Empirical portfolio selection with transaction cost

László Györfi and István Vajda

Budapest University of Technology and Economics

Department of Computer Science and Information Theory

e-mail: gyorfi@szit.bme.hu and istvajda@yahoo.com

www.szit.bme.hu/~gyorfi

Portfolio selection

investment in the stock market

return vector $x = (x^{(1)}, \dots x^{(d)})$

j-th component $x^{(j)}$ is the factor by which capital invested in stock j grows during the market period

a portfolio vector $b = (b^{(1)}, \dots b^{(d)})$

j-th component $b^{(j)}$ of which gives the proportion of the investor's capital invested in stock j

S_0 denotes the initial capital

$$S_1 = S_0 \sum_{j=1}^{d} b^{(j)} x^{(j)} = S_0 (b, x)$$

long run investment, initial capital S_0

x_i the return vector on day i, $b = b_1$ is the portfolio vector for the first day

$$S_1 = S_0 \cdot (b_1, x_1)$$

for the second day, S_1 new initial capital, the portfolio vector $b_2 = b(x_1)$

$$S_2 = S_0 \cdot (b, x_1) \cdot (b(x_1), x_2).$$

nth day a portfolio strategy $b_n = b(x_1^{n-1})$

$$S_n = S_0 \prod_{i=1}^{n} (b(x_1^{i-1}), x_i) = S_0 e^{nW_n(B)}$$

with the average growth rate

$$W_n(B) = \frac{1}{n} \sum_{i=1}^{n} \log(b(x_1^{i-1}), x_i).$$

log-optimum portfolio

$X_1, X_2, \ldots$ drawn from the vector valued stationary and ergodic process

log-optimum portfolio $B^* = \{b^*(\cdot)\}$

$$\mathbf{E}\{\log\langle b^*(X_1^{n-1}), X_n\rangle \mid X_1^{n-1}\} = \max_{b(\cdot)} \mathbf{E}\{\log\langle b(X_1^{n-1}), X_n\rangle \mid X_1^{n-1}\}.$$

If $S_n^* = S_n(B^*)$ denotes the capital after day n achieved by a log-optimum portfolio strategy B^*, then for any portfolio strategy B with capital $S_n = S_n(B)$ and for any stationary ergodic process $\{X_n\}_{-\infty}^{\infty}$,

$$\limsup_{n\to\infty} \frac{1}{n} \log \frac{S_n}{S_n^*} \leq 0 \quad \text{almost surely}$$

and

$$\lim_{n\to\infty} \frac{1}{n} \log S_n^* = W^* \quad \text{almost surely,}$$

where

$$W^* = \mathbf{E}\left\{ \max_{b(\cdot)} \mathbf{E}\{\log(b(X_{-\infty}^{-1}), X_0)) \mid X_{-\infty}^{-1}\} \right\}$$

is the maximal growth rate of any portfolio.

Elementary portfolio

$H^{(k,\ell)} = \{h^{(k,\ell)}(\cdot)\}$, $k, \ell = 1, 2, \ldots$

$\mathcal{P}_\ell = \{A_{\ell,j}, j = 1, 2, \ldots, m_\ell\}$ finite partitions of $\mathbb{R}^d$,

G_ℓ be the corresponding quantizer:

$$G_\ell(x) = j, \text{ if } x \in A_{\ell,j}.$$

$x_1^n \in \mathbb{R}^{dn}$, $G_\ell(x_1^n)$

Fix k, ℓ

for each k-long string s of positive integers, define the elementary portfolio $h^{(k,\ell)}$ by

$$h^{(k,\ell)}(x_1^{n-1}) = \arg\max_b \prod_{\{k<i<n: G_\ell(x_{i-k}^{i-1}) = G_\ell(x_{n-k}^{n-1})\}} (b, x_i)$$

That is, $h_n^{(k,\ell)}$ quantizes the sequence x_1^{n-1} according to the partition $\mathcal{P}_\ell$, and browses through all past appearances of the last seen quantized string $G_\ell(x_{n-k}^{n-1})$ of length k. Then it designs a fixed portfolio vector according to the returns on the days following the occurence of the string.

Combining elementary portfolios

Finally, let $\{q_{k,\ell}\}$ be a probability distribution on the set of all pairs (k,ℓ) of positive integers such that for all k,ℓ, $q_{k,\ell} > 0$. The strategy B then arises from weighing the elementary portfolio stratgies $H^{(k,\ell)}$ according to their past performances and $\{q_{k,\ell}\}$ such that, after day n, the investor's capital becomes

$$S_n(B) = \sum_{k,\ell} q_{k,\ell} S_n(H^{(k,\ell)}).$$

Theorem on universal consistency

Assume that

(a) the sequence of partitions is nested, that is, any cell of $\mathcal{P}_{\ell+1}$ is a subset of a cell of $\mathcal{P}_\ell$, $\ell = 1, 2, \ldots$;

(b) if $\operatorname{diam}(A) = \sup_{x,y\in A} \|x - y\|$ denotes the diameter of a set, then for any sphere S centered at the origin

$$\lim_{\ell\to\infty} \max_{j:A_{\ell,j}\cap S\neq\emptyset} \operatorname{diam}(A_{\ell,j}) = 0 \ .$$

If $\{X_n\}_{-\infty}^{\infty}$ is a stationary and ergodic process such that

$\mathbf{E}\{|\log X^{(j)}| < \infty$, for $j = 1, 2, \ldots, d$

then

$$\lim_{n\to\infty} \frac{1}{n} \log S_n(B) = W^* \quad \text{almost surely.}$$

L. Györfi, D. Schäfer (2003) "Nonparametric prediction", in *Advances in Learning Theory: Methods, Models and Applications*, J. A. K. Suykens, G. Horváth, S. Basu, C. Micchelli, J. Vandevalle (Eds.), IOS Press, NATO Science Series, pp. 341-356.
www.szit.bme.hu/∼gyorfi/histog.ps

L. Györfi, G. Lugosi, F. Udina (2005) "Nonparametric kernel-based sequential investment strategies", *Mathematical Finance*, .., pp. ...-...
www.szit.bme.hu/∼gyorfi/kernel.ps

Experiments on NYSE data

Transaction cost

S_n: the wealth at the close of market day n

N_n: the net wealth at the close of market day n

$$S_n = N_{n-1} \langle \mathbf{b}_n \, , \, \mathbf{x}_n \rangle \, ,$$

The transaction cost will be paid only for selling an asset.

$$N_n = S_n - c \sum_{j=1}^{d} \left(b_{n,j} x_{n,j} N_{n-1} - b_{n+1,j} N_n \right)^+ ,$$

Introducing ratio

$$w_n = \frac{N_n}{S_n},$$

we get

$$1 = w_n + \sum_{j=1}^{d} c \left(\frac{b_{n,j} x_{n,j}}{\langle \mathbf{b}_n \, , \, \mathbf{x}_n \rangle} - b_{n+1,j} w_n \right)^+ .$$

Then

$$w_n = w(\mathbf{b}_n, \mathbf{b}_{n+1}, \mathbf{x}_n),$$

for some function w. For $S_0 = 1$,

$$S_n = N_{n-1} \langle \mathbf{b}_n \, , \, \mathbf{x}_n \rangle = S_{n-1} w_{n-1} \langle \mathbf{b}_n \, , \, \mathbf{x}_n \rangle = \prod_{i=1}^{n} w_{i-1} \langle \mathbf{b}_i \, , \, \mathbf{x}_i \rangle .$$

Introduce the notation

$$h(\mathbf{b}_{i-1}, \mathbf{b}_i, \mathbf{x}_{i-1}, \mathbf{x}_i) = \log(w(\mathbf{b}_{i-1}, \mathbf{b}_i, \mathbf{x}_{i-1}) \langle \mathbf{b}_i \, , \, \mathbf{x}_i \rangle),$$

then the average growth rate becomes

$$\frac{1}{n} \log S_n \quad = \quad \frac{1}{n} \sum_{i=1}^{n} h(\mathbf{b}_{i-1}, \mathbf{b}_i, \mathbf{x}_{i-1}, \mathbf{x}_i).$$

Our aim is to maximize the average growth rate.

In the sequel $\mathbf{x}_i$ will be random variable and is denoted by $\mathbf{X}_i$.

We use the decomposition

$$\frac{1}{n}\log S_n = I_n + J_n,$$

where

$$I_n = \frac{1}{n}\sum_{i=1}^{n}(h(\mathbf{b}_{i-1},\mathbf{b}_i,\mathbf{X}_{i-1},\mathbf{X}_i) - \mathbf{E}\{h(\mathbf{b}_{i-1},\mathbf{b}_i,\mathbf{X}_{i-1},\mathbf{X}_i)|\mathbf{X}_1,\dots,\mathbf{X}_{i-1}\})$$

and

$$J_n = \frac{1}{n}\sum_{i=1}^{n}\mathbf{E}\{h(\mathbf{b}_{i-1},\mathbf{b}_i,\mathbf{X}_{i-1},\mathbf{X}_i)|\mathbf{X}_1,\dots,\mathbf{X}_{i-1}\}.$$

I_n is an average of martingale differences, which converges to 0 almost surely. Therefore the maximization of the average growth rate is asymptotically equivalent to the maximization of J_n.

Consider a recursive portfolio selection:

$$\mathbf{b}_i = \mathbf{b}_i(\mathbf{x}_1^{i-1}) = \mathbf{b}_i(\mathbf{b}_{i-1}, \mathbf{x}_{i-1})$$

and suppose that $\mathbf{X}_i$, $i = 1, 2\dots$, is a first order Markov process.

$$J_n = \frac{1}{n} \sum_{i=1}^{n} v(\mathbf{b}_{i-1}, \mathbf{b}_i, \mathbf{X}_{i-1}).$$

Optimization: Bellmann, complicated

Schäfer, D. (2002) Nonparametric Estimation for Financial Investment under Log-Utility. PhD Dissertation, Mathematical Institute, University Stuttgart, Shaker Verlag, Aachen. www.szit.bme.hu/~gyorfi/dominik.ps

Non-Bellmann strategy

$$J_n = \frac{1}{n} \sum_{i=1}^{n} v(\mathbf{b}_{i-1}, \mathbf{b}_i, \mathbf{X}_{i-1}).$$

Recursive non-Bellman algorithm with known theoretical distribution

$$\mathbf{b}_i^* = \arg\max_{\mathbf{b}} v(\mathbf{b}_{i-1}^*, \mathbf{b}, \mathbf{X}_{i-1})$$

Empirical non-Bellmann strategy

Non-Bellmann algorithm for data dependent learning

At day n, we are given $\mathbf{b}_{n-1}$ and $\mathbf{X}_{n-1}$ and let $\tilde{\mathbf{X}}_1, \ldots, \tilde{\mathbf{X}}_N$ be the return vectors following the matches (elementary portfolios).

Then

$$\mathbf{b}_n = \arg\max_{\mathbf{b}} \frac{1}{N} \sum_{i=1}^{N} h(\mathbf{b}_{n-1}, \mathbf{b}, \mathbf{X}_{n-1}, \tilde{\mathbf{X}}_i)$$

Empirical Bellmann strategy

Iteration for the value function $F(\mathbf{b},\mathbf{x})$

At day n, we are given $\mathbf{b}_{n-1}$ and $\mathbf{X}_{n-1}$ and let $\tilde{\mathbf{X}}_1,\ldots,\tilde{\mathbf{X}}_N$ be the return vectors following the matches. Then:

$$F_m(\mathbf{b},\mathbf{x}) \quad = \quad \max_{\mathbf{b'}} \frac{1}{N}\sum_{i=1}^{N}\Big(h(\mathbf{b},\mathbf{b'},\mathbf{x},\tilde{\mathbf{X}}_i)+F_{m-1}(\mathbf{b'},\tilde{\mathbf{X}}_i)\Big)$$

Then we have the convergences

$$F_m(\mathbf{b},\mathbf{x}) \to F(\mathbf{b},\mathbf{x})$$

$$\mathbf{b}_n \quad = \quad \arg\max_{\mathbf{b'}} \frac{1}{N}\sum_{i=1}^{N}\Big(h(\mathbf{b}_{n-1},\mathbf{b'},\mathbf{X}_{n-1},\tilde{\mathbf{X}}_i)+F(\mathbf{b'},\tilde{\mathbf{X}}_i)\Big)$$

Cost subtract

Algorithm without any optimization

with known theoretical distribution

$$\mathbf{b}_n^* = \arg\max_{\mathbf{b}} \mathbf{E} \log \left\langle \mathbf{b}\, ,\, \tilde{\mathbf{X}}_i \right\rangle$$

in each step subtract the transaction cost

Empirical cost subtract

data dependent version

At day n, let $\tilde{\mathbf{X}}_1, \ldots, \tilde{\mathbf{X}}_N$ be the return vectors following by the matches.

$$\mathbf{b}_n^* = \arg\max_{\mathbf{b}} \frac{1}{N} \sum_{i=1}^{N} \log(\mathbf{b}, \tilde{\mathbf{X}}_i)$$

in each step subtract the transaction cost

Experiments

i.) Cover example: the portfolio consists of cash and one stock. The market vector on day i becomes $\mathbf{x}_i = (1, y_i)$, where y_i is the return on stock. The stock returns can take only two values, 2 and $\frac{1}{2}$ with equal probability.
memoryless model on the stock: independent, identically distributed

ii.) bull market model on the stock: Markov-chain

$$\begin{pmatrix} P(2 \to 2) = 0.55 & P(2 \to 1/2) = 0.45 \\ P(1/2 \to 2) = 0.5 & P(1/2 \to 1/2) = 0.5 \end{pmatrix},$$

iii.) bear market model on the stock: Markov-chain

$$\begin{pmatrix} P(2 \to 2) = 0.45 & P(2 \to 1/2) = 0.55 \\ P(1/2 \to 2) = 0.5 & P(1/2 \to 1/2) = 0.5 \end{pmatrix}.$$

iv.) symmetrical dependence model on the stock: Markov-chain

$$\begin{pmatrix} P(2 \to 2) = 0.45 & P(2 \to 1/2) = 0.55 \\ P(1/2 \to 2) = 0.55 & P(1/2 \to 1/2) = 0.45 \end{pmatrix}.$$

market type	cost (%)	Bellmann (%)	non-Bellmann (%)	cost-subtract (%)
bear	0	4.479	4.479	4.479
	1	4.359	4.359	4.049
	2.5	4.184	4.182	3.320
	5	3.938	3.893	1.986
i.i.d	0	5.889	5.889	5.889
	1	5.805	5.805	5.388
	2.5	5.687	5.685	4.546
	5	5.535	5.496	3.012
bull	0	7.990	7.990	7.990
	1	7.945	7.945	7.408
	2.5	7.877	7.866	6.436
	5	7.770	7.704	4.602
symmetric	0	6.427	6.427	6.427
	1	6.269	6.269	5.924
	2.5	6.034	6.034	5.080
	5	5.684	5.646	3.546

Future work

- Experiments on NYSE data
- Bound the difference between the performances of Bellmann and non-Bellmann strategies
- Search for good empirical portfolio strategies