IT and the Stock Market

The Stock Vector

- A stock market is represented as a vector of stocks $\mathbf{X} = (X_1, \ldots, X_m)$ where m is the number of stocks and X_i represents the price relative, i.e.

$$X_i = \frac{\text{price closed}}{\text{price open}}, \quad \text{for stock } i.$$

  E.g., $X_i = 1.05 \Leftrightarrow$ the price of stock i is up by 5%.

- Apparently $\mathbf{X}$ is a random vector. Let $\mathbf{X} \sim F(\mathbf{x})$.

The Portfolio

- A **portfolio** $\mathbf{b} = (b_1, \ldots, b_m)$, is an allocation of wealth across the stocks, where b_i is the fraction of wealth invested in stock i.
- The **wealth relative** is defined by

$$S = \frac{\text{wealth at the end of the day}}{\text{wealth at the beginning of the day}}$$

- When using a portfolio $\mathbf{b}$ on stock vector $\mathbf{X}$,

$$S = \mathbf{b}^{\mathbf{t}}\mathbf{X}.$$

The Meaning of Optimum

- We want to maximize S in some sense. Since S is a random variable, we need to be careful about our definition of optimum.
- The standard theory of stock market investment is based on maximizing the expected value of S subject to a constraint on the variance.
- Since in stock markets one reinvests every day, the wealth is the product of wealth relatives. The behavior of this product is determined by the expected logarithm of the wealth relative rather than the expected value.

Doubling Rate

- The **doubling rate** of a stock market and portfolio $\mathbf{b}$ is defined by

$$W(\mathbf{b}, F) = E(\log \mathbf{b}^{\mathbf{t}}\mathbf{X})$$

- The **optimal** doubling rate is defined by

$$W^*(F) = \max_{\mathbf{b}} W(\mathbf{b}, F)$$

- The **log-optimal portfolio** is the portfolio that achieves the optimal doubling rate

$$\mathbf{b}^*(F) = \arg\max_{\mathbf{b}} W(\mathbf{b}, F)$$

Theorem

- Let $\mathbf{X_1}, \ldots, \mathbf{X_n}$ be i.i.d. $\sim F(\mathbf{x})$. Define

$$S_n^* = \prod_{i=1}^{n} \mathbf{b}^{*\mathbf{t}}\mathbf{X}.$$

to be the wealth relative after n days. Then

$$\frac{1}{n}\log S_n^* \to W^*(F).$$

In other words, $S_n^* \doteq 2^{nW^*}$.

- **Proof:**

$$\frac{1}{n}\log S_n^* = \frac{1}{n}\sum_{i=1}^{n}\log \mathbf{b}^{*\mathbf{t}}\mathbf{X} \to E(\log \mathbf{b}^{*\mathbf{t}}\mathbf{X}) = W^*(F).$$

Properties of the Doubling Rates

- $W(\mathbf{b}, F)$ is linear in F and concave in $\mathbf{b}$.
  - This follows from the definition of doubling rate

$$W(\mathbf{b}, F) = E(\log \mathbf{b}^{\mathbf{t}}\mathbf{X}) = \int \log \mathbf{b}^{\mathbf{t}}\mathbf{X}\, dF(\mathbf{x}).$$

    Simply plug in $aF_1 + bF_2$ and $\lambda \mathbf{b}_1 + (1-\lambda)\mathbf{b}_2$.

- $W^*(F)$ is convex in F.
  - Let F_1, F_2 be distributions. Let $F_\lambda = \lambda F_1 + (1-\lambda)F_2$ with optimum $\mathbf{b}_\lambda^*$.

$$\begin{aligned} W^*(F_\lambda) &= W(\mathbf{b}_\lambda^*, F_\lambda) = \lambda W(\mathbf{b}_\lambda^*, F_1) + (1-\lambda)W(\mathbf{b}_\lambda^*, F_2) \\ &\leq \lambda W^*(F_1) + (1-\lambda)W^*(F_2) \end{aligned}$$

Kuhn-Tucker Characterization

- Determining $\mathbf{b}^*$ given F is a problem of maximizing a concave function over a convex set. The maximum may lie on the boundary or be an interior point. We can use Kuhn-Tucker conditions to characterize the maximum.
- The log-optimal portfolio $\mathbf{b}^*$ for a stock market $\mathbf{X}$ satisfies the conditions

$$E\left(\frac{X_i}{\mathbf{b}^{*t}\mathbf{X}}\right)\begin{cases} = 1, & \text{if } b_i^* > 0; \\ \leq 1, & \text{if } b_i^* = 0. \end{cases}$$

Proof of the Conditions for Log-Optimum

- If $\mathbf{b}^*$ is an optimal point for $W(\mathbf{b})$, the directional derivative of W in a "feasible direction" at $\mathbf{b}^*$ must be non-positive. Let $\mathcal{B}$ be the feasible set and $\mathbf{b} \in \mathcal{B}$. Let

$$\mathbf{b}_\lambda = (1-\lambda)\mathbf{b}^* + \lambda\mathbf{b} = \mathbf{b}^* + \lambda(\mathbf{b} - \mathbf{b}^*).$$

  Since $\mathbf{b}^*$ is an optimal point, we have

$$\frac{d}{d\lambda} W(\mathbf{b}_\lambda)|_{\lambda=0^+} \leq 0 \ \ \forall\, \mathbf{b} \in \mathcal{B}.$$

  Carrying out the one-sided derivative at $\lambda = 0^+$ yields

$$E\left(\frac{\mathbf{b}^{\mathbf{t}}\mathbf{X}}{\mathbf{b}^{*\mathbf{t}}\mathbf{X}}\right) - 1 \leq 0.$$

- Set b to be the extreme points.

Properties of Log-Optimal Portfolio

- Let S^* and S be the wealth relatives from $\mathbf{b}^*$ and an arbitrary $\mathbf{b}$ respectively. Then $E\frac{S}{S^*} \le 1$.

$$E\left(\frac{X_i}{\mathbf{b}^{*t}\mathbf{X}}\right) \le 1 \Rightarrow \sum_i b_i E\left(\frac{X_i}{\mathbf{b}^{*t}\mathbf{X}}\right) = E\left(\frac{\mathbf{b^t X}}{\mathbf{b}^{*t}\mathbf{X}}\right) \le 1$$

  Therefore optimizing expected log ratio also optimizes expected ratio.

- The expected proportion of wealth in each stock is constant under the log-optimal portfolio.

$$E\left(\frac{b_i^* X_i}{\mathbf{b}^{*t}\mathbf{X}}\right) = b_i^* E\left(\frac{X_i}{\mathbf{b}^{*t}\mathbf{X}}\right) = b_i^*.$$

Comparing Log-Optimal and Causal Portfolios

- Let $\mathbf{X_1}, \ldots, \mathbf{X_n}$ be i.i.d. $\sim F(\mathbf{x})$. Let S_n^* be the wealth after n days for an investor using $\mathbf{b}^*$,

$$S_n^* = \prod_{i=1}^{n} \mathbf{b}^{*t} \mathbf{X}_i$$

and let S_n be the wealth of another investor using any causal strategy $\mathbf{b}_i$ on day i.

$$S_n = \prod_{i=1}^{n} \mathbf{b}_i^t \mathbf{X}_i$$

Log Optimal of Wealth

- The expectation of $\log S_n^*$ is optimal.

$$E \log S_n^* = nW^* \geq E \log S_n.$$

- Proof:

$$\begin{aligned}\max_{\mathbf{b_1 \ldots b_n}} E \log S_n &= \max_{\mathbf{b_1 \ldots b_n}} E \sum_{i=1}^{n} \log \mathbf{b}_i^t \mathbf{X} \\ &= \sum_{i=1}^{n} \max_{\mathbf{b}_i} E \log \mathbf{b}_i^t (\mathbf{X}_1, \ldots \mathbf{X}_{i-1}) \mathbf{X} \\ &= \sum_{i=1}^{n} E \log \mathbf{b}^{*t} \mathbf{X} = nW^*\end{aligned}$$

Asymptotic Optimality

- Using log-optimal portfolio will not do any worse than using any causal strategy for almost every sequence of outcomes of the stock market. That is,

$$\lim_{n\to\infty} \sup \frac{1}{n} \log \frac{S_n}{S_n^*} \leq 0 \text{ with probability } 1.$$

Side Information

- Let $\mathbf{X}_1 \ldots \mathbf{X}_n$ be drawn i.i.d. $\sim F(\mathbf{x})$. Define $\mathbf{b}^*_F = \arg\max_{\mathbf{b}} W(\mathbf{b}, F)$ and $\mathbf{b}^*_G = \arg\max_{\mathbf{b}} W(\mathbf{b}, G)$. Then

$$\Delta W \triangleq W(\mathbf{b}^*_F, F) - W(\mathbf{b}^*_G, F) \le D(F||G).$$

- Proof

$$\begin{aligned}
\Delta W &= \int \log \mathbf{b}^{*t}_F \mathbf{x} \, F(\mathbf{x}) d\mathbf{x} - \int \log \mathbf{b}^{*t}_G \mathbf{x} \, F(\mathbf{x}) d\mathbf{x} \\
&= \int \log \frac{\mathbf{b}^{*t}_F \mathbf{x}}{\mathbf{b}^{*t}_G \mathbf{x}} \, F(\mathbf{x}) d\mathbf{x} = \int \log \left(\frac{\mathbf{b}^{*t}_F \mathbf{x}}{\mathbf{b}^{*t}_G \mathbf{x}} \frac{G(\mathbf{x})}{F(\mathbf{x})} \frac{F(\mathbf{x})}{G(\mathbf{x})} \right) F(\mathbf{x}) d\mathbf{x} \\
&\le D(F||G) + \log \int \frac{\mathbf{b}^{*t}_F \mathbf{x}}{\mathbf{b}^{*t}_G \mathbf{x}} \frac{G(\mathbf{x})}{F(\mathbf{x})} F(\mathbf{x}) d\mathbf{x} \\
&\le D(F||G) + \log 1 = D(F||G).
\end{aligned}$$

Side Information

- The increase in the doubling rate due to side information Y is bounded by

$$\Delta W \leq I(\mathbf{X}; Y).$$

- Proof

$$\Delta W_{Y=y} \leq D(F_{\mathbf{X}|Y=y} || F_{\mathbf{X}}) = \int F(\mathbf{x}|Y=y) \log \frac{F(\mathbf{x}|Y=y)}{F_{\mathbf{X}}(\mathbf{x})} d\mathbf{x}$$

$$\begin{aligned}
\Delta W &\leq \int F_Y(y) \int F(\mathbf{x}|Y=y) \log \frac{F(\mathbf{x}|Y=y)}{F_{\mathbf{X}}(\mathbf{x})} d\mathbf{x} dy \\
&= \int \int F_Y(y) F(\mathbf{x}|Y=y) \log \frac{F(\mathbf{x}|Y=y)}{F_{\mathbf{X}}(\mathbf{x})} d\mathbf{x} dy \\
&= \int \int F(\mathbf{x}, y) \log \frac{F(\mathbf{x}, y)}{F_{\mathbf{X}}(\mathbf{x}) F_Y(y)} d\mathbf{x} dy = I(\mathbf{X}; Y)
\end{aligned}$$

Stationary Markets

- We extend the discussion from i.i.d. markets to non-i.i.d. markets.
  - The distribution of stock vector $\mathbf{X_t}$ is generally dependent on the previous stock vectors.
  - The optimal portfolio is also dependent on previous stock vectors.
- The wealth relative using a causal strategy is

$$S_n = \prod_{i=1}^{n} \mathbf{b_i^t}(\mathbf{X_1}, \ldots, \mathbf{X_{i-1}})\mathbf{X_i}.$$

Log-Optimal Portfolio

- Define the conditional log-optimum

$$W^*(\mathbf{X_i}|\mathbf{x_1},\ldots,\mathbf{x_{i-1}}) \triangleq \max_{\mathbf{b}} E_{\mathbf{X_i}|\mathbf{x_1},\ldots,\mathbf{x_{i-1}}} \log \mathbf{b^t X_i}$$

- Taking the expectation over $\mathbf{X_1},\ldots,\mathbf{X_{i-1}}$,

$$W^*(\mathbf{X_i}|\mathbf{X_1},\ldots,\mathbf{X_{i-1}}) \triangleq E \max_{\mathbf{b}} E_{\mathbf{X_i}|\mathbf{X_1},\ldots,\mathbf{X_{i-1}}} \log \mathbf{b^t X_i}$$

- It follows that

$$W^*(\mathbf{X_1},\ldots,\mathbf{X_n}) \triangleq \max_{\mathbf{b_1}\ldots\mathbf{b_n}} E \log S_n = \sum_{i=1}^{n} W^*(\mathbf{X_i}|\mathbf{X_1},\ldots,\mathbf{X_{i-1}})$$

- This is the chain rule for W^*. In some ways, W^* is the dual of H.

Doubling Rate

- The doubling rate is defined by

$$W_{\infty}^{*} \triangleq \lim_{n\to\infty} \frac{W^{*}(\mathbf{X_1}, \ldots, \mathbf{X_n})}{n}$$

- Theorem: For a stationary market, we have

$$W_{\infty}^{*} = \lim_{n\to\infty} W^{*}(\mathbf{X_n} | \mathbf{X_1}, \ldots, \mathbf{X_{n-1}})$$

This follows the theorem of the Cesáro mean.

Asymptotic Optimality

- Let $\{\mathbf{X}_i\}$ be a stationary stock market. Let S_n^* be the wealth resulting from conditional log-optimal portfolios and S_n be the wealth resulting from any other causal portfolios. Then

$$\lim_{n\to\infty} \sup \frac{1}{n} \log \frac{S_n}{S_n^*} \leq 0.$$

AEP for Stock Markets

- Let $\{\mathbf{X}_i\}$ be a stationary ergodic stock market. Let S_n^* be the wealth resulting from conditional log-optimal portfolios. Then

$$\frac{1}{n}\log S_n^* \to W^* \ \text{ with probability } 1.$$

Competitive Optimality

- Is the log-optimal portfolio outperforms alternative portfolios in a finite period of n days?
- From the KT condition, we have

$$E\frac{S_n}{S_n^*} \leq 1,$$

therefore, by the Markov inequality,

$$p(S_n > tS_n^*) \leq \frac{1}{t}.$$

- But this does not provide useful information for $p(S_n > S_n^*)$.

Fair Randomization

- Let S^* be the wealth at the end of one period of investment in a stock market $\mathbf{X}$ with the log-optimal portfolios, and S be the wealth of another strategy. Let $U \sim U[0,2]$ be independent of $\mathbf{X}$ and V be any other r.v. independent of U and $\mathbf{X}$ with $V \geq 0$ and $EV = 1$. Then

$$p(VS \geq US^*) \leq \frac{1}{2}.$$

- This result provides a short-term justification for the use of the log-optimal portfolio when fair randomization of initial wealth is allowed.

Horse Races

- The horse race is a special case of stock market, where there are m stocks. At the end of day, the stock price r_i for stock i and 0 for all other stocks.
- The log-optimal portfolio is proportional betting

$$b_i^* = p_i.$$

- In the case of uniform fair odds,

$$W^* = \log m - H(\mathcal{X}),$$

and

$$S_n^* \doteq 2^{nW^*}.$$

AEP for Ergodic Process

- If $\{X\}$ is a stationary and ergodic source, then

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to EX \text{ w.p. } 1.$$

- Thus the strong law of large numbers holds for a stationary and ergodic source.

The Shannon-McMillan-Breiman Theorem

- If H is the entropy rate of a finite-valued stationary ergodic process $\{X_n\}$, then

$$-\frac{1}{n}\log p(X_0, X_1, \ldots, X_{n-1}) \rightarrow H \ \text{ w.p. } 1$$

- We wish to conclude that

$$\begin{aligned}-\frac{1}{n}\log p(X_0, \ldots, X_{n-1}) &= -\frac{1}{n}\sum_{i=0}^{n-1}\log p(X_i|X_0^{i-1}) \\ &\rightarrow \lim_{n\rightarrow\infty} E[-\log p(X_n|X_0^{n-1})].\end{aligned}$$

But $X_n|X_0^{n-1}$ is not stationary.

Outline of Proof

- While $X_n|X_0^{n-1}$ is not ergodic, $X_n|X_{n-k}^{n-1}$, $X_n|X_{-\infty}^{n-1}$ are. Define

$$H^k \triangleq E[-\log p(X_k|X_{k-1},\ldots,X_0)]$$
$$H^\infty \triangleq E[-\log p(X_0|X_{-1},X_{-2},\ldots)]$$

- Note that

$$H = H^\infty = \lim_{k\to\infty} H^k = \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} H^k.$$

- The main idea is the conditional proportional gambling:
  - A gambler knowing k past has a growth rate $1 - H^k$.
  - Another knowing infinite past has $1 - H^\infty$.
  - Our gambler knowing X_0^{n-1} has a rate in between.
  - Since the gap closes, all must be $1 - H$.