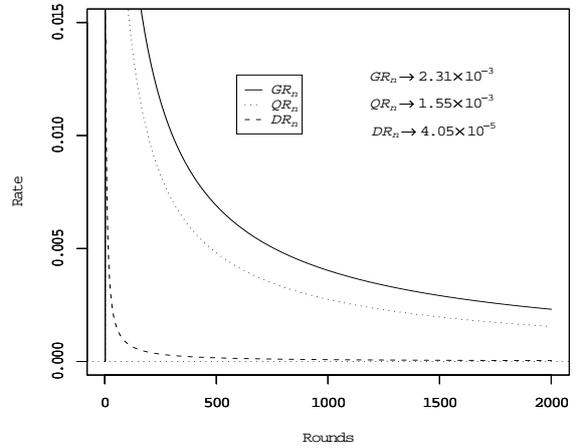Figure 21: Log deficiency processes

Figure 22: Rate processes

25

and y_n, $n = 1, 2, \ldots$, are bounded.

Now we argue by contradiction. Suppose that y_n, $n = 1, 2, \ldots$, do not converge to zero. Then there exists a subsequence n_k, $k = 1, 2, \ldots$ such that $y_{n_k} \to a \neq 0$, $(k \to \infty)$. In view of (30), if $u_{n_k} \to 0$ then $y_{n_k} \to 0$, which is a contradiction. Therefore u_{n_k}, $k = 1, 2, \ldots$, do not converge to 0. Then there exists a further subsequence $\{\tilde{n}_k\} \subset \{n_k\}$ such that $u_{\tilde{n}_k} \to b \neq 0$. Then $y_{\tilde{n}_k} \to a$, $u_{\tilde{n}_k} \to b$. Consider

$$(u_1 u_1^{\mathrm{t}} + \cdots + u_{\tilde{n}_k - 1} u_{\tilde{n}_k - 1}^{\mathrm{t}}) y_{\tilde{n}_k} = u_{\tilde{n}_k}.$$

Then

$$(u_1 u_1^{\mathrm{t}} + \cdots + u_{\tilde{n}_k - 1} u_{\tilde{n}_k - 1}^{\mathrm{t}}) y_{\tilde{n}_k} \to b.$$

Multiplying by $y_{\tilde{n}_k}^{\mathrm{t}}$ from the left we have

$$y_{\tilde{n}_k}^{\mathrm{t}} (u_1 u_1^{\mathrm{t}} + \cdots + u_{\tilde{n}_k - 1} u_{\tilde{n}_k - 1}^{\mathrm{t}}) y_{\tilde{n}_k} = y_{\tilde{n}_k}^{\mathrm{t}} u_{\tilde{n}_k} \to a^{\mathrm{t}} b.$$

Now the left-hand side is written as

$$(y_{\tilde{n}_k}^{\mathrm{t}} u_1)^2 + \cdots + (y_{\tilde{n}_k}^{\mathrm{t}} u_{\tilde{n}_k - 1})^2.$$

Note that for sufficiently large k, k', $(y_{\tilde{n}_k}^{\mathrm{t}} u_{\tilde{n}_{k'}})^2$ are all close to $(b^{\mathrm{t}} a)^2$. Since we have infinitely many such terms, the left-hand side diverges to ∞ if $b^{\mathrm{t}} a \neq 0$. This contradicts the fact that the right-hand side converges to a finite value. Therefore $b^{\mathrm{t}} a = 0$. But then

$$\liminf (y_{\tilde{n}_k}^{\mathrm{t}} u_1)^2 + \cdots + (y_{\tilde{n}_k}^{\mathrm{t}} u_{\tilde{n}_k - 1})^2 \geq (y_{\tilde{n}_k}^{\mathrm{t}} u_1)^2 + \cdots + (y_{\tilde{n}_k}^{\mathrm{t}} u_d)^2$$
$$\to (a^{\mathrm{t}} u_1)^2 + \cdots + (a^{\mathrm{t}} u_d)^2 > 0,$$

which is again a contradiction. □

We also present the following corollary of the above lemma.

Corollary A.1. *With the same notation and conditions as in Lemma 3.1*

$$\tilde{y}_n = (u_1 u_1^{\mathrm{t}} + u_2 u_2^{\mathrm{t}} + \cdots + u_{n-1} u_{n-1}^{\mathrm{t}})^{-1/2} u_n \to 0, \quad (n \to \infty).$$

This corollary follows easily from the fact that $\|\tilde{y}_n\|^2 = u_n^{\mathrm{t}} y_n$ and u_n is bounded.

Based on the above corollary we give a proof of Lemma 3.1. Before going into the proof, we summarize some facts on matrix inequalities. For a symmetric matrix A, let $A > 0$ mean that A is positive definite. If $A \geq B > 0$, then $B^{-1} \geq A^{-1} > 0$ (Lemma 4.2 of [2]). Note that $A \geq B \geq 0$ does not imply $A^2 \geq B^2$ (e.g. Chapter 1 of [17]), which complicates our proof.

Proof of Lemma 3.1. By the definition of C_1 in (15) we have

$$V_{0,n-1}(\alpha_{n-1}^*, \alpha_n^*) \geq \frac{1}{C_1^2} V_{0,n-1}(0,0),$$

where $V_{0,n-1}(0,0) = \sum_{i=-n_0+1}^{n} x_i x_i^{\mathrm{t}}$ is positive definite because of the training data. Write

$$\Delta \alpha_n^* = V_{0,n-1}(\alpha_{n-1}^*, \alpha_n^*)^{-1/2} V_{0,n-1}(\alpha_{n-1}^*, \alpha_n^*)^{-1/2} x_n(\alpha_{n-1}^*).$$

Then

$$\|\Delta \alpha_n^*\|^2 \leq \frac{x_n(\alpha_n^*)^{\mathrm{t}} V_{0,n-1}(\alpha_{n-1}^*, \alpha_n^*)^{-1} x_n(\alpha_n^*)}{\lambda_{\min,0,n-1}(\alpha_{n-1}^*, \alpha_n^*)},$$

where $\lambda_{\min,0,n-1}(\alpha_{n-1}^*, \alpha_n^*)$ is the minimum eigenvalue of $V_{0,n-1}(\alpha_{n-1}^*, \alpha_n^*)$. Let $\lambda_{\min,0,0}$ denote the minimum eigenvalue of $V_{0,0}$. Then $\lambda_{\min,0,n-1}(\alpha_{n-1}^*, \alpha_n^*) \geq \lambda_{\min,0,0}/C_1^2$ for all $n \geq 1$ and

$$\|\Delta \alpha_n^*\|^2 \leq \frac{C_1^2}{\lambda_{\min,0,0}} x_n(\alpha_n^*)^{\mathrm{t}} V_{0,n-1}(\alpha_{n-1}^*, \alpha_n^*)^{-1} x_n(\alpha_n^*).$$

For $n \geq 1$, $1 + \alpha_n^* \cdot x_n \geq \epsilon_0$. Hence

$$\|\Delta \alpha_n^*\|^2 \leq \frac{C_1^2}{\epsilon_0^2 \lambda_{\min,0,0}} x_n^{\mathrm{t}} V_{0,n-1}(\alpha_{n-1}^*, \alpha_n^*)^{-1} x_n \leq \frac{C_1^4}{\epsilon_0^2 \lambda_{\min,0,0}} x_n^{\mathrm{t}} V_{0,n-1}(0,0)^{-1} x_n.$$

The right-hand side converges to 0 by Corollary A.1. $\qquad\square$

References

[1] T. W. Anderson. *An Introduction to Multivariate Statistical Analysis.* 3rd ed., Wiley, Hoboken, New Jersey, 2003.

[2] T. W. Anderson and A. Takemura. A new proof of admissibility of tests in the multivariate analysis of variance. *Journal of Multivariate Analysis*, **12**, 457–478, 1982.

[3] Thomas M. Cover. Universal portfolios. *Mathematical Finance*, **1**, (1), 1-29, 1991.

[4] Thomas M. Cover and E. Ordentlich. Universal portfolios with side information. *IEEE Trans. Inf. Theory*, **IT-42**, 348-363, 1996.

[5] E. Ordentlich and Thomas M. Cover. The cost of achieving the best portfolio in hindsight. *Math. Operations Res.*, **23**, (4), 960-982, 1998.

[6] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory.* 2nd ed., Wiley, New York, 2006.

[7] John L. Kelly. A new interpretation of information rate. *Bell System Technical Journal*, **35**, 917–26, 1956.

[8] Masayuki Kumon and Akimichi Takemura. On a simple strategy weakly forcing the strong law of large numbers in the bounded forecasting game. *Annals of the Institute of Statistical Mathematics*, **60**, 801–812, 2008.

[9] Masayuki Kumon, Akimichi Takemura and Kei Takeuchi. Capital process and optimality properties of a Bayesian Skeptic in coin-tossing games. *Stochastic Analysis and Applications*, **26**, 1161–1180, 2008.

[10] Glenn Shafer and Vladimir Vovk. *Probability and Finance: It's Only a Game!* Wiley, New York, 2001.

[11] Albert N. Shiryaev. *Probability.* Second edition. Springer, New York, 1996.

[12] Kei Takeuchi, Masayuki Kumon and Akimichi Takemura. A new formulation of asset trading games in continuous time with essential forcing of variation exponent. `arXiv:0708.0275v1`, 2007. To appear in *Bernoulli.*

[13] Kei Takeuchi, Masayuki Kumon and Akimichi Takemura. Multistep Bayesian strategy in coin-tossing games and its application to asset trading games in continuous time. `arXiv:0802.4311v2`, 2008. Submitted for publication.

[14] Vladimir Vovk. Continuous-time trading and the emergence of randomness. *Stochastics*, **81**, 455–466, 2009.

[15] Vladimir Vovk. Continuous-time trading and the emergence of volatility *Elect. Comm. in Probab.* **13**, 319–324, 2008.

[16] Vladimir Vovk. Continuous-time trading and the emergence of probability. `arXiv:0904.4364v1`, 2009.

[17] Xingzhi Zhan. *Matrix Inequalities.* Springer, Berlin, 2002.