Control and Cybernetics

vol. **32** (2003) No. 4

Information pricing for portfolio optimization

by

Tadeusz Banek[1,2] and Roman Kulikowski[2]

[1]Department of Operations Research
Faculty of Management
Technical University of Lublin
Nadbystrzycka 38, 20-618 Lublin, Poland

[2]Systems Research Institute PAN
Newelska 6, 01-447 Warsaw, Poland

Abstract: We consider the following problem: is there a rational or fair price for the reports made by analysts, experts, investor advisers concerning the rate of return (RR) of investments? We define the notion of the value of information included in the family of probability distributions of the RR. Next, we illustrate this notion for a linear-quadratic utility function.

Keywords: value of information, martingale, portfolio optimization.

1. Introduction

According to a popular and almost universally accepted opinion, information is one of the most important and desired goods. In the world of economy a person possessing information has an advantage over worse informed rivals. That person can use opportunities not known to others, or avoid errors which they will probably make. The larger and better the information, the greater the possible profit of its user. That is why one of ways of risk reduction is enlarging information. It is done at the price of expensive investigation, provided, of course, that its cost does not exceed the resulting advantages. However, even though the saying "time is money" is widely accepted, the saying "information is money" is not in common use. Moreover, the problem of money-information exchange and substitution, and, in general, pricing and trading information is almost absent in the economic literature (though prominent economists such as e.g. Kenneth Arrow tried to face those problems, see for instance Arrow, 1970) and in the operations research literature. This is clearly in conflict with know how to estimate or measure that value or compare it to other values. In the mathematical literature this problem has been investigated from a different angle. See for instance the papers of M.H.A. Davis, M.A.H. Dempster and R.J. Elliott (1991), M.H.A. Davis, I. Karatzas (1994) and references given there.

Definitions of measures of information characterize the amount of information contained in a message with the known probability distribution, completely ignoring what the information pertains to and how, and with what result, it will possibly be used in a decision process. In order to understand better this complicated matter, to limit the area of considerations and to pose some questions, let us consider a stock exchange.

What value does the information included in so-called "historical data" have for a stock-exchange investor? Here we have in mind the information included in data pertaining to previous economic performance of stock exchange companies, information about performance of their competitors, co-operators, and other companies in the branch, not necessarily present on the stock exchange. How will the value of information change if we increase the set of data by taking into account more and more companies and branches and go farther back in the past? Obtaining information interesting for our investor from an increasing set of data (as we simultaneously go back into the past and increase the field of observation by analyzing data from larger and larger economic areas) will probably require more efficient methods of data analysis, better computer hardware and software, but mainly, and maybe most of all, a coherent economic theory explaining and systematizing the registered data; a theory, which is not only internally consistent but, also coherent with data. This is a domain of activity not only for theoreticians of economics but also for analysts, experts and tax advisors. Investigations if this kind are probably expensive. First of all they contribute a lot to our understanding of economic phenomena and processes and that is why they are indispensable. On the other hand, results of such investigations constitute attractive material for investors, enabling them, e.g. to estimate more precisely the returns of specific companies, and thus to make better decisions.

Is there any relationship between money spent on work of theoreticians, analysts, experts and investment advisors and the advantage for the investor financing these investigations? The purpose of this paper is to attempt answering some of the above questions. We focus in particular on a fundamental question: what value for an investor does the information about statistics of returns of specific companies present and how does this value change when the statistics change?

The plan of the paper is as follows. In Section 2 we present assumptions concerning the way of parametrization of statistics of distributions. In Section 3 we consider a decision problem for an investor having an additional option of purchasing information just before making an investment decision. Analysis of this "thought experiment" allows us to define a notion of information value.

Section 5 we briefly present conditions of trading information.

This paper continues and extends the ideas and results published earlier in Banek (2000, 2002) and discussed later in Banek (2001).

2. Assumptions

For simplicity, we assume in the entire paper that the family of probability distribution functions obtained as a result of analytic investigation concerning the vector ξ of returns is Gaussian; as is commonly known, this reduces the problem to investigating only two statistics: the vector m of mean values and the covariance matrix Q.

Let us assume that the above statistics are parametrized by $t \geq 0$, i.e., $\{(m_t, Q_t); 0 \leq t \leq T\}$, $T < \infty$.

In the classical models of portfolio selection (Markowitz, Roy) an investor possessing a cash amount $M > 0$ selects a portfolio $x = col\,(x_1, ..., x_n)$, $x_1 + ... + x_n = M$, whose expected return and covariance are $\langle x, m\rangle$, $x^T Q x$ respectively, where $\langle , \rangle$ means scalar product, and T stands for transposition, i.e., $x^T = (x_1, ..., x_n)$. In this paper we shall consider an extended version of the problem in which the investor has an extra option: he can buy the results $\{(m_s, Q_s); 0 \leq s \leq t\}$, $(m_0, Q_0) = (m, Q)$ of experts' investigation at a price c_t, under the conditions which we shall describe later. Using the "best" estimators (m_t, Q_t) of ξ, he will select the portfolio whose expected return and covariance are mow

$$\begin{aligned} E_t \langle x, \xi\rangle &= \langle x, E_t\xi\rangle = \langle x, m_t\rangle, & (1)\\ E_t \langle x, \xi - m_t\rangle^2 &= E_t \langle x, \xi - m_t\rangle \langle \xi - m_t, x\rangle \\ &= E_t x^T [\xi - m_t][\xi - m_t]^T x \\ &= x^T E_t \left([\xi - m_t][\xi - m_t]^T\right) x \\ &= x^T Q_t x & (2) \end{aligned}$$

where E_t means the conditional expected value relative to the P measure on the probability space $(\Omega, \mathcal{F}, P)$ on which all random objects in this paper are defined, i.e., for a bounded Borel function f we have

$$E_t [f(\xi)] = E\{f(\xi) | \mathcal{F}_t\}$$

where $\mathcal{F}_t$ means a sigma-sub-field of $\mathcal{F}$ containing all data used by analysts to evaluate $\{(m_s, Q_s); 0 \leq s \leq t\}$. We assume that $\mathcal{F}_s \subset \mathcal{F}_t$ for $s \leq t$, and that $\mathcal{F}_s = \cap_{t>s} \mathcal{F}_t$ for $0 \leq s \leq T$.

REMARK 1 *Note that the family $(\mathcal{F}_t)$ depends very much on the experts', statisticians', or investor advisers' choices and decisions as to how to select material from historical data, and how to parametrize it. The parameter t can be inter-*

included in $\mathcal{F}_t$, or their salary, or something else. The choice: go deeper into the past, or increase the area of observation, or in other words: what data have to be included in $\mathcal{F}_t$ (a control problem for data mining) is of independent interest, but will not be pursued here.

Let us assume that the parametrization of the statistics $\{(m_t, Q_t); 0 \leq t \leq T\}$ was selected in such a way that for $t \geq s$, the increase $I_F(t) - I_F(s)$ of the Fisher measure

$$I_F(t) = \int_{R^n} \frac{\|\nabla \rho_n(t,x)\|^2}{\rho_n(t,x)} dx, \tag{3}$$

where

$$\rho_n(t,x) = \frac{1}{\sqrt{(2\pi)^n \det Q_t}} \exp\left\{-\frac{1}{2}(x - m_t)^T Q_t^{-1}(x - m_t)\right\}, \tag{4}$$

is proportional to $t - s$, i.e.,

$$I_F(t) - I_F(s) \sim t - s. \tag{5}$$

Because for the density ρ_n given by formula (4) we have

$$I_F(t) = TrQ_t^{-1}, \tag{6}$$

it was proposed in (Banek, 2000) to write

$$Q_t^{-1} = Q_0^{-1} + \int_0^t H_s^T H_s ds \tag{7}$$

where $\{H_s; s \in [0,T]\}$ is an $\mathcal{F}_t$-adapted stochastic process with values in the set of square-integrable matrices with elements (h_s^{ij}), i. e.,

$$P\left(\int_0^T (h_s^{ij})^2 ds < \infty\right) = 1.$$

If $H_s = H$ and $TrH^T H > 0$, then

$$\begin{aligned} I_F(t) - I_F(s) &= Tr\left(Q_t^{-1} - Q_s^{-1}\right) = Tr\int_s^t H^T H ds \\ &= (t-s)\, TrH^T H. \end{aligned} \tag{8}$$

In general

In conclusion, in order to satisfy the requirement (5), it is necessary to assume that Q_t^{-1} is given by formula (7), or, equivalently, to assume that

$$Q_t = \left(I + Q_0 \int_0^t H_s^T H_s ds\right)^{-1} Q_0, \tag{10}$$

$$\int_0^t Tr\left(H_s^T H_s\right) ds > 0 \text{ for } t \geq 0, \tag{11}$$

$$Q_0 = Q_0^T, Q_0 > 0 \text{ (positive definite)} \tag{12}$$

Let us remark that matrix (10) satisfies the following Riccati differential equation:

$$\frac{d}{dt} Q_t = -Q_t^T H_t^T H_t Q_t, \quad Q_{t=0} = Q_0, \tag{13}$$

and that the derivative of the portfolio's variance is equal to

$$\begin{aligned} \frac{d}{dt} x^T Q_t x &= -x^T Q_t^T H_t^T H_t Q_t x \\ &= -\left\|H_t Q_t x\right\|^2 \leq 0 \end{aligned} \tag{14}$$

Property (14) shows that the family $\{(m_t, Q_t); 0 \leq t \leq T\}$ of statistics is well (correctly) parametrized in the sense that the greater the value of the parameter, the smaller the portfolio variance (i.e., risk in Markowitz theory).

For the statistic m_t we see that it is a martingale relative to the sigma-subfield $\mathcal{F}_t$, since $m_t = E_t[\xi]$, which means that for $0 \leq s \leq t \leq T$ we have

$$E_s[m_t] = E[m_t | \mathcal{F}_s] = m_s. \tag{15}$$

This statement is supported by the argument that there is no reason to suppose a priori that during analytical (and statistical) investigations a trend distinguishing the conditional returns m_t from the value m_0 will appear.

We adopt in this paper the following convention. In order to distinguish deterministic functions from stochastic processes, we shall always use the notation $c(t)$, $Q(t)$, $H(t)$, etc., for functions and c_t, Q_t, H_t for processes.

EXAMPLE 1 *For a matrix* $H = (h^{ij})$, $h^{ij} \in L_2[0,T]$ *define the observation process*

$$y_t = \int_0^t H(t) \xi ds + w_t \tag{16}$$

where $\{w_t; 0 \leq t \leq T\}$ *is a Wiener process independent of* ξ. *Then, from the Kalman-Bucy filtering theory (see Liptser, Shiryaev, 1977, for example) it follows that the conditional distribution of* ξ *is given by the formulae*

$$P(\xi \in A | y_s, 0 \leq s \leq t) = \int \ldots$$

where $\rho(t,z)$ is Gaussian with the mean

$$m_t = m_0 + \int_0^t Q(s)\, H^T(s)\, d\nu_s$$

with the innovation process (a Wiener process)

$$\nu_t = \int_0^t \left[dy_s - H(s)\, m_s ds\right]$$

and the covariance matrix

$$\frac{d}{dt} Q(t) = -Q^T(t)\, H^T(t)\, H(t)\, Q(t), \quad Q(0) = Q_0.$$

In this example $F_t = \sigma\{y_s; 0 \le s \le t\}$, and (16) have a nice interpretation as a simple linear econometric model build by the experts.

EXAMPLE 2 *Denote by $C([0,T], R^m)$ a space of continuous function on $[0,T]$ with values in R^m, and introduce a matrix $H = \left(h^{ij}\right)$, where $h^{ij} :.[0,T] \times C([0,T], R^m) \to R$, are bounded and non-anticipative, i.e., if y^1, y^2 belong to $C([0,T], R^m)$, and $y^1(s) = y^2(s)$ for $0 \le s \le t$, then $h^{ij}\left(s, y^1\right) = h^{ij}\left(s, y^2\right)$ for $0 \le s \le t$. Define the observation process*

$$y_t = \int_0^t H_s \xi ds + w_t \tag{17}$$

where $H_t = \left(h^{ij}(t,y)\right)$, with $\{w_t; 0 \le t \le T\}$ a Wiener process independent of ξ. Then, from the Liptser-Shiryaev filtering theory (see Liptser, Shiryaev, 1977) it follows that the conditional distribution of ξ is given by the formulae

$$P(\xi \in A\,|y_s; 0 \le s \le t) = \int_A \rho(t,z)\, dz$$

where $\rho(t,z)$ is Gaussian with the mean

$$m_t = m + \int_0^t Q_s H_s^T d\nu_s$$

with the innovation process (a Wiener process)

$$\nu_t = \int_0^t \left[dy_s - H_s m_s ds\right]$$

and the covariance matrix

$$\frac{d}{dt} Q_t = -Q_t^T H_t^T H_t Q_t, \quad Q_{t=0} = Q_0.$$

As above $F_t = \sigma\{y_s; 0 \le s \le t\}$, and (17) describes a linear (more advanced) econometric model build by the experts. The explanatory variables h^{ij} used in

3. Value of information

Let us consider the situation of an investor maximizing his utility function

$$U(x, m, Q, M) \tag{18}$$

subject to the constraints

$$\Phi_i(x, m, Q, M) = 0,\ i = 1, \ldots, p, \tag{19}$$

where x, m, Q, M denote, as before, an investment vector, the vector of mean values, the covariance matrix and the amount of cash, respectively. Let us assume that a solution $x^\star$ of the above problem exists and set

$$W(m, Q, M) = U(x^\star, m, Q, M). \tag{20}$$

Next, assume that the investor has the option of purchasing information before making investment decisions: he can buy a segment

$$S_t \triangleq \{(m_s, Q_s); 0 \le s \le t\}$$

at a price c_t, where $\{c_t;\ t \ge 0\}$, $c_0 = 0$, is a G_t-adapted stochastic process on $(\Omega, \mathcal{F}, P)$ with continuous, increasing realizations. Here $G_t = \sigma\{(m_t, Q_t);$ $0 \le t \le T\}$ is a sigma-sub-field of $\mathcal{F}_t$.

Remark 2 *Since* $\mathcal{F}_t \supset G_t$, *possibly* $\mathcal{F}_t \supsetneq G_t$, *the condition that* c_t, *is* $\mathcal{F}_t-$, *instead of* G_t- *adapted, could mean that* c_t *contains some additional information not included in* S_t.

Technically the transaction can be performed as follows: an information seller shows to an information buyer a band with a record of the statistics forming the segment S_t, along with a record of realization of the relevant price $\{c_t;\ t \ge 0\}$. The buyer pays c_t and immediately decides whether to look at the band (and to pay) any more or to stop the process. Such a procedure is necessary because of the specific properties of information as an object of trade (e.g. it is impossible to see a piece of information and then to refuse purchasing it).

The investor, viewing segments $\{S_t; t \ge 0\}$, solves for each t a problem (18)(19) with the couple (m, Q) replaced by (m_t, Q_t), and the cash amount M replaced after payment by $M - c_t$.

A generalized investor problem has the following form: find

$$\sup E\left[U(x, m_\tau, Q_\tau, M - c_\tau)\right] \tag{21}$$

subject to the constraints

$$\Phi_i(x, m_\tau, Q_\tau, M - c_\tau) = 0, \ i = 1, ..., p, \tag{23}$$

where $\tau \geq 0$ is a Markov stopping time relative to G_t, i.e., $\{\tau(\omega) \leq t\} \in G_t$ for $t \in [0, T]$.

Let us notice that for fixed $t \geq 0$,

$$\begin{aligned} & \sup\left[U(x, m_t, Q_t, M - c_t); \Phi_i(x, m_t, Q_t, M - c_t) = 0, \ i = 1, ..., p\right] \\ = \ & W(m_t, Q_t, M - c_t). \end{aligned}$$

So, the problem (21)(22)(23) reduces to the following optimal stopping problem:

PROBLEM 1 *Find*

$$\sup E\left[W(m_\tau, Q_\tau, M - c_\tau); \tau \geq 0, 0 \leq c_\tau \leq M\right] \tag{24}$$

relative to G_t -Markov stopping time $\tau \geq 0$, satisfying the condition

$$0 \leq c_\tau \leq M.$$

DEFINITION 1 ***(Value of Information)*** *Let $V = \{V_t, t \geq 0\}$, $V_0 = 0$, be a G_t-adapted stochastic process on (Ω, F, P) with continuous, non-decreasing trajectories, such that the process*

$$u_t \triangleq W(m_t, Q_t, M - V_t) \tag{25}$$

is a G_t-martingale. The random variable V_t, which is the value of the process V for the parameter t, is called the value of the information included in the segment $S_t = \{(m_s, Q(s)); 0 \leq s \leq t\}$ defined for the investor (21)(22)(23), or equivalently (24).

REMARK 3 *If the charge for the information included in the segment S_t, equal to c_t, is set correctly then we have a "fair game" between the seller and the buyer, in the sense that the buyer gains nothing (on average) and loses nothing (on average). Moreover, his (average) anticipations of future growths and drops of the criteria index are also equal to zero. This means that we have to do with an equivalent information-money exchange. Such a "just" charge is called the information value and denoted by V_t. In the language of mathematics that means that the process u_t is a G_t-martingale.*

Considering the investor's problem of purchasing information from historical data we see that there are two possible extreme cases and many others in between. The first is when the analysts offering the segments $\{S_t; t \geq 0\}$ have done their job earlier. In this case the square error matrix Q_t of estimation of ξ is known to them and can, or even should be known to buyers in order to convince them of the quality of the job done. Consequently, the functions $t \to Q(t)$

the analysts' job will be done in the future and will continue as a function of financial support. In this case, $t \to Q_t$ and $t \to c_t$ are random processes, c_t being $\sigma\{(m_s, Q_s); 0 \leq s \leq t\}$ adapted. Between these two extreme cases there are also possible "mixed cases", for instance a part of the job, say S_t, is already done for some horizon $t > 0$, i.e., $Q(s)$ is a known deterministic function on $[0, t]$ and $\{Q_s; s \geq t\}$ is a stochastic process with initial value $Q_t = Q(t)$, and so on.

Some remarks are now in order. Why would the analysts sell S_t if they actually have S_T, $T > t$, available? In most countries this is known as withholding material information by an investment adviser and is expressly prohibited by regulatory bodies.

In order to explain our reasoning, let us consider many advisers, say N, working in different offices (shops). Each k-th offers for sale $S^k \equiv S_{t_k}$ only. The advisers work on the same material (available to the public opinion). They work by using standard procedures, so the possible differences in the results obtained come from the possible differences, for instance, in computational power, which depends on the equipment in the offices and so, generally, depends on previous investments in the offices. Thus, taking into account Remark 1, we are led to the conclusion that $S^k \subset S^{k+1}$, $k = 1, ..., N-1$. The investor goes to $k = 1$ and decides if the information he bought is enough for him to construct a portfolio. If so, he stops the process. If not, he goes to $k = 2$ and first negotiates the payment arguing that he already has a piece of knowledge S^1 and so he is going to pay for the increment only. Here the situation is quite similar to the case of a shoemaker, when the customer wishes to order one shoe only, since he already has one from an other shoemaker. Next, the investor decides if the information he has bought from $k = 2$ is enough or not, etc. For N large enough this process can be idealized by the continuous process of buying S_t, $t \geq 0$ from one source. Moreover, the process can be stopped at arbitrary $t \geq 0$.

In contrast to the square error matrix Q, the mean m_t is always a stochastic process for buyers. It does not depend on whether the analysts' job was done in the past or will be done in the future.

In this paper we shall deal with the first case only.

PROPOSITION 1 *Assume that*

(I) H_t *is deterministic:* $H_t = H(t)$, $t \geq 0$,

(II) *the scalar valued function* $\vartheta(t, m, M) \triangleq W(m, Q(t), M)$ *is of class* $C^{1,2,1}$ $(R_+ \times R^n \times R_+)$,

(III) m_t *is a continuous, square-integrable martingale with the representation*

$$m_t = m_0 + \int_0^t \sigma(s)\, db_s, \quad t \geq 0, \tag{26}$$

where (σ^{ij}), $\sigma^{ij} \in C(R_+)$, *is a matrix satisfying the condition*

$$\sum \int_0^\infty [\sigma^{ij}(s)]^2\, ds < \infty$$

and b_t is a G_t-adapted, vector valued standard Brownian motion,

(IV) *for some G_t-measurable process V_t, the process $u_t \triangleq \vartheta(t, m_t, M - V_t)$ is a continuous, square-integrable G_t-martingale,*

(V) *$f(t, V, m) \geq 0$ where*

$$f(t, V, m) \triangleq \begin{cases} \frac{L_t \vartheta(t,m,M-V)}{\partial_M \vartheta(t,m,M-V)} & \text{for } (t, V, m) \in R_+ \times [0, M] \times R^n, \\ 0 & \text{for } \quad V \geq M, \end{cases} \tag{27}$$

$$L_t \triangleq \frac{\partial}{\partial t} + \frac{1}{2} \sum_{ij} \sigma^{ij}(t) \sigma^{ji}(t) \frac{\partial^2}{\partial m_i \partial m_j}, \quad \partial_M \triangleq \frac{\partial}{\partial M}.$$

Then V_t is a pathwise solution of the following stochastic ODE:

$$\frac{dV_t}{dt} = f(t, V_t, m_t), \quad V_0 = 0, \tag{28}$$

and so it is a non-negative G_t-adapted process with non-decreasing C^1 trajectories (except possibly at the random point $T = \min\{t \geq 0; V_T = M\}$).

Proof. Equation (28) follows from Ito's formula (see Karatzas, Shreve, 1991, for example) for the process

$$u_t = \vartheta(t, m_t, M - V_t).$$

Indeed,

$$du_t = L_t \vartheta(t, m_t, M - V_t) dt - \partial_M \vartheta(t, m_t, M - V_t) dV_t + \textit{ a martigale term} \tag{29}$$

and to make u_t a martingale, V_t must annihilate the first two terms on the right in (29). Thus, if V_t satisfies (28), then (V) implies it must be non-decreasing. Since (II) and (III) implies existence of a C^1 solution of (28), the result follows. ■

The usefulness of the value of information concept follows from the observation that the simple scalar process V_t divides the "big" space $R_+ \times R^n \times R_+$ of triples (t, m, c) into two regions: $R_+ \times R^n \times [0, V_t)$ and $R_+ \times R^n \times (V_t, M]$. If $(t, m_t, c(t))$ belongs to the first subset, then the purchase at the time t is reasonable. If it belongs to the second, then it is not. Hence, the subset $R_+ \times R^n \times \{V_t\}$ separates the purchase and "non-purchase" regions. In some cases the usefulness of the concept is immediate as the next result shows.

PROPOSITION 2 *Assume (I) - (V) and additionally (VI)*

$$E[\vartheta(t, m_t, M - V_t + \delta_t) | G_s] \geq \vartheta(s, m_s, M - V_s + \delta_s) \tag{30}$$

$$\textit{when } E[\delta_t | G_s] \geq \delta_s, \tag{31}$$

$$E[\vartheta(t, m_t, M - V_t + \delta_t) | G_s] \leq \vartheta(s, m_s, M - V_s + \delta_s) \tag{32}$$

for any G_t-adapted process δ_t. Then the optimal stopping time for the problem

$$\sup E\left[\left(V_\tau - c_\tau\right); \tau \geq 0, 0 \leq c_\tau \leq M\right]$$

is also optimal for the problem

$$\sup E\left[\vartheta\left(\tau, m_\tau, M - c_\tau\right); \tau \geq 0, 0 \leq c_\tau \leq M\right].$$

Proof. Let $\delta_t \triangleq V_t - c_t$. Then $\eta_t \triangleq \vartheta\left(t, m_t, M - c_t\right) = \vartheta\left(t, m_t, M - V_t + \delta_t\right)$ is a supermartingale (submartingale) if δ_t is a supermartingale (submartingale). From the optional sampling theorem it follows therefore that the implications (31)(30) and (33)(32) hold for the stopping times $\tau \geq 0, 0 \leq c_\tau \leq M$, assuming they hold for the ordinary ones. From the well known properties of the so-called Snell envelope in the optimal stopping theory (see Kazatzas, Shreve, 1998, Appendix D, for instance) follows that (31) and (33) hold for the best stopping time τ_1 of the first problem, i.e., for $t = \tau_1$ in (31) and for $s = \tau_1$ in (33). Set $t = \tau_1$ and $s = \tau \leq \tau_1$ in (30)(31) and integrate both sides. Set $s = \tau_1$ and $t = \tau \geq \tau_1$ in (32)(33) and integrate both sides. The resulting inequalities show that τ_1 is optimal for the second problem as well. ∎

We are now in a position to state an important problem.

PROBLEM 2 *What should in fact the information buyer's strategy look like? Indeed, to solve the optimal stopping problem, the investor has to know the process (V_t, c_t), but then he or she has to know (m_t, Q_t), and hence does not need to buy anything. Vicious circle !*

Indeed, one reason for introducing the concept of information value process V_t is to answer the question: when one should stop the buying process. In this paper we are dealing only with the case where $c(t)$, $Q(t)$ are deterministic and known to the buyer. From (28) it follows that V_t is G_t-adapted, hence the stopping problem

$$\sup E\left[\vartheta\left(\tau, m_\tau, M - V_\tau + \delta_\tau\right)\right]$$

over all G_t-stopping times $\tau \geq 0, 0 \leq c(\tau) \leq M$ is well posed and the buyer stops the process at the optimal time $\tau_\star$ for which the triple $(\tau_\star, m_{\tau_\star}, c(\tau_\star))$ belongs to the purchase region, thus getting the segment $S_\star = \{(m_s, Q(s)); 0 \leq s \leq \tau_\star\}$ at the price $c(\tau_\star)$.

4. Application in portfolio optimization

In this section we will illustrate the concept of information value with an example based on portfolio theory.

Let $m \in R^n$, $Q = Q^T$ denote, as before, the vector of mean values and the an investor, $r > 0$ the largest risk-free interest rates, and $\beta > 0$ the investor's risk aversion coefficient.

A linear-quadratic utility function has the form

$$U(x_0, x, m, Q, M) = rx_0 + \langle x, m\rangle - \beta x^T Q x \tag{34}$$

in which x is a portfolio of risky assets and x_0 is a risk-free investment. Since the risk-free investments are included in (34) separately, we may assume without loss of generality that

$$Q > 0. \tag{35}$$

We introduce notation:

$$\begin{aligned} p(t,m) &= m^T Q^{-1}(t)\, m, \\ q(t) &= J^T Q^{-1}(t)\, J, \\ \rho(t,m) &= m^T Q^{-1}(t)\, J, \end{aligned}$$

$$\phi(t,m) = \frac{p(t,m)\, q(t) - \rho^2(t,m) + 4\beta\rho(t,m)}{4\beta q(t)} + \frac{r^2 q(t)}{4\beta}.$$

The main result in this section gives an explicit representation of the information value process.

THEOREM 1 *If $\phi(t,m)$ and the matrix $\left(\sigma^{ij}(t)\right)$ are such that*

$$L_t\phi(t,m) \geq 0, \quad (t,m) \in R_+ \times R^n, \tag{36}$$

then (i) the information value process V_t has the representation

$$V_t = \int_0^t v(s, m_s)\, ds, \quad t \leq T = \inf\{s; V_s = M\}, \tag{37}$$

where

$$v(t,m) = \frac{1}{r} L_t\phi(t,m), \tag{38}$$

(ii) the function

$$\vartheta(t,m,M) = \phi(t,m) + rM \tag{39}$$

satisfies the hypothesis (VI) of Proposition 2.

Proof. From the second lemma in the Appendix we have (39). Hence from (27) it follows that

$$f(t,V,m) = \begin{cases} \frac{1}{r} L_t\phi(t,m) & \text{for } (t,V,m) \in R_+ \times [0,M] \times R^n, \end{cases}$$

proving (37). Since $u_t = \vartheta(t, m_t, M - V_t) = \phi(t, m_t) + r(M - V_t)$ is a G_t–martingale, and

$$\begin{aligned}\vartheta(t, m_t, M - c(t)) &= \vartheta(t, m_t, M - V_t + \delta_t) \\ &= \phi(t, m_t) + r(M - V_t) + r\delta_t \\ &= \text{a martingale } + r\delta_t,\end{aligned}$$

(ii) is obvious. ■

5. Price of information

The price for the information included in a segment S_t is a result of a bargaining process (or game) between the seller and the buyer. If

$$c_t < V_t, \tag{40}$$

then the purchase is reasonable. If, on the contrary,

$$c_t > V_t \tag{41}$$

then it is not. When

$$c_t = V_t \tag{42}$$

then we say that the information is "of value".

Consider, for instance, the situation of a seller who guessed (or knew from a buyer) the values r and β appearing in (34). If he decides to set the highest possible price: $c_t = V_t$, for $t \geq 0$, then from Theorem 1 and Proposition 2 follows that the optimal $\tau^* = 0$. But this choice (optimal for the buyer) is totally unsatisfactory for the seller, scice he earns nothing (!). Clearly, this is a consequence of an asymmetric information structure and shows once more the dominant role the information structures play in the games and bargaining problems.

The information value V as defined in the Section 2 is an individual characteristic of the segment S_t depending on the particular investor, his subjective risk estimation, risk aversion (utility function), the cash amount M, etc. We describe it by introducing a notion of the information value $V_t(a)$ for an investor with parameter a, $a = (U(\cdot), M)$.

The average value of the information included in the segment S_t may be defined as

$$\overline{V_t} = \int V_t(a)\, dp(a) \tag{43}$$

where $p(a)$ is an appropriate measure on $C(R_+) \times R_+$.

In the simple market of one seller and many buyers the price c_t of the segment theoretical matter even to estimate the range of these fluctuations. The general case of many buyers and sellers will be still more difficult.

Summary We have presented the concept of information value as a property which is jointly attributed to: (1) the parametrized family of probability distribution functions of the investment returns, and (2) a specific investor with his oun preferences and possibilities. This is the concept of an equivalent money - information exchange. In the mathematic language this last requirement is expressed by the property of being a martingale: when the price for information equals its value, the utility function of the investor is a martingale. For a particular utility function of linear-quadratic form we have expressed the information value explicitly. Possible extensions of our results to the free-market theory of information value require further studies.

Acknowledgement The authors are grateful to the anonymous referees for criticism and remarks. In particular Problem 2 was posed by one of them.

References

ARROW, K. (1970) *Essays in the Theory of Risk Bearing.* North-Holland, London.

BANEK, T. (2000) *Risk Calculus.* CBS WSZiA, Seria Monografie Nr 2, Lublin (in Polish).

BANEK, T. (2002) Value of information. *Proc. Conf. on Modelling Preferences vs. Risk*, Ustroń (in Polish).

BANEK, T. (2001) On the value of information. Talk at R. Kulikowski's seminar IBS PAN, Warsaw, November 2001.

DAVIS, M.H.A., DEMPSTER, M.A.H. and ELLIOTT, R.J. (1991) On the value of information in controlled diffusion processes. *Liber Amicorum for M. Zakai*, Academic Press, 125–138.

DAVIS, M.H.A. and KARATZAS, I. (1994) A deterministic approach to optimal stopping Probability. In: F.P. Kelly, ed., *Statistics and Optimisation.* John Wiley & Sons, 1994, 455 – 466.

KARATZAS, I. and SHREVE, S.E. (1998) *Methods of Mathematical Finance.* Springer-Verlag.

KARATZAS, I. and SHREVE, S.E. (1991) *Brownian Motion and Stochastic Calculus.* Springer-Verlag.

LIPTSER, R.S. and SHIRYAEV, A.N. (1977) *Statistics of Random Processes.*

Appendix

With the notations

$$\begin{aligned} p &= m^T Q^{-1} m, \\ q &= J^T Q^{-1} J, \\ \rho &= m^T Q^{-1} J, \end{aligned}$$

we have two elementary lemmas:

LEMMA A1

$$\begin{aligned} &\max_{\langle x,J\rangle = a} \left[\langle x,m\rangle - \beta x^T Q x\right] \\ = &\ \frac{pq-\rho^2+4\beta\rho}{4\beta q} - \frac{\beta a^2}{q}. \end{aligned}$$

Proof. Since $Q > 0$ we may define $Q^{-1/2}$ (the square root of Q^{-1}) and

$$\begin{aligned} \widetilde{m} &= Q^{-1/2} m, \quad \widetilde{J} = Q^{-1/2} J, \\ y &= Q^{1/2}\left[x - \frac{1}{2\beta} Q^{-1} m\right]. \end{aligned}$$

Then

$$\begin{aligned} \langle x,m\rangle - \beta x^T Q x &= \frac{p}{4\beta} - \beta \left\|y\right\|^2, \\ \langle x, J\rangle &= \left\langle y, \widetilde{J}\right\rangle + \frac{\rho}{2\beta}. \end{aligned}$$

Hence

$$\begin{aligned} &\max_{\langle x,J\rangle = a} \left[\langle x,m\rangle - \beta x^T Q x\right] \\ = &\ \max_{\langle y,\widetilde{J}\rangle = b} \left[\frac{p}{4\beta} - \beta\left\|y\right\|^2\right] \qquad \left(b = a - \frac{\rho}{2\beta}\right) \\ = &\ \frac{p}{4\beta} - \beta \min_{\langle y,\widetilde{J}\rangle = b} \left\|y\right\|^2 \\ = &\ \frac{p}{4\beta} - \beta\frac{b^2}{q}. \end{aligned}$$

■

LEMMA A2

$$\max_{x_0 + \langle x,J\rangle = M} \left[r x_0 + \langle x,m\rangle - \beta x^T Q x\right]$$

$$pq - \rho^2 + 4\beta\rho \quad qr^2$$

Proof. Since

$$\begin{aligned} & \max_{x_0+\langle x,J\rangle=M} \left[rx_0 + \langle x,m\rangle - \beta x^T Qx\right] \\ = \; & \max_{x_0} \left\{ rx_0 + \max_{\langle x,J\rangle=M-x_0} \left[\langle x,m\rangle - \beta x^T Qx\right] \right\} \end{aligned}$$

it is enough to apply the previous lemma. ■