数理解析研究所講究録
第 1899 巻 2014 年 11-31

Continuous-time linear-quadratic dynamic optimization — evaluation/optimization and Bellman equation —

Seiichi Iwamoto

Department of Economic Engineering
Graduate School of Economics, Kyushu University

JEL classification : C61, D81

Mathematics Subject Classification (2010): 90C39, 90C40, 49L20, 91B55

Abstract. This paper considers three continuous-time dynamic optimization problems on one-dimensional state and control spaces. The three have a common feature – linear dynamics and discounted quadratic criterion (LQ) –. The first problem is on deterministic dynamics. The second and the third are on stochastic dynamics. The second dynamics is an Ornstein-Uhlenbeck process. The third is a geometric Brownian motion. We discuss the optimal solution from two reciprocal points of view. One is dynamics; from deterministic to stochastic. The other is approach; evaluation-optimization versus Bellman equation. A complete optimal solution is given. Each solution is expressed in terms of three parameters — (1) discount-rate, (2) characteristics of dynamics and (3) diffusion coefficient —. The optimal solutions have a common feature, too. The optimal control is proportional and the optimal value functions are quadratic. Both the optimal proportional rate and the optimal value functions are explicitly specified. Further we show a zero-sum property between optimal value function and optimal proportional control. Sum of the optimal value and the optimal rate is zero.

Key words: proportional policy, proportional rate, evaluation-optimization, Bellman equation, zero-sum, continuous-time, certainty equivalence principle

1 Introduction

This paper discusses a class of infinite-horizon discounted quadratic dyamic optimization problems on one-dimensional state and control spaces. The class is classified under dynamics and approach. The dynamics are (a) deterministic and (b) stochastic. Two of the stochastic dyanmics are (b-1) Ornstein-Uhlenbeck process and (b-2) geometric Brownian motion. Approaches are (i) evaluation-optimization and (ii) Bellman equation.

We are concerned with optimality of proportional policy, which is a stationary one. Section 2 lists three dynamic optimization problems. The first problem is on a deterministic dynamics. The second is on an Ornstein-Uhlenbeck process. The third is on a geometric Brownian motion.

Section 3 gives explicit solutions of deterministic control problem through (i) evaluation-optimization and (ii) Bellman equation. Each solution is expressed in terms of discount rate, characteristic of dynamics and diffusion coefficient. Sections 4 and 5 solve the contol problem on the Ornstein-Uhlenbeck process and on the geometric Brownian motion, respectively. Section 6 derives Bellman equation both for deterministic control process and for stochastic one.

It is shown that two approaches yield the same optimal solutions. A zero-sum property between the quadratic coefficient of value function and optimal proportional rate is derived. The property claims that the higher the optimal rate is in absolute value, the higher the optimal value. This property is common to three optimal solutions.

2 Linear Quadratic Models

This section specifies three dynamic optimization problems we shall consider in the paper. Throughout the paper, let $\rho > 0$ be a discount rate on ontinuous-time process (as for discrete-time model, see [1–5, 8–12]).

The deterministic problem is minimization of discounted quadratic criterion

$$\int_0^\infty e^{-\rho t}\left(x^2 + u^2\right) dt$$

under a linear dynamics

$$\dot{x} = bx + u \quad 0 \le t < \infty, \qquad x(0) = c$$

where $b\,(\in R^1)$ represents a characteristic of dynamics and $c\,(\in R^1)$ is an initial state. Let C be the set of all continuous functions on the one-dimensional Euclidean space R^1 :

$$C = \{x = x(t) \mid x : R^1 \to R^1 \text{ continuous}\}.$$

For the sake of simplicity, we take trajectory $x = x(\cdot)$ in C^1 and control function $u = u(\cdot)$ in R^1, respectively.

The stochastic problem is minimization of expected value of discounted quadratic criterion

$$E_x\left[\int_0^\infty e^{-\rho t}\left(x^2 + u^2\right) dt\right]$$

under a stochastic dynamics

$$dx(t) = (bx(t) + u(t))dt + \sigma(x(t))dw(t) \quad 0 \le t < \infty, \qquad x(0) = x$$

where $\{w(\cdot)\}$ is the standard one-dimensional Brownian motion. Here $\sigma(x)$ is a nonnegative continuous function of x. We take two cases: (i) $\sigma(x) = \sigma$ and (ii) $\sigma(x) = \sigma x$, where σ is a nonnegative constant. The cases (i) and (ii) lead an Ornstein-Uhlenbeck process and a geometric Brownian motion, respectively.

Thus we take three problems as follows.

$$
\mathrm{D}(c) \qquad
\begin{array}{lll}
\text{minimize} & \displaystyle\int_0^\infty e^{-\rho t}\left(x^2+u^2\right)dt & \\
\text{subject to} & \text{(i)} \quad \dot{x} = bx + u & \\
 & \text{(ii)} \quad x \in C^1,\ u(t) \in R^1 & 0 \le t < \infty \\
 & \text{(iii)} \quad x(0) = c &
\end{array}
$$

$$
\mathrm{O}(x) \qquad
\begin{array}{lll}
\text{minimize} & E_x\left[\displaystyle\int_0^\infty e^{-\rho t}\left(x^2+u^2\right)dt\right] & \\
\text{subject to} & \text{(i)} \quad dx(t) = (bx+u)dt + \sigma dw(t) & \\
 & \text{(ii)} \quad x \in C,\ u(t) \in R^1 & 0 \le t < \infty \\
 & \text{(iii)} \quad x(0) = x &
\end{array}
$$

$$
\mathrm{G}(x) \qquad
\begin{array}{lll}
\text{minimize} & E_x\left[\displaystyle\int_0^\infty e^{-\rho t}\left(x^2+u^2\right)dt\right] & \\
\text{subject to} & \text{(i)} \quad dx(t) = (bx+u)dt + \sigma x dw(t) & \\
 & \text{(ii)} \quad x \in C,\ u(t) \in R^1 & 0 \le t < \infty \\
 & \text{(iii)} \quad x(0) = x. &
\end{array}
$$

3 Deterministic dynamics

In this section, we solve a continuous-time dynamic optimization problem D(c) through two methods — (i) evaluation-optimization and (ii) dynamic programming —. The evaluation-optimization method consists of two steps. At the first step we evaluates any proportional policy. At the second, of all the proportional policies, we find an optimal solution by solving an associated one-variable fractional minimization problem. The dynamic programming method solves Bellam equation in an analytic form.

Consider the deterministic dynamic optimization problem:

$$
\mathrm{D}(c) \qquad
\begin{array}{lll}
\text{minimize} & \displaystyle\int_0^\infty e^{-\rho t}\left(x^2+u^2\right)dt & \\
\text{subject to} & \text{(i)} \quad \dot{x} = bx + u & \\
 & \text{(ii)} \quad x \in C^1,\ u(t) \in R^1 & 0 \le t < \infty \\
 & \text{(iii)} \quad x(0) = c. &
\end{array}
$$

3.1 Evaluation-optimization

Any proportional control is specified by a *control function*

$$u(t) = ux(t) \quad (u \in R^1)$$

where u is called a *proportional rate.* There exists a one-to-one correspondence $u(\cdot) \longleftrightarrow u$ between the set of all proportional control functions and the set of all proportional rates. The latter constitutes one-dimensional Euclidean space R^1. Thus any proportional control function $u(t) = ux(t)$ is identified by a real number $u \in R^1$ and vice versa.

We evaluate any proportional control and minimize the evaluated value over the set of all proportional rates. The evaluation problem is written as follows.

$$\mathrm{D}(c;u) \qquad \begin{array}{ll} \text{evaluate} & \displaystyle\int_0^\infty e^{-\rho t}\left(x^2 + u^2x^2\right) dt \\ \text{subject to} & \text{(i)} \quad \dot{x} = bx + ux \qquad 0 \le t < \infty \\ & \text{(ii)} \quad x(0) = c. \end{array}$$

The proportional control $u(x) = ux$ is evaluated as follows. Let $V_c(u)$ denote the evaluated value:

$$V_c(u) \;=\; \int_0^\infty e^{-\rho t}\left(x^2 + u^2x^2\right) dt.$$

Then we have

Lemma 3.1

$$V_c(u) \;=\; \begin{cases} \infty & \textit{for} \;\; \rho - 2b - 2u \le 0 \\ \dfrac{1+u^2}{\rho - 2b - 2u}c^2 & \textit{for} \;\; \rho - 2b - 2u > 0. \end{cases}$$

Proof. First we note that the control $u(x) = ux$ yields

$$V_c(u) \;=\; (1+u^2)\int_0^\infty e^{-\rho t}x^2 dt.$$

Second the linear dynamics (i), (ii) is reduced to

$$\dot{x} = (b+u)x, \quad x(0) = c.$$

This has a unique solution

$$x(t) = ce^{(b+u)t}.$$

Hence

$$V_c(u) \;=\; c^2\left(1+u^2\right)\int_0^\infty e^{-\gamma t} dt$$

where $\gamma = \rho - 2b - 2u$. Thus the control u is evaluated as follows.

$$V_c(u) = \begin{cases} \infty & \text{for } \gamma \leq 0 \\ \dfrac{1+u^2}{\gamma}c^2 & \text{for } \gamma > 0. \end{cases}$$

□

Since our concern is the minimization, it is enough to restrict u to $\rho - 2b - 2u > 0$. Now let us consider the ratio minimization problem

$$\text{(R1)} \qquad \begin{array}{l} \text{minimize} \quad \dfrac{1+u^2}{\rho - 2b - 2u} \\ \text{subject to} \quad \text{(i)} \ \rho - 2b - 2u > 0. \end{array}$$

Lemma 3.2 *(See Fig.1) The problem* (R1) *has the minimum value m at $\hat{u}$, where*

$$m = -\hat{u} = b - \frac{\rho}{2} + \sqrt{\left(b - \frac{\rho}{2}\right)^2 + 1}\,. \tag{1}$$

We call m and $\hat{u}$ *optimal value* and *optimal rate*, respectively.

Proof. Let us take

$$g(u) = \frac{1+u^2}{\eta - 2u} \qquad \eta := \rho - 2b. \tag{2}$$

Then

$$\begin{aligned} g(u) &= -\frac{1}{2}u - \frac{1}{4}\eta + \frac{\eta^2/4+1}{\eta - 2u} \\ &= -\frac{1}{2}\eta + \frac{1}{4}(\eta - 2u) + \frac{\eta^2/4+1}{\eta - 2u}. \end{aligned}$$

Thus hyperbolic curve $y = g(u)$ has a unique minimum for $-\infty < u < \dfrac{\eta}{2}$. Differentiating g, we get

$$g'(u) = -2\frac{u^2 - \eta u - 1}{(\eta - 2u)^2}, \qquad g''(u) = 2\frac{\eta^2 + 4}{(\eta - 2u)^3}.$$

Letting the numerator of $g'(u)$ be zero, we have the quadratic equation

$$u^2 - \eta u - 1 = 0. \tag{3}$$

Solving this yields a minimum point

$$\hat{u} = \frac{\eta}{2} - \sqrt{\eta^2/4 + 1}\,.$$

From (3) we have the minimum value

$$g(\hat{u}) = \frac{2 + \eta\hat{u}}{\eta - 2\hat{u}} = -\hat{u}. \tag{4}$$

$\square$

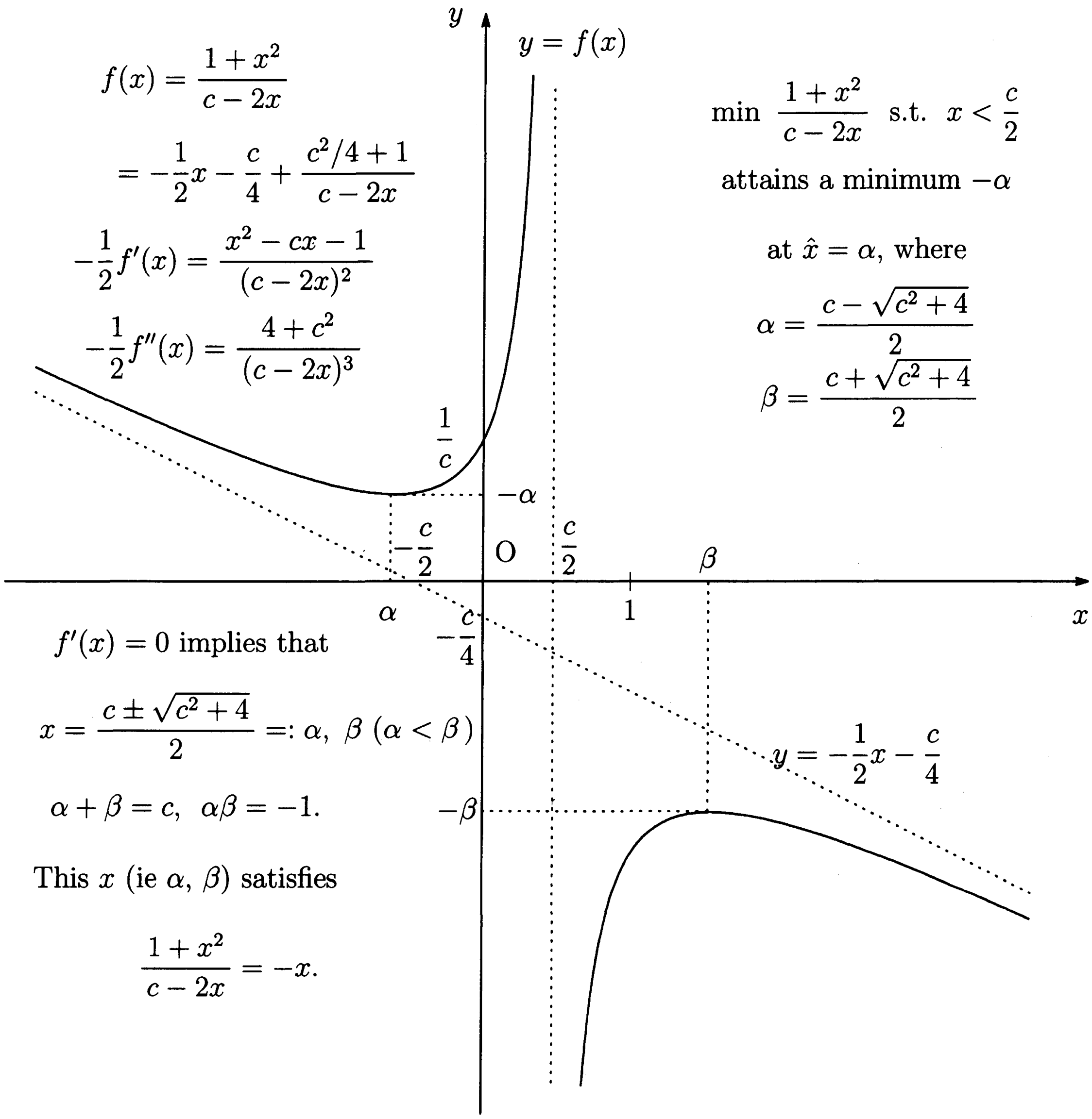

Fig.1 $\quad \min \dfrac{1+x^2}{c-2x}$ s.t. $x < \dfrac{c}{2}$ $\quad (c = \rho - 2b$ or $\rho - \sigma^2 - 2b)$

Thus we have the optimal control function with rate $\hat{u}$:

$$\hat{u}(x) \; = \; \hat{u}x$$

where

$$\hat{u} \; = \; \frac{\rho}{2} - b - \sqrt{\left(\frac{\rho}{2} - b\right)^2 + 1}.$$

The value function

$$v(x) = \hat{v}x^2 \quad (\hat{v} := -\hat{u})$$

is optimal in the class of proportional policies.

Thus we have a remarkable property between optimal value $\hat{v}$ and optimal rate $\hat{u}$:

Proposition 3.1 *(Zero-sum property) It holds that*

$$\hat{u} + \hat{v} = 0. \tag{5}$$

3.2 Dynamic programming

Let $v(c)$ be the minimum value. Then the value function $v : R^1 \to R^1$ satisfies the Bellman equation (which is derived in Section 4)

$$\rho v(x) = \min_{u \in R^1} \left[x^2 + u^2 + v'(x)(bx + u)\right]. \tag{6}$$

We solve (6). From $\frac{d}{du}[\cdots] = 0$, we get

$$\rho v(x) = x^2 - \frac{1}{4}v'^2(x) + bxv'(x), \quad \hat{u}(x) = -\frac{1}{2}v'(x).$$

The linear-quadratic scheme enables us to assume that v is quadratic $v(x) = vx^2$ $(v \geq 0)$. Substituting $v'(x) = 2vx$, we have

$$\rho v x^2 = x^2 - v^2x^2 + 2bvx^2 \quad \text{i.e.} \quad \rho v = 1 - v^2 + 2bv.$$

This yields the quadratic equation

$$v^2 - (2b - \rho)v - 1 = 0, \tag{7}$$

which has a unique positive solution

$$\hat{v} = b - \frac{\rho}{2} + \sqrt{\left(b - \frac{\rho}{2}\right)^2 + 1}. \tag{8}$$

, which is also called *optimal value.* We have the desired optimal solution

$$v(x) = \hat{v}x^2, \quad \hat{u}(x) = -\hat{v}x \quad (\hat{u} := -\hat{v}) \tag{9}$$

Thus the zero-sum property between optimal value $\hat{v}$ and optimal rate $\hat{u}$ holds true.

4 Ornstein-Uhlenbeck processs

In this section, we solve the stochastic dynamic optimization problem $\mathrm{O}(x)$ through the two methods.

Consider a dynamic optimization on an Ornstein-Uhlenbeck process as follows.

$$\mathrm{O}(x) \qquad \begin{array}{ll} \text{minimize} & E_x\left[\int_0^\infty e^{-\rho t}\left(x^2+u^2\right)dt\right] \\ \text{subject to} & \text{(i)} \quad dx(t) = (bx+u)dt + \sigma dw(t) \\ & \text{(ii)} \quad x \in C,\ u(t) \in R^1 \\ & \text{(iii)} \quad x(0) = x. \end{array} \qquad 0 \le t < \infty$$

Here and frequently in the following we use a notation x with double meaning. One is an initial state $x(0) = x$. The other is a process $x = x(t)$. This double usage does not matter.

4.1 Evaluation-optimization

We evaluate any proportional control $u(x) = ux$ with proportional rate u and minimize the expected value over all proportional rates.

Our evaluation problem is

$$\mathrm{O}(x;u) \qquad \begin{array}{ll} \text{evaluate} & E_x\left[\int_0^\infty e^{-\rho t}\left(x^2+u^2x^2\right)dt\right] \\ \text{subject to} & \text{(i)} \quad dx(t) = (b+u)xdt + \sigma dw(t) \quad 0 \le t < \infty \\ & \text{(ii)} \quad x(0) = x. \end{array}$$

Then (i), (ii) is an Ornstein-Uhlenbeck process ([6, p.358])

$$dx(t) \;=\; \mu x dt + \sigma dw(t) \quad x(0) = x \quad (\mu = b + u). \tag{10}$$

This has a unique solution

$$x(t) \;=\; e^{\mu t}\left(x + \sigma\int_0^t e^{-\mu s}dw(s)\right). \tag{11}$$

Thus the proportional control $f(x) = ux$ is evaluated as follows. Let $V_x(u)$ denote the evaluated value:

$$V_x(u) \;=\; E_x\left[\int_0^\infty e^{-\rho t}\left(x^2+u^2x^2\right)dt\right].$$

Then we have

Lemma 4.1

$$V_x(u) \;=\; \begin{cases} \infty & \text{for } \rho - 2b - 2u \le 0 \\ \dfrac{1+u^2}{\rho-2b-2u}\left(x^2+\dfrac{\sigma^2}{\rho}\right) & \text{for } \rho - 2b - 2u > 0. \end{cases}$$

Proof. First we note that the control $u(x) = ux$ yields

$$V_x(u) \ = \ (1+u^2)\int_0^\infty E_x\left[e^{-\rho t}x^2\right]dt.$$

Second we evaluate the discounted squared process $e^{-\rho t}x^2 = e^{-\rho t}x^2(t)$. Taking expectation of both sides

$$\left(x+\sigma\int_0^t e^{-\mu s}dw(s)\right)^2 \ = \ x^2 + 2\sigma x\int_0^t e^{-\mu s}dw(s) + \sigma^2\left(\int_0^t e^{-\mu s}dw(s)\right)^2,$$

we have

$$E_x\left(x+\sigma\int_0^t e^{-\mu s}dw(s)\right)^2 \ = \ x^2+\sigma^2\int_0^t e^{-2\mu s}ds.$$

Here are two cases (i) $\mu \neq 0$ and (ii) $\mu = 0$.

First we assume that (i) $\mu \neq 0$. Then

$$E_x\left(x+\sigma\int_0^t e^{-\mu s}dw(s)\right)^2 \ = \ \left(x^2+\frac{\sigma^2}{2\mu}\right) - \frac{\sigma^2}{2\mu}e^{-2\mu t}.$$

Thus the expected value of $e^{-\rho t}x^2(t)$ is

$$E_x\left[e^{-\rho t}x^2\right] \ = \ e^{-(\rho-2\mu)t}\left(x^2+\frac{\sigma^2}{2\mu}\right) - \frac{\sigma^2}{2\mu}e^{-\rho t}.$$

The integral part is evaluated as follows.

$$\begin{aligned}\int_0^\infty E_x\left[e^{-\rho t}x^2\right]dt \ &= \ \left(x^2+\frac{\sigma^2}{2\mu}\right)\int_0^\infty e^{-(\rho-2\mu)t}dt - \frac{\sigma^2}{2\mu}\int_0^\infty e^{-\rho t}dt \\ &= \ \begin{cases} \infty & \text{for } \rho-2\mu \le 0 \\ \dfrac{1}{\rho-2\mu}\left(x^2+\dfrac{\sigma^2}{\rho}\right) & \text{for } \rho-2\mu > 0. \end{cases}\end{aligned}$$

Second we take (ii) $\mu = 0$. Then

$$E_x\left(x+\sigma\int_0^t e^{-\mu s}dw(s)\right)^2 \ = \ E_x\left(x+\sigma w(t)\right)^2 \ = \ x^2+\sigma^2 t.$$

Thus

$$E_x\left[e^{-\rho t}x^2\right] \ = \ e^{-\rho t}\left(x^2+\sigma^2 t\right).$$

Thus the integral part is

$$\begin{aligned}\int_0^\infty E_x\left[e^{-\rho t}x^2\right]dt \ &= \ x^2\int_0^\infty e^{-\rho t}dt + \sigma^2\int_0^\infty te^{-\rho t}dt \\ &= \ \frac{1}{\rho}\left(x^2+\frac{\sigma^2}{\rho}\right).\end{aligned}$$

Consequently in either case we have

$$\int_0^\infty E_x\left[e^{-\rho t}x^2\right]dt = \begin{cases} \infty & \text{for } \rho - 2\mu \le 0 \\ \dfrac{1}{\rho - 2\mu}\left(x^2 + \dfrac{\sigma^2}{\rho}\right) & \text{for } \rho - 2\mu > 0. \end{cases}$$

Finally the control yields the desired evaluation:

$$V_x(u) = \begin{cases} \infty & \text{for } \rho - 2\mu \le 0 \\ \dfrac{1+u^2}{\rho - 2\mu}\left(x^2 + \dfrac{\sigma^2}{\rho}\right) & \text{for } \rho - 2\mu > 0. \end{cases}$$

□

We reconsider the ratio minimization problem

$$\text{(R1)} \qquad \begin{array}{ll} \text{minimize} & \dfrac{1+u^2}{\rho - 2b - 2u} \\ \text{subject to} & \text{(i)} \;\; \rho - 2b - 2u > 0. \end{array}$$

From Lemma 3.2, (R1) has the minimum value m at $\hat{u}$, where

$$m = -\hat{u} = b - \frac{\rho}{2} + \sqrt{\left(b - \frac{\rho}{2}\right)^2 + 1}\,.$$

Thus we have the optimal decision function with rate $\hat{u}$:

$$\hat{u}(x) = \hat{u}x.$$

The value function

$$v(x) = m\left(x^2 + \frac{\sigma^2}{\rho}\right)$$

is optimal in the class of proportional policies.

As Proposition 3.1 claims, zero-sum property holds:

$$\hat{u} + m = 0.$$

4.2 Dynamic programming

Let $v(x)$ be the minimum value. Then the value function $v : R^1 \to R^1$ satisfies the Bellman equation (which is derived in Section 4)

$$\rho v(x) = \min_{u \in R^1}\left[x^2 + u^2 + (bx+u)v'(x) + \frac{\sigma^2}{2}v''(x)\right]. \tag{12}$$

Eq.(12) is solved as follows. $\dfrac{d}{du}[\cdots] = 0$ implies

$$\rho v(x) \;=\; x^2 - \frac{1}{4}v'^2(x) + bxv'(x) + \frac{\sigma^2}{2}v''(x), \qquad \hat{u}(x) \;=\; -\frac{1}{2}v'(x).$$

The stochastic dynamics enables us to assume that v is quadratic $v(x) = vx^2 + w$ $(v, w \geq 0)$. Substituting $v'(x) = 2vx,\ v''(x) = 2v$, we have

$$\rho(vx^2 + w) \;=\; x^2 - v^2x^2 + 2bvx^2 + \sigma^2 v$$

i.e.

$$\rho v \;=\; 1 - v^2 + 2bv, \qquad \rho w \;=\; \sigma^2 v.$$

This yields the quadratic equation (7) once again

$$v^2 - (2b - \rho)\, v - 1 \;=\; 0. \tag{13}$$

But this time with the additional linear relation. Eq. (13) has a unique positive solution

$$\hat{v} \;=\; b - \frac{\rho}{2} + \sqrt{\left(b - \frac{\rho}{2}\right)^2 + 1}\,.$$

Hence

$$\hat{w} \;=\; \frac{\sigma^2}{\rho}\hat{v}.$$

Thus we have the desired optimal solution

$$v(x) \;=\; \hat{v}x^2 + \hat{w}, \quad \hat{u}(x) \;=\; -\hat{v}x.$$

We note that the optimal control function $\hat{u} = \hat{u}(x)$ is identical with the optimal one for the corresponding deterministic problem. This is called *certainty equivalence principle* (as for discrete-time model, see [4,11]). This principle comes from the stochastic dynamics (i) and the linear-quadratic scheme.

5 Geometric Brownian motion

In this section, we solve the stochastic dynamic optimization problem $\mathrm{G}(x)$ through the two methods.

Let us now consider the dynamic optimization on geometric Brownian motion:

$$\mathrm{G}(x) \quad \begin{array}{lll} \text{minimize} & E_x\left[\displaystyle\int_0^\infty e^{-\rho t}\left(x^2 + u^2\right)dt\right] & \\ \text{subject to} & \text{(i)} \;\; dx(t) = (bx + u)dt + \sigma x dw(t) & \\ & \text{(ii)} \;\; x \in C,\ u(t) \in R^1 & 0 \leq t < \infty \\ & \text{(iii)} \;\; x(0) = x. & \end{array}$$

5.1 Evaluation-optimization

First we evaluate any proportional control $f(x) = ux$ with proportional rate u. Second we minimize the expected value over all rates.

Now our problem is

$$
\mathrm{G}(x;u) \qquad \begin{array}{ll} \text{evaluate} & E_x\left[\int_0^\infty e^{-\rho t}\left(x^2+u^2x^2\right)dt\right] \\ \text{subject to} & \text{(i)} \quad dx(t) = (b+u)xdt + \sigma x dw(t) \quad 0 \le t < \infty \\ & \text{(ii)} \quad x(0) = x. \end{array}
$$

Then (i), (ii) is a geometric Brownian process ([6, p.349], [7])

$$
dx(t) = \mu x dt + \sigma x dw(t) \quad x(0) = x \quad (\mu = b + u). \tag{14}
$$

This has a unique solution

$$
x(t) = x e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma w(t)}. \tag{15}
$$

Thus the proportional control $u(x) = ux$ yields the evaluated value:

$$
V_x(u) = E_x\left[\int_0^\infty e^{-\rho t}\left(x^2+u^2\right)dt\right].
$$

Then

Lemma 5.1

$$
V_x(u) = \begin{cases} \infty & \textit{for } \rho - \sigma^2 - 2b - 2u \le 0 \\ \dfrac{1+u^2}{\rho - \sigma^2 - 2b - 2u}x^2 & \textit{for } \rho - \sigma^2 - 2b - 2u > 0. \end{cases}
$$

Proof. As in Ornstein-Uhlenbeck process, the control f with rate u yields

$$
V_x(u) = (1+u^2)\int_0^\infty E_x\left[e^{-\rho t}x^2\right]dt.
$$

The discounted squared process $e^{-\rho t}x^2 = e^{-\rho t}x^2(t)$ is evaluated as follows. Since

$$
E_x\left[e^{2\sigma w(t)}\right] = e^{2\sigma^2 t}
$$

we have the expected value of $x^2(t)$ as follows.

$$
E_x\left[x^2\right] = x^2 e^{(2\mu - \sigma^2)t + 2\sigma^2 t}.
$$

The integral part becomes

$$
\begin{aligned}
\int_0^\infty e^{-\rho t}E_x[x^2]dt &= x^2\int_0^\infty e^{-(\rho-\sigma^2-2\mu)t}dt \\
&= \begin{cases} \infty & \text{for } \rho - \sigma^2 - 2\mu \le 0 \\ \dfrac{1}{\rho-\sigma^2-2\mu}x^2 & \text{for } \rho - \sigma^2 - 2\mu > 0. \end{cases}
\end{aligned}
$$

Thus we have the desired evaluation

$$V_x(u) \;=\; \begin{cases} \infty & \text{for } \rho-\sigma^2-2\mu \le 0 \\ \dfrac{1+u^2}{\rho-\sigma^2-2\mu}x^2 & \text{for } \rho-\sigma^2-2\mu > 0. \end{cases}$$

□

Now let us consider the ratio minimization problem

$$\text{(R2)} \qquad \begin{array}{ll} \text{minimize} & \dfrac{1+u^2}{\rho-\sigma^2-2b-2u} \\ \text{subject to} & \text{(i)} \;\; \rho-\sigma^2-2b-2u > 0. \end{array}$$

Lemma 5.2 *(See Fig.1) The problem* (R2) *has the minimum value m at $\breve{u}$, where*

$$m = -\breve{u} \;=\; b+\frac{\sigma^2}{2}-\frac{\rho}{2}+\sqrt{\left(b+\frac{\sigma^2}{2}-\frac{\rho}{2}\right)^2+1}\,.$$

Proof. The proof is the same as in (R1). A difference is the appearance of constant σ^2. The fractional scheme is unchanged. □

We note that $\breve{u}$ is the negative solution to

$$u^2+(2b+\sigma^2-\rho)u-1 \;=\; 0.$$

Thus we have the optimal control function with rate $\breve{u}$:

$$\hat{u}(x) \;=\; \breve{u}x.$$

The value function

$$v(x) \;=\; mx^2$$

is optimal in the class of proportional controls.

We note that zero-sum property holds true even now:

$$\breve{u}+m \;=\; 0.$$

5.2 Dynamic programming

The value function $v : R^1 \to R^1$ satisfies the Bellman equation

$$\rho v(x) \;=\; \min_{u\in R^1}\left[x^2+u^2+(bx+u)v'(x)+\frac{\sigma^2x^2}{2}v''(x)\right]. \tag{16}$$

Eq.(16) is solved as follows. First $\frac{d}{du}[\cdots] = 0$ implies that

$$\rho v(x) = x^2 - \frac{1}{4}v'^2(x) + bxv'(x) + \frac{\sigma^2 x^2}{2}v''(x), \qquad \hat{u}(x) = -\frac{1}{2}v'(x).$$

This linear-quadratic scheme enables us to assume that v is quadratic $v(x) = vx^2$ ($v \geq 0$). Substituting $v'(x) = 2vx$, $v''(x) = 2v$, we have

$$\rho v x^2 = x^2 - v^2x^2 + 2bvx^2 + \sigma^2 v x^2$$

i.e.

$$\rho v = 1 - v^2 + (2b + \sigma^2)v.$$

This yields the quadratic equation

$$v^2 - (2b + \sigma^2 - \rho)v - 1 = 0. \tag{17}$$

Here we note that this equation is similar to Eqs.(7) in D(c) and (13) in O(x). A difference is the appearance of σ^2.

Eq.(17) has a unique positive solution

$$\tilde{v} = b + \frac{\sigma^2}{2} - \frac{\rho}{2} + \sqrt{\left(b + \frac{\sigma^2}{2} - \frac{\rho}{2}\right)^2 + 1}\,.$$

Thus we have the desired optimal solution

$$v(x) = \tilde{v}x^2, \quad \hat{u}(x) = -\tilde{v}x.$$

We note that the certainty equivalence principle does not holds true for $\sigma > 0$. When in particular $\sigma = 0$, it holds that $\tilde{v} = \hat{v}$, where

$$\hat{v} = b - \frac{\rho}{2} + \sqrt{\left(b - \frac{\rho}{2}\right)^2 + 1}$$

is given both in deterministic dynamics and in Ornstein-Uhlenbeck process.

6 Bellman Equation

Let us now derive Bellman equation both for deterministic control process and for stochastic one under existence of optimal process. In this section we assume that f, $g : R^2 \to R^1$ are continuous.

6.1 Deterministic control process

We consider a general control process with discounted cost function :

$$
\mathrm{D}(x) \qquad \begin{array}{ll} \text{minimize} & \displaystyle\int_0^\infty e^{-\rho t} f(x,u)dt \\ \text{subject to} & \text{(i)} \quad \dot{x} = g(x,u) \\ & \text{(ii)} \quad x \in C_p^1,\ u \in U(x) \\ & \text{(iii)} \quad x(0) = x \end{array}
$$

where C_p^1 is the set of all continuously differentiable functions except for a finite set of points. Let $v(x)$ be the minimum value. Then the value function $v : R^1 \to R^1$ satisfies the Bellman equation :

$$
\rho v(x) = \min_{u \in U(x)} [f(x,u) + v'(x)g(x,u)] \quad x \in R^1. \tag{18}
$$

This has been derived by applying intuitively Priciple of Optimality (see [1–3]).

Now we derive Eq.(18) under assumption:

1. $v \in C^1$.
2. There exists a feasible process (x,u) such that

$$
v(x) = \int_0^\infty e^{-\rho s} f(x,u)ds \quad \forall x \in R^1. \tag{19}
$$

The feasibility denotes a solution to differential equation (i) – (iii). A process (x,u) satisfying (19) is called *optimal process.*

We take any feasible paired process (x,u). Let us take any small $\Delta > 0$. Then we define a new process (y,z) as follows :

$$
y(t) := x(t+\Delta), \ z(t) := u(t+\Delta), \quad t \in [0,\infty).
$$

Then the process $y = \{y(\cdot)\}_{[0,\infty)}$ satisfies

$$
\begin{array}{ll} \text{(i)}' & \dot{y} = g(y,w) \\ \text{(ii)}' & y \in C_p^1,\ z \in U(y) \\ \text{(iii)}' & y(0) = x(\Delta). \end{array} \qquad 0 \le t < \infty
$$

Conversely, concatenating the process (x,u) on time-interval $[0,\Delta]$ for any process(y,z) satisfying (i)$'$ – (iii)$'$, we can construct a (x,u)-process on the interval $[0,\infty)$ satisfying conditions (i) – (iii). Then the x is in C_p^1.

First we take the feasible process (x, u) in (19). From the discounted stationary accumulation, we get for any $\Delta(>0)$

$$\begin{aligned} v(x) &= \int_0^\Delta e^{-\rho s} f(x,u)ds + \int_\Delta^\infty e^{-\rho s} f(x,u)ds \\ &= \int_0^\Delta e^{-\rho s} f(x,u)ds + e^{-\rho\Delta} \int_0^\infty e^{-\rho s} f(y,w)ds \\ &= \int_0^\Delta e^{-\rho s} f(x,u)ds + e^{-\rho\Delta} v(x(\Delta)). \end{aligned} \tag{20}$$

From the mean value theorem, there exists $\theta(0 \leq \theta \leq 1)$ satisfying

$$\int_0^\Delta e^{-\rho s} f(x,u)ds = h(\theta\Delta)\Delta$$

where $h(s) = e^{-\rho s} f(x(s), u(s))$. It holds that

$$e^{-\rho\Delta} = 1 - \rho\Delta + o(\Delta).$$

We note that

$$x(\Delta) = x + g(x,u)\Delta + o(\Delta)$$

where $x = x(0),\ u = u(0)$. This implies that

$$v(x(\Delta)) = v(x) + v'(x)\,[\,g(x,u)\Delta + o(\Delta)\,] + o(\Delta).$$

Hence we obtain

$$\begin{aligned} &\int_0^\Delta e^{-\rho s} f(x,u)ds + e^{-\rho\Delta} v(x(\Delta)) \\ &\quad = h(\theta\Delta)\Delta + v(x)(1-\rho\Delta) + v'(x)g(x,u)\Delta + o(\Delta). \end{aligned} \tag{21}$$

Combining (20) and (21), we get

$$v(x) = h(\theta\Delta)\Delta + v(x)(1-\rho\Delta) + v'(x)g(x,u)\Delta + o(\Delta).$$

Subtracting $v(x)$, dividing it by Δ and letting Δ tend to zero, we have the equality

$$\rho v(x) = f(x,u) + g(x,u)v'(x). \tag{22}$$

On the other hand, let (x, u) be *any* feasible process. We take any $\Delta(>0)$. From the definition of $v(x)$, we have

$$v(x) \leq \int_0^\Delta e^{-\rho s} f(x,u)ds + e^{-\rho\Delta} \int_0^\infty e^{-\rho s} f(y,w)ds \tag{23}$$

The inequality (23) holds for any feasible process $(x,\ u)$ on $[0,\ \infty)$. We note that

$$v(y) \ = \ \min\left[\int_0^{\infty} e^{-\rho s} f(y,w)ds \ \mid \ y(0)=y\right]$$

where the minimization is over all feasible processes on $[0,\ \infty)$. We have assumed the existence of a "minimum" process. A monotonicity works. From (23), we have

$$v(x) \ \le \ \int_0^{\Delta} e^{-\rho s} f(x,u)ds + e^{-\rho\Delta} v(x(\Delta)). \tag{24}$$

Then again, we get

$$\begin{aligned}&\int_0^{\Delta} e^{-\rho s} f(x,u)ds + e^{-\rho\Delta} v(x(\Delta))\\ &\quad = \ h(\theta\Delta)\Delta + v(x)(1-\rho\Delta) + v'(x)g(x,u)\Delta + o(\Delta).\end{aligned} \tag{25}$$

From (24) and (25), we have

$$v(x) \ \le \ h(\theta\Delta)\Delta + v(x)(1-\rho\Delta) + v'(x)g(x,u)\Delta + o(\Delta) \qquad \forall u = u(0).$$

This in turn leads to

$$\rho v(x) \ \le \ f(x,u) + g(x,u)v'(x) \qquad \forall u. \tag{26}$$

A combination of (22) and (26) yields the desired forward equation. □

6.2 Stochastic control process

Let $\{w(\cdot)\}$ be the one-dimensional standard Brownian motion and $\sigma : R^1 \to [0,\ \infty)$ be continuous. We consider a general control process with discounted criterion $e^{-\rho t} f = e^{-\rho t} f(x,u)$ and stochastic dynamics $dx(t) = g(x,u)dt + \sigma(x)dw(t)$ on an infinite time-period $[0,\infty)$:

$$S(x) \qquad \begin{array}{ll} \text{minimize} & E_x\left[\displaystyle\int_0^{\infty} e^{-\rho t} f(x,u)dt\right] \\ \text{subject to} & \text{(i)} \quad dx(t) = g(x,u)dt + \sigma(x)dw(t) \\ & \text{(ii)} \quad x \in C, \ \ u \in U(x) \\ & \text{(iii)} \quad x(0) = x \end{array}$$

where $x \in R^1$ is a given initial state.

Now we derive a Bellman equation for S(x) through forward approach. Let $v(x)$ be the minimum value of $S(x)$. Then the value function $v : R^1 \to R^1$ satisfies the Bellman equation

$$\rho v(x) \ = \ \min_{u\in U(x)}\left[f(x,u) + g(x,u)v'(x) + \frac{1}{2}\sigma^2(x)v''(x)\right] \quad x \in R^1. \tag{27}$$

Eq.(27) is derived under assumption:

1. $v \in C^2$.

2. There exists a feasible policy function $u : R^1 \to R^1$ such that the feasible process $(x, u) = (x(t), u(x(t)))$ satisfies

$$v(x) = E_x \left[\int_0^\infty e^{-\rho s} f(x, u) ds \right] \qquad \forall x \in R^1. \tag{28}$$

The feasible policy function denotes $u(x) \in U(x)$ for any $x \in R^1$. The feasible process denotes a solution to stochastic differential equation (i) – (iii).

Let us take any small $\Delta > 0$. For any feasible paired process (x, u) for $S(x)$, we define a paired process (y, z) on $[0, \ \infty)$ by

$$y(s) := x(s + \Delta), \ \ z(s) := u(s + \Delta), \quad s \in [0, \infty).$$

Then the stochastic process $y = \{y(\cdot)\}_{[0,\infty)}$ with any fixed initial state $y \in R^1$ satisfies

$$\begin{aligned} &(\mathrm{i})' \quad dy(s) = g(y, z)ds + \sigma(y)dw(s) \\ &(\mathrm{ii})' \quad y \in C, \ z \in U(y) \\ &(\mathrm{iii})' \quad y(0) = y. \end{aligned}$$

The process (x, u) on $[0, \infty)$ induces a family of of processes (y, z) on $[0, \infty)$ with initial state $y(0) = y$, where the family is the set of all paired processes parametrized with $y \in R^1$. Conversely, let a family of processes (y, z) satisfying (i)′ – (iii)′ be given. Then, concatenating the process (x, u) on interval $[0, \Delta)$ for the family, we can construct a process (x, u) on $[0, \infty)$ satisfying conditions (i) – (iii).

First we take the feasible process (x, u) in (28). From the Markov property and the discounted stationary accumulation, we get for any $\Delta(> 0)$

$$\begin{aligned} v(x) &= E_x \left[\int_0^\Delta e^{-\rho s} f(x, u) ds + \int_\Delta^\infty e^{-\rho s} f(x, u) ds \right] \\ &= E_x \left[\int_0^\Delta e^{-\rho s} f(x, u) ds + e^{-\rho \Delta} E \left[\int_0^\infty e^{-\rho s} f(y, z) ds \,\middle|\, x(\Delta) \right] \right] \\ &= E_x \left[\int_0^\Delta e^{-\rho s} f(x, u) ds + e^{-\rho \Delta} v(x(\Delta)) \right]. \end{aligned} \tag{29}$$

From the mean value theorem, there exists $\theta (0 \le \theta \le 1)$ satisfying

$$\int_0^\Delta e^{-\rho s} f(x, u) ds = h(\theta \Delta) \Delta \qquad \text{a.s. } P_x$$

where $h(s) = e^{-\rho s} f(x(s), u(s))$. It holds that

$$e^{-\rho \Delta} = 1 - \rho \Delta + o(\Delta).$$

Since

$$x(\Delta) = x + g(x,u)\Delta + \sigma(x)(w(\Delta) - w(0)) + o(\Delta) \quad \text{a.s. } P_x \quad (x = x(0),\ u = u(0))$$

it follows that

$$\begin{aligned} v(x(\Delta)) &= v(x) + v'(x)\,[\,g(x,u)\Delta + \sigma(x)w(\Delta)\,] \\ &\quad + \frac{1}{2}v''(x)\,[\,g(x,u)\Delta + \sigma(x)w(\Delta)\,]^2 + o(\Delta) \quad \text{a.s. } P_x. \end{aligned}$$

From

$$E_x\,[\,w(\Delta)\,] = E_x\,[\,w(\Delta)\Delta\,] = 0, \qquad E_x\,[\,w(\Delta)w(\Delta)\,] = \Delta,$$

we obtain

$$\begin{aligned} & E_x\left[\int_0^\Delta e^{-\rho s} f(x,u)ds + e^{-\rho\Delta}v(x(\Delta))\right] \\ &= f(x,u)\Delta + v(x)(1-\rho\Delta) + v'(x)g(x,u)\Delta + \frac{1}{2}v''(x)\sigma^2(x)\Delta + o(\Delta). \end{aligned} \tag{30}$$

Combining (29) and (30), we get

$$v(x) = f(x,u)\Delta + v(x)(1-\rho\Delta) + v'(x)g(x,u)\Delta + \frac{1}{2}v''(x)\sigma^2(x)\Delta + o(\Delta).$$

Subtracting $v(x)$, dividing it by Δ and letting Δ tend to zero, we have the equality

$$\rho v(x) = f(x,u) + g(x,u)v'(x) + \frac{1}{2}\sigma^2(x)v''(x). \tag{31}$$

On the other hand, let (x, u) be *any* feasible process. We take any $\Delta(>0)$. From the definition of $v(x)$ and Markov property, we have

$$v(x) \leq E_x\left[\int_0^\Delta e^{-\rho s} f(x,u)ds + e^{-\rho\Delta} E\left[\int_0^\infty e^{-\rho s} f(y,z)ds \mid x(\Delta)\right]\right]. \tag{32}$$

This inequality holds for any feasible process $(x,\ u)$ on $[0,\ \infty)$. We note that

$$v(y) = \min E\left[\int_0^\infty e^{-\rho s} f(x,u)ds \mid x(\Delta) = y\right]$$

where the minimization is over all feasible processes on $[0,\ \infty)$. We have assumed the existence of a "minimum" process. A monotonicity works as follows. If $X \leq Y$, then $E[c+X] \leq E[c+Y]$. From (32), we have

$$v(x) \leq E_x\left[\int_0^\Delta e^{-\rho s} f(x,u)ds + e^{-\rho\Delta}v(x(\Delta))\right]. \tag{33}$$

Then again, we get

$$\begin{aligned} & E_x\left[\int_0^{\Delta} e^{-\rho s} f(x,u)ds + e^{-\rho\Delta} v(x(\Delta))\right] \\ = \; & f(x,u)\Delta + v(x)(1-\rho\Delta) + v'(x)g(x,u)\Delta + \frac{1}{2}v''(x)\sigma^2(x)\Delta + o(\Delta). \qquad (34) \end{aligned}$$

From (33) and (34), we have

$$v(x) \;\le\; f(x,u)\Delta + v(x)(1-\rho\Delta) + v'(x)g(x,u)\Delta + \frac{1}{2}v''(x)\sigma^2(x)\Delta + o(\Delta).$$

This in turn leads to

$$\rho v(x) \;\le\; f(x,u) + g(x,u)v'(x) + \frac{1}{2}\sigma^2(x)v''(x) \qquad \forall u. \qquad (35)$$

A combination of (31) and (35) yields the desired forward equation. □

References

[1] R.E. Bellman, *Dynamic Programming*, Princeton Univ. Press, NJ, 1957.

[2] R.E. Bellman, *Introduction of the Mathematical Theory of Control Processes, Vol.I, Linear Equations and Quadratic Criteria; Vol.II, Nonlinear Processes*, Academic Press, NY, 1967; 1971.

[3] R.E. Bellman, *Methods of Nonlinear Analysis, Vol.I, Vol.II,* Academic Press, New York, 1969, 1972.

[4] D.P. Bertsekas, *Dynamic Programming and Stochastic Control* Academic Press, New York, 1976.

[5] D.P. Bertsekas and S.E. Shreve, *Stochastic Optimal Control: The Discrete-time Case* Academic Press, New York, 1978.

[6] I. Karatzas and S.E. Shreve, *Brownian Motion and Stochastic Calculus. 2nd ed.*, Springer-Verlag, NY, 1991.

[7] I. Karatzas and S.E. Shreve, *Methods of Mathematical Finance.* Springer-Verlag, NY, 1998.

[8] S. Iwamoto, *Theory of Dynamic Program: Japanese*, Kyushu Univ. Press, Fukuoka, 1987.

[9] S. Iwamoto and M. Yasuda, "Dynamic programming creates the Golden Ratio, too," *Proc. of the Sixth Intl Conference on Optimization: Techniques and Applications (ICOTA 2004)*, Ballarat, Australia, December 2004.

[10] S. Iwamoto and M. Yasuda, Golden optimal path in discrete-time dynamic optimization processes, Ed. S. Elaydi, K. Nishimura, M. Shishikura and N. Tose, Advanced Studies in Pure Mathematics 53, 2009, Advances in Discrete Dynamic Systems, pp.77–86.

[11] L. Ljungqvist and T.J. Sargent, *Recursive Macroeconomic Theory*, 2nd Edition, MIT Press, Mass., 2004.

[12] N.L. Stokey and R.E. Lucas, *Recursive Methods in Economic Dynamics.* Harvard University Press, Cambridge, Mass., 1989.

Department of Economic Engineering
Graduate School of Economics, Kyushu University
Fukuoka 813-0012, Japan
E-mail address: iwamotodp@kyudai.jp

九州大学・大学院 経済学研究院　岩本 誠一