Numerical Methods for Universal Portfolios

Hiroshi Ishijima *

December, 2001

Abstract

We construct easy-to-use numerical methods for computing universal portfolios. The universal portfolio pioneered by Cover (1991) has a nice property that it can learn the optimal growth rate of portfolio value asymptotically, under incomplete information. The incomplete information here is defined as the sample path of asset prices alone, and the assumption of underlying asset price processes is not required. Up to this time, practical numerical methods for universal portfolios are not presented. Universal portfolios can be expressed in the expectation of constant portfolios weighted by Dirichlet measure. From this viewpoint, we apply Monte Carlo methods for calculating universal portfolios. By the virtue of Monte Carlo, they are not time-consuming even if we increase the universe of assets. We first show how to generate Dirichlet variates in the feasible region of constant portfolios. Our attention is especially paid to the generation of uniform variates which is a special case of Dirichlet variates. Then we test its property via numerical experiments. We empirically found that the suitable sampling method should be adopted according to the property of asset price processes.

Key Words: Universal Portfolios, Monte Carlo Methods, Uniform Sampling, Dirichlet Sampling

JEL Classification: G11

1 Introduction

In this paper, we propose numerical methods for computing universal portfolios which can be simply and efficiently implemented for the practical use. The universal portfolio is

*Faculty of Policy Management, Keio University. Address: 5332, Endoh, Fujisawa 252-8520, Japan. e-Mail : hiroshi_ishijima@hi-ho.ne.jp . Phone & Fax Number: +81-466-49-3444

pioneered by Cover in 1991 [5]. The main property of the universal portfolio is summarized as follows. First we derive the *best constant rebalanced portfolio (BCRP)*, which provides the best growth rate of the portfolio value, given the entire observation of asset prices a priori. Of course, this portfolio uses the future information which is *hindsight* and no investor in the market can obtain this. Hence every actual portfolio constructed must be predictable from available information set. Under this incomplete information, the universal portfolio is able to achieve the asymptotic growth rate of BCRP. Cover [5] and Cover-Ordentlich [6] have proved this property for very general asset price processes in discrete-time framework. Blum-Kalai [3] gives another proof for this. Jamshidian [11] extended the Cover's result [5] to the continuous-time framework. Ishijima [10] provides an interpretation of the universal portfolios in the Bayes' sense in the continuous-time framework under incomplete information. The result there is that the universal portfolio simultaneously estimates the drift parameter and controls the portfolio continuously, as the asymptotically optimal Bayes portfolio.

On the other hand, the universal portfolio can be interpreted from the numerical point of view. Roughly speaking, the universal portfolio is *the portfolio of constant portfolios*. Hence one natural numerical method for the universal portfolio is the familiar Monte Carlo. The key issue in this method is addressed on sampling constant portfolios efficiently, from the feasible region of portfolio weights. To do this, by employing the definition of universal portfolios in the expectation form with respect to the Dirichlet measure F [6], several methods for generating independent constant portfolios from the identical F are introduced. In contrast to our approach, the recent work of Kalai-Vempala [12] adopts another measure from the different viewpoint on the definition of universal portfolios and obtain an efficient algorithm. In their work, however, numerical experiments are not examined. Our numerical methods for computing universal portfolios developed can be simply and efficiently implemented for the practical use. Moreover, we find one should carefully adopt the sampling method according to the property of underlying asset price processes and BCRPs.

This paper is organized as follows; In Section 2, we briefly mention the universal portfolio and address the aim of this paper. In Section 3, numerical methods for computing universal portfolios are described in detail. In Section 4, we conduct two numerical experiments to show the empirical property of the numerical methods developed. In Section 5, we provide the conclusion and the direction of the future research.

2 Preliminaries

We consider a market in which n assets are traded. Let the asset prices be $S_t \stackrel{\triangle}{=} (S_{1,t}, \ldots, S_{n,t})'$ $(t = 0, \ldots, T)$, where the superscript $'$ denotes the transpose. We define the return from the investment in these assets as their price-relatives, $R_{i,t} \stackrel{\triangle}{=} \frac{S_{i,t}}{S_{i,t-1}}$ $(i = 1, \ldots, n)$, and write $\boldsymbol{R}_t \stackrel{\triangle}{=} (R_{1,t}, \ldots, R_{n,t})'$. The important point here is that we make *no* assumption on the asset price processes. The information about asset prices available by time t is denoted by $\mathcal{G}_t \stackrel{\triangle}{=} \sigma(\boldsymbol{R}_u; u = 1, \ldots, t)$. Investors then select their portfolio $\boldsymbol{b}_{t+1} = (b_{1,t+1}, \ldots, b_{n,t+1})$ by exploiting the available information $\mathcal{G}_t$. Here $b_{i,t+1}$ $(i = 1, \ldots, n)$ represents the ratio of amounts invested in the i-th asset to the whole investment. The portfolio is selected within the simplex:

$$\mathbf{D} \stackrel{\triangle}{=} \{\boldsymbol{b} \in \boldsymbol{R}^n | \boldsymbol{b}'\mathbf{1} = 1, \boldsymbol{b} \geq \mathbf{0}\}. \tag{1}$$

where $\mathbf{1}$ is a vector of ones and $\mathbf{0}$ is a vector of zeros. Given the initial bankroll V_0, the portfolio value at time t is:

$$V_t = V_0 \prod_{u=1}^{t} \boldsymbol{b}_u' \boldsymbol{R}_u = V_0 \exp\left(t \cdot G_t(b_\bullet)\right) ,$$

where

$$G_t(\boldsymbol{b}_\bullet) \stackrel{\triangle}{=} \frac{1}{t} \sum_{u=1}^{t} \log \boldsymbol{b}_u' \boldsymbol{R}_u . \tag{2}$$

We call this quantity the *mean growth rate.*

Define the set of *hindsight* information as $\mathcal{G}_T \stackrel{\triangle}{=} \sigma(\boldsymbol{R}_u; u = 1, \ldots, T)$ where T is the terminal-time of the investment horizon. We consider the following problem:

$$\mathbf{P}_0 \left| \begin{array}{ll} \underset{\mathbf{b}}{\text{maximize}} & G_T(\boldsymbol{b}) = \frac{1}{T} \sum_{u=1}^{T} \log \boldsymbol{b}' \boldsymbol{R}_u \\ \text{subject to} & \boldsymbol{b} \text{ is constant, } \mathcal{G}_T\text{-measurable portfolios ,} \\ & \textit{and } \boldsymbol{b} \in \mathbf{D} \text{ .} \end{array} \right.$$

Let the solution to problem $\mathbf{P}_0$ be $\boldsymbol{b}^*$ and call it the *best constant rebalanced portfolio (BCRP).* No one in the market, however, cannot obtain the BCRP, since $\mathcal{G}_T$ is not available until time T arrives. The only information provided to investors at time t is $\mathcal{G}_t$.

Under the setting above, the aim of this paper is to propose numerical methods for universal portfolios which attain the asymptotic mean growth rate of BCRP. The notion of *universality* is first posed by Cover [5], defined as:

Definition 1 (Universality)
$\mathcal{G}_t$-predictable portfolio $\boldsymbol{b}_\bullet$ is said to be universal, if the gap in the mean growth rates between BCRP and $\boldsymbol{b}_\bullet$ vanishes asymptotically, i.e.

$$\lim_{T\to\infty} \sup\left(G_T(\boldsymbol{b}^*) - G_T(\boldsymbol{b}_\bullet)\right) = \lim_{T\to\infty} \sup \frac{1}{T} \log \frac{V_T(\boldsymbol{b}^*)}{V_T(\boldsymbol{b}_\bullet)} = 0 .$$

Since the original universal portfolio initiated by Cover in 1991 [5], several universal portfolios are proposed [8, 12]. We adopt the definition by Cover-Ordentlich [6].

Definition 2 (Universal Portfolio)
Define the universal portfolio at time t as

$$\hat{\boldsymbol{b}}_t \stackrel{\triangle}{=} \frac{\int_{\boldsymbol{b}\in\mathbf{D}} \boldsymbol{b} V_{t-1}(\boldsymbol{b}) dF(\boldsymbol{b})}{\int_{\boldsymbol{b}\in\mathbf{D}} V_{t-1}(\boldsymbol{b}) dF(\boldsymbol{b})} = \int_{\boldsymbol{b}\in\mathbf{D}} \boldsymbol{b} w_{t-1}(\boldsymbol{b}) d\boldsymbol{b} \ , \tag{3}$$

where $F(\cdot)$ is the probability measure which will be defined later and

$$w_{t-1}(\boldsymbol{b}) \stackrel{\triangle}{=} \frac{V_{t-1}(\boldsymbol{b}) f(\boldsymbol{b})}{\int_{\boldsymbol{b}\in\mathbf{D}} V_{t-1}(\boldsymbol{b}) f(\boldsymbol{b}) d\boldsymbol{b}} \ ,$$

will be called the weighting density function.

Remark 1
Since $V_t(\boldsymbol{b}) \geq 0$, then $w_t(\boldsymbol{b}) \geq 0$. From $\int_{\boldsymbol{b}\in\mathbf{D}} w_t(\boldsymbol{b}) d\boldsymbol{b} = 1$, $w_t(\boldsymbol{b})$ can be regarded as density. Also $w_t(\boldsymbol{b})$ weights more to the constant portfolio which has grown its portfolio value most by time t. As a result, the universal portfolio $\hat{\boldsymbol{b}}_t$ can be interpreted as the average of constant portfolios, weighted by $w_t(\boldsymbol{b})$. We also note that, by simple calculation, the value of the universal portfolio is given by

$$V_t(\hat{\boldsymbol{b}}_\bullet) = \int_{\boldsymbol{b}\in\mathbf{D}} V_t(\boldsymbol{b}) dF(\boldsymbol{b}) \ .$$

Concerning the measure $dF(\boldsymbol{b})$, Cover-Ordentlich [6] adopted the Dirichlet distribution $Dir(\nu_1, \ldots, \nu_{n-1}; \nu_n)$ of the form:

$$dF^D(\boldsymbol{b}) = \frac{\Gamma(\nu_1 + \ldots + \nu_n)}{\Gamma(\nu_1) \ldots \Gamma(\nu_n)} b_1^{\nu_1 - 1} \ldots b_{n-1}^{\nu_{n-1}-1} b_n^{\nu_n - 1} db_1 \ldots db_n \ (\boldsymbol{b} \in \mathbf{D}) \ . \tag{4}$$

This exhibit one reason why we employ the Dirichlet distribution as F. That is, the support of the Dirichlet distribution coincides with the simplex $\mathbf{D}$. Or by redefining the simplex as

$$\mathbf{D}' \stackrel{\triangle}{=} \{\boldsymbol{b} \in \boldsymbol{R}^{n-1} | \boldsymbol{b}'\mathbf{1} \leq 1, \boldsymbol{b} \geq \mathbf{0}\} \ ,$$

it can be written as

$$dF^D(\boldsymbol{b}) = \frac{\Gamma(\nu_1 + \ldots + \nu_n)}{\Gamma(\nu_1) \ldots \Gamma(\nu_n)} b_1^{\nu_1 - 1} \ldots b_{n-1}^{\nu_{n-1}-1} (1 - b_1 - \ldots - b_{n-1})^{\nu_n - 1} db_1 \ldots db_{n-1} \ (\boldsymbol{b} \in \mathbf{D}') \ . \tag{5}$$

For a special case, $Dir(1, \ldots, 1; 1)$ is the uniform distribution:

$$dF^D(\boldsymbol{b}) \ = \ (n-1)! \ db_1 \ldots db_n \ (\boldsymbol{b} \in \mathbf{D}) \ . \tag{6}$$

We also pay special attention to $Dir\left(\frac{1}{2},\ldots,\frac{1}{2};\frac{1}{2}\right)$ of the form:

$$dF^D(\boldsymbol{b}) = \frac{\Gamma\left(\frac{n}{2}\right)}{\left(\Gamma\left(\frac{1}{2}\right)\right)^n}\prod_{i=1}^{n} b_i^{-\frac{1}{2}} db_1 \ldots db_n \ (\boldsymbol{b} \in \mathbf{D}) \ . \tag{7}$$

For $n = 2$, this is the beta distribution.

At the end of this section, we note that if F is symmetric as $Dir(1,\ldots,1;1)$ and $Dir\left(\frac{1}{2},\ldots,\frac{1}{2};\frac{1}{2}\right)$, the initial universal portfolio can be written as

$$\hat{\boldsymbol{b}}_1 = \left(\frac{1}{n},\ldots,\frac{1}{n}\right)' \ .$$

3 Numerical Methods

We describe numerical methods for universal portfolios in detail. Rewrite the universal portfolio of Eq. (3) as

$$\hat{\boldsymbol{b}}_t = \frac{\int_{\boldsymbol{b}\in\mathbf{D}} \boldsymbol{b} V_{t-1}(\boldsymbol{b}) dF(\boldsymbol{b})}{\int_{\boldsymbol{b}\in\mathbf{D}} V_{t-1}(\boldsymbol{b}) dF(\boldsymbol{b})} = \frac{E_F[\boldsymbol{b} V_{t-1}(\boldsymbol{b})]}{E_F[V_{t-1}(\boldsymbol{b})]} \ , \tag{8}$$

where $E_F[\cdot]$ is the expectation operator with respect to the measure F. It is assumed that $\int_{\boldsymbol{b}\in\mathbf{D}} V_{t-1}(\boldsymbol{b}) d\boldsymbol{b}$ is finite, which can be expected in the practical market. With $\boldsymbol{b} \in \mathbf{D}$, Eq. (8) is well-defined.

To employ Monte Carlo methods, let $\boldsymbol{b}^{(k)}$ $(k = 1,\ldots,K)$ be the random numbers of constant portfolios which is independently drawn from the identical distribution F, where K is the number of samplings. The universal portfolio is then approximated by

$$\hat{b}_{i,t}^{(K)} = \frac{\frac{1}{K}\sum_{k=1}^{K} b_i^{(k)} V_{t-1}(\boldsymbol{b}^{(k)})}{\frac{1}{K}\sum_{k=1}^{K} V_{t-1}(\boldsymbol{b}^{(k)})} = \frac{\sum_{k=1}^{K} b_i^{(k)} V_{t-1}(\boldsymbol{b}^{(k)})}{\sum_{k=1}^{K} V_{t-1}(\boldsymbol{b}^{(k)})} \ (i = 1,\ldots,n) \ . \tag{9}$$

By the law of large numbers, the approximation of Eq. (9) is expected to converge the universal portfolio as the number of samples become large enough.

Thus the key issue in this numerical method is addressed on how to generate constant portfolios as the variates which are independently drawn from the identical distribution F. In the following two subsections, we describe in detail how to sample constant portfolios $\boldsymbol{b}^{(k)}$ from F numerically. Following Cover-Ordentlich [6], we adopt the Dirichlet distribution $Dir(\nu_1,\ldots,\nu_{n-1};\nu_n)$ for F. We first describe the uniform $Dir(1,\ldots,1;1)$ sampling as a special case. We then focus on the general Dirichlet $Dir(\nu_1,\ldots,\nu_{n-1};\nu_n)$ sampling.

3.1 Uniform samplings

We propose a method for generating constant portfolios on the simplex $\mathbf{D}$, as the uniform variates $Dir(1,\ldots,1;1)$.

Assume $\{X_1,\ldots,X_{n-1}\}$ are mutually independently distributed as the identical uniform distribution $U(0,1)$. The p.d.f. and the c.d.f. are denoted by $f^U(x)=1$ $(0\le x\le 1)$ and $F^U(x)=x$ $(0\ \le x\le 1)$, respectively. We let the *order statistics* of $\{X_1,\ldots,X_{n-1}\}$, sorted from the smallest to the largest, be $\left\{X_{(1)}\le\ldots\le X_{(n-1)}\right\}$. By the result of Takeuchi [15], the density $f_k(x)$ of $X_{(k)}$ can be written as:

$$\begin{aligned} f_k(x) &= \frac{1}{B(k,n-k)} f^U(x)\left\{F^U(x)\right\}^{k-1}\left\{1-F^U(x)\right\}^{n-k-1} \\ &= \frac{1}{B(k,n-k)} x^{k-1}(1-x)^{n-k-1}, \end{aligned} \tag{10}$$

where $B(\cdot,\cdot)$ is the beta function.

Next we consider the two order statistics $X_{(m)},X_{(l)}$ $(m<l)$. Given $X_{(l)}=x_{(l)}$, we introduce the conditional random variable $X|X\le x_l$, with the p.d.f. and the c.d.f. of

$$\begin{aligned} f(x|X\le x_l) &= \frac{f^U(x)}{F^U(x_l)}, \\ F(x|X\le x_l) &= \frac{F^U(x)}{F^U(x_l)}, \ . \end{aligned}$$

Then the density of the m-th order statistics, among mutually independent $(l-1)$ order statistics drawn from $X|X\le x_l$, can be written as:

$$\begin{aligned} f_m(x|X_{(l)}=x_l) &= \frac{1}{B(m,l-m)}\left\{\frac{f^U(x)}{F^U(x_l)}\right\}\left\{\frac{F^U(x)}{F^U(x_l)}\right\}^{m-1}\left\{1-\frac{F^U(x)}{F^U(x_l)}\right\}^{l-m-1} \\ &= \frac{1}{B(m,l-m)} x^{m-1}(x_l-x)^{l-m-1}x_l^{-(l-1)}\ (0\le x\le x_l\le 1)\ . \end{aligned} \tag{11}$$

This can be derived by Eq. (10) with the following substitution:

$$\begin{aligned} n=l,\ f^U(x) &= \frac{f^U(x)}{F^U(x_l)}\ , \\ k=m,\ F^U(x) &= \frac{F^U(x)}{F^U(x_l)}\ . \end{aligned}$$

Hence we have the following results as a summary of the above discussion. From Eq. (10), the p.d.f. and the c.d.f. of $X_{(n-1)}$ are

$$\begin{array}{ll} f_{n-1}(x)=(n-1)x^{n-2} & (0\le x\le 1), \\ F_{n-1}(x)=x^{n-1} & (0\le x\le 1), \\ F_{n-1}^{-1}(x)=\sqrt[n-1]{x} & (0\le x\le 1)\ . \end{array} \tag{12}$$

From Eq. (11), the conditional p.d.f. and the c.d.f. of $X_{(k)}$ $(k = 1, \ldots, n-2)$ are

$$\begin{aligned}
f_k(x|X_{(k+1)} = x_{k+1}) &= \tfrac{k}{x_{k+1}^k} x^{k-1} && (0 \le x \le x_{k+1}), \\
F_k(x|X_{(k+1)} = x_{k+1}) &= \tfrac{x^k}{x_{k+1}^k} && (0 \le x \le x_{k+1}), \\
F_k^{-1}(x|X_{(k+1)} = x_{k+1}) &= x_{k+1}\sqrt[k]{x} && (0 \le x \le 1) \ .
\end{aligned} \tag{13}$$

We transform these order statistics $X_{(1)}, \ldots, X_{(n-1)}$ into constant portfolios as follows:

$$\begin{aligned}
b_1 &= X_{(1)} \ , \\
b_2 &= X_{(2)} - X_{(1)} \ , \\
&\vdots \\
b_{n-1} &= X_{(n-1)} - X_{(n-2)} \ , \\
b_n &= 1 - X_{(n-1)} \ .
\end{aligned} \tag{14}$$

We can easily check that $\{b_1, \ldots, b_n\} \in \mathbf{D}$. Finally we should confirm whether the transformation of Eq. (14) is uniform sampling.

Theorem 1 (Uniform Variates)
$\{b_1, \ldots, b_n\}$ defined as Eq. (14) is uniformly distributed on the simplex $\mathbf{D}$.

Proof. We should prove $b_1, \ldots, b_{n-1}$ is uniformly distributed on $\mathbf{D}'$. From (12) and (13), the p.d.f. of $X_{(1)}, \ldots, X_{(n-1)}$ is given by

$$\begin{aligned}
f(x_1, \ldots, x_{n-1}) &= f(x_{n-1}) f(x_{n-2}|x_{n-1}) \ldots f(x_1|x_{n-1}, x_{n-2}, \ldots, x_2) \\
&= (n-1)\,! \ .
\end{aligned}$$

From Eq. (14), the order statistics can be written as $\left\{X_{(1)} = b_1, \ \ldots \ , \ X_{(n-1)} = \sum_{i=1}^{n-1} b_i\right\}$, and the Jacobian is $|J| = 1$. Hence the probability element (p.e.) of $b_1, \ldots, b_{n-1}$ is

$$\begin{aligned}
g(b_1, \ldots, b_{n-1}) db_1 \ldots db_{n-1} &= f\left(b_1, \ldots, \sum_{i=1}^{n-1} b_i\right) |J| db_1 \ldots db_{n-1} \\
&= n\,!\ db_1 \ldots db_{n-1} \ .
\end{aligned}$$

□

From this theorem, we obtain an algorithm for generating constant portfolios distributed on the simplex.

Algorithm 1 (Uniform Sampling 1)

1. *Generate a uniform variate as u_{n-1}, and set $x_{n-1} = \sqrt[n-1]{u_{n-1}}$.*
2. *Iterate the following steps for $k = n-2, \ldots, 1$;*
   Generate a uniform variate u_k independent of u_{k+1}, and set $x_k = x_{k+1}\sqrt[k]{u_k}$.

3. Transform $\{x_1, \ldots, x_{n-1}\}$ to $\{b_1, \ldots, b_n\}$ as follows:
$\{b_1 = x_1 \ , b_2 = x_2 - x_1 \ , \ldots \ , b_{n-1} = x_{n-1} - x_{n-2} \ , b_n = 1 - x_{n-1}\}$.

Remark 2
One may think that there is a more easy way to generate uniform variates on **D**. *One possible algorithm can be stated in the following steps; First generate $(n-1)$ mutually independent uniform variates and then sort them from the smallest to the largest to obtain $\{x_1, \ldots, x_{n-1}\}$. The next step is the transformation into $\{b_1, \ldots, b_n\}$ which is the same as Algorithm 1. The difference between the above algorithm and Algorithm 1 is that the former requires the sorting procedure and the latter does not.*

3.2 Dirichlet Samplings

We mention how to sample constant portfolios as the Dirichlet variates. We start with the well-known theorem concerning the Dirichlet distribution [16].

Theorem 2 (Dirichlet Variates)
Assume x_i $(i = 1, \ldots, n)$ are mutually independent gamma variates $G(\nu_i)$ and let the p.d.f be $f_i^G(x_i) = \frac{1}{\Gamma(\nu_i)} x_i^{\nu_i - 1} e^{-x_i}$ $(x_i > 0)$, where $\Gamma(\cdot)$ is the gamma function. Then

$$b_i \stackrel{\triangle}{=} \frac{x_i}{\sum_{j=1}^{n} x_j} \ (i = 1, \ldots, n-1) \ ,$$

is the Dirichlet variate $Dir(\nu_1, \ldots, \nu_{n-1}; \nu_n)$.

Proof. Since the proof for this theorem is omitted in Wilks[16], we provide it here. Let the p.e. of $\{x_1, \ldots, x_n\}$ be

$$f^{(1)}(x_1, \ldots, x_n) dx_1 \ldots dx_n = \frac{1}{\prod_{i=1}^{n} \Gamma(\nu_i)} x_1^{\nu_1 - 1} \ldots x_n^{v_n - 1} e^{-(x_1 + \ldots + x_n)} dx_1 \ldots dx_n \ .$$

We then introduce the following transformation:

$$\begin{cases} b_i = \frac{x_i}{\sum_{j=1}^{n} x_j} \ (i = 1, \ldots, n-1) \ , \\ z = x_n \ , \end{cases}$$

Or this can be written as

$$\begin{cases} x_i = \frac{z b_i}{1 - N} \ (i = 1, \ldots, n-1) \ , \\ x_n = z \ , \end{cases} \quad \text{where } N \stackrel{\triangle}{=} \sum_{i=1}^{n-1} b_i \ .$$

Then the Jacobian is $|J_n| = z^{n-1}(1-N)^{-n}$. Hence the p.e. of $\{b_1, \ldots, b_{n-1}, z\}$ is

$$\begin{aligned} f^{(2)}(b_1, \ldots, b_{n-1}, z) db_1 \ldots db_{n-1} dz &= f^{(1)}\left(\frac{zb_1}{1-N}, \ldots, \frac{zb_{n-1}}{1-N}, z\right) |J_n| db_1 \ldots db_{n-1} dz \\ &= \frac{1}{\prod_{i=1}^{n} \Gamma(\nu_1)} (1-N)^{-(\nu_1 + \ldots + \nu_{n-1} + 1)} b_1^{\nu_1 - 1} \ldots b_{n-1}^{\nu_{n-1}-1} \\ &\quad \times z^{(\nu_1 + \ldots + \nu_n) - 1} e^{-\frac{z}{1-N}} db_1 \ldots db_{n-1} dz. \end{aligned}$$

Finally we obtain the joint density function of $\{b_1, \ldots, b_{n-1}\}$ is

$$\begin{aligned} f^{(3)}(y_1, \ldots, y_n) &= \frac{1}{\prod_{i=1}^{n} \Gamma(\nu_i)} (1-N)^{-(\nu_1 + \ldots + \nu_{n-1} + 1)} b_1^{\nu_1 - 1} \ldots b_{n-1}^{\nu_{n-1}-1} \\ &\quad \times \int_0^\infty z^{\nu_1 + \ldots + \nu_n - 1} e^{-\frac{z}{1-N}} dz \\ &= \frac{\Gamma(\nu_1 + \ldots + \nu_n)}{\prod_{i=1}^{n} \Gamma(\nu_i)} b_i^{\nu_1 - 1} \ldots b_{n-1}^{\nu_{n-1}-1} \left(1 - \sum_{j=1}^{n-1} b_j\right)^{\nu_n - 1} . \end{aligned}$$

And clearly $\{b_1, \ldots, b_{n-1}\}$ belongs to the simplex $\mathbf{D}'$. □

Based on this theorem, it is shown that the Dirichlet variates $Dir(\nu_1, \ldots, \nu_{n-1}; \nu_n)$ can be generated efficiently, by Arnason et al [1, 2, 7]. In order to implement this, one should generate the gamma variates $G(\nu)$. For various parameter ν, there are well-established methods in the literature, and refer to Gentle for details [7].

We summarize the method for generating the general Dirichlet variates as follows:

Algorithm 2 (Dirichlet Sampling)

1. *Generate the independent gamma variates* $G(\nu_i)$ $(i = 1, \ldots, n)$ *as* $\{x_1^{(k)}, \ldots x_n^{(k)}\}$.

2. *Transform these gamma variates to Dirichlet variates as follows:*

$$\begin{aligned} b_i &= \tfrac{x_i}{\sum_{j=1}^{n} x_j} \ (i = 1, \ldots, n-1) \ , \\ b_n &= 1 - \textstyle\sum_{i=1}^{n-1} b_i \ . \end{aligned} \tag{15}$$

As a special case, we obtain the second method for uniform samplings:

Algorithm 3 (Uniform Sampling 2)

1. *Generate* n *mutually independent and identical exponential variates* $G(1)$ *as* $\{x_1^{(k)}, \ldots x_n^{(k)}\}$.

2. *Transform these exponential variates so as to be uniformly distributed on the simplex* $\mathbf{D}$*:*

$$\begin{aligned} b_i &= \tfrac{x_i}{\sum_{j=1}^{n} x_j} \ (i = 1, \ldots, n-1) \ , \\ b_n &= 1 - \textstyle\sum_{i=1}^{n-1} b_i \ . \end{aligned} \tag{16}$$

3.3 Numerical Examples

We proceed to implement some numerical examples. First we show visually how our Algorithms 1 and 3 generates constant portfolios on the simplex and also verify whether they are uniformly distribued as claimed by Theorem 1 and 2. Concerning the uniform variate generator, we adopt two mothods; the *ran1* of Press et al [14] and the *Mersenne-Twister* of Matsumoto-Nishimura [13]. Hereafter R is the abbreviation for the ran1 and M is the one for the Mersenne-Twister. Also two methods for transforming uniform variates into constant portfolios are tested; namely Algorithm 1 and 3 which are abbreviated by U and D respectively. Hence we tested $2 \times 2 = 4$ sampling methods.

To grasp our sampling methods visually, we show the results with $n = 3$ assets. First we show in Figure 1 how the uniform sampling of constant portfolios progresses on the simplex, for the case of $R \times U$.

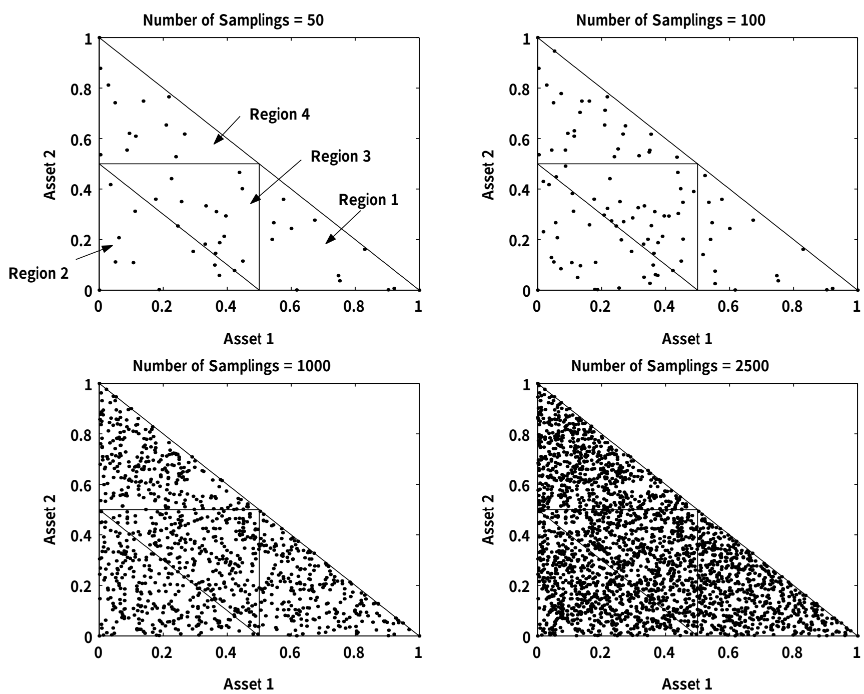

Figure 1: The progress of the uniform sampling: The case of $n = 3$.

Next we statistically verify whether the constant portfolio samplings are uniformly distributed, as claimed by Theorem 1 and 2. As shown by Figure 1, we divide the simplex **D** into four regions, and implement the test for goodness of fit with the hypothesis if the frequencies of the samplings from each of four regions are equal. We show the result in Table 1 and Figure 2. The result supports the hypothesis for all the four methods against the

0.05 level of significance. This confirms any of four methods can be safely used for uniform samplings.

# Generation	U-R	U-M	D-R	D-M
50	1.68	0.08	5.52	2.00
100	0.32	3.44	5.04	1.52
250	0.98	2.06	2.45	1.97
500	4.66	0.98	1.30	2.32
750	3.46	1.29	0.30	4.72
1000	5.34	1.64	1.00	1.99
2500	0.57	0.17	1.53	0.49
5000	0.82	1.84	0.50	1.34
7500	0.47	1.73	0.96	1.79
10000	1.60	1.13	1.55	0.83

Table 1: The test for goodness of fit, where $\chi^2(3; 0.05) = 7.81$.

We proceed to briefly mention the Dirichlet sampling. For $n = 3$ assets, the p.d.f. of $Dir(1/2, 1/2; 1/2)$ has the shape of Figure 3. Hence the Dirichlet sampling is concentrated toward the boundary of $\mathbf{D}$ as shown in Figure 4. Intuitively, if BCRP lies near the boundary of $\mathbf{D}$, the Dirichlet sampling is expected to effectively sample the constant portfolios which will achieve high growth. This intuition is formally validated by Cover-Ordentlich [6]. The upper bound of the gap in the mean growth rate between BCRP $\boldsymbol{b}^*$ and the universal portfolio $\hat{\boldsymbol{b}}_\bullet$ is given as follows:

$$\frac{1}{t} \log \frac{V_t(\boldsymbol{b}^*)}{V_t(\hat{\boldsymbol{b}}_\bullet)} \leq \begin{cases} (n-1) \cdot \dfrac{\log(t+1)}{t} & \text{(Uniform } Dir(1, \ldots, 1; 1) \text{ sampling)} , \\ (n-1) \cdot \dfrac{\log\{2(t+1)\}}{2t} & \text{(Dirichlet } Dir(1/2, \ldots, 1/2; 1/2) \text{ sampling)} . \end{cases} \tag{17}$$

Hence the universal portfolio with the Dirichlet $Dir(1/2, \ldots, 1/2; 1/2)$ sampling is superior in the convergence to BCRP.

3.4 Sampling Termination Criteria

In the foregoing three subsections, we provide several methods for generating constant portfolios. The remaining numerical procedure is to obtain the universal portfolio by Eq. (9) with generated constant portfolios. As mentioned earlier, the theoretical background for the approximation Eq. (9) is the law of large numbers. Thus Eq. (9) is expected to converge the universal portfolio a.s., as the number of samplings, K, becomes large enough. Also it

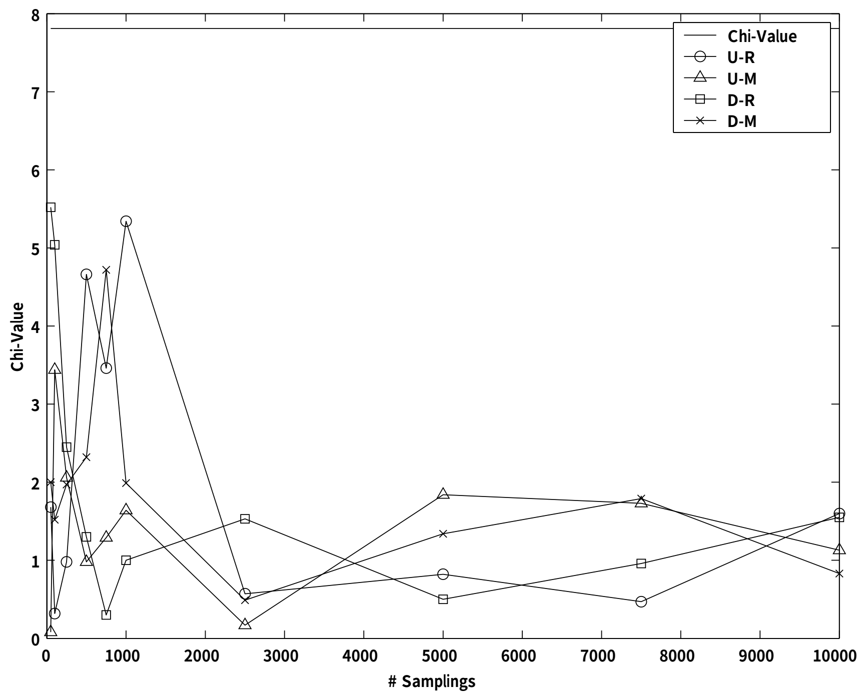

Figure 2: Visualization of the test for goodness of fit.

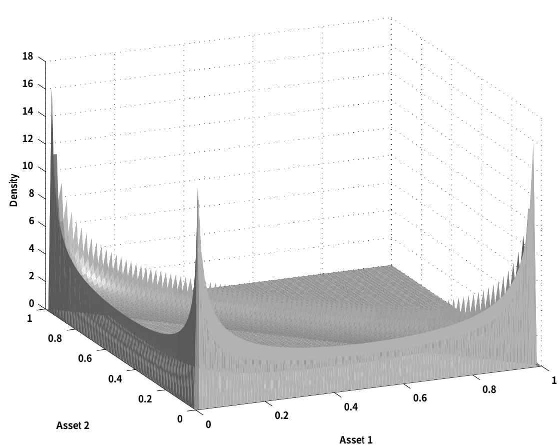

Figure 3: The p.d.f. of $Dir(1/2, 1/2; 1/2)$.

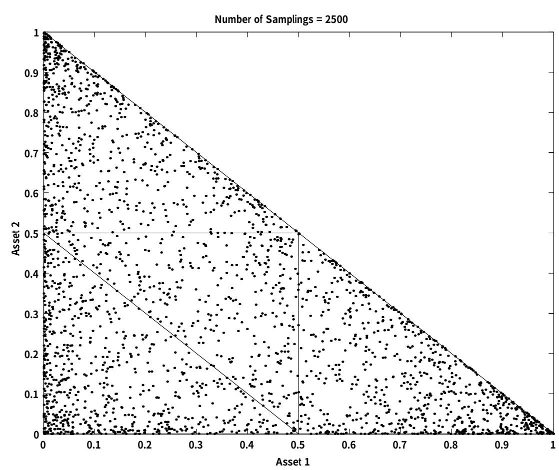

Figure 4: The $Dir(1/2, 1/2; 1/2)$ sampling.

is well-known that the Monte Carlo methods decreases the standard deviation (error) of the approximation Eq. (9) with the order $O(\frac{1}{\sqrt{K}})$. We cannot, however, determine theoretically the number of samplings within the admissible error, since it is difficult to evaluate Eq. (8) analytically. This is why the numerical methods for universal portfolios are needed. Of course, if we assume some suitable processes for asset prices and relax the constraints on portfolio weights, we are able to derive the universal portfolio in closed-form [10]. Our aim here is to propose the numerical methods for universal portfolios which can be used for very general discrete processes of asset prices.

We propose a valid criterion to terminate samplings. The criterion is to use the time-average of the Kullback-Leibler information number (K-L), between two universal portfolios of the $(k+1)$-th and the k-th samplings. Samplings are terminated if K-L becomes small enough. Formally, this can be expressed as:

$$\overline{D\left(\hat{\boldsymbol{b}}_t^{(k+1)} \| \hat{\boldsymbol{b}}_t^{(k)}\right)} = \frac{1}{T-1} \sum_{t=2}^{T} \sum_{i=1}^{n} \hat{b}_{i,t}^{(k+1)} \log\left(\frac{\hat{b}_{i,t}^{(k+1)}}{\hat{b}_{i,t}^{(k)}}\right) < \epsilon \ , \tag{18}$$

where we set $\epsilon = 1.0 \times 10^{-8}$ in the following numerical experiments. K-L can be interpreted as the distance between $\hat{\boldsymbol{b}}_t^{(k+1)}$ and $\hat{\boldsymbol{b}}_t^{(k)}$, and exhibits the increase in information introduced by a new sampling. Hence K-L is a valid criterion to terminate samplings.

4 Numerical Experiments

In this section, we conduct two numerical experiments to show some empirical properties of our methods developed. The purpose of the first experiment is to verify whether the universal portfolio is numerically obtained and it converges to BCRP. The second experiment aims to show how much time it is required to obtain the universal portfolio numerically. The result is shown by varying the number of assets included in the portfolio.

4.1 Experiment 1

This experiment is designed as follows.

Asset prices:

We adopt three assets which are included in the portfolio. We denote the investment horizon to be T and define $h \triangleq \dfrac{T}{N}$. It is assumed that the rates of return of asset prices are generated by

$$\left(\operatorname{diag}\left(\boldsymbol{S}_t\right)\right)^{-1}\left(\boldsymbol{S}_{t+h} - \boldsymbol{S}_t\right) = \boldsymbol{\mu} h + \boldsymbol{\Sigma}\sqrt{h}\boldsymbol{Z}_t \ , \tag{19}$$

where $\boldsymbol{\mu}$ is the constant drift parameter, $\boldsymbol{\Sigma}$ is also the constant diffusion parameter, and $\boldsymbol{Z}_t$ is a random vector which elements follow a mutually independent standard normal distribution. Eq. (19) is a discrete expression of the following stochastic differential equation which is familiar in the financial economics literature:

$$(\mathrm{diag}\,(\boldsymbol{S}_t))^{-1}\, d\boldsymbol{S}_t = \boldsymbol{\mu} dt + \boldsymbol{\Sigma} d\mathbf{W}_t \ , \tag{20}$$

where $\mathbf{W}_t$ is an n-dimensional standard Brownian motion. We artificially generate the $25,000$ rates of return according to Eq. (19), with $h = \frac{1}{12}$.

Tracking objective:

We track that the universal portfolio is numerically obtained in the sense of the terminating criterion Eq. (18), as the sampling iteration progresses. In other words, we trace that numerical approximation $\hat{\boldsymbol{b}}_t^{(K)}$ converge to the true expectation $\hat{\boldsymbol{b}}_t$ of Eq. (8). Also we verify if the universal portfolio converges to BCRP as the number of the return observation increases. BCRP is set as follows in this this experiment. We consider the value process of the constant portfolio under Eq. (20)

$$dV_t(\boldsymbol{b}) = V_t(\boldsymbol{b}) \cdot \boldsymbol{b}' \left(\boldsymbol{\mu} dt + \boldsymbol{\Sigma} d\mathbf{W}_t \right) \ .$$

Then BCRP is set as the optimal solution $\boldsymbol{b}^*$ for the following problem:

$$\mathbf{P} \left| \begin{array}{ll} \text{maximize} & E\left[\log V_T(\boldsymbol{b})\right] = \left(\boldsymbol{b}'\boldsymbol{\mu} - \frac{1}{2}\boldsymbol{b}'\boldsymbol{\Sigma}\boldsymbol{\Sigma}'\boldsymbol{b} \right) T \\ \text{subjectto} & \boldsymbol{b} \in \mathbf{D} \ . \end{array} \right.$$

We remark that $\boldsymbol{b}^*$ for Problem $\mathbf{P}$ is the optimal portfolio among all the admissible portfolios [10]. We consider three cases for $\boldsymbol{b}^*$. In Case 1, BCRP coincides with an extreme point. In Case 2, BCRP is on the boundary of $\mathbf{D}$. In Case 3, BCRP is the center of gravity of $\mathbf{D}$. According to these BCRPs, we prepare drift $\boldsymbol{\mu}$ and diffusion $\boldsymbol{\Sigma}$ as in Table 2. Concerning the method for sampling constant portfolios, we test the uniform and the Dirichlet samplings. Thus we execute $3 \times 2 = 6$ experiments.

Result:

Figure 5, 6, and 7 show the progress to the convergence in the sense of Eq. (18) for each of three cases. Along the time-average of K-L, we also track the mean growth rate of Eq. (2). All cases achieve the convergence steadily within 15,000 samplings at most. Also the mean growth rate seems to converge into some value for all cases. Hence it can be said that our numerical methods for computing universal portfolios with the termination criterion of Eq. (18) guarantee the convergence. Besides the confirmation of the convergence, we empirically find that the Dirichlet sampling incurs more samples than the uniform one for all cases.

	BCRP $\boldsymbol{b}^*$	drift $\boldsymbol{\mu}$	diffusion $\boldsymbol{\Sigma}$
Case 1	$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0.2934 \\ 0.0245 \\ 0.1782 \end{pmatrix}$	$\begin{pmatrix} 0.4194 & -0.0279 & 0.1291 \\ -0.0279 & 0.2076 & -0.0145 \\ 0.1291 & -0.0145 & 0.3227 \end{pmatrix}$
Case 2	$\begin{pmatrix} 0.75 \\ 0.25 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0.24 \\ 0.20 \\ 0.15 \end{pmatrix}$	$\begin{pmatrix} 0.2968 & 0.0369 & 0.0234 \\ 0.0369 & 0.2608 & -0.0252 \\ 0.0234 & -0.0252 & 0.1697 \end{pmatrix}$
Case 3	$\begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix}$	$\begin{pmatrix} 0.17 \\ 0.17 \\ 0.17 \end{pmatrix}$	$\begin{pmatrix} 0.2 & 0 & 0 \\ 0 & 0.2 & 0 \\ 0 & 0 & 0.2 \end{pmatrix}$

Table 2: Three cases for numerical experiment 1.

Next we track the gap in the mean growth rate between BCRP and the universal portfolio vanishes asymptotically as claimed by Eq. (17). Figure 8 shows this gap in percentage for each case. The gap is gradually diminishing as the time progresses. After 20,000 days (80 years) of learning, the universal portfolio achieve the mean growth rate which is only a few percent less than the one of BCRP. Also we empirically find that the Dirichlet sampling is superior to the uniform one for the first two cases, but the former deteriorates its performance for Case 3. This result is interesting, since we know from Eq. (17) that the Dirichlet sampling provides better upper bound of the gap in the mean growth rate. In actual, however, there are some cases in which the Dirichlet sampling does not provide the less gap than the uniform sampling. Thus this result implies that the suitable sampling method should be carefully adopted according to the property of underlying asset price processes and BCRPs.

4.2 Experiment 2

Experiment 2 is designed as follows.

Asset prices:

We vary the number of assets included in the portfolio as 10, 20, 30, 40, and 50. For each universe, it is assumed that asset price processes follow Eq. (19) as in Experiment 1. The drift $\boldsymbol{\mu}$ and diffusion $\boldsymbol{\Sigma}$ used are the estimated values from daily stock prices which are randomly selected from the Nikkei 225 composing stocks, by setting $h = 1/250$. These data is obtained from Nikkei Quick AMSUS and has the length of 368 days, from Oct. 1998 to Mar. 2000. According to Eq. (19), we artificially generate 25,000 rates of return as in Experiment 1.

Tracking objective:

Helmbold et al report that the numerical method for computing universal portfolios pro-

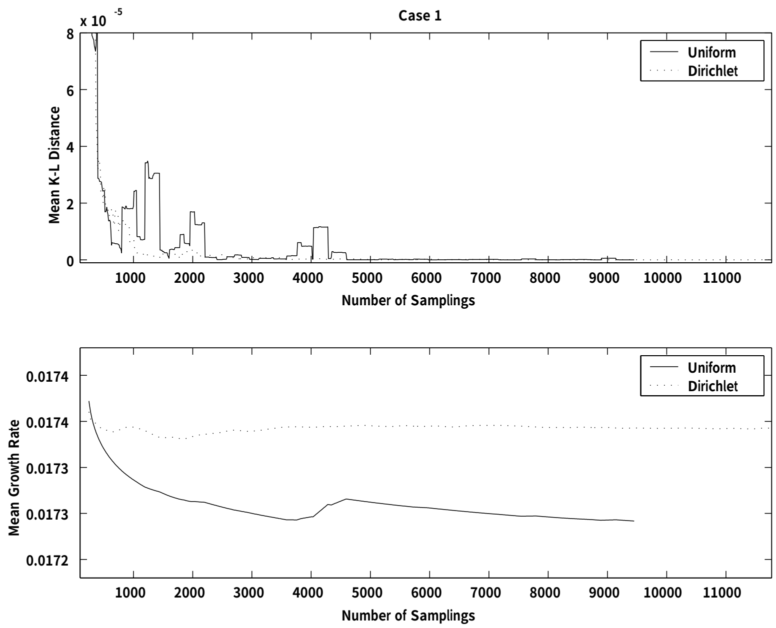

Figure 5: Case 1: Progress to the convergence.

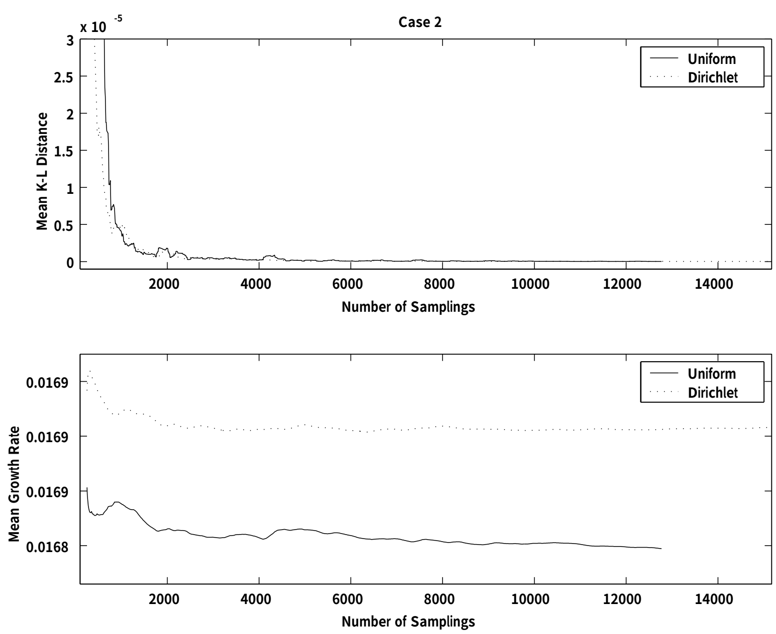

Figure 6: Case 2: Progress to the convergence.

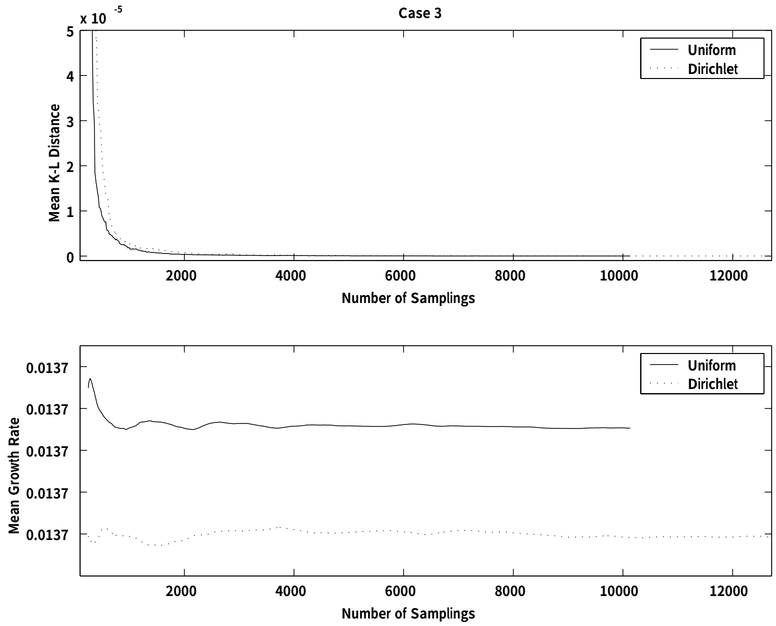

Figure 7: Case 3: Progress to the convergence.

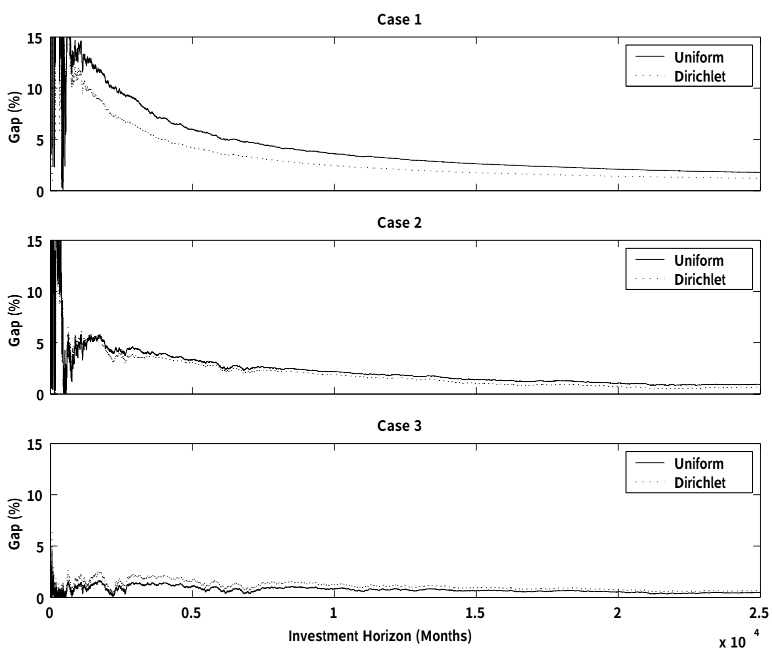

Figure 8: The gap in the growth rates.

posed by Cover [5] is intractable beyond nine assets [8]. The Cover's method is summarized as follows; If one is to sample constant portfolios with precision $1/m$, the number of samplings is equivalent to the number of cases of how to allocate m indiscriminate balls with the size of $1/m$ (portfolio weights) into n discriminated boxes (assets) with the redundancy allowed. Thus the number of samplings is $\frac{(m+n-1)!}{m!(n-1)!}$. For the case of $n = 50$ assets with $m = 10$ precision, the number of samplings is 6.28×10^{10}. Hence this method is hopeless in computing the larger sized problem.

On the other hand, our method can expect that $\hat{\boldsymbol{b}}_t^{(K)}$ of Eq. (9) converges to $\hat{\boldsymbol{b}}_t$ of Eq. (8) at the speed of $O(\frac{1}{\sqrt{K}})$. That is, the speed of convergence does not depend on the universe of assets as the Cover's method. Hence our method is expected to be efficient. Based on the above discussion, we track the elapsed time to compute the universal portfolios for each of five universes, using the uniform sampling ($U \times M$).

Result:

As a preliminary experiment, we track the number of samplings which attains the convergence defined by Eq. (18). We verify that 26,000 samplings are enough for all the universe. With this number of samplings held fixed, we track the elapsed time to compute the universal portfolios in the entire investment horizon of 25,000 days. We iterate this procedure for each universe. All the experiments (including Experiment 1) are executed on a normal PC machine running at 866 MHz with Microsoft Windows Me. The result is shown in Figure 9. As we expected, the required time to compute the universal portfolios does not diverge as the universe increases. Seeing this result, the required time increases linearly and it is about 2.86 hours to compute the entire universal portfolios composed from 50 assets.

4.3 Conclusion and the Direction of the Future Research

We have developed several methods for computing universal portfolios and conducted numerical experiments to show its possibility for the practical use. The empirical findings through these experiments are summarized as follows; First, these methods are not time-consuming even if we increase the universe of assets. Second, one should carefully adopt the valid sampling method according to the property of asset price processes. As far as our knowledge, our work is the first trial to treat the middle-sized problem for computing universal portfolios. In this sense, we have contributed to the research on the universal portfolio selection to some extent.

This paper leaves, however, two open problems for the practical use. First, our numerical methods for computing universal portfolios incur about 80 years to learn BCRP. This time-consumption may be rather unrealistic for the practical use. Hence we should develop some methods to accelerate the convergence to BCRP. Second, we should modify our methods

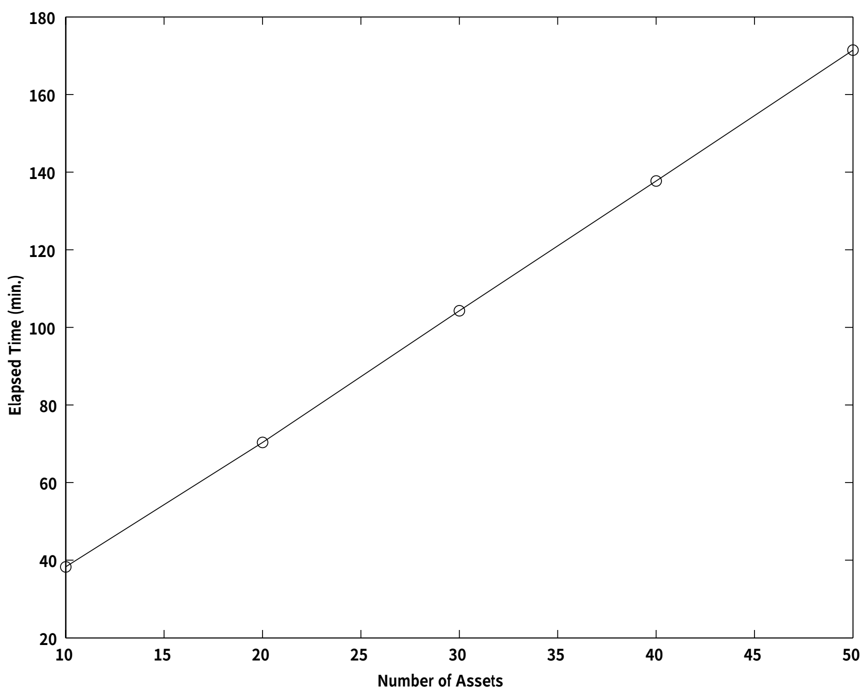

Figure 9: Elapsed time in minutes

developed to implement the large-scaled universal portfolio selection.

Acknowledgment

I would like to thank Hiroshi Shirakawa at Tokyo Institute of Technology and all the participants at the 24th Nihon Keiei Zaimu Gakkai.

References

[1] Arnason, A.N. (1972), "Simple, exact, efficient methods for generating Beta and Dirichlet variates" *Utilitas Mathematica*, 1, 249–290.

[2] Arnason, A.N. and L. Baniuk (1978), "A computer generation of Dirichlet variates", *Proceedings of the Eighth Manitoba Conference on Numerical Mathematics and Computing*, Utilitas Mathematica Publishing, Winnipeg, 97–105.

[3] Blum, A. and A. Kalai (1999), "Universal Portfolios With and Without Transaction Costs" *Machine Learning*, 35(3), 193-205.

[4] Cover, T.M. and J.A. Thomas (1991), *Elements of Information Theory*, John Wiley & Sons.

[5] Cover, T.M. (1991), "Universal Portfolios" *Mathematical Finance*, 1(1), 1-29.

[6] Cover, T.M. and E. Ordentlich (1996), "Universal Portfolios with Side Information" *IEEE Transactions on Information Theory*, 42(2), 348-363.

[7] Gentle, J.E. (1998), *Random Number Generation and Monte Carlo Methods*, Springer.

[8] Helmbold, D.P., R.E. Schapire, Y. Singer, and M.K. Warmuth (1998), "On-line portfolio selection using multiplicative updates" *Mathematical Finance*, 8(4), 325–347.

[9] Ishijima, H. (1999), "The Dynamic Portfolio Management under Incomplete Information" Doctoral Dissertation, Tokyo Institute of Technology.

[10] Ishijima, H. (2001), "Bayesian Interpretation of Continuous-time Universal Portfolios" unpublished manuscript.

[11] Jamshidian, F. (1992), "Asymptotically Optimal Portfolios" *Mathematical Finance*, 2(2), 131-150.

[12] Kalai, A. and S. Vempala (2000), "Efficient algorithms for universal portfolios" *Proceedings of the 41st Annual Symposium on the Foundations of Computer Science*.

[13] Matsumoto, M. and T. Nishimura (1998), "Mersenne Twister: A 623-dimensionally equidistributed uniform pseudorandom number generator", *ACM Trans. on Modeling and Computer Simulation*, 8(1), 3–30.

[14] Press, W.H., S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery (1992), *Numerical Recipes in C, Second Edition*, Cambridge University Press.

[15] Takeuchi, K. (1963), *Mathematical Statistics*, Tōyoukeizaishinpou-sya.

[16] Wilks, S.S. (1962), *Mathematical Statistics*, John Wiley & Sons.