The initial efficiency of investment for the general market

Elza Erkip[1] and Thomas M. Cover
elza@ece.rice.edu, cover@isl.stanford.edu

Abstract — **We investigate how one invests in the stock market X when there is a rate constraint on the side information V. The doubling function $\Delta(R)$ is the maximum increase in the doubling rate when V is described to the investor at rate R.**

The initial efficiency $\Delta'(0)$ is the largest possible increase in the doubling function per bit of description. We introduce the linearized maximal correlation and use it to provide a lower bound for the initial efficiency. We also show that the initial efficiency is bounded above by the square of the Hirschfeld-Gebelein-Rényi maximal correlation between the side information V and the market X. We can use these bounds to find the initial efficiency for $V = \mathbf{X}$ and for the horse race market.

I. Summary

Suppose one invests in the stock market $\mathbf{X} \in \mathbf{R}^m$ and wishes to maximize the doubling rate, the rate at which the wealth grows in repeated investments. The portfolio which achieves this maximum growth rate of wealth is denoted by $\mathbf{b}^*$. Now suppose a description of side information V is available to the investor. We let the *doubling function* $\Delta(R)$ be the maximum increase in the doubling rate when V is described at rate R.

After observing that the doubling function is a concave and nondecreasing function of R, we turn our attention to $\Delta'(0)$, the *initial efficiency*. Initial efficiency tells how much the first few bits about V increase the doubling rate for investment in market X. A single-letter characterization for the doubling function and the initial efficiency for the horse race market were established in [1, 2]. In this paper we extend these results by finding bounds on the initial efficiency for the general market.

First we define the Hirschfeld-Gebelein-Rényi maximal correlation between two random variables.

Definition: The *Hirschfeld-Gebelein-Rényi maximal correlation* $\rho_m(V, X)$ between two random variables V and X is defined by

$$\rho_m(V, X) = \sup Eg(V)h(X),$$

where the supremum is over all functions g and h such that

$$Eg(V) = Eh(X) = 0, \; Eg^2(V) = Eh^2(X) = 1.$$

Next we introduce another measure of dependence which we call the *linearized maximal correlation*.

Definition: The *linearized maximal correlation* between a random variable V and a random vector $\mathbf{X}$ is given by

$$\rho_L(V, \mathbf{X}) = \sup E(\mathbf{c}^t \mathbf{X} g(V)), \tag{1}$$

where the supremum is taken over all functions g such that $Eg(V) = 0$, $Eg^2(V) = 1$, and over all vectors $\mathbf{c}$ such that $\mathbf{c}^t\mu = 0$, $\mathbf{c}^t J\mathbf{c} = 1$, where $\mu = E(\mathbf{X})$ is the mean of $\mathbf{X}$ and $J = E\mathbf{X}\mathbf{X}^t$ is the second moment matrix of $\mathbf{X}$.

For ρ_L we extremize only over linear functions of $\mathbf{X}$ rather than all functions $h(\mathbf{X})$. The constraints $\mathbf{c}^t\mu = 0$ and $\mathbf{c}^t J\mathbf{c} = 1$ correspond to zero-mean and variance-one constraints. Hence

$$\rho_L(V, \mathbf{X}) \le \rho_m(V, \mathbf{X}).$$

Definition: The *active* stocks in a stock market $\mathbf{X}$ are the stocks included in the log optimum portfolio $\mathbf{b}^*$. That is, stock i is active only if $b_i^* > 0$. We denote the collection of active stocks by the vector $\mathbf{X}_A$.

The next theorem gives bounds on the initial efficiency for the general market.

Theorem 1 *For the general market*

$$\rho_L^2(V, \frac{1}{S^*}\mathbf{X}_A) \le \Delta'(0) \le \rho_m^2(V, \frac{1}{S^*}\mathbf{X})$$

where $S^* = \mathbf{b}^{*t}\mathbf{X}$ is the log optimum wealth.

Examples demonstrate that the bounds in Theorem 1 are not necessarily tight. However when $V = \mathbf{X}$, both of the bounds are equal to 1.

Corollary 1 *When $V = \mathbf{X}$,*

$$\Delta'(0) = 1,$$

provided that the active set consists of at least two stocks.

When the encoder can observe the stock vector outcomes, the initial efficiency achieves its largest possible value, which is 1.

Another important special case where the bounds coincide is the horse race market, where $\mathbf{X}$ takes values in the set of basis vectors of $\mathbf{R}^m$.

Corollary 2 *For the horse race market, the bounds in Theorem 1 coincide and the initial efficiency is given by*

$$\Delta'(0) = \rho_m^2(V, \mathbf{X}).$$

Therefore for the horse race market, the most efficient description pays off as the square of the maximal correlation in improving the doubling function.

References

[1] E. Erkip and T. M. Cover, "The efficiency of information in investment," submitted to *IEEE Transactions on Information Theory*. Stanford University Statistics Department Technical Report #93.

[2] E. Erkip and T. M. Cover, "The relation of description rate and investment growth rate," *Proceedings of 1995 IEEE International Symposium on Information Theory*, p. 73.

[1] E. Erkip is with Department of Electrical and Computer Engineering, MS-366, Rice University, 6100 S. Main, Houston, TX 77005. T. Cover is with the Departments of Statistics and Electrical Engineering, Stanford University, Stanford, CA 94305-4055.
This work was partially supported by NSF under Grant NCR-9205663, by JSEP under Contract DAAH04-94-G-0058 and by ARPA under Contract J-FBI-94-218-2.