THE CODING THEOREM AND ORDINARY LEAST-SQUARES MODELS IN ECONOMICS

W. D. O'NEILL

Department of Electrical Engineering and Computer Science, University of Illinois at Chicago,
P.O. Box 4348, Chicago, IL 60680, U.S.A.

Abstract—This paper considers ordinary least-squares (OLS) models of competitive markets wherein the independent variable represents trader information which should be a predictor of the market clearing price. The market context of the OLS model draws a parallel with an additive noise channel having noiseless feedback. The feedback link represents the information gained by the trader from observing present prices and using this information to modify his next round of market activity. Several models from the economic literature are shown to fit this prescribed information channel format and information theory constructs are used to interpret their economic results.

1. INTRODUCTION

We consider linear, memoryless, additive noise models as representations of information channels in the Shannon [1] sense. The supposition is made that the model population coefficients satisfy the conditions of an ordinary, least-squares (OLS), regression. This supposition places a variance constraint on the transmitted signals and leads to new results when a noiseless feedback path exists around the channel. The example of OLS models of competitive markets is especially amenable to information channel analysis because the OLS model and its attendant assumptions are, in an economic context, an integral part of the model.

Several levels of models are considered. Single and multivariable OLS models provide the theoretical structure of the economic results. The Grossman–Stiglitz [2], continuous state, additive Gaussian noise market model is shown to be the link between the concept of a competitive economic equilibrium and single and multiple access channels of information theory. The rational expectations hypothesis model of Muth [3] is then shown to be a discrete state information channel for which codes based on noiseless price feedback achieve channel capacity. The final economic model analyzed combines the multivariable OLS structure with multiple access channel theory as an explanation of the dependence of long-term interest rates on related economic variables.

2. THE OLS CHANNEL WITH FEEDBACK

The information source, whose samples are to be transmitted by the OLS channel, produces a random variable \( x_r \) every period \( T = 0, 1, 2, \ldots \), and \( x_r \sim N(0, \sigma^2_x) \). In any period the channel has the characterization

\[
y_k = \beta x_k + n_k, \quad k = 1, 2, \ldots, N
\]

and \( n_k \sim N(0, \sigma^2_n) \) independent of \( x_k \). The coding problem is to select the \( x_k \) based on \( \hat{x}_r \) and past estimates thereof so as to optimally encode \( x_r \) according to prescribed criteria. A constraint on (1) is that channel users do not know \( \beta \) but will use an OLS estimate of this parameter. From OLS theory the population parameters constraining the channel are

\[
\rho = E(xy)/\sigma_x \sigma_y, \quad \beta = \rho \sigma_y/\sigma_x, \quad \sigma^2 = \sigma^2_x(1 - \rho^2).
\]

From (2) the signal \( \beta x_k \) must satisfy for every \( k \) the constraint

\[
\text{Var}(\beta x_k) = \beta^2 \sigma^2_x = \rho^2 \sigma^2_x(1 - \rho^2).
\]
The well-known Shannon result for a Gaussian noise channel with signal variance constrained by $E > 0$ gives the capacity expression (in natural logarithms)

$$\begin{align*}
C &= \frac{N}{2} \log(1 + \frac{E}{\sigma^2}) \text{ nats per source sample} \\
&= \frac{1}{2} \log(1 + \frac{E}{\sigma^2}) \text{ nats per channel use.} \quad (4)
\end{align*}$$

Applying (2) and (3) to (4) yields for the OLS channel

$$C = -\frac{1}{2} \log(1 - \rho^2) \text{ nats per channel use.} \quad (5)$$

This result was apparently first found by Kullback [4] in a non-OLS context.

Feedback is employed by the channel user in such a way that with knowledge of the present source sample $\hat{x}_0$, his past estimate $\hat{x}_{k-1}$ of $\hat{x}_0$, and $y_k$ he can optimally select $x_k$ so that $\hat{x}_k \to \hat{x}_0$ as $k \to N$ in the mean-square sense. This procedure will produce optimal OLS encoding.

3. OPTIMAL OLS CODES

It is well-known [5] that the minimum distortion, $D_{\text{min}}$, any coding scheme can achieve for the linear additive noise channel satisfies

$$R(D_{\text{min}}) = C \quad (6)$$

where $R(\cdot)$ is the rate distortion function of the source. When $x$ is Gaussian,

$$R(D) = \max \{0, \frac{1}{2} \log(\sigma^2/D)\}. \quad (7)$$

From (4) and (7),

$$D_{\text{min}} = \sigma^2(1 - \rho^2)^N. \quad (8)$$

An OLS channel code is optimal if $D_{\text{min}}$ in (8) is achieved by the code. $D_{\text{min}}$ is of practical concern in economic markets since

$$D_{\text{min}} = \min [\text{var}(\hat{x}_0 - \hat{x}_k)] \quad (9)$$

where the minimum is over choices of $\hat{x}_k$. We note that if the distortion is small enough it can be said that all of the market traders have “caught on” to the latest information driving the market.

The optimal estimates $\hat{x}_k$ are found by a trivial extension of the Schalkwijk–Bluestein [6] results. Define

$$\gamma_k = \beta x_k \quad (10a)$$

and the sequences $Q_k$ and $B_k$ by the recursive formulae

$$\begin{align*}
\gamma_k &= Q_k(\hat{x}_0 - \hat{x}_{k-1}), \quad \hat{x}_0 = 0, \quad k = 1, 2, \ldots, N; \\
\hat{x}_k &= \hat{x}_{k-1} + B_k(\gamma_k + n_k), \quad k = 1, 2, \ldots, N. \quad (10b)
\end{align*}$$

The sequence $Q_k$ is selected so that constraint (3) is met for all $k$ and $B_k$ is selected so that $\hat{x}_k$ is a minimum mean-square-error estimate of $\hat{x}_0$ given $\gamma_m + n_m, m = 1, 2, \ldots k$. These conditions immediately yield from the Schalkwijk–Bluestein formulae

$$\begin{align*}
D_k &= \sigma^2(1 - \rho^2)^k, \quad k = 1, 2, \ldots, N; \\
Q_k &= \sigma_x \rho (1 - \rho^2)^{-k/2}/\sigma_x, \quad k = 1, 2, \ldots, N; \\
B_k &= \sigma_x \rho (1 - \rho^2)^k/\sigma_x, \quad k = 1, 2, \ldots, N. \quad (11)
\end{align*}$$

From (8) and the first equation of (11) evaluated at $k = N$ the optimality of the recursive codes in (10b) is confirmed.
4. MARKET APPLICATIONS

The work of Grossman and Stiglitz [2] assumes that market agents hold differing information sets about future prices. Informed agents pay a fee for "the latest forecast" of future prices while uninformed agents select to observe only market clearing prices to uncover the costly information. The authors argue: if the market conveys information perfectly, then there is no incentive to purchase information. Conversely, if all agents are uninformed it obviously pays some agents to become informed. Hence, perfect market information transfer through prices implies a competitive price equilibrium does not exist. The authors further argue that a price equilibrium has the property that prices reflect the costly information of informed traders but only partially so.

We briefly develop the Grossman–Stiglitz model. Agents can choose a safe asset with a fixed return $R$ and/or a risky asset with a future return $u$, and by paying a fee $c$ they can observe a sample of $\theta$ where

$$u_i = \theta_i + \epsilon_i, \quad (12)$$

with $\epsilon \sim N(0, \sigma_\epsilon^2)$ independent of $\theta \sim N(0, \sigma_\theta^2)$. Present price of the risky asset is $P$, and all agents maximize the expected value of an exponential utility function having a coefficient of absolute risk aversion $\alpha$. Uninformed agents only observe $P$, and no agents observe the per capita supply of the risky asset $x \sim N(0, \sigma_x^2)$ independent of $\theta$ and $\epsilon$. Maximizing expected utility results in the informed agent demand for risky assets

$$d_i = (\theta_i - RP_i)/\alpha \sigma_\theta^2, \quad (13)$$

with a similar expression (with $\theta$, replaced by $E[u | P]$ and $\sigma_\theta^2$ by $\text{var}(u | P)$) for uninformed demand. A parameter $\lambda$, $0 \leq \lambda \leq 1$, is defined as the ratio of informed to the total number of agents. The following economic description of the product $ac$ suggests that it is the rate distortion function of the source $\theta$. If $c$ is high, and there are informed agents, then their act of purchasing the information they are signaling that $\theta$ is a reliable measure of future returns and conversely. If $\alpha$ is large then any agent takes a large market position thereby revealing this diminished risk aversion. Such positions signal directly his regard for the reliability of $\theta$ if he has purchased a sample thereof and indirectly (through $P$) if not. The converse holds if $\alpha$ is small.

A stationary, competitive equilibrium is said to exist if $0 < \lambda < 1$ and $z$, the ratio of the expected utility of informed agents to that of the uninformed, is unity or if $z > 1$ when $\lambda = 0$ or if $z < 1$ when $\lambda = 1$. Grossman and Stiglitz show that

$$z = e^{\alpha} \left[ \frac{\text{var}(u | \theta)}{\text{var}(u | P)} \right]^{1/2}, \quad (14)$$

with

$$e^{\alpha} = \left(\frac{m + 1}{nm + m + 1}\right)^{1/2} \quad (13)$$

and

$$n = \frac{\sigma_\theta^2}{\sigma_\epsilon^2}, \quad m = \left[\frac{a \sigma_\epsilon^2}{\lambda \sigma_x^2} \frac{\sigma_\theta^2}{\sigma_x^2}\right].$$

Their main equilibrium result is: (a) if $\sigma_\epsilon^2 = 0$, an equilibrium never exists; (b) if $\sigma_x^2 = 0$, an equilibrium does not exist if and only if

$$e^{\alpha} < \left(1 + \frac{\sigma_\theta^2}{\sigma_\epsilon^2}\right)^{1/2}. \quad (14)$$

We observe that informed traders use the channel given by (12) while uninformed traders use the OLS channel

$$P_i = xu_i + n_i. \quad (15)$$
Note that the optimal OLS codes developed above apply to this channel. Let $C_{\text{th}}$ be the capacity of (12) and $C_{\text{up}}$ the capacity of (13). Taking logarithms of (13) under the condition $z = 1$, $0 < \lambda < 1$ yields

$$ac = C_{\text{th}} - C_{\text{up}}.$$  

Equations (6) and (7) applied to the $(\theta, u)$ channel allow a simple proof of the Grossman-Stiglitz equilibrium result.

**Proof.** Part (a): if $\sigma_{\lambda}^2 = 0$, $D_{\text{min}} = \sigma_{\lambda}^2 / 1 + n = 0$ so the $(\theta, u)$ channel is perfectly revealing and therefore $\lambda = 0$. But $z(0) = 0$ for $n = \infty$ so there never is an equilibrium.

Part (b): if $ac < \frac{1}{2} \log(1 + n)$ the $(\theta, u)$ channel is not perfectly revealing because $D > D_{\text{min}} > 0$. Thus, it pays to become informed so $\lambda > 0$. But $\lambda > 0$ and $\sigma_{\lambda}^2 = 0$ imply $m = 0$ and therefore $C_{\text{up}} = C_{\text{th}}$. Thus $z(\lambda) = 1$ implies $ac = 0$. Hence there is no equilibrium if information is costly. The converse is readily proved with a contradiction argument.

The next market application is the landmark model of Muth [3] with which he introduced the rational expectation hypothesis (REH). The REH is this: market agents are assumed to hold certain subjective probability distributions of price outcomes based upon the market price they expect to prevail in period $t$ conditional on available information through period $t-1$. The alignment of these subjective distributions with the realized distributions of the market place is called the rational expectations hypothesis. Included in the agent information set is the $(t-1)$st market clearing price. This establishes a noiseless feedback path around the market channel. The Muth model has supply and demand sectors and an equilibrium condition,

\[
\begin{align*}
\text{DEMAND} & \quad D(t) = -\beta p(t) + D_0 + v(t) \\
\text{SUPPLY} & \quad S(t) = \gamma p_e(t) + S_0 + w(t) \\
\text{EQUILIBRIUM} & \quad D(t) = S(t) + r(t). \quad (17)
\end{align*}
\]

$p(t)$ is the market price in period $t$, $p_e(t)$ is the market price expected in period $t$ conditional on available information through period $t-1$, and $v$, $w$ and $r$ are zero mean, independent and individually identically distributed random error terms. The REH applied to (17) implies that agents act so as to achieve the equilibrium result

\[E[p(t)] = p_e(t) = \bar{p},\]  

where $\bar{p}$ is the static ($v = w = r = 0$) equilibrium price

\[\bar{p} = (D_0 - S_0) / (\gamma + \beta).\]  

Equations (17) and (19) readily produce the reduced form model

\[p(t) - \bar{p} = \frac{\gamma}{\beta} [p_e(t) - \bar{p}] - \frac{u(t)}{\beta},\]  

where the density of $u$ is the convolution of the density of $w + r$ with that of $v$. The supply and demand sectors are subject to random disturbances in $D_0$, $S_0$, $\gamma$ and $\beta$ so that $\bar{p}$ is a sample of a random variable with $M$ possible values. This is the subjective distribution of the agents. Over a given trading epoch $\bar{p}$ is fixed. The agents’ expectations are realized and an REH equilibrium exists if it is possible, using price measurements alone, to estimate $\bar{p}$ with arbitrarily small error probability. Specifically, if $\bar{p}$ is one of $M$ possible price expectations and this $\bar{p}$ elicits a vector $p = [p(1), p(2), \ldots, p(N)]$ of market clearing prices then a REH equilibrium exists if there exists a sequence of estimates $p_e = [p_e(1), p_e(2), \ldots, p_e(N)]$ depending only on $p$ such that for any $\bar{p}$,

\[\lim_{N \to \infty} \Pr[p_e(N) \neq \bar{p}] = 0.\]  

We prove below that this definition of an REH equilibrium includes (18) since we show that

\[\lim_{N \to \infty} [p(N)] = p \lim_{N \to \infty} p_e(N) = \bar{p}.\]
Define \( x(t), y(t) \) and \( n(t) \) as
\[
\begin{align*}
x(t) &= p_e(t) - \bar{p}, \\
y(t) &= \bar{p} - p(t), \\
n(t) &= -\frac{1}{\beta} u(t),
\end{align*}
\]
and (21) becomes the OLS model
\[
y(t) = \frac{\bar{y}}{\bar{p}} x(t) + n(t).
\] (24)

The results of Schalkwijk and Kailath [7] are used to generate the optimal coding scheme for (24) that achieves an REH equilibrium. The stochastic approximation algorithm summarized below is due to Robbins and Monro [8] and is the basis of the asymptotic estimates derived by Schalkwijk and Kailath.

If
\[
\begin{align*}
x(t + 1) &= x(t) - \frac{1}{a t} y(t), \\
y(t) &= a[x(t) - \bar{y}] + z(t),
\end{align*}
\]
with \( z(t) \) zero mean and i.i.d., then
\[
x(t + 1) \sim N[\theta, \sigma_z^2/a^2 t].
\] (25)
The definitions (23) and (24) give directly
\[
p_e(t) \sim N[\bar{p}, \sigma_e^2/\gamma^2 t]
\] (26)
and the difference equation
\[
p_e(t + 1) = p_e(t) + \frac{\beta}{\gamma t} [p(t) - \bar{p}].
\] (27)

Equations (20) and (27) are, respectively, the instantaneous price clearing mechanism and price feedback rule by which agents adjust their price expectations to bring prices to equilibrium at \( \bar{p} \). Observe that (21) is obviously true by (26). Solving (20) and (27) simultaneously,
\[
p(t + 1) = \bar{p} + \frac{1}{\beta} \left[ \frac{1}{t} \sum_{i=1}^{t} u(i) - u(t + 1) \right],
\]
so that (22) is proved. It is shown in O’Neill [9] that the feedback coding rule given in (27) achieves capacity for the Muth channel.

The final economic model we present relates changes in the interest rate of long-term bonds to the difference between long- and short-term rates (the spread) and the premium paid (capital gain) for holding a long-term versus a short-term bond, Mankiw [10]. The REH theory states that the linear (approximate) relationship should be
\[
y(t) = \beta_1 x_1(t) + \beta_2 x_2(t) + u(t),
\] (28)
with \( y(t) \) the long-term interest rate difference, \( x_1(t) \) the premium, \( x_2(t) \) the spread, and \( u(t) \) a zero mean, Gaussian, i.i.d. random disturbance term. This relationship is evidently a three variable, linear regression model with the well-known population constraints
\[
\begin{align*}
\beta_1 &= \left[ \frac{\rho_1 - \rho_2 \rho_{12}}{1 - \rho_{12}^2} \right] \sigma_2 \sigma_1, \\
\beta_2 &= \left[ \frac{\rho_2 - \rho_1 \rho_{12}}{1 - \rho_{12}^2} \right] \sigma_2 \sigma_1, \\
\sigma_u^2 &= \sigma_z^2 [1 - \rho^2],
\end{align*}
\] (29)
where $\rho_1 = \rho_{x_1x_1}$, $\rho_2 = \rho_{x_2x_2}$, $\rho_{12} = \rho_{x_1x_2}$, $\sigma_1 = \sigma_{x_1}$, $\sigma_2 = \sigma_{x_2}$ and $\rho$ is the correlation coefficient of the regression. The economic question asked of the model is: does the change in long-term rates simultaneously reflect information in the capital markets ($x_1$) and the bond markets ($x_2$)? A nuance in this question arises because the interest rate of long-term bonds, though nominally fixed, changes with the price of bonds as determined in the capital markets. When $u(t)$ is as specified then (29) is a white, Gaussian, multiple access channel with noiseless feedback since all investors can observe published rates on long-term bonds. Sources $x_1$ and $x_2$ compete for information transfer through the market. Optimal codes operating in the achievable rate region of the channel should permit users to estimate $x_1$ and $x_2$ reliably from measurements on $y(t)$ alone. Ozarow [11] has found the capacity region for the model $\beta_1 = \beta_2 = 1$. The OLS constraints

\[
\text{var} [\beta_1, x_1] = \left( \frac{\rho_1 - \rho_2 \rho_{12}}{1 - \rho_{12}^2} \right) \sigma_1^2 \Delta p_1,
\]

\[
\text{var} [\beta_2, x_2] = \left( \frac{\rho_2 - \rho_1 \rho_{12}}{1 - \rho_{12}^2} \right) \sigma_2^2 \Delta p_2,
\]

(30)

give the capacity region for (28) as

\[
0 \leq R_i \leq \frac{1}{2} \log \left[ 1 + \rho_i \frac{1 - \rho_{12}^2}{\sigma_i^2} \right], \quad i = 1, 2;
\]

\[
0 \leq R_1 + R_2 \leq \frac{1}{2} \log \left[ 1 + \rho_1 + \rho_2 + \rho_{12}\sqrt{\rho_1 \rho_2} \right] \sigma_u^2.
\]

(31)

The last equation of (29) indicates this region depends only on $\rho_1$, $\rho_2$, $\rho_{12}$ and $\sigma_u^2$. Specific capacity producing codes are given by Ozarow for $\beta_1 = \beta_2 = 1$ and are shown to depend on a nonlinear, recursive, estimate of $\rho_{12}$.

5. SUMMARY

We have shown that Shannon channels with feedback are significant models of economic markets. Known results for feedback channels, extended to include OLS constraints, provide reputable models of competitive markets. The rational expectations hypothesis is shown to be expressible by Shannon's coding theorem when competitive equilibrium is cast in terms of market agents being able to predict input information from market clearing prices.

It is evident that the theory of competitive markets can benefit from the results presently being produced by multiple access channel research. Particularly intriguing is the economic interpretation of multiple access codes. Is the $n$-user multiple access channel a useful model of the $n$-agent competitive market where every agent can be both a seller and a buyer of commodities depending on how he encodes his personal information or on how he decodes his competitor's information from market clearing prices?

REFERENCES