Empirical portfolio selections

László (Laci) Györfi[1]

[1]Department of Computer Science and Information Theory
Budapest University of Technology and Economics
Budapest, Hungary

September 1, 2013

e-mail: gyorfi@cs.bme.hu
www.szit.bme.hu/~gyorfi

Portfolio selections

- static portfolio selection (single period)

Portfolio selections

- static portfolio selection (single period)
- constantly rebalanced portfolio

Portfolio selections

- static portfolio selection (single period)
- constantly rebalanced portfolio
- general rebalancing

Portfolio selections

- static portfolio selection (single period)
- constantly rebalanced portfolio
- general rebalancing

- rebalancing

Portfolio selections

- static portfolio selection (single period)
- constantly rebalanced portfolio
- general rebalancing

- rebalancing
- multi-asset, multi-period

Portfolio selections

- static portfolio selection (single period)
- constantly rebalanced portfolio
- general rebalancing

- rebalancing
- multi-asset, multi-period
- empirical (nonparametric statistics, machine learning)

Growth rate

investment in the stock market

d assets

Growth rate

investment in the stock market
d assets
$S_n^{(j)}$ price of asset j at the end of trading period (day) n,
$j = 1, \ldots, d$

Growth rate

investment in the stock market

d assets

$S_n^{(j)}$ price of asset j at the end of trading period (day) n,

$j = 1, \dots, d$

$S_0^{(j)} = 1$

Growth rate

investment in the stock market

d assets

$S_n^{(j)}$ price of asset j at the end of trading period (day) n,

$j = 1, \ldots, d$

$S_0^{(j)} = 1$

$$S_n^{(j)} = e^{nW_n^{(j)}} \approx e^{nW^{(j)}}$$

Growth rate

investment in the stock market

d assets

$S_n^{(j)}$ price of asset j at the end of trading period (day) n,

$j = 1, \dots, d$

$S_0^{(j)} = 1$

$$S_n^{(j)} = e^{nW_n^{(j)}} \approx e^{nW^{(j)}}$$

asymptotic growth rate

$$W^{(j)} = \lim_{n\to\infty} W_n^{(j)} = \lim_{n\to\infty} \frac{1}{n} \ln S_n^{(j)}$$

Static portfolio selection: single period investment

the aim is to achieve $\max_j W^{(j)}$

Static portfolio selection: single period investment

the aim is to achieve $\max_j W^{(j)}$
static portfolio selection

Static portfolio selection: single period investment

the aim is to achieve $\max_j W^{(j)}$

static portfolio selection

a portfolio vector $\mathbf{b} = (b^{(1)}, \dots b^{(d)})$, $b^{(j)} \geq 0$, $\sum_j b^{(j)} = 1$

Static portfolio selection: single period investment

the aim is to achieve $\max_j W^{(j)}$
static portfolio selection
a portfolio vector $\mathbf{b} = (b^{(1)}, \dots b^{(d)})$, $b^{(j)} \geq 0$, $\sum_j b^{(j)} = 1$
initial capital $S_0 = 1$

Static portfolio selection: single period investment

the aim is to achieve $\max_j W^{(j)}$
static portfolio selection
a portfolio vector $\mathbf{b} = (b^{(1)}, \dots b^{(d)})$, $b^{(j)} \geq 0$, $\sum_j b^{(j)} = 1$
initial capital $S_0 = 1$

$$S_n = \sum_j b^{(j)} S_n^{(j)}$$

Static portfolio selection: single period investment

the aim is to achieve $\max_j W^{(j)}$
static portfolio selection
a portfolio vector $\mathbf{b} = (b^{(1)}, \ldots b^{(d)})$, $b^{(j)} \geq 0$, $\sum_j b^{(j)} = 1$
initial capital $S_0 = 1$

$$S_n = \sum_j b^{(j)} S_n^{(j)}$$

$$\max_j b^{(j)} S_n^{(j)} \leq S_n \leq d \max_j b^{(j)} S_n^{(j)}$$

Static portfolio selection: single period investment

the aim is to achieve $\max_j W^{(j)}$
static portfolio selection
a portfolio vector $\mathbf{b} = (b^{(1)}, \ldots b^{(d)})$, $b^{(j)} \geq 0$, $\sum_j b^{(j)} = 1$
initial capital $S_0 = 1$

$$S_n = \sum_j b^{(j)} S_n^{(j)}$$

$$\max_j b^{(j)} S_n^{(j)} \leq S_n \leq d \max_j b^{(j)} S_n^{(j)}$$

If $b^{(j)} > 0$ then

$$\lim_{n\to\infty} \frac{1}{n} \ln S_n = \lim_{n\to\infty} \max_j \frac{1}{n} \ln S_n^{(j)} = \max_j W^{(j)}$$

Static portfolio selection: single period investment

the aim is to achieve $\max_j W^{(j)}$
static portfolio selection
a portfolio vector $\mathbf{b} = (b^{(1)}, \dots b^{(d)})$, $b^{(j)} \geq 0$, $\sum_j b^{(j)} = 1$
initial capital $S_0 = 1$

$$S_n = \sum_j b^{(j)} S_n^{(j)}$$

$$\max_j b^{(j)} S_n^{(j)} \leq S_n \leq d \max_j b^{(j)} S_n^{(j)}$$

If $b^{(j)} > 0$ then

$$\lim_{n\to\infty} \frac{1}{n} \ln S_n = \lim_{n\to\infty} \max_j \frac{1}{n} \ln S_n^{(j)} = \max_j W^{(j)}$$

we can do much better, applying dynamic portfolio selection

Relative prices

$$x_i^{(j)} = \frac{S_i^{(j)}}{S_{i-1}^{(j)}}$$

Relative prices

$$x_i^{(j)} = \frac{S_i^{(j)}}{S_{i-1}^{(j)}}$$

$\mathbf{x}_i = (x_i^{(1)}, \dots x_i^{(d)})$ the return vector on day i

Relative prices

$$x_i^{(j)} = \frac{S_i^{(j)}}{S_{i-1}^{(j)}}$$

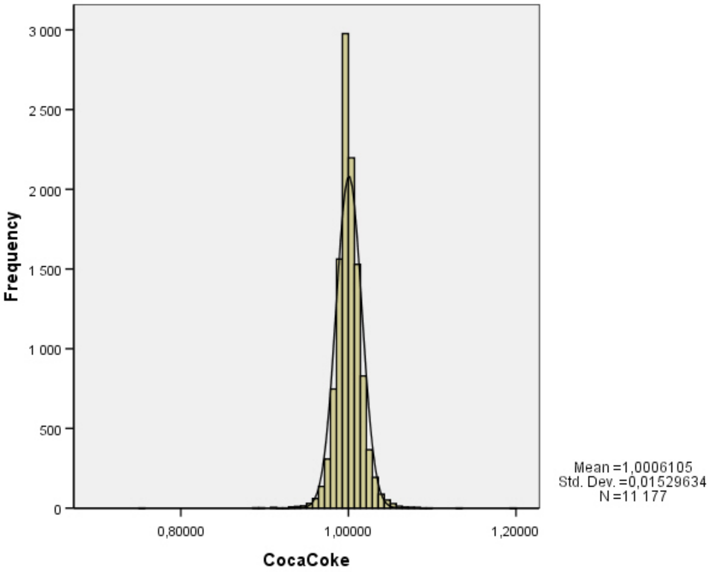
$\mathbf{x}_i = (x_i^{(1)}, \dots x_i^{(d)})$ the return vector on day i
$x_i^{(j)}$ is the factor by which capital invested in stock j grows during the market period i

Relative prices

$$x_i^{(j)} = \frac{S_i^{(j)}}{S_{i-1}^{(j)}}$$

$\mathbf{x}_i = (x_i^{(1)}, \dots x_i^{(d)})$ the return vector on day i
$x_i^{(j)}$ is the factor by which capital invested in stock j grows during the market period i
For NYSE data

$$0.7 \leq x_i^{(j)} \leq 1.2$$


3 000
2 500
2 000
1 500
1 000
500
0
Frequency
0,80000
1,00000
1,20000
CocaCoke
Mean =1,0006105
Std. Dev. =0,01529634
N =11 177

Dynamic portfolio selection: multi-period investment

multi-period investment

multi-period investment
Constantly Re-balanced Portfolio (CRP)

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

multi-period investment
Constantly Re-balanced Portfolio (CRP)

a portfolio vector $\mathbf{b} = (b^{(1)}, \ldots b^{(d)})$
$b^{(j)}$ gives the proportion of the investor's capital invested in stock j

Dynamic portfolio selection: multi-period investment

multi-period investment
Constantly Re-balanced Portfolio (CRP)

a portfolio vector $\mathbf{b} = (b^{(1)}, \dots b^{(d)})$
$b^{(j)}$ gives the proportion of the investor's capital invested in stock j
$\mathbf{b}$ is the constant portfolio vector for each trading day

Evolution

for the first day S_0 denotes the initial capital

$$S_1 = S_0 \sum_{j=1}^{d} b^{(j)} x_1^{(j)} = S_0 \langle \mathbf{b} \, , \, \mathbf{x}_1 \rangle$$

Evolution

for the first day S_0 denotes the initial capital

$$S_1 = S_0 \sum_{j=1}^{d} b^{(j)} x_1^{(j)} = S_0 \langle \mathbf{b} \, , \, \mathbf{x}_1 \rangle$$

for the second day, S_1 new initial capital

$$S_2 = S_1 \cdot \langle \mathbf{b} \, , \, \mathbf{x}_2 \rangle = S_0 \cdot \langle \mathbf{b} \, , \, \mathbf{x}_1 \rangle \cdot \langle \mathbf{b} \, , \, \mathbf{x}_2 \rangle \, .$$

Evolution

for the first day S_0 denotes the initial capital

$$S_1 = S_0 \sum_{j=1}^{d} b^{(j)} x_1^{(j)} = S_0 \langle \mathbf{b} \, , \, \mathbf{x}_1 \rangle$$

for the second day, S_1 new initial capital

$$S_2 = S_1 \cdot \langle \mathbf{b} \, , \, \mathbf{x}_2 \rangle = S_0 \cdot \langle \mathbf{b} \, , \, \mathbf{x}_1 \rangle \cdot \langle \mathbf{b} \, , \, \mathbf{x}_2 \rangle \, .$$

for the nth day:

$$S_n = S_{n-1} \langle \mathbf{b} \, , \, \mathbf{x}_n \rangle = S_0 \prod_{i=1}^{n} \langle \mathbf{b} \, , \, \mathbf{x}_i \rangle$$

Evolution

for the first day S_0 denotes the initial capital

$$S_1 = S_0 \sum_{j=1}^{d} b^{(j)} x_1^{(j)} = S_0 \langle \mathbf{b} \, , \, \mathbf{x}_1 \rangle$$

for the second day, S_1 new initial capital

$$S_2 = S_1 \cdot \langle \mathbf{b} \, , \, \mathbf{x}_2 \rangle = S_0 \cdot \langle \mathbf{b} \, , \, \mathbf{x}_1 \rangle \cdot \langle \mathbf{b} \, , \, \mathbf{x}_2 \rangle \, .$$

for the nth day:

$$S_n = S_{n-1} \langle \mathbf{b} \, , \, \mathbf{x}_n \rangle = S_0 \prod_{i=1}^{n} \langle \mathbf{b} \, , \, \mathbf{x}_i \rangle \; = S_0 e^{n W_n(\mathbf{b})}$$

Evolution

for the first day S_0 denotes the initial capital

$$S_1 = S_0 \sum_{j=1}^{d} b^{(j)} x_1^{(j)} = S_0 \langle \mathbf{b}\,,\,\mathbf{x}_1 \rangle$$

for the second day, S_1 new initial capital

$$S_2 = S_1 \cdot \langle \mathbf{b}\,,\,\mathbf{x}_2 \rangle = S_0 \cdot \langle \mathbf{b}\,,\,\mathbf{x}_1 \rangle \cdot \langle \mathbf{b}\,,\,\mathbf{x}_2 \rangle\,.$$

for the nth day:

$$S_n = S_{n-1} \langle \mathbf{b}\,,\,\mathbf{x}_n \rangle = S_0 \prod_{i=1}^{n} \langle \mathbf{b}\,,\,\mathbf{x}_i \rangle \; = S_0 e^{nW_n(\mathbf{b})}$$

with the average growth rate

$$W_n(\mathbf{b}) = \frac{1}{n} \sum_{i=1}^{n} \ln \langle \mathbf{b}\,,\,\mathbf{x}_i \rangle\,.$$

log-optimum portfolio

Special market process: $\mathbf{X}_1, \mathbf{X}_2, \dots$ is independent and identically distributed (i.i.d.)

log-optimum portfolio

Special market process: $\mathbf{X}_1, \mathbf{X}_2, \ldots$ is independent and identically distributed (i.i.d.)

log-optimum portfolio $\mathbf{b}^*$

$$\mathbf{E}\{\ln \langle \mathbf{b}^* , \mathbf{X}_1 \rangle\} = \max_{\mathbf{b}} \mathbf{E}\{\ln \langle \mathbf{b} , \mathbf{X}_1 \rangle\}$$

Special market process: $\mathbf{X}_1, \mathbf{X}_2, \ldots$ is independent and identically distributed (i.i.d.)

log-optimum portfolio $\mathbf{b}^*$

$$\mathbf{E}\{\ln \langle \mathbf{b}^* \, , \, \mathbf{X}_1 \rangle\} = \max_{\mathbf{b}} \mathbf{E}\{\ln \langle \mathbf{b} \, , \, \mathbf{X}_1 \rangle\}$$

Best Constantly Re-balanced Portfolio (BCRP)

log-optimum portfolio

Special market process: $\mathbf{X}_1, \mathbf{X}_2, \dots$ is independent and identically distributed (i.i.d.)

log-optimum portfolio $\mathbf{b}^*$

$$\mathbf{E}\{\ln \langle \mathbf{b}^* \, , \, \mathbf{X}_1 \rangle\} = \max_{\mathbf{b}} \mathbf{E}\{\ln \langle \mathbf{b} \, , \, \mathbf{X}_1 \rangle\}$$

Best Constantly Re-balanced Portfolio (BCRP)
universal portfolio

log-optimum portfolio

Special market process: $\mathbf{X}_1, \mathbf{X}_2, \dots$ is independent and identically distributed (i.i.d.)

log-optimum portfolio $\mathbf{b}^*$

$$\mathbf{E}\{\ln \langle \mathbf{b}^*, \mathbf{X}_1 \rangle\} = \max_{\mathbf{b}} \mathbf{E}\{\ln \langle \mathbf{b}, \mathbf{X}_1 \rangle\}$$

Best Constantly Re-balanced Portfolio (BCRP)
universal portfolio

for dependent market process we can do even better

History

gambling,

gambling, horse racing,

gambling, horse racing, information theory

History

gambling, horse racing, information theory

Algoet

History

gambling, horse racing, information theory

Algoet
Barron

History

gambling, horse racing, information theory

Algoet
Barron
Breiman

History

gambling, horse racing, information theory

Algoet
Barron
Breiman
Cover

History

gambling, horse racing, information theory

Algoet
Barron
Breiman
Cover
Kelly

History

gambling, horse racing, information theory

Algoet
Barron
Breiman
Cover
Kelly
Latané

Chapter 15 of D. G. Luenberger, *Investment Science*. Oxford University Press, 1998.

Chapter 15 of D. G. Luenberger, *Investment Science*. Oxford University Press, 1998.

"Volatility is NOT the same as Risk. Volatility is Opportunity"

Chapter 15 of D. G. Luenberger, *Investment Science*. Oxford University Press, 1998.

"Volatility is NOT the same as Risk. Volatility is Opportunity"

"Conclusions about multiperiod investment situations are not mere variations of single-period conclusions – rather they often *reverse* those earlier conclusions.

Chapter 15 of D. G. Luenberger, *Investment Science*. Oxford University Press, 1998.

"Volatility is NOT the same as Risk. Volatility is Opportunity"

"Conclusions about multiperiod investment situations are not mere variations of single-period conclusions – rather they often *reverse* those earlier conclusions. This makes the subject exiting, both intellectually and in practice.

Chapter 15 of D. G. Luenberger, *Investment Science*. Oxford University Press, 1998.
"Volatility is NOT the same as Risk. Volatility is Opportunity"
"Conclusions about multiperiod investment situations are not mere variations of single-period conclusions – rather they often *reverse* those earlier conclusions. This makes the subject exiting, both intellectually and in practice. Once the subtleties of multiperiod investment are understood, the reward in terms of enhanced investment performance can be substantial."

Chapter 15 of D. G. Luenberger, *Investment Science*. Oxford University Press, 1998.

"Volatility is NOT the same as Risk. Volatility is Opportunity"

"Conclusions about multiperiod investment situations are not mere variations of single-period conclusions – rather they often *reverse* those earlier conclusions. This makes the subject exiting, both intellectually and in practice. Once the subtleties of multiperiod investment are understood, the reward in terms of enhanced investment performance can be substantial."

"Fortunately the concepts and the methods of analysis for multiperiod situation build on those of earlier chapters. Internal rate of return, present value, the comparison principle, portfolio design, and lattice and tree valuation all have natural extensions to general situations.

Chapter 15 of D. G. Luenberger, *Investment Science*. Oxford University Press, 1998.

"Volatility is NOT the same as Risk. Volatility is Opportunity"

"Conclusions about multiperiod investment situations are not mere variations of single-period conclusions – rather they often *reverse* those earlier conclusions. This makes the subject exiting, both intellectually and in practice. Once the subtleties of multiperiod investment are understood, the reward in terms of enhanced investment performance can be substantial."

"Fortunately the concepts and the methods of analysis for multiperiod situation build on those of earlier chapters. Internal rate of return, present value, the comparison principle, portfolio design, and lattice and tree valuation all have natural extensions to general situations. But conclusions such as volatility is "bad" or diversification is "good" are no longer universal truths. The story is much more interesting."

Dynamic portfolio selection: general case

$\mathbf{x}_i = (x_i^{(1)}, \dots x_i^{(d)})$ the return vector on day i

Dynamic portfolio selection: general case

$\mathbf{x}_i = (x_i^{(1)}, \ldots x_i^{(d)})$ the return vector on day i
$\mathbf{b} = \mathbf{b}_1$ is the portfolio vector for the first day

Dynamic portfolio selection: general case

$\mathbf{x}_i = (x_i^{(1)}, \dots x_i^{(d)})$ the return vector on day i
$\mathbf{b} = \mathbf{b}_1$ is the portfolio vector for the first day
initial capital S_0

$$S_1 = S_0 \cdot \langle \mathbf{b}_1 \, , \, \mathbf{x}_1 \rangle$$

Dynamic portfolio selection: general case

$\mathbf{x}_i = (x_i^{(1)}, \dots x_i^{(d)})$ the return vector on day i
$\mathbf{b} = \mathbf{b}_1$ is the portfolio vector for the first day
initial capital S_0

$$S_1 = S_0 \cdot \langle \mathbf{b}_1 \, , \, \mathbf{x}_1 \rangle$$

for the second day, S_1 new initial capital, the portfolio vector $\mathbf{b}_2 = \mathbf{b}(\mathbf{x}_1)$

$$S_2 = S_0 \cdot \langle \mathbf{b}_1 \, , \, \mathbf{x}_1 \rangle \cdot \langle \mathbf{b}(\mathbf{x}_1) \, , \, \mathbf{x}_2 \rangle \, .$$

Dynamic portfolio selection: general case

$\mathbf{x}_i = (x_i^{(1)}, \dots x_i^{(d)})$ the return vector on day i
$\mathbf{b} = \mathbf{b}_1$ is the portfolio vector for the first day
initial capital S_0

$$S_1 = S_0 \cdot \langle \mathbf{b}_1 \, , \, \mathbf{x}_1 \rangle$$

for the second day, S_1 new initial capital, the portfolio vector $\mathbf{b}_2 = \mathbf{b}(\mathbf{x}_1)$

$$S_2 = S_0 \cdot \langle \mathbf{b}_1 \, , \, \mathbf{x}_1 \rangle \cdot \langle \mathbf{b}(\mathbf{x}_1) \, , \, \mathbf{x}_2 \rangle \, .$$

nth day a portfolio strategy $\mathbf{b}_n = \mathbf{b}(\mathbf{x}_1, \dots, \mathbf{x}_{n-1}) = \mathbf{b}(\mathbf{x}_1^{n-1})$

$$S_n = S_0 \prod_{i=1}^{n} \left\langle \mathbf{b}(\mathbf{x}_1^{i-1}) \, , \, \mathbf{x}_i \right\rangle$$

Dynamic portfolio selection: general case

$\mathbf{x}_i = (x_i^{(1)}, \dots x_i^{(d)})$ the return vector on day i
$\mathbf{b} = \mathbf{b}_1$ is the portfolio vector for the first day
initial capital S_0

$$S_1 = S_0 \cdot \langle \mathbf{b}_1 \, , \, \mathbf{x}_1 \rangle$$

for the second day, S_1 new initial capital, the portfolio vector $\mathbf{b}_2 = \mathbf{b}(\mathbf{x}_1)$

$$S_2 = S_0 \cdot \langle \mathbf{b}_1 \, , \, \mathbf{x}_1 \rangle \cdot \langle \mathbf{b}(\mathbf{x}_1) \, , \, \mathbf{x}_2 \rangle \, .$$

nth day a portfolio strategy $\mathbf{b}_n = \mathbf{b}(\mathbf{x}_1, \dots, \mathbf{x}_{n-1}) = \mathbf{b}(\mathbf{x}_1^{n-1})$

$$S_n = S_0 \prod_{i=1}^{n} \left\langle \mathbf{b}(\mathbf{x}_1^{i-1}) \, , \, \mathbf{x}_i \right\rangle = S_0 e^{nW_n(\mathbf{B})}$$

with the average growth rate

$$W_n(\mathbf{B}) = \frac{1}{n} \sum_{i=1}^{n} \ln \left\langle \mathbf{b}(\mathbf{x}_1^{i-1}) \, , \, \mathbf{x}_i \right\rangle .$$

$\mathbf{X}_1, \mathbf{X}_2, \dots$ drawn from the vector valued stationary and ergodic process

$\mathbf{X}_1, \mathbf{X}_2, \ldots$ drawn from the vector valued stationary and ergodic process
log-optimum portfolio $\mathbf{B}^* = \{\mathbf{b}^*(\cdot)\}$

$$\mathbf{E}\{\ln \left\langle \mathbf{b}^*(\mathbf{X}_1^{n-1}), \, \mathbf{X}_n \right\rangle \mid \mathbf{X}_1^{n-1}\} = \max_{\mathbf{b}(\cdot)} \mathbf{E}\{\ln \left\langle \mathbf{b}(\mathbf{X}_1^{n-1}), \, \mathbf{X}_n \right\rangle \mid \mathbf{X}_1^{n-1}\}$$

Optimality

Algoet and Cover (1988): If $S_n^* = S_n(\mathbf{B}^*)$ denotes the capital after day n achieved by a log-optimum portfolio strategy $\mathbf{B}^*$, then for any portfolio strategy $\mathbf{B}$ with capital $S_n = S_n(\mathbf{B})$ and for any stationary ergodic process $\{\mathbf{X}_n\}_{-\infty}^{\infty}$,

$$\limsup_{n\to\infty}\left(\frac{1}{n}\ln S_n - \frac{1}{n}\ln S_n^*\right) \le 0 \quad \text{almost surely}$$

Algoet and Cover (1988): If $S_n^* = S_n(\mathbf{B}^*)$ denotes the capital after day n achieved by a log-optimum portfolio strategy $\mathbf{B}^*$, then for any portfolio strategy $\mathbf{B}$ with capital $S_n = S_n(\mathbf{B})$ and for any stationary ergodic process $\{\mathbf{X}_n\}_{-\infty}^{\infty}$,

$$\limsup_{n\to\infty} \left(\frac{1}{n} \ln S_n - \frac{1}{n} \ln S_n^* \right) \le 0 \quad \text{almost surely}$$

and

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

where

$$W^* = \mathbf{E}\left\{ \max_{\mathbf{b}(\cdot)} \mathbf{E}\{\ln \left\langle \mathbf{b}(\mathbf{X}_{-\infty}^{-1}) \, , \, \mathbf{X}_0 \right\rangle \mid \mathbf{X}_{-\infty}^{-1}\} \right\}$$

is the maximal growth rate of any portfolio.

Proof

$$\frac{1}{n}\ln S_n = \frac{1}{n}\sum_{i=1}^{n}\ln\left\langle \mathbf{b}(\mathbf{X}_1^{i-1}),\, \mathbf{X}_i\right\rangle$$

Proof

$$
\begin{aligned}
\frac{1}{n}\ln S_n &= \frac{1}{n}\sum_{i=1}^{n} \ln \left\langle \mathbf{b}(\mathbf{X}_1^{i-1}), \mathbf{X}_i \right\rangle \\
&= \frac{1}{n}\sum_{i=1}^{n} \mathbf{E}\{\ln \left\langle \mathbf{b}(\mathbf{X}_1^{i-1}), \mathbf{X}_i \right\rangle \mid \mathbf{X}_1^{i-1}\} \\
&+ \frac{1}{n}\sum_{i=1}^{n} \left(\ln \left\langle \mathbf{b}(\mathbf{X}_1^{i-1}), \mathbf{X}_i \right\rangle - \mathbf{E}\{\ln \left\langle \mathbf{b}(\mathbf{X}_1^{i-1}), \mathbf{X}_i \right\rangle \mid \mathbf{X}_1^{i-1}\} \right)
\end{aligned}
$$

Proof

$$\frac{1}{n}\ln S_n = \frac{1}{n}\sum_{i=1}^{n}\ln\left\langle \mathbf{b}(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle$$

$$= \frac{1}{n}\sum_{i=1}^{n}\mathbf{E}\{\ln\left\langle \mathbf{b}(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle \mid \mathbf{X}_1^{i-1}\}$$

$$+ \frac{1}{n}\sum_{i=1}^{n}\left(\ln\left\langle \mathbf{b}(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle - \mathbf{E}\{\ln\left\langle \mathbf{b}(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle \mid \mathbf{X}_1^{i-1}\}\right)$$

and

$$\frac{1}{n}\ln S_n^* = \frac{1}{n}\sum_{i=1}^{n}\mathbf{E}\{\ln\left\langle \mathbf{b}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle \mid \mathbf{X}_1^{i-1}\}$$

$$+ \frac{1}{n}\sum_{i=1}^{n}\left(\ln\left\langle \mathbf{b}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle - \mathbf{E}\{\ln\left\langle \mathbf{b}^*(\mathbf{X}_1^{i-1}), \mathbf{X}_i\right\rangle \mid \mathbf{X}_1^{i-1}\}\right)$$

Universally consistent portfolio

These limit relations give rise to the following definition:

Definition

A portfolio strategy $\mathbf{B}$ is called **universally consistent with respect to a class $\mathcal{C}$ of stationary and ergodic processes** $\{\mathbf{X}_n\}_{-\infty}^{\infty}$, if for each process in the class,

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

$$\mathbf{E}\{\ln\left\langle \mathbf{b}^*(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle \mid \mathbf{X}_1^{n-1}\} = \max_{\mathbf{b}(\cdot)} \mathbf{E}\{\ln\left\langle \mathbf{b}(\mathbf{X}_1^{n-1}), \mathbf{X}_n\right\rangle \mid \mathbf{X}_1^{n-1}\}$$

Empirical portfolio selection

$$\mathbf{E}\{\ln \left\langle \mathbf{b}^*(\mathbf{X}_1^{n-1}), \mathbf{X}_n \right\rangle \mid \mathbf{X}_1^{n-1}\} = \max_{\mathbf{b}(\cdot)} \mathbf{E}\{\ln \left\langle \mathbf{b}(\mathbf{X}_1^{n-1}), \mathbf{X}_n \right\rangle \mid \mathbf{X}_1^{n-1}\}$$

fixed integer $k > 0$

$$\mathbf{E}\{\ln \left\langle \mathbf{b}(\mathbf{X}_1^{n-1}), \mathbf{X}_n \right\rangle \mid \mathbf{X}_1^{n-1}\} \approx \mathbf{E}\{\ln \left\langle \mathbf{b}(\mathbf{X}_{n-k}^{n-1}), \mathbf{X}_n \right\rangle \mid \mathbf{X}_{n-k}^{n-1}\}$$

Empirical portfolio selection

$$\mathbf{E}\{\ln \left\langle \mathbf{b}^*(\mathbf{X}_1^{n-1}), \mathbf{X}_n \right\rangle \mid \mathbf{X}_1^{n-1}\} = \max_{\mathbf{b}(\cdot)} \mathbf{E}\{\ln \left\langle \mathbf{b}(\mathbf{X}_1^{n-1}), \mathbf{X}_n \right\rangle \mid \mathbf{X}_1^{n-1}\}$$

fixed integer $k > 0$

$$\mathbf{E}\{\ln \left\langle \mathbf{b}(\mathbf{X}_1^{n-1}), \mathbf{X}_n \right\rangle \mid \mathbf{X}_1^{n-1}\} \approx \mathbf{E}\{\ln \left\langle \mathbf{b}(\mathbf{X}_{n-k}^{n-1}), \mathbf{X}_n \right\rangle \mid \mathbf{X}_{n-k}^{n-1}\}$$

and

$$\mathbf{b}^*(\mathbf{X}_1^{n-1}) \approx \mathbf{b}_k(\mathbf{X}_{n-k}^{n-1}) = \arg\max_{\mathbf{b}(\cdot)} \mathbf{E}\{\ln \left\langle \mathbf{b}(\mathbf{X}_{n-k}^{n-1}), \mathbf{X}_n \right\rangle \mid \mathbf{X}_{n-k}^{n-1}\}$$

because of stationarity

$$\mathbf{b}_k(\mathbf{x}_1^k) = \arg\max_{\mathbf{b}(\cdot)} \mathbf{E}\{\ln \left\langle \mathbf{b}(\mathbf{X}_{n-k}^{n-1}), \mathbf{X}_n \right\rangle \mid \mathbf{X}_{n-k}^{n-1} = \mathbf{x}_1^k\}$$

because of stationarity

$$
\begin{aligned}
\mathbf{b}_k(\mathbf{x}_1^k) &= \arg\max_{\mathbf{b}(\cdot)} \mathbf{E}\{\ln \left\langle \mathbf{b}(\mathbf{X}_{n-k}^{n-1}) \,,\, \mathbf{X}_n \right\rangle \mid \mathbf{X}_{n-k}^{n-1} = \mathbf{x}_1^k\} \\
&= \arg\max_{\mathbf{b}(\cdot)} \mathbf{E}\{\ln \left\langle \mathbf{b}(\mathbf{x}_1^k) \,,\, \mathbf{X}_n \right\rangle \mid \mathbf{X}_{n-k}^{n-1} = \mathbf{x}_1^k\}
\end{aligned}
$$

because of stationarity

$$\begin{aligned}
\mathbf{b}_k(\mathbf{x}_1^k) &= \arg\max_{\mathbf{b}(\cdot)} \mathbf{E}\{\ln \left\langle \mathbf{b}(\mathbf{X}_{n-k}^{n-1}) \, , \, \mathbf{X}_n \right\rangle \mid \mathbf{X}_{n-k}^{n-1} = \mathbf{x}_1^k\} \\
&= \arg\max_{\mathbf{b}(\cdot)} \mathbf{E}\{\ln \left\langle \mathbf{b}(\mathbf{x}_1^k) \, , \, \mathbf{X}_n \right\rangle \mid \mathbf{X}_{n-k}^{n-1} = \mathbf{x}_1^k\} \\
&= \arg\max_{\mathbf{b}(\cdot)} \mathbf{E}\{\ln \left\langle \mathbf{b}(\mathbf{x}_1^k) \, , \, \mathbf{X}_{k+1} \right\rangle \mid \mathbf{X}_1^k = \mathbf{x}_1^k\}
\end{aligned}$$

because of stationarity

$$
\begin{aligned}
\mathbf{b}_k(\mathbf{x}_1^k) &= \arg\max_{\mathbf{b}(\cdot)} \mathbf{E}\{\ln \left\langle \mathbf{b}(\mathbf{X}_{n-k}^{n-1}) \, , \, \mathbf{X}_n \right\rangle \mid \mathbf{X}_{n-k}^{n-1} = \mathbf{x}_1^k\} \\
&= \arg\max_{\mathbf{b}(\cdot)} \mathbf{E}\{\ln \left\langle \mathbf{b}(\mathbf{x}_1^k) \, , \, \mathbf{X}_n \right\rangle \mid \mathbf{X}_{n-k}^{n-1} = \mathbf{x}_1^k\} \\
&= \arg\max_{\mathbf{b}(\cdot)} \mathbf{E}\{\ln \left\langle \mathbf{b}(\mathbf{x}_1^k) \, , \, \mathbf{X}_{k+1} \right\rangle \mid \mathbf{X}_1^k = \mathbf{x}_1^k\} \\
&= \arg\max_{\mathbf{b}} \mathbf{E}\{\ln \left\langle \mathbf{b} \, , \, \mathbf{X}_{k+1} \right\rangle \mid \mathbf{X}_1^k = \mathbf{x}_1^k\},
\end{aligned}
$$

because of stationarity

$$\begin{aligned}
\mathbf{b}_k(\mathbf{x}_1^k) &= \arg\max_{\mathbf{b}(\cdot)} \mathbf{E}\{\ln \langle \mathbf{b}(\mathbf{X}_{n-k}^{n-1}) \,,\, \mathbf{X}_n \rangle \mid \mathbf{X}_{n-k}^{n-1} = \mathbf{x}_1^k\} \\
&= \arg\max_{\mathbf{b}(\cdot)} \mathbf{E}\{\ln \left\langle \mathbf{b}(\mathbf{x}_1^k) \,,\, \mathbf{X}_n \right\rangle \mid \mathbf{X}_{n-k}^{n-1} = \mathbf{x}_1^k\} \\
&= \arg\max_{\mathbf{b}(\cdot)} \mathbf{E}\{\ln \left\langle \mathbf{b}(\mathbf{x}_1^k) \,,\, \mathbf{X}_{k+1} \right\rangle \mid \mathbf{X}_1^k = \mathbf{x}_1^k\} \\
&= \arg\max_{\mathbf{b}} \mathbf{E}\{\ln \langle \mathbf{b} \,,\, \mathbf{X}_{k+1} \rangle \mid \mathbf{X}_1^k = \mathbf{x}_1^k\},
\end{aligned}$$

which is the maximization of the regression function

$$m_{\mathbf{b}}(\mathbf{x}_1^k) = \mathbf{E}\{\ln \langle \mathbf{b} \,,\, \mathbf{X}_{k+1} \rangle \mid \mathbf{X}_1^k = \mathbf{x}_1^k\}$$

Regression function

Y real valued
X observation vector

Regression function

Y real valued
X observation vector
Regression function

$$m(x) = \mathbf{E}\{Y \mid X = x\}$$

i.i.d. data: $D_n = \{(X_1, Y_1), \ldots, (X_n, Y_n)\}$

Regression function

Y real valued
X observation vector
Regression function

$$m(x) = \mathbf{E}\{Y \mid X = x\}$$

i.i.d. data: $D_n = \{(X_1, Y_1), \ldots, (X_n, Y_n)\}$
Regression function estimate

$$m_n(x) = m_n(x, D_n)$$

Regression function

Y real valued
X observation vector
Regression function

$$m(x) = \mathbf{E}\{Y \mid X = x\}$$

i.i.d. data: $D_n = \{(X_1, Y_1), \ldots, (X_n, Y_n)\}$
Regression function estimate

$$m_n(x) = m_n(x, D_n)$$

Kernel regression estimate with window kernel

Regression function

Y real valued
X observation vector
Regression function

$$m(x) = \mathbf{E}\{Y \mid X = x\}$$

i.i.d. data: $D_n = \{(X_1, Y_1), \ldots, (X_n, Y_n)\}$
Regression function estimate

$$m_n(x) = m_n(x, D_n)$$

Kernel regression estimate with window kernel
Bandwidth $r > 0$

Regression function

Y real valued

X observation vector

Regression function

$$m(x) = \mathbf{E}\{Y \mid X = x\}$$

i.i.d. data: $D_n = \{(X_1, Y_1), \ldots, (X_n, Y_n)\}$

Regression function estimate

$$m_n(x) = m_n(x, D_n)$$

Kernel regression estimate with window kernel

Bandwidth $r > 0$

$$m_n(x) = \frac{\sum_{i=1}^n Y_i I_{\{\|x - X_i\| \le r\}}}{\sum_{i=1}^n I_{\{\|x - X_i\| \le r\}}}$$

Regression function

Y real valued
X observation vector
Regression function

$$m(x) = \mathbf{E}\{Y \mid X = x\}$$

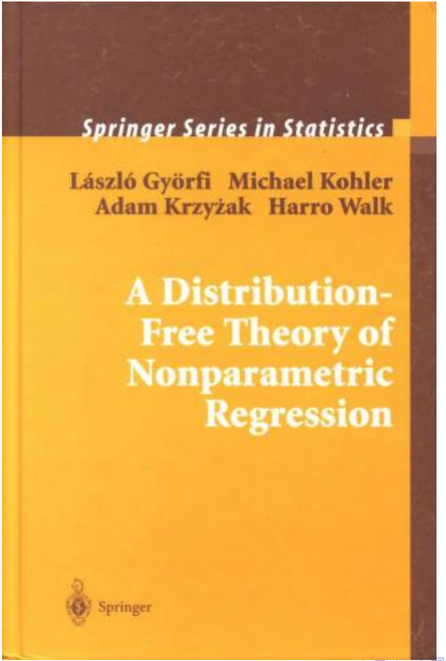
i.i.d. data: $D_n = \{(X_1, Y_1), \ldots, (X_n, Y_n)\}$
Regression function estimate

$$m_n(x) = m_n(x, D_n)$$

Kernel regression estimate with window kernel
Bandwidth $r > 0$

$$m_n(x) = \frac{\sum_{i=1}^n Y_i I_{\{\|x - X_i\| \le r\}}}{\sum_{i=1}^n I_{\{\|x - X_i\| \le r\}}}$$

L. Györfi, M. Kohler, A. Krzyzak, H. Walk (2002) *A Distribution-Free Theory of Nonparametric Regression*, Springer-Verlag, New York.


Springer Series in Statistics
László Györfi Michael Kohler
Adam Krzyżak Harro Walk
A Distribution-Free Theory of Nonparametric Regression
Springer

Correspondence

$$X \quad \sim \quad \mathbf{X}_1^k$$

Correspondence

$$
\begin{aligned}
X &\sim \mathbf{X}_1^k \\
Y &\sim \ln \langle \mathbf{b}\,, \mathbf{X}_{k+1} \rangle
\end{aligned}
$$

Correspondence

$$\begin{aligned} X &\sim \mathbf{X}_1^k \\ Y &\sim \ln \langle \mathbf{b}\,,\, \mathbf{X}_{k+1} \rangle \\ m(x) = \mathbf{E}\{Y \mid X = x\} &\sim m_{\mathbf{b}}(\mathbf{x}_1^k) = \mathbf{E}\{\ln \langle \mathbf{b}\,,\, \mathbf{X}_{k+1} \rangle \mid \mathbf{X}_1^k = \mathbf{x}_1^k\} \end{aligned}$$

Kernel-based portfolio selection

choose the radius $r_{k,\ell} > 0$ such that for any fixed k,

$$\lim_{\ell \to \infty} r_{k,\ell} = 0.$$

choose the radius $r_{k,\ell} > 0$ such that for any fixed k,

$$\lim_{\ell \to \infty} r_{k,\ell} = 0.$$

for $n > k + 1$, define the expert $\mathbf{b}^{(k,\ell)}$ by

$$\mathbf{b}^{(k,\ell)}(\mathbf{x}_1^{n-1}) = \arg\max_{\mathbf{b}} \sum_{i \in J_n^{(k,\ell)}} \ln \langle \mathbf{b} \, , \, \mathbf{x}_i \rangle \, ,$$

Kernel-based portfolio selection

choose the radius $r_{k,\ell} > 0$ such that for any fixed k,

$$\lim_{\ell \to \infty} r_{k,\ell} = 0.$$

for $n > k + 1$, define the expert $\mathbf{b}^{(k,\ell)}$ by

$$\mathbf{b}^{(k,\ell)}(\mathbf{x}_1^{n-1}) = \arg\max_{\mathbf{b}} \sum_{i \in J_n^{(k,\ell)}} \ln \langle \mathbf{b}\,,\ \mathbf{x}_i \rangle\,,$$

where

$$J_n^{(k,\ell)} = \left\{ k < i < n : \|\mathbf{x}_{i-k}^{i-1} - \mathbf{x}_{n-k}^{n-1}\| \le r_{k,\ell} \right\}$$

if the sum is non-void, and $\mathbf{b}_0 = (1/d, \ldots, 1/d)$ otherwise.

Combining elementary portfolios

for fixed $k, \ell = 1, 2, \ldots$,
$\mathbf{B}^{(k,\ell)} = \{\mathbf{b}^{(k,\ell)}(\cdot)\}$, are called elementary portfolios

Combining elementary portfolios

for fixed $k, \ell = 1, 2, \ldots,$
$\mathbf{B}^{(k,\ell)} = \{\mathbf{b}^{(k,\ell)}(\cdot)\}$, are called elementary portfolios

How to choose k, ℓ

- small k or large $r_{k,\ell}$: large bias
- large k and small $r_{k,\ell}$: few matching, large variance

Combining elementary portfolios

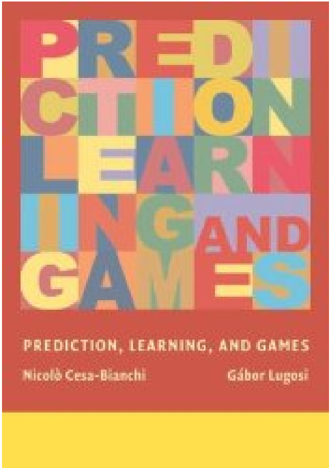
for fixed $k, \ell = 1, 2, \dots,$
$\mathbf{B}^{(k,\ell)} = \{\mathbf{b}^{(k,\ell)}(\cdot)\}$, are called elementary portfolios

How to choose k, ℓ

- small k or large $r_{k,\ell}$: large bias
- large k and small $r_{k,\ell}$: few matching, large variance

Machine learning: combination of experts

N. Cesa-Bianchi and G. Lugosi, *Prediction, Learning, and Games.* Cambridge University Press, 2006.


PREDICTION LEARNING AND GAMES
PREDICTION, LEARNING, AND GAMES
Nicolò Cesa-Bianchi
Gábor Lugosi

Aggregation

combine the elementary portfolio strategies $\mathbf{B}^{(k,\ell)} = \{\mathbf{b}_n^{(k,\ell)}\}$

combine the elementary portfolio strategies $\mathbf{B}^{(k,\ell)} = \{\mathbf{b}_n^{(k,\ell)}\}$
let $\{q_{k,\ell}\}$ be a probability distribution on the set of all pairs (k,ℓ) such that for all k,ℓ, $q_{k,\ell} > 0$.

Aggregation

combine the elementary portfolio strategies $\mathbf{B}^{(k,\ell)} = \{\mathbf{b}_n^{(k,\ell)}\}$
let $\{q_{k,\ell}\}$ be a probability distribution on the set of all pairs (k,ℓ) such that for all k,ℓ, $q_{k,\ell} > 0$.
put

$$w_{n,k,\ell} = q_{k,\ell} S_{n-1}(\mathbf{B}^{(k,\ell)})$$

Aggregation

combine the elementary portfolio strategies $\mathbf{B}^{(k,\ell)} = \{\mathbf{b}_n^{(k,\ell)}\}$
let $\{q_{k,\ell}\}$ be a probability distribution on the set of all pairs (k, ℓ) such that for all k, ℓ, $q_{k,\ell} > 0$.
put

$$w_{n,k,\ell} = q_{k,\ell} S_{n-1}(\mathbf{B}^{(k,\ell)})$$

the aggregated portfolio $\mathbf{b}$:

$$\mathbf{b}_n(\mathbf{x}_1^{n-1}) = \frac{\sum_{k,\ell} w_{n,k,\ell} \mathbf{b}_n^{(k,\ell)}(\mathbf{x}_1^{n-1})}{\sum_{k,\ell} w_{n,k,\ell}}.$$

Theorem

The kernel-based portfolio scheme is universally consistent with respect to the class of all ergodic processes such that $\mathbf{E}\{|\ln X^{(j)}|\} < \infty$, for $j = 1, 2, \ldots, d$.

L. Györfi, G. Lugosi, F. Udina (2006) "Nonparametric kernel-based sequential investment strategies", *Mathematical Finance*, 16, pp. 337-357
www.szit.bme.hu/~gyorfi/kernel.pdf

Proof

$$S_n(\mathbf{B}) = \prod_{i=1}^{n} \left\langle \mathbf{b}_i(\mathbf{x}_1^{i-1}) \, , \, \mathbf{x}_i \right\rangle$$

Proof

$$
\begin{aligned}
S_n(\mathbf{B}) &= \prod_{i=1}^{n} \left\langle \mathbf{b}_i(\mathbf{x}_1^{i-1}) \, , \, \mathbf{x}_i \right\rangle \\
&= \prod_{i=1}^{n} \frac{\sum_{k,\ell} w_{i,k,\ell} \left\langle \mathbf{b}_i^{(k,\ell)}(\mathbf{x}_1^{i-1}) \, , \, \mathbf{x}_i \right\rangle}{\sum_{k,\ell} w_{i,k,\ell}}
\end{aligned}
$$

Proof

$$
\begin{aligned}
S_n(\mathbf{B}) &= \prod_{i=1}^{n} \left\langle \mathbf{b}_i(\mathbf{x}_1^{i-1}) \,,\, \mathbf{x}_i \right\rangle \\
&= \prod_{i=1}^{n} \frac{\sum_{k,\ell} w_{i,k,\ell} \left\langle \mathbf{b}_i^{(k,\ell)}(\mathbf{x}_1^{i-1}) \,,\, \mathbf{x}_i \right\rangle}{\sum_{k,\ell} w_{i,k,\ell}} \\
&= \prod_{i=1}^{n} \frac{\sum_{k,\ell} q_{k,\ell} S_{i-1}(\mathbf{B}^{(k,\ell)}) \left\langle \mathbf{b}_i^{(k,\ell)}(\mathbf{x}_1^{i-1}) \,,\, \mathbf{x}_i \right\rangle}{\sum_{k,\ell} q_{k,\ell} S_{i-1}(\mathbf{B}^{(k,\ell)})}
\end{aligned}
$$

Proof

$$
\begin{aligned}
S_n(\mathbf{B}) &= \prod_{i=1}^{n} \left\langle \mathbf{b}_i(\mathbf{x}_1^{i-1}) \, , \, \mathbf{x}_i \right\rangle \\
&= \prod_{i=1}^{n} \frac{\sum_{k,\ell} w_{i,k,\ell} \left\langle \mathbf{b}_i^{(k,\ell)}(\mathbf{x}_1^{i-1}) \, , \, \mathbf{x}_i \right\rangle}{\sum_{k,\ell} w_{i,k,\ell}} \\
&= \prod_{i=1}^{n} \frac{\sum_{k,\ell} q_{k,\ell} S_{i-1}(\mathbf{B}^{(k,\ell)}) \left\langle \mathbf{b}_i^{(k,\ell)}(\mathbf{x}_1^{i-1}) \, , \, \mathbf{x}_i \right\rangle}{\sum_{k,\ell} q_{k,\ell} S_{i-1}(\mathbf{B}^{(k,\ell)})} \\
&= \prod_{i=1}^{n} \frac{\sum_{k,\ell} q_{k,\ell} S_i(\mathbf{B}^{(k,\ell)})}{\sum_{k,\ell} q_{k,\ell} S_{i-1}(\mathbf{B}^{(k,\ell)})}
\end{aligned}
$$

Proof

$$
\begin{aligned}
S_n(\mathbf{B}) &= \prod_{i=1}^{n} \left\langle \mathbf{b}_i(\mathbf{x}_1^{i-1}) \,,\, \mathbf{x}_i \right\rangle \\
&= \prod_{i=1}^{n} \frac{\sum_{k,\ell} w_{i,k,\ell} \left\langle \mathbf{b}_i^{(k,\ell)}(\mathbf{x}_1^{i-1}) \,,\, \mathbf{x}_i \right\rangle}{\sum_{k,\ell} w_{i,k,\ell}} \\
&= \prod_{i=1}^{n} \frac{\sum_{k,\ell} q_{k,\ell} S_{i-1}(\mathbf{B}^{(k,\ell)}) \left\langle \mathbf{b}_i^{(k,\ell)}(\mathbf{x}_1^{i-1}) \,,\, \mathbf{x}_i \right\rangle}{\sum_{k,\ell} q_{k,\ell} S_{i-1}(\mathbf{B}^{(k,\ell)})} \\
&= \prod_{i=1}^{n} \frac{\sum_{k,\ell} q_{k,\ell} S_i(\mathbf{B}^{(k,\ell)})}{\sum_{k,\ell} q_{k,\ell} S_{i-1}(\mathbf{B}^{(k,\ell)})} \\
&= \sum_{k,\ell} q_{k,\ell} S_n(\mathbf{B}^{(k,\ell)}),
\end{aligned}
$$

Equivalent form of aggregation

The strategy $\mathbf{B}$ then arises from weighing the elementary portfolio strategies $\mathbf{B}^{(k,\ell)} = \{\mathbf{b}_n^{(k,\ell)}\}$ such that the investor's capital becomes

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

We have to prove that

$$\liminf_{n\to\infty} W_n(\mathbf{B}) = \liminf_{n\to\infty} \frac{1}{n} \ln S_n(\mathbf{B}) \geq W^* \quad \text{a.s.}$$

W.l.o.g. we may assume $S_0 = 1$, so that

$$W_n(\mathbf{B}) = \frac{1}{n} \ln S_n(\mathbf{B})$$

We have to prove that

$$\liminf_{n\to\infty} W_n(\mathbf{B}) = \liminf_{n\to\infty} \frac{1}{n} \ln S_n(\mathbf{B}) \geq W^* \quad \text{a.s.}$$

W.l.o.g. we may assume $S_0 = 1$, so that

$$\begin{aligned} W_n(\mathbf{B}) &= \frac{1}{n} \ln S_n(\mathbf{B}) \\ &= \frac{1}{n} \ln \left(\sum_{k,\ell} q_{k,\ell} S_n(\mathbf{B}^{(k,\ell)}) \right) \end{aligned}$$

We have to prove that

$$\liminf_{n\to\infty} W_n(\mathbf{B}) = \liminf_{n\to\infty} \frac{1}{n} \ln S_n(\mathbf{B}) \geq W^* \quad \text{a.s.}$$

W.l.o.g. we may assume $S_0 = 1$, so that

$$\begin{aligned} W_n(\mathbf{B}) &= \frac{1}{n} \ln S_n(\mathbf{B}) \\ &= \frac{1}{n} \ln \left(\sum_{k,\ell} q_{k,\ell} S_n(\mathbf{B}^{(k,\ell)}) \right) \\ &\geq \frac{1}{n} \ln \left(\sup_{k,\ell} q_{k,\ell} S_n(\mathbf{B}^{(k,\ell)}) \right) \end{aligned}$$

We have to prove that

$$\liminf_{n\to\infty} W_n(\mathbf{B}) = \liminf_{n\to\infty} \frac{1}{n} \ln S_n(\mathbf{B}) \geq W^* \quad \text{a.s.}$$

W.l.o.g. we may assume $S_0 = 1$, so that

$$\begin{aligned}
W_n(\mathbf{B}) &= \frac{1}{n} \ln S_n(\mathbf{B}) \\
&= \frac{1}{n} \ln \left(\sum_{k,\ell} q_{k,\ell} S_n(\mathbf{B}^{(k,\ell)}) \right) \\
&\geq \frac{1}{n} \ln \left(\sup_{k,\ell} q_{k,\ell} S_n(\mathbf{B}^{(k,\ell)}) \right) \\
&= \frac{1}{n} \sup_{k,\ell} \left(\ln q_{k,\ell} + \ln S_n(\mathbf{B}^{(k,\ell)}) \right)
\end{aligned}$$

We have to prove that

$$\liminf_{n\to\infty} W_n(\mathbf{B}) = \liminf_{n\to\infty} \frac{1}{n} \ln S_n(\mathbf{B}) \geq W^* \quad \text{a.s.}$$

W.l.o.g. we may assume $S_0 = 1$, so that

$$\begin{aligned}
W_n(\mathbf{B}) &= \frac{1}{n} \ln S_n(\mathbf{B}) \\
&= \frac{1}{n} \ln \left(\sum_{k,\ell} q_{k,\ell} S_n(\mathbf{B}^{(k,\ell)}) \right) \\
&\geq \frac{1}{n} \ln \left(\sup_{k,\ell} q_{k,\ell} S_n(\mathbf{B}^{(k,\ell)}) \right) \\
&= \frac{1}{n} \sup_{k,\ell} \left(\ln q_{k,\ell} + \ln S_n(\mathbf{B}^{(k,\ell)}) \right) \\
&= \sup_{k,\ell} \left(W_n(\mathbf{B}^{(k,\ell)}) + \frac{\ln q_{k,\ell}}{n} \right).
\end{aligned}$$

Thus

$$\liminf_{n\to\infty} W_n(\mathbf{B}) \quad \geq \quad \liminf_{n\to\infty} \sup_{k,\ell} \left(W_n(\mathbf{B}^{(k,\ell)}) + \frac{\ln q_{k,\ell}}{n} \right)$$

Thus

$$\begin{aligned}
\liminf_{n\to\infty} W_n(\mathbf{B}) &\geq \liminf_{n\to\infty} \sup_{k,\ell} \left(W_n(\mathbf{B}^{(k,\ell)}) + \frac{\ln q_{k,\ell}}{n} \right) \\
&\geq \sup_{k,\ell} \liminf_{n\to\infty} \left(W_n(\mathbf{B}^{(k,\ell)}) + \frac{\ln q_{k,\ell}}{n} \right)
\end{aligned}$$

Thus

$$
\begin{aligned}
\liminf_{n\to\infty} W_n(\mathbf{B}) &\geq \liminf_{n\to\infty} \sup_{k,\ell} \left(W_n(\mathbf{B}^{(k,\ell)}) + \frac{\ln q_{k,\ell}}{n} \right) \\
&\geq \sup_{k,\ell} \liminf_{n\to\infty} \left(W_n(\mathbf{B}^{(k,\ell)}) + \frac{\ln q_{k,\ell}}{n} \right) \\
&= \sup_{k,\ell} \liminf_{n\to\infty} W_n(\mathbf{B}^{(k,\ell)})
\end{aligned}
$$

Thus

$$\begin{aligned}
\liminf_{n\to\infty} W_n(\mathbf{B}) &\geq \liminf_{n\to\infty} \sup_{k,\ell} \left(W_n(\mathbf{B}^{(k,\ell)}) + \frac{\ln q_{k,\ell}}{n} \right) \\
&\geq \sup_{k,\ell} \liminf_{n\to\infty} \left(W_n(\mathbf{B}^{(k,\ell)}) + \frac{\ln q_{k,\ell}}{n} \right) \\
&= \sup_{k,\ell} \liminf_{n\to\infty} W_n(\mathbf{B}^{(k,\ell)}) \\
&= \sup_{k,\ell} \epsilon_{k,\ell}
\end{aligned}$$

Because of $\lim_{\ell\to\infty} r_{k,\ell} = 0$, we have that

$$\sup_{k,\ell} \epsilon_{k,\ell} = \lim_{k\to\infty} \lim_{l\to\infty} \epsilon_{k,\ell} = W^*.$$

Semi-log-optimal (mean-variance) portfolio

empirical log-optimal:

$$\mathbf{b}^{(k,\ell)}(\mathbf{x}_1^{n-1}) = \arg\max_{\mathbf{b}} \sum_{i \in J_n^{(k,\ell)}} \ln \langle \mathbf{b} \, , \, \mathbf{x}_i \rangle$$

Semi-log-optimal (mean-variance) portfolio

empirical log-optimal:

$$\mathbf{b}^{(k,\ell)}(\mathbf{x}_1^{n-1}) = \arg\max_{\mathbf{b}} \sum_{i \in J_n^{(k,\ell)}} \ln \langle \mathbf{b} \, , \, \mathbf{x}_i \rangle$$

Taylor expansion: $\ln z \approx h(z) = z - 1 - \frac{1}{2}(z-1)^2$

Semi-log-optimal (mean-variance) portfolio

empirical log-optimal:

$$\mathbf{b}^{(k,\ell)}(\mathbf{x}_1^{n-1}) = \arg\max_{\mathbf{b}} \sum_{i \in J_n^{(k,\ell)}} \ln \langle \mathbf{b} \, , \, \mathbf{x}_i \rangle$$

Taylor expansion: $\ln z \approx h(z) = z - 1 - \frac{1}{2}(z-1)^2$ empirical semi-log-optimal:

$$\tilde{\mathbf{b}}^{(k,\ell)}(\mathbf{x}_1^{n-1}) = \arg\max_{\mathbf{b}} \sum_{i \in J_n^{(k,\ell)}} h(\langle \mathbf{b} \, , \, \mathbf{x}_i \rangle) = \arg\max_{\mathbf{b}} \{ \langle \mathbf{b} \, , \, \mathbf{m} \rangle - \langle \mathbf{b} \, , \, \mathbf{C}\mathbf{b} \rangle \}$$

Semi-log-optimal (mean-variance) portfolio

empirical log-optimal:

$$\mathbf{b}^{(k,\ell)}(\mathbf{x}_1^{n-1}) = \arg\max_{\mathbf{b}} \sum_{i \in J_n^{(k,\ell)}} \ln \langle \mathbf{b}\,,\, \mathbf{x}_i \rangle$$

Taylor expansion: $\ln z \approx h(z) = z - 1 - \frac{1}{2}(z-1)^2$ empirical semi-log-optimal:

$$\tilde{\mathbf{b}}^{(k,\ell)}(\mathbf{x}_1^{n-1}) = \arg\max_{\mathbf{b}} \sum_{i \in J_n^{(k,\ell)}} h(\langle \mathbf{b}\,,\, \mathbf{x}_i \rangle) = \arg\max_{\mathbf{b}} \{\langle \mathbf{b}\,,\, \mathbf{m} \rangle - \langle \mathbf{b}\,,\, \mathbf{C}\mathbf{b} \rangle\}$$

Connection to the Markowitz theory

empirical log-optimal:

$$\mathbf{b}^{(k,\ell)}(\mathbf{x}_1^{n-1}) = \arg\max_{\mathbf{b}} \sum_{i \in J_n^{(k,\ell)}} \ln \langle \mathbf{b}\,,\,\mathbf{x}_i \rangle$$

Taylor expansion: $\ln z \approx h(z) = z - 1 - \frac{1}{2}(z-1)^2$ empirical semi-log-optimal:

$$\tilde{\mathbf{b}}^{(k,\ell)}(\mathbf{x}_1^{n-1}) = \arg\max_{\mathbf{b}} \sum_{i \in J_n^{(k,\ell)}} h(\langle \mathbf{b}\,,\,\mathbf{x}_i \rangle) = \arg\max_{\mathbf{b}} \{ \langle \mathbf{b}\,,\,\mathbf{m} \rangle - \langle \mathbf{b}\,,\,\mathbf{C}\mathbf{b} \rangle \}$$

Connection to the Markowitz theory

L. Györfi, A. Urbán, I. Vajda (2007) "Kernel-based semi-log-optimal portfolio selection strategies", *International Journal of Theoretical and Applied Finance*, 10, pp. 505-516. www.szit.bme.hu/∼gyorfi/semi.pdf

Conditions of the model:

Assume that

- the assets are arbitrarily divisible,
- the assets are available in unbounded quantities at the current price at any given trading period,
- there are no transaction costs,
- the behavior of the market is not affected by the actions of the investor using the strategy under investigation.

NYSE data sets

At www.szit.bme.hu/~oti/portfolio there are two benchmark data set from NYSE:

- The first data set consists of daily data of 36 stocks with length 24 years (1962–1985)
- The second data set contains 19 stocks and has length 45 years (1962–2006)

At `www.szit.bme.hu/~oti/portfolio` there are two benchmark data set from NYSE:

- The first data set consists of daily data of 36 stocks with length 24 years (1962–1985)
- The second data set contains 19 stocks and has length 45 years (1962–2006)

Our experiment is on the second data set

Experiments on average annual yields (AAY) for kernel based experts

Kernel based semi-log-optimal portfolio selection with

Experiments on average annual yields (AAY) for kernel based experts

Kernel based semi-log-optimal portfolio selection with finite array of size $K \times L$ such that $K = 5$ and $L = 10$. Choose the uniform distribution $\{q_{k,\ell}\} = 1/(KL)$ over the experts in use.

Experiments on average annual yields (AAY) for kernel based experts

Kernel based semi-log-optimal portfolio selection with finite array of size $K \times L$ such that $K = 5$ and $L = 10$. Choose the uniform distribution $\{q_{k,\ell}\} = 1/(KL)$ over the experts in use. $k = 1, \ldots, 5$ and $l = 1, \ldots, 10$

$$r_{k,l}^2 = 0.0002 \cdot d \cdot k + 0.00002 \cdot d \cdot k \cdot \ell,$$

Experiments on average annual yields (AAY) for kernel based experts

Kernel based semi-log-optimal portfolio selection with finite array of size $K \times L$ such that $K = 5$ and $L = 10$. Choose the uniform distribution $\{q_{k,\ell}\} = 1/(KL)$ over the experts in use. $k = 1, \ldots, 5$ and $l = 1, \ldots, 10$

$$r_{k,l}^2 = 0.0002 \cdot d \cdot k + 0.00002 \cdot d \cdot k \cdot \ell,$$

AAY of kernel based semi-log-optimal portfolio is 31%

Experiments on average annual yields (AAY) for kernel based experts

Kernel based semi-log-optimal portfolio selection with finite array of size $K \times L$ such that $K = 5$ and $L = 10$. Choose the uniform distribution $\{q_{k,\ell}\} = 1/(KL)$ over the experts in use.
$k = 1, \dots, 5$ and $l = 1, \dots, 10$

$$r_{k,l}^2 = 0.0002 \cdot d \cdot k + 0.00002 \cdot d \cdot k \cdot \ell,$$

AAY of kernel based semi-log-optimal portfolio is 31%
MORRIS had the best AAY, 20%

Experiments on average annual yields (AAY) for kernel based experts

Kernel based semi-log-optimal portfolio selection with finite array of size $K \times L$ such that $K = 5$ and $L = 10$. Choose the uniform distribution $\{q_{k,\ell}\} = 1/(KL)$ over the experts in use. $k = 1, \ldots, 5$ and $l = 1, \ldots, 10$

$$r_{k,l}^2 = 0.0002 \cdot d \cdot k + 0.00002 \cdot d \cdot k \cdot \ell,$$

AAY of kernel based semi-log-optimal portfolio is 31%
MORRIS had the best AAY, 20%
the BCRP had average AAY 21%

The average annual yields of the of the kernel based experts.

k / ℓ	1	2	3	4	5
1	31%	30%	24%	21%	26%
2	34%	31%	27%	25%	22%
3	35%	29%	26%	24%	23%
4	35%	30%	30%	32%	27%
5	34%	29%	33%	24%	24%
6	35%	29%	28%	24%	27%
7	33%	29%	32%	23%	23%
8	34%	33%	30%	21%	24%
9	37%	33%	28%	19%	21%
10	34%	29%	26%	20%	24%

Experiments on average annual yields for nearest neighbor based experts

Experiments on average annual yields for nearest neighbor based experts

We performed some experiments using nearest neighbor strategy.

Experiments on average annual yields for nearest neighbor based experts

We performed some experiments using nearest neighbor strategy. take a finite array of size $K \times L$ such that $K = 5$ and $L = 10$. Choose the uniform distribution $\{q_{k,\ell}\} = 1/(KL)$ over the experts in use.

Experiments on average annual yields for nearest neighbor based experts

We performed some experiments using nearest neighbor strategy.
take a finite array of size $K \times L$ such that $K = 5$ and $L = 10$.
Choose the uniform distribution $\{q_{k,\ell}\} = 1/(KL)$ over the experts in use.
Experts are indexed by $k = 1 \dots 5$ in columns and
$\ell = 50, 100, \dots, 500$ in rows, where ℓ is the number of nearest neighbors.

Experiments on average annual yields for nearest neighbor based experts

We performed some experiments using nearest neighbor strategy. take a finite array of size $K \times L$ such that $K = 5$ and $L = 10$. Choose the uniform distribution $\{q_{k,\ell}\} = 1/(KL)$ over the experts in use.

Experts are indexed by $k = 1 \dots 5$ in columns and $\ell = 50, 100, \dots, 500$ in rows, where ℓ is the number of nearest neighbors.

The average annual yield of nearest neighbor portfolio is 35%

Experiments on average annual yields for nearest neighbor based experts

We performed some experiments using nearest neighbor strategy. take a finite array of size $K \times L$ such that $K = 5$ and $L = 10$. Choose the uniform distribution $\{q_{k,\ell}\} = 1/(KL)$ over the experts in use.
Experts are indexed by $k = 1 \ldots 5$ in columns and $\ell = 50, 100, \ldots, 500$ in rows, where ℓ is the number of nearest neighbors.
The average annual yield of nearest neighbor portfolio is 35%
Comparing the tables, one can conclude that the nearest neighbor strategy is more robust.

The average annual yields of the nearest neighbor based experts.

k / ℓ	1	2	3	4	5
50	31%	33%	28%	24%	35%
100	33%	32%	25%	29%	28%
150	38%	33%	26%	32%	27%
200	38%	28%	32%	32%	24%
250	37%	31%	37%	28%	26%
300	41%	35%	35%	30%	29%
350	39%	36%	31%	34%	32%
400	39%	35%	33%	32%	35%
450	39%	34%	34%	35%	37%
500	42%	36%	33%	38%	35%