Information Theory and Networks

Lecture 18: Information Theory and the Stock Market

Paul Tune

<paul.tune@adelaide.edu.au>

http://www.maths.adelaide.edu.au/matthew.roughan/Lecture_notes/InformationTheory/

School of Mathematical Sciences,
University of Adelaide

September 18, 2013

Part I

The Stock Market

Put all your eggs in one basket and then watch that basket.

Mark Twain, Pudd'nhead Wilson and Other Tales

Section 1

Basics of the Stock Market

Stock Market

- "Market" referred to is really the *secondary market*
  - *primary market* deals with the issuance of stock
- Consider m assets
  - one asset has the *risk-free rate*: theoretical zero risk
  - our goal: construct a portfolio i.e. allocation of assets with exponential wealth growth
- We assume no
  - Short selling
  - Leveraging

Some Definitions

- We look at day-to-day fluctuations of the stock prices
- Stock market $\mathbf{X} = (X_1, X_2, \cdots, X_m)$, $X_i \geq 0$
  - our universe of stocks is m
  - X_i price relative: (price at start of day)/(price at end of day)
  - $F(\mathbf{x})$: underlying distribution of X_is
- The portfolio $\mathbf{b} = (b_1, b_2, \cdots, b_m)$, $b_i \geq 0$, $\sum_{i=1}^{m} b_i = 1$
- The wealth relative $S = \mathbf{b}^T \mathbf{X}$
- Investment period n days results in $S_n = \prod_{i=1}^{n} \mathbf{b}_i^T \mathbf{X}_i$

Section 2

Log-Optimal Portfolios

Optimising Growth Rate

- Want to maximise $W(\mathbf{b}, F) := E[\log S]$
  - $W^*(F) := \max_{\mathbf{b}} W(\mathbf{b}, F)$
  - portfolio $\mathbf{b}^*$ achieving $W^*(F)$ is the *log-optimal portfolio*
- Suppose price relatives are i.i.d. according to $f(\mathbf{x})$. Assume constant rebalancing with allocation $\mathbf{b}^*$, so $S_n^* = \prod_{i=1}^n \mathbf{b}^{*T} \mathbf{X}_i$. Then,

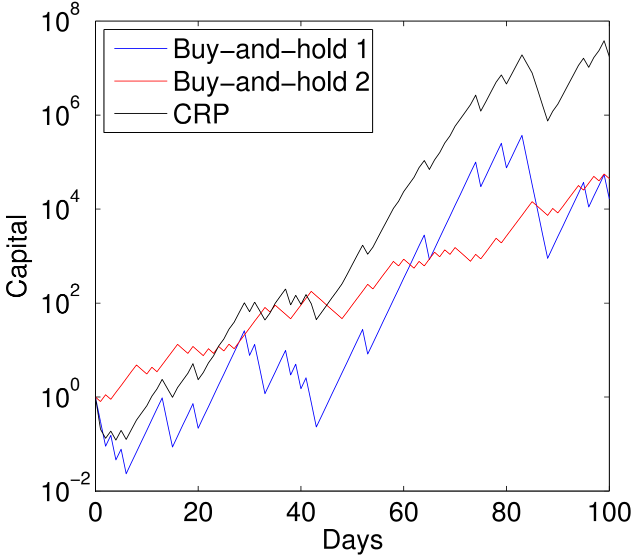
$$\frac{1}{n} \log S_n^* \to W^* \text{ with probability } 1.$$

- **Implication**: regardless of current wealth, keep allocations between assets constant!
- Can we justify constant rebalancing portfolios beyond i.i.d.? Yes, for stationary markets, conditional allocation

Shannon's Volatility Pumping

- Constant rebalancing portfolio (CRP): suggested by Shannon in a lecture at MIT in the 1960s
- Shannon used geometric Wiener to model the price relatives
- CRPs essentially exploit volatility of the price relatives
  - the higher the price volatilty between assets, the higher the excess returns

CRP vs. Buy and Hold



Karush-Kuhn-Tucker Characterisation

- Observe the admissible portfolios form an m-simplex $\mathcal{B}$
- Karush-Kuhn-Tucker (KKT) conditions yield:

$$E\left[\frac{X_i}{\mathbf{b}^{*T}\mathbf{X}}\right] = \begin{cases} 1 & \text{if } b_i^* > 0 \\ 0 & \text{if } b_i^* = 0 \end{cases}$$

- **Implication**: portfolio at least as good as best stock return on average
- KKT conditions also imply:

$$E\left[\log\frac{S}{S^*}\right] \leq 0 \text{ for all } S \text{ iff } E\left[\frac{S}{S^*}\right] \leq 1 \text{ for all } S.$$

- Also, $E\left[\frac{b_i^* X_i}{\mathbf{b}^{*T}\mathbf{X}}\right] = b_i^* E\left[\frac{X_i}{\mathbf{b}^{*T}\mathbf{X}}\right] = b_i^*$ (c.f. Kelly criterion)

Wrong Belief

- In horse racing, side information improves wealth growth rate
- Suppose investor believes underlying distribution is $G(\mathbf{x})$ instead of $F(\mathbf{x})$: what is the impact?
  - end up using allocation $\mathbf{b}_G$ instead of $\mathbf{b}_F$
  - characterise *increase in growth rate*

$$\Delta W = W(\mathbf{b}_F, F) - W(\mathbf{b}_G, G)$$

- Turns out $\Delta W \leq D(F \,||\, G)$ (proof: Jensen's inequality and KKT condition)

Side Information

- Result can be used to show $\Delta W \leq I(\mathbf{X}; Y)$, equality holds if it is the horse race i.e. return due to win or loss
- In real life: private insider trading can significantly increase wealth
  - e.g. buying stock before press release of profit upgrades or sensitive announcement
  - practice is banned in most developed countries
  - insider trading must be declared in public records
- Information asymmetry lead to significant (dis)advantages, not just wealth-wise

Causality

- Nothing said about causal strategies: in real life, not possible to invest in hindsight
- *Nonanticipating* or *causal* portfolio: sequence of mappings $b_i : \mathbb{R}^{m(i-1)} \to \mathcal{B}$, with the interpretation $b_i(\mathbf{x}_1, \cdots, \mathbf{x}_{i-1})$ used on day i
- Suppose $\mathbf{X}_i$ drawn i.i.d. from $F(\mathbf{x})$, S_n is wealth relative from any causal strategy,

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

- **Caveat**: theorem does not say for a fixed n, log-optimal portfolio does better than any strategy

Part II

Universal Portfolios

Background

- Previous discussions assume F is known
- What's the best we can do, if F is not known?
  - use best CRP based on hindsight as benchmark
  - think of something (clever) to approach this benchmark
- Needs to be (somewhat) practical
  - causal strategy
  - universal: distribution free strategy
- **Solution**: adaptive strategy

Finite Horizon

- Assume n is known in advance, $\mathbf{x}^n = (\mathbf{x}_1, \cdots, \mathbf{x}_n)$ is the stock market sequence
- **Theorem**: For any causal strategy $\hat{\mathbf{b}}_i(\cdot)$,

$$\max_{\hat{\mathbf{b}}_i(\cdot)} \min_{\mathbf{x}_1, \cdots, \mathbf{x}_n} \frac{\hat{S}_n(\mathbf{x}^n)}{S_n^*(\mathbf{x}^n)} = V_n$$

- V_n is the normalisation factor, for reasons clearer later on
- Nothing said about the underlying distribution: distribution free!

Finite Horizon: Big Picture

- **Big Picture**: look at all the outcomes length n, allocate wealth in hindsight, then construct best causal strategy from the optimal
- Has to perform close to optimal under "adversarial" outcomes
  - if $m = 2$, outcomes are $((1,0)^T, (1,0)^T, \cdots, (1,0)^T)$, clearly best hindsight strategy is to allocate only to stock 1
  - without hindsight, might want to "spread" allocation to maximise return, minimise loss
  - $\hat{b} = (1/2, 1/2)$ but will be 2^n away from best strategy, need some form of adaptation

Finite Horizon: Construction

- By optimality of CRPs, only need to compare the best CRP to the causal strategies
- Consider the case $m = 2$, can generalise from this case
- **Key idea**: convert $S_n(\mathbf{x}^n) = \prod_{i=1}^n \mathbf{b}^T \mathbf{X}_i$ to

$$S_n(\mathbf{x}^n) = \sum_{j^n \in \{1,2\}^n} \prod_{i=1}^n b_{i,j_i} \prod_{i=1}^n x_{i,j_i} = \sum_{j^n \in \{1,2\}^n} w(j^n) x(j^n)$$

- Now, problem is about determining allocation $w(j^n)$ to 2^n stocks

Finite Horizon: Construction II

- With 2 stocks, $w(j^n) = \prod_{i=1}^{n} b^k (1-b)^{n-k}$, k number of times stock 1 price > stock 2 price
  - what is the optimal allocation b^* for this?
- $\sum_{j^n} w^*(j^n) > 1$ because best CRP has benefit of hindsight: can allocate more to the best sequences
  - causal strategy does not have this hindsight
  - make $\hat{w}(j^n)$ proportional to $w^*(j^n)$ by normalisation (using V_n)
- Then, find the optimal allocation for adversarial sequences
  - what is the best allocation, if at each time step in a sequence, exactly one stock yields non-zero return?
- Putting these two together can show

$$V_n \leq \max_{\hat{\mathbf{b}}_i(\cdot)} \min_{\mathbf{x}_1, \cdots, \mathbf{x}_n} \frac{\hat{S}_n(\mathbf{x}^n)}{S_n^*(\mathbf{x}^n)} \leq V_n$$

Finite Horizon: Sequential

- Finally, need to convert back to the causal portfolio mapping
- For allocation to stock 1 at day i, sum over all sequences with 1 in position i

$$\hat{\mathbf{b}}_{i,1}(\mathbf{x}^{i-1}) = \frac{\sum_{j^{i-1} \in m^{i-1}} \hat{w}(j^{i-1}) x(j^{i-1})}{\sum_{j^i \in m^i} \hat{w}(j^i) x(j^{i-1})}$$

- Algorithm enumerates over all m^n sequences: computationally prohibitive
- Asymptotics yield, for $m = 2$ and all n, $\frac{1}{2\sqrt{n+1}} \le V_n \le \frac{2}{\sqrt{n+1}}$
- Observe:

$$\lim_{n \to \infty} \frac{1}{n} \log \frac{\hat{S}_n(\mathbf{x}^n)}{S_n^*(\mathbf{x}^n)} = \lim_{n \to \infty} \frac{1}{n} \log V_n = 0$$

for any $\mathbf{x}^n$

Horizon-Free

- Two tier process: think of all CRPs with various $\mathbf{b}$ as mutual funds
- Now, we allocate our wealth according to a distribution $\mu(\mathbf{b})$ to all these funds
  - each fund gets $d\mu(\mathbf{b})$ of wealth
  - some will perform better than others, one is the best CRP in hindsight
- What kind of distribution should one choose? (Hint: think adversarial)

Horizon-Free

- **Idea**: Choose a distribution $\mu(\mathbf{b})$ that spreads over all CRPs to maximise

$$\hat{S}(\mathbf{x}^n) = \int_{\mathcal{B}} S_n(\mathbf{b}, \mathbf{x}^n) d\mu(\mathbf{b})$$

- Choose allocation $\hat{\mathbf{b}}_{i+1}(\mathbf{x}^i) = \frac{\int_{\mathcal{B}} \mathbf{b} S_i(\mathbf{b}, \mathbf{x}^i) d\mu(\mathbf{b})}{\int_{\mathcal{B}} S_i(\mathbf{b}, \mathbf{x}^i) d\mu(\mathbf{b})}$
  - interpretation: numerator is weighted performance of the fund, denominator is total wealth
  - best performing CRP dominates overall, especially as $n \to \infty$
- Allocation results in

$$\frac{\hat{S}_n(\mathbf{x}^n)}{S_n^*(\mathbf{x}^n)} \geq \min_{j^n} \frac{\int_{\mathcal{B}} \prod_{i=1}^n b_{j_i} d\mu(\mathbf{b})}{\prod_{i=1}^n b_{j_i}^*}$$

- With the right distribution, for e.g. the Dirichlet$(\frac{1}{2}, \frac{1}{2})$ for $m = 2$,

$$\lim_{n \to \infty} \frac{1}{n} \log \frac{\hat{S}_n(\mathbf{x}^n)}{S_n^*(\mathbf{x}^n)} = 0$$

Caveats

- There is no assumption on brokerage fees
  - in real life, a commission is charged by the broker for any trade
  - CRP relies on daily(!) rebalancing for best performance
- Optimal for a long enough investment horizon
- Relies on the volatility between stocks
  - simulations show that it performs poorly otherwise
  - need the daily rebalancing to exploit volatility
  - longer horizons such as a month or year less volatile (in general)
- Computationally impractical
  - finite horizon: need to evaluate over all possible m^i sequences on day i, combinatorial explosion
  - horizon free: need to work out the integral of returns over the simplex $\mathcal{B}$

Further reading I

Thomas M. Cover and Joy A. Thomas, *Elements of information theory*, John Wiley and Sons, 1991.