Shannon Day
May 18, 1998

Shannon Theory and Investment

Tom Cover

Shannon '48
Kullback Leibler '51
Kelly '56
Breiman '61

Shannon MIT talk 60's
Thorp '69
Shannon & Samuelson '60's

Competitive optimality '80, '92
Side information '88, '96
Growth rate optimality, AEP '88
Universal data compression '68
Universal investment '86, '91, '96, '98

Shannon, 1948

Entropy: $H(X) = -\sum p(x) \log p(x)$

Mutual information: $I(X;Y)$

AEP: $\frac{1}{n} \log \frac{1}{p(X^n)} \longrightarrow H$

Kullback and Leibler, 1951

Relative entropy: $D(p \parallel q) = \sum p(x) \log \frac{p(x)}{q(x)}$

$$I(X;Y) = D(p(x,y) \parallel p(x)p(y))$$

Kelly, 1956

Horse race

$$\Pr\{X = i\} = p(i), \quad i = 1, 2, \ldots, m$$

Odds: m for 1

1. Growth rate of wealth:

$$W(X) = \max_{b(\cdot)} \textstyle\sum_x p(x) \log b(x) m$$

$$\Longrightarrow b^*(x) = p(x)$$

$$W(X) = \log m - H(X)$$

2. Side information Y :

$$b^*(x|y) = p(x|y)$$

$$W(X|Y) = \log m - H(X|Y)$$

3. Increase in growth rate:

$$\Delta W = I(X;Y)$$

Breiman, 1961

$$S_n = \prod_{i=1}^{n} (\sum_{i=1}^{m} b_i X_i)$$

$$X_1, X_2, \ldots \text{ i.i.d. } \sim p(x)$$

$$b^* : \quad \max_b E \ln \sum b_i X_i$$

Then

1. $\frac{1}{n} \log S_n^* \longrightarrow W(X)$

2. $\limsup \frac{1}{n} \log \frac{S_n}{S_n^*} \leq 0$

3. Minimum time to achieve wealth A.

Shannon MIT lecture

Wyner

Inside information

Teledyne

Thorp, 1969

Growth optimal portfolio not on efficient frontier

Thorp and Shannon

Blackjack article

Shannon and Samuelson

Evening discussion meetings in '60's

"...information theory of the Shannon type"

Brighton

Competitive optimality

Horse race (Bell and TC, '80):

Competitive optimality of log optimal gambling?

$$S^*(x) = b^*(x)m, \quad p(x)$$

$$\Pr\{S \geq cS^*\} \leq \frac{1}{c}$$

General market (Bell and TC, '88):

Log optimal portfolio **b** wrt $F(x)$:

$$\begin{aligned} W(X) &= \max \textstyle\int \ln(b^t x) dF(x) \\ S^* &= b^{*t} X \end{aligned}$$

Then $\Pr\{S \geq cS^*\} \leq \frac{1}{c}$

Shannon code (TC '92):

The description lengths

$$l^*(x) = \log \frac{1}{p(x)}$$

are competitively optimal in the sense that

$$\Pr\{l(X) \leq l^*(X) - c\} \leq 2^{-c}$$

Side information Y:

The increase in growth rate is $\Delta W = I(X;Y)$ for the horse race market.

What is ΔW for the general market?

Let

$$b_f : \max_b \int \ln(b^t x) f(x) dx.$$

Thus b_f is *log optimal* wrt f.

Define the *relative growth* rate

$$\Delta(f \parallel g) = \int \ln \frac{b_f^t x}{b_g^t x} f(x) dx$$

Then (Barron - TC, 1988)

1. $\Delta(f \parallel g) \geq 0$
2. $\Delta(f \parallel g) \leq D(f \parallel g)$
3. $\Delta W \leq I(X;Y)$

Asymptotic growth rate

$\mathbf{X}_1, \mathbf{X}_2, \ldots$ *ergodic* market process, $\mathbf{X}_i \in \mathbb{R}^m$.

Let $b_k^* : \max_b E[\ln b^t X_k | X^{k-1}]$. Thus $b_k^*(X^{k-1})$ is conditionally log optimal.

Wealth: $S_n^* = \Pi_{k=1}^n b_k^{*t}(X^{k-1})X_k$

Then (Algoet and TC, 1988)

1. $\Pr\{S_n \geq cS_n^*\} \leq \frac{1}{c}$
2. $E\frac{S_n}{S_n^*} \leq 1$
3. $\limsup \frac{1}{n} \ln \frac{S_n}{S_n^*} \leq 0$, a.e.
4. $\frac{1}{n} \ln S_n^* \longrightarrow W_\infty$ (AEP for Market)

This is proved with a "sandwich proof" that also provides a proof of the AEP (Algoet & TC, 1988):

$$-\frac{1}{n} \log p(X^n) = -\frac{1}{n} \sum_{i=1}^{n} \log p(X^i | X^{i-1})$$

Thus $-\frac{1}{n} \log p(X^n) \longrightarrow H$, for all ergodic processes.

Universality

Data compression:

Davisson, Gallager, Ryabko, Shtarkov, Krischevsky

$$\{p_\theta(x)\} \qquad \begin{aligned} l_\theta(x) &= \log \tfrac{1}{p_\theta(x)} \\ l(x) &= \log \tfrac{1}{p(x)} \end{aligned}$$

Regret: $E_\theta(l(X) - l_\theta(X)) = D(p_\theta \parallel p)$

$$D^* = \min_p \max_\theta D(p_\theta \parallel p)$$

$$D^* = C \qquad \text{(Gallager, 1968)}$$

Excess growth rate: (TC and Ordentlich, 1998)

$$\Delta(p_\theta \parallel p) = \int p_\theta \ln \frac{b_\theta^t x}{b_p^t x}$$

$$\Delta^* = \max_p \min_\theta \Delta(p_\theta \parallel p)$$

Then

$$\Delta^* \leq C,$$

and

$$\Delta_n^*(p_\theta^n \parallel p_U) = \frac{m-1}{2} \log n + o(\log n)$$

Universal portfolios

Let $\mathbf{x}_1, \mathbf{x}_2, \ldots$ arbitrary sequence of stock vectors $\in \mathbb{R}^m$.

Wealth from rebalanced portfolio b:

$$S_n(b) = \prod_{i=1}^{n} b^t X_i$$

Wealth from best constant rebalanced portfolio in hindsight:

$$S_n^* = \max_b S_n(b)$$

Theorem (TC-Gluss, 1986): There exists a causal portfolio $\hat{b}_i(X^{i-1})$ such that

$$\frac{1}{n} \log \hat{S}_n - \frac{1}{n} \log S_n^* \longrightarrow 0.$$

Theorem (TC, 1991): The universal portfolio

$$\hat{b}_i = \frac{\int b S_{i-1}(b) db}{\int S_{i-1}(b) db},$$

yields wealth

$$\hat{S}_n = \prod_{i-1}^{n} \hat{b}_i^t X_i,$$

satisfying $\frac{\hat{S}_n}{S_n^*} \geq \frac{c}{(\sqrt{n})^{m-1}}$, for all stock sequences $x_1, x_2, \ldots$ with bounded components.

Theorem: (Ordentlich and TC, 1996, 1998):

$$\frac{\hat{S}_n}{S_n^*} \geq \left(\sum_{k=0}^{n} \binom{n}{k} 2^{-nH(k/n)} \right)^{-1} \geq \frac{2}{\sqrt{n+1}}$$

Conclusions

Information Theory	**Investment**
Entropy rate $H = -\sum p(x)\log p(x)$	**Doubling rate** $W = \max_{\mathbf{b}} \int \log \mathbf{b}^t \mathbf{x} dF(\mathbf{x})$
AEP $-\frac{1}{n}\log p(X_1,\ldots,X_n) \longrightarrow H$	$\frac{1}{n}\log S^*(X_1,\ldots,X_n) \longrightarrow W^*$
Competitive optimality $P\{L^* \geq L + c\} \leq 2^{-c}$	$P\{S \geq cS^*\} \leq \frac{1}{c}$
Side information $H(X) \longrightarrow H(X\|Y)$, $I(X;Y)$	$W(X) \longrightarrow W(X\|Y)$, $\Delta(X;Y)$

$$\Delta \leq I$$

Universal data compression	**Universal portfolios**
$\frac{1}{n}\hat{L}_n \longrightarrow H$	$\frac{1}{n}\log \hat{S}_n \longrightarrow W^*$