Information Theory in Economics and Investment

Preston Hansen
University of Illinois at Chicago
E-mail: phanse4@uic.edu

***Abstract*—**

I. Introduction

Big results in information theory applied to economics and investing. Hopefully a look at open problems.

II. Information Theory in Economics

Many traditional economic models utilize the concept of rational agents to develop various results. While it is analytically convenient to assume that every consumer has access to perfect information about innumerable state variables representing their environment, it is intuitive that such models are not great (and indeed are often terrible) approximations of human behavior.

It turns out that information theoretic concepts can help to better approximate two realities that we observe in economic decision making - limited information processing capacity, and distrust or lack of access to a perfect model of the environment.

Sims [9] [10] aims to address the first point in his work on rational inattention.

A. Rational Inattention

In [9], rational inattention is developed for linear quadratic (LQ) control with multivariate state. The full setup for this case is lengthy, so for brevity the simpler optimization problem is given:

$$\min_q \left\{ E\left[(Y-X)^2\right] \right\}$$

Given the mutual information constraint $I(X;Y) \leq K$ where X is the environment and Y is some related signal.

For another perspective, the problem can be reframed as such: An agent aims to make a decision $u \epsilon U$ based on some signal $y \epsilon Y$ with the constraint that $I(X;Y) \leq K$. Assuming that Y is beyond the control of the agent, the agent aims to optimize the cost of their decision $P_{U|Y}$, and their observation channel $P_{Y|X}$. For an arbitrary cost function we have

$$U^* = \min_u \left\{ E\left[C(X,u)|Y\right] \right\}$$

and now an intuitive goal is to maximize the difference between this decision and the best decision possible without knowing Y:

$$P^*_{Y|X} = \max_{P_{Y|X}: I(X;Y) \leq K} \left\{ \min_u E\left[C(X,u)\right] - E\left[U^*\right] \right\}$$

From [10], development of this model has interesting implications. First, RI smooths responses and injects effective signal processing noise into responses to changes in the environment. This explains the observed phenomena of “price stickiness”, which is the resistance/delay in reaction to a change in price of something.

B. Applications of Rational Inattention

The Home Bias Puzzle (in trade) is based around the fact that domestic trade is typically much higher than international trade. In [6], it is argued that domestic and foreign investors start with asymmetric prior knowledge of home and abroad. Given a limited information processing capacity (an idea from rational inattention), and the assumption that specialization (further knowledge of domestic assets) has increasing returns, this paper explains why domestic investors would choose to ignore information about foreign investments and further increase the asymmetry of knowledge about domestic assets between themselves and foreign investors.

[6] defines the capacity constraint as

$$\prod_i \widehat{\Lambda_i} \geq \exp(-2K) \prod_i \Lambda_i$$

For capacity limit K, where Λ_i and $\widehat{\Lambda_i}$ are the variance of a risk factor i, and so the difference $(\Lambda_i - \widehat{\Lambda_i})$ is the reduction in variance in a risk factor i (i.e., learning about that risk factor).

It is then shown that given this constraint (and others), that home bias increases with information mobility.

In a similar vein, [5] finds that rational inattention helps to explain the discrepancy between expected and observed correlation of consumption between consumptions. Previous models predicted that consumers would hedge their domestic investments with international investments, but this correlation is measured to not be nearly as strong as predicted.

C. Robustness

Robustness exists in a similar space as rational inattention. The basic idea is that instead of composing an information capacity constraint, instead a parameter of uncertainty or distrust of the model of the environment.

To simplify the setup (skipping significant background in control theory), the problem setup can again be viewed similar to above. An agent aims to make a decision $u \epsilon U$ based on some signal $y \epsilon Y$. Assuming that the true distribution describing the environment is Q, the agent looks to optimize the expected cost of their deicison given the constraint $D(Q||P) \leq R$. Thus a lower value of R implies a higher level of certainty or trust of the model P of the real distribution Q.

D. Applications of Robustness

[2] uses the idea of robustness to model the price and quantity of items in an economy where decision makers have fear of model misspecification. They find that

In [4], the Sargent and Hansen's ideas of distrust of model parameters (desire for robustness) are used to develop optimal investment portfolios. Further analysis suggests that portfolios chosen based on the foundations of robustness to model uncertainty are more stable over time.

For a simpler (non-robust) introduction to portfolio theory and optimal portfolios, see the discussion of Cover's work below.

E. Akaike Information Criterion

In a highly influential paper [1], a metric for comparing statistical models (performing model fitting) is given that is defined using KL-divergence. The AIC metric is stated as

$$E_{p(X,\hat{\theta})}\left[\log\left(f(X|\hat{\theta}\right)\right]$$

which is shown to be equivalent to maximization of

$$E\left[\log\left(\frac{f(X|\hat{\theta})}{f(X|\theta)}\right)\right]$$

which we can recognize as the KL-Divergence $D(f(X|\theta)||f(X|\hat{\theta}))$. Here $\hat{\theta}$ is an estimate of a random variable of interest θ, and X is a random variable distributed according to $f(x|\theta)$. It is important to note that (as the paper's title says)this is not just the maximum likelihood principle, since the AIC depend on the true value of the variable of interest, θ_0, whereas the maximum likelihood would seek the function $\theta(\hat{Z})$ which maximizes $f(z|\theta)$, which does not depend on θ_0.

As the AIC can be written using KL-Divergence, it is not surprising that in terms of comparing statistical model fit, AIC is a relative measure. That is, if all of the models of interest are bad, or all of them are good, not much insight can be gained through use of AIC. It follows that many applications use AIC against known good models to provide useful insight.

Statistical model fitting has a huge array of applications, but for the purpose of this survey it is worth noting that AIC could be used to, for example, perform model fitting for credit risk of a given consumer.

As another example, [7] uses AIC (among other methods) to quantify the predictability of US stock market returns for a range of years.

III. Information Theory in Investment

Some of the results below deal with the value or influence of information on invenstment in a somewhat less information-theoretic centered approach (though in the abstract they still present a quantification of information, and so they are worth considering/contrasting with the other results).

A. Value of Information

B. Value of Information in Biology

The results above are extended and used to model a population as a financial portfolio in [8]. The growth rate of a population is, using virtually the same setup, bounded by the mutual information between a set of variables representing the environment, and some signal in the environment, $I(X_t;Y_t)$. Notably, this bound is shown to not hold when considering that in contrast to financial models, biological populations must process information at an individual level. This leads to the result that

$$I(X_t;Y_t) \leq I_p^{q_{env},q_{in}} \leq I(X_t;X_t')$$

i.e. in general the information gathered by any member of a population is less than the collective information gathered by the entire population. Note that X_t' here refers to a component of the environment defined at the population level.

C. Universal Portfolios

Consider a stock market vector given by

$$\boldsymbol{x} = (x_1, x_2, ..., x_m)^t$$

where x_i is the price relative of a given stock - the ratio of its closing to opening price - for a given day.

A portfolio is defined as

$$\boldsymbol{b} = (b_1, b_2, ..., b_m)^t, b_i \geq 0, \Sigma b_i = 1.$$

where each b_i represents the proportion of current wealth invested in stock i.

Finally, the factor by which wealth increases in a given investment period, S, can then be defined as $S = \Sigma b_i x_i$ A straightforward comparison can be made between any two investment strategies by comparing S for various scenarios.

In [3], a strategy is shown that achieves S_n (S over a sequence of n stock vectors) equal to,in the first order of the exponent, the best constant rebalanced portfolio S_n^*.

D. *Influence of Side Information in Investment*

additional result from optimal portfolios above

E. *Cost of Achieving the Best Portfolio in Hindsight*

References

[1] H. Akaike, *Information Theory and an Extension of the Maximum Likelihood Principle*, E. Parzen, K. Tanabe, and G. Kitagawa, Eds. New York, NY: Springer New York, 1998. [Online]. Available: https://doi.org/10.1007/978-1-4612-1694-0_15

[2] M. Cagetti, L. P. Hansen, T. Sargent, and N. Williams, "Robustness and pricing with uncertain growth," *The Review of Financial Studies*, vol. 15, no. 2, pp. 363–404, 2002. [Online]. Available: + http://dx.doi.org/10.1093/rfs/15.2.363

[3] T. M. Cover, "Universal portfolios," *Mathematical Finance*, vol. 1, no. 1, pp. 1–29, 1991. [Online]. Available: http://dx.doi.org/10.1111/j.1467-9965.1991.tb00002.x

[4] L. Garlappi, R. Uppal, and T. Wang, "Portfolio selection with parameter and model uncertainty: A multi-prior approach," *The Review of Financial Studies*, vol. 20, no. 1, pp. 41–81, 2007. [Online]. Available: + http://dx.doi.org/10.1093/rfs/hhl003

[5] W. Li, Y. Luo, and J. Nie, "Elastic attention, risk sharing, and international comovements," *Journal of Economic Dynamics and Control*, vol. 79, no. C, pp. 1–20, 2017. [Online]. Available: https://ideas.repec.org/a/eee/dyncon/v79y2017icp1-20.html

[6] S. V. Nieuwerburgh and L. Veldkamp, "Information immobility and the home bias puzzle," National Bureau of Economic Research, Working Paper 13366, September 2007. [Online]. Available: http://www.nber.org/papers/w13366

[7] M. H. Pesaran and A. Timmermann, "Predictability of stock returns: Robustness and economic significance," *The Journal of Finance*, vol. 50, no. 4, pp. 1201–1228, 1995. [Online]. Available: http://www.jstor.org/stable/2329349

[8] O. Rivoire and S. Leibler, "The value of information for populations in varying environments," *Journal of Statistical Physics*, vol. 142, no. 6, pp. 1124–1166, Apr 2011. [Online]. Available: https://doi.org/10.1007/s10955-011-0166-2

[9] C. A. Sims, "Implications of rational inattention," *Journal of Monetary Economics*, vol. 50, no. 3, pp. 665–690, April 2003. [Online]. Available: https://ideas.repec.org/a/eee/moneco/v50y2003i3p665-690.html

[10] ——, "Rational Inattention and Monetary Economics," in *Handbook of Monetary Economics*, ser. Handbook of Monetary Economics, B. M. Friedman and M. Woodford, Eds. Elsevier, 2010, vol. 3, ch. 4, pp. 155–181. [Online]. Available: https://ideas.repec.org/h/eee/monchp/3-04.html