Golden optimal processes on three dynamics: deterministic, stochastic and non-deterministic (Mathematical Economics)

Author(s): IWAMOTO, Seiichi

Citation: 数理解析研究所講究録 (2009), 1654: 1-18

Issue Date: 2009-06

URL: http://hdl.handle.net/2433/140847

Type: Departmental Bulletin Paper

Textversion: publisher: Kyoto University
Golden optimal processes on three dynamics: deterministic, stochastic and non-deterministic

Seiichi IWAMOTO

Department of Economic Engineering
Graduate School of Economics, Kyushu University
Fukuoka 812-8581, Japan
tel&fax. +81(92)642-2488, email: iwamoto@en.kyushu-u.ac.jp

February 23, 2009

Abstract

This paper discusses a common criterion on three different dynamics. The criterion is discounted quadratic. The dynamics are deterministic, stochastic and non-deterministic. We consider two problems from a viewpoint of Golden optimality. The first problem is to find an optimal solution – value function and optimal policy –. The second problem is to discuss whether the optimal solution is Golden or not. Is the value function Golden? Is the optimal policy Golden? We give a complete solution to the first problem through two approaches – evaluation-optimization method and dynamic programming method –. The solution of the second depends on the discount rate $\beta (0 < \beta < \infty)$. We show that both – deterministic and non-deterministic – dynamics allow the Golden optimal solution for $\beta = 1$. Further all the three dynamics allow the Golden optimal policy for $\beta = \frac{1}{\sqrt{5}}$.

Keywords: golden, optimal, policy, deterministic, stochastic, non-deterministic.
JEL classification: C61

1 Introduction

The Golden ratio is the symbol of beauty and practical use. It has been utilized in architecture, art, design, biology, science, engineering, and others [13]. Recently it has been incorporated into optimization problems. There a new – Golden (and) optimal – solution is obtained. Both static problems [5–7] and dynamic problems [8, 10, 11] are studied from the Golden optimality. The static optimization is two-variable. The dynamic
one is infinite variable — discrete-horizon [10,11] and continuous-time [8] —. All of them are on deterministic system.

This paper minimizes a discounted quadratic criterion on three — (1) deterministic, (2) stochastic and (3) non-deterministic — dynamics. We consider two problems from a viewpoint of Golden optimality. The first problem is to find an optimal solution — (i) value function, (ii) optimal policy, (iii) minimum value —. The second problem is to discuss whether the optimal solution is Golden or not. Our approaches are evaluation-optimization method and dynamic programming method. We give an optimal solution to the first. The solution of the second depends on the discount rate $\beta (0 < \beta < \infty)$. We show that both — (1) deterministic and (3) non-deterministic — dynamics allow the Golden optimal solution for $\beta = 1$. Further all the three dynamics allow the Golden optimal policy for $\beta = \frac{1}{\sqrt{5}} \approx 0.4772$.

2 Golden Paths

A real number

$$\phi = \frac{1 + \sqrt{5}}{2} \approx 1.618$$

is called Golden number. It is the larger of the two solutions to quadratic equation

$$x^2 - x - 1 = 0. \tag{1}$$

Sometimes (1) is called Fibonacci quadratic equation. The Fibonacci quadratic equation has two real solutions: $\phi$ and its conjugate $\overline{\phi} := 1 - \phi$. We note that

$$\phi + \overline{\phi} = 1, \quad \phi \cdot \overline{\phi} = -1.$$

Further we have

$$\phi^{-1} = \phi - 1 \approx 0.618, \quad \phi^{-2} = 2 - \phi \approx 0.382$$

$$\phi^{-1} + \phi^{-2} = 1.$$

A point $\phi^{-2}x$ splits an interval $[0, x]$ into two intervals $[0, \phi^{-2}x]$ and $[\phi^{-2}x, x]$. A point $\phi^{-1}x$ splits the interval into $[0, \phi^{-1}x]$ and $[\phi^{-1}x, x]$. In either case, the length constitutes the Golden ratio $\phi^{-2} : \phi^{-1} = 1 : \phi$. Thus both divisions are the Golden section.

**Definition 2.1** A sequence $\{x_n\}_0^\infty$ is called Golden if and only if either

$$x_{n+1} = \phi^{-1}x_n \quad n \geq 0 \quad \text{or} \quad x_{n+1} = \phi^{-2}x_n \quad n \geq 0.$$

**Lemma 2.1** A Golden sequence $\{x_n\}_0^\infty$ is either

$$x_n = c\phi^{-n} \quad n \geq 0 \quad (\text{Fig. 2}) \quad \text{or} \quad x_n = c\phi^{-2n} \quad (\text{Fig. 1}),$$

where $c$ is a real constant.
Definition 2.2 A sequence of Markov random variables $\{X_n\}_{0}^{\infty}$ with $X_0 = x_0$ is called Golden if and only if either

$$E[X_{n+1} | x_n] = \phi^{-1}x_n \ n \geq 0 \ or \ E[X_{n+1} | x_n] = \phi^{-2}x_n \ n \geq 0.$$
We remark that either Golden sequence is supermartingale. In either case, $E[X_{n+1} | x_n]$ generates a Golden section of interval $[0, x_n]$ for $x_n \geq 0$ and does a Golden section of interval $[x_n, 0]$ for $x_n \leq 0$.

Let $\{\epsilon_n\}^\infty_1$ be a sequence of independent and identical random variables with the standard normal distribution. Then

$$E[\epsilon_n] = 0, \quad E[\epsilon_n^2] = 1.$$ 

For given $x_0$ and $y_0$ we define two sequences of Markov random variables $\{X_n\}$ and $\{Y_n\}$ by

$$X_{n+1} = \phi^{-1}X_n - \epsilon_{n+1}, \quad X_0 = x_0$$
$$Y_{n+1} = \phi^{-2}Y_n - \epsilon_{n+1}, \quad Y_0 = y_0.$$ 

**Lemma 2.2** Then $\{X_n\}$ and $\{Y_n\}$ are Golden.

## 3 Three Dynamics

We consider three dynamic optimization problems with a common discounted quadratic criterion.

The first is a deterministic dynamics on which we minimizes a typical quadratic function. The problem is called linear-quadratic (LQ) [2, 3]:

\[
\text{minimize } \sum_{n=0}^{\infty} \beta^n (x_n^2 + u_n^2) \\
\text{subject to } (i) \quad x_{n+1} = x_n - u_n, \quad n \geq 0 \\
(ii) \quad u_n \in R^1 \\
(iii) \quad x_0 = c,
\]

where $c \in R^1$. Here (i) denotes that next state $x_{n+1}$ turns out to be $x_n - u_n$ with certainty from state $x_n$ under decision $u_n$. This dynamics together with immediate cost is depicted as

$$R^1 \ni x_n \downarrow u_n \in R^1 \mapsto x_n^2 + u_n^2$$

a unique $x_{n+1} := x_n - u_n \in R^1$,

where $\mapsto$ denotes that state $x_n$ under decision $u_n$ yields the stage-cost $x_n^2 + u_n^2$.

The second is a stochastic dynamics on which we minimizes the expected value of the
same quadratic function as in deterministic one:

$$\text{minimize } E_{x_0} \left[ \sum_{n=0}^{\infty} \beta^n (x_n^2 + u_n^2) \right]$$

subject to \( (i) \ x_{n+1} = x_n - u_n - \epsilon_{n+1} \)

\( (ii) \ u_n \in R^1 \)

\( (iii) \ x_0 = c \)

where \( c \in R^1 \). The problem \( (S) \) is called stochastic linear-quadratic (LQ). Here \( \{\epsilon_n\}_{1}^{\infty} \) is a sequence of independently and identically distributed random variables with the standard normal distribution. Thus \( (i) \) denotes that \( x_{n+1} \) appears on \( R^1 \) with transition probability \( q(x_{n+1} | x_n, u_n) = \frac{1}{\sqrt{2\pi}} e^{-(x_{n+1} - x_n + u_n)^2/2} \) from \( x_n \) under \( u_n \). This dynamics is depicted as

\[
\begin{align*}
R^1 \ni x_n \downarrow u_n \in R^1 & \Rightarrow x_{n+1} \text{ w.p. } q(x_{n+1} | x_n, u_n), \text{ for any } x_{n+1} \in R^1.
\end{align*}
\]

The third is on non-deterministic dynamics. There we minimize a total discounted weighted value of quadratic cost:

$$\text{minimize } \sum_{n=0}^{\infty} \beta^n W_{x_0} [x_n^2 + u_n^2]$$

subject to \( (i) \ 0 < x_{n+1} < x_n - u_n \) with weight \( 2/x_{n+1} \)

\( (ii) \ u_n \in R^1 \)

\( (iii) \ x_0 = c, \)

where \( c > 0 \). We call the problem \( (N) \) is non-deterministic quadratic (Q). Here the infinite series is defined in Section 6. The successive constraint \( (i) \) denotes that \( x_{n+1} \) appears on the open interval \((0, x_n - u_n)\) with transition weight \( 2/x_{n+1} \) from \( x_n \) under \( u_n \). This dynamics is depicted as

\[
(0, \infty) \ni x_n \downarrow u_n \in (-\infty, x_n) & \Rightarrow x_{n+1} \text{ w.w. } \frac{2}{x_{n+1}}, \text{ for any } x_{n+1} \in (0, x_n - u_n).
\]

A characteristic feature of the dynamics is as follows. As next state degenerates small, its weight grows unboudedly large. The total weight from any state \( x_n \) under any decision \( u_n \) diverges to \( \infty \) as long as \( x_n - u_n > 0 \):

$$\int_{0}^{x_n-u_n} \frac{2}{x_{n+1}} dx_{n+1} = \infty.$$
4 Deterministic Dynamics

Let us consider the discounted quadratic criterion on deterministic dynamics:

\[
\text{minimize } \sum_{n=0}^{\infty} \beta^n (x_n^2 + u_n^2) \\
\text{subject to } \begin{align*}
(i) & \quad x_{n+1} = x_n - u_n, & n \geq 0 \\
(ii) & \quad u_n \in R^1 \\
(iii) & \quad x_0 = c,
\end{align*}
\]

where \(c \in R^1\) is a given constant.

4.1 Evaluation-optimization

Let us solve (D) through evaluation-optimization method, which has two stages. The first stage evaluates any policy in a class of policies and the second minimizes the evaluated value over the class.

A stationary policy \(f^\infty\) is called \textit{proportional} if the decision function is specified by \(f(x) = px\), where \(p\) is a real constant. Then \(p\) is called a \textit{proportional rate}. In this subsection, we consider the set of all proportional policies whose rate \(p\) satisfies \(\beta(1-p)^2 < 1\).

**Lemma 4.1** A proportional policy \(f^\infty, f(x) = px\), yields the objective value

\[
\sum_{n=0}^{\infty} \beta^n (x_n^2 + u_n^2) = \frac{r}{1 - \beta q} x_0^2,
\]

where \(r = 1 + p^2, \quad q = (1 - p)^2\).

Now let us consider the ratio minimization problem

\[
\text{minimize } \frac{r}{1 - \beta q} \quad \text{subject to } \beta q < 1.
\]

This is expressed as a single-variable problem:

\[
\text{(C}_\beta) \quad \text{minimize } \frac{1 + p^2}{1 - \beta(1-p)^2} \\
\text{subject to } 1 - \frac{1}{\sqrt{\beta}} < p < 1 + \frac{1}{\sqrt{\beta}}.
\]

**Lemma 4.2** The problem \((C_\beta)\) has the minimum value

\[
m = \frac{2\beta - 1 + \sqrt{4\beta^2 + 1}}{2\beta} \quad \text{at } \hat{p} = \frac{\sqrt{4\beta^2 + 1} - 1}{2\beta}.
\]
Thus we have the optimal policy \( \hat{f}^\infty \);
\[
\hat{f}(x) = \hat{p}x, \quad \hat{p} = \frac{\sqrt{4\beta^2 + 1} - 1}{2\beta}
\]
in the proportional policy class and the value function
\[
v(x) = mx^2, \quad m = \frac{2\beta - 1 + \sqrt{4\beta^2 + 1}}{2\beta}.
\]
We remark that
\[
m = 1 + \hat{p}.
\]

### 4.2 Dynamic programming

In this subsection, we apply dynamic programming to optimize the infinite stage problem \([1, 4, 8, 12]\).

Let \( v(c) \) be the minimum value for \( c \in \mathbb{R}^1 \). Then \( v : \mathbb{R}^1 \to \mathbb{R}^1 \) is called a value function. The value function \( v \) satisfies the Bellman equation:
\[
v(x) = \min_{-\infty < u < \infty} [x^2 + u^2 + \beta v(x-u)], \quad v(0) = 0. \tag{2}
\]

**Lemma 4.3** The control process (D) has the proportional optimal policy \( f^\infty, f(x) = px \), and the quadratic value function \( v(x) = vx^2 \), where
\[
v = \frac{2\beta - 1 + \sqrt{4\beta^2 + 1}}{2\beta}, \quad p = \frac{\sqrt{4\beta^2 + 1} - 1}{2\beta}.
\]

The proportional optimal policy \( f^\infty \) splits at any time an interval \( [0, x] \) into \( [0, (1-p)x] \) and \( \left[ \frac{x}{1 + \beta v}, x \right] \). When, in particular, \( \beta = 1 \), the quadratic coefficient \( v \) is reduced to the Golden number
\[
\phi = \frac{1 + \sqrt{5}}{2} \approx 1.618
\]
and the proportional rate \( p \) is reduced to its inverse number
\[
\phi^{-1} = \phi - 1 = \frac{\sqrt{5} - 1}{2} \approx 0.618
\]
Further the division of \([0, x]\) into \([0, \phi^{-2}x]\) and \([\phi^{-2}x, x]\) is Golden. That is, the ratio of length of two intervals constitutes the Golden ratio:
\[
\phi^{-2} : \phi^{-1} = 1 : \phi.
\]
A quadratic function \( w(x) = ax^2 \) is called Golden if \( a = \phi \).

**Theorem 4.1** The control process (D) with unit discount rate \( \beta = 1 \) has a Golden optimal policy \( f^\infty, f(x) = \phi^{-1}x \), and the Golden quadratic value function \( v(x) = \phi x^2 \).
5 Stochastic Dynamics

Let us consider the stochastic dynamic process under the condition that the discount rate $\beta$ should be $0 < \beta < 1$. Soon it will be clarified that the expected value diverges for the case $\beta \geq 1$. Our stochastic dynamic minimization problem is

$$\text{minimize } E_{x_0} \left[ \sum_{n=0}^{\infty} \beta^n (x_n^2 + u_n^2) \right]$$

subject to

(i) $x_{n+1} = x_n - u_n - \epsilon_{n+1}$

(ii) $u_n \in \mathbb{R}^1$

(iii) $x_0 = c$,

where an initial state $c \in \mathbb{R}^1$ is given, and $\{\epsilon_n\}$ is a sequence of random variables that is independently and identically distributed through time and obeys the standard normal distribution. Thus

$$E[\epsilon_n] = 0, \quad E[\epsilon_n^2] = 1.$$  

We note that $\epsilon_n$ has the probability density function

$$p(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \quad -\infty < z < \infty.$$  

The next state (random variable) $x_{n+1}$ obeys the normal distribution with mean $x_n - u_n$ and unit variance, provided that a decision $u_n$ is taken at state $x_n$ on stage $n$. When a decision maker adopts a decision $u$ on state $x$, the system will go to state (scalar) $y$ with probability $q(y|x,u) = p(y-x+u)$:

$$q(y|x,u) = \frac{1}{\sqrt{2\pi}} e^{-(y-x+u)^2/2} \quad -\infty < y < \infty.$$  

We depict this dynamics as

$$(-\infty, \infty) \ni x \downarrow u \in (-\infty, \infty) \mapsto x^2 + u^2 \ni y \text{ w.p. } q(y|x,u) \text{ for any } y \in (-\infty, \infty).$$

5.1 Evaluation-optimization

Let us evaluate any proportional policy $f^\infty, f(x) = px$ for $0 < p < 2$. The decision maker adopts the decision $u_n = f(x_n) = px_n$ on state $x_n$. Hence

$$E_{x_0} \left[ \sum_{n=0}^{\infty} \beta^n (x_n^2 + u_n^2) \right] = r \sum_{n=0}^{\infty} \beta^n E_{x_0} [x_n^2],$$

where $r = 1 + p^2$. The controlled dynamics $x_{n+1} = x_n - u_n - \epsilon_{n+1}$ is reduced to

$$x_{n+1} = (1-p)x_n - \epsilon_{n+1}, \quad x_0 = c.$$ (3)

Here we note that $|1-p| < 1.$
Lemma 5.1 It follows that under (3)
\[ E[x_n^2] = \frac{1}{1-q} + \left( x_0^2 - \frac{1}{1-q} \right) q^n, \tag{4} \]
where \( q = (1-p)^2 \).

Lemma 5.2 A proportional policy \( f^\infty, f(x) = px \), yields the expected value
\[ E_{x_0} \left[ \sum_{n=0}^{\infty} \beta^n (x_n^2 + u_n^2) \right] = \frac{r}{1-\beta q} \left( x_0^2 + \frac{\beta}{1-\beta} \right), \]
where \( r = 1 + p^2, \ q = (1 - p)^2 \).

Note that the term \( x_0^2 + \frac{\beta}{1-\beta} \) is independent of \( p \). We have reached the ratio minimization problem \((C_{\beta})\) in the deterministic dynamics. Lemma 4.2 gives the minimum solution of \((C_{\beta})\).

Thus we have the optimal policy \( \hat{f}^\infty \);
\[ \hat{f}(x) = \hat{p}x, \quad \hat{p} = \frac{\sqrt{4\beta^2 + 1} - 1}{2\beta} \]
in the proportional policy class and the value function
\[ v(x) = mx^2 + \rho, \quad m = \frac{2\beta - 1 + \sqrt{4\beta^2 + 1}}{2\beta}, \quad \rho = \frac{2\beta - 1 + \sqrt{4\beta^2 + 1}}{2(1-\beta)}. \]

5.2 Dynamic programming

Let \( v(c) \) be the minimum value. Then the value function \( v : R^1 \to R^1 \) satisfies the Bellman equation:
\[ v(x) = \min_{-\infty<u<\infty} \left[ x^2 + u^2 + \beta E_x [v(x-u-\epsilon)] \right]. \tag{5} \]

This is also written as the controlled integral equation
\[ v(x) = \min_{-\infty<u<\infty} \left[ x^2 + u^2 + \frac{\beta}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(y-z)^2} v(y) dy \right]. \]

Lemma 5.3 The control process \((S)\) has a proportional optimal policy \( f^\infty, f(x) = px \), and a quadratic value function \( v(x) = vx^2 + \rho \), where
\[ v = \frac{2\beta - 1 + \sqrt{4\beta^2 + 1}}{2\beta}, \quad \rho = \frac{2\beta - 1 + \sqrt{4\beta^2 + 1}}{2(1-\beta)} \tag{6} \]
\[ p = \frac{\sqrt{4\beta^2 + 1} - 1}{2\beta}. \]

Thus we see that the stochastic dynamic system \((S)\) has the same optimal policy as deterministic dynamic system \((D)\). This is what we call certainty equivalence principle. The value function has a difference \( \rho \) which comes from the discounted total noise under uncertainty.
6 Non-deterministic Dynamics

Now we consider the minimization problem on non-deterministic dynamics:

\[
\text{minimize } \sum_{n=0}^{\infty} \beta^n W_{x_0} [x_n^2 + u_n^2] \\
\text{subject to } (i) \quad 0 < x_{n+1} < x_n - u_n \quad \text{with weight } \frac{2}{x_{n+1}} \quad n \geq 0 \\
(ii) \quad u_n \in R^1 \\
(iii) \quad x_0 = c,
\]

where \( c > 0 \) is a given constant. The constraints (i), (ii) yields the feasibily \(-\infty < u_n < x_n\). Here the \( n \)-th term is defined as follows.

\[
W_{x_0} [x_n^2 + u_n^2] = \int \cdots \int_R \gamma_0 \cdot \gamma_1 \cdots \gamma_{n-1} r_n \, dx_1 dx_2 \cdots dx_n \\
= \int \cdots \int_R 2^n (x_n^2 + u_n^2) x_1 x_2 \cdots x_n \, dx_1 dx_2 \cdots dx_n,
\]

where the transition weight function and cost function are stationary:

\[
\gamma_m = \gamma(x_m, u_m, x_{m+1}) = \frac{2}{x_{m+1}} \\
r_n = r(x_n, u_n) = x_n^2 + u_n^2.
\]

The integral domain \( R \) is determined through the sequence of decision functions \( f_0, f_1, \ldots, f_{n-1} \):

\[
R = \{(x_1, x_2, \ldots, x_n) | 0 < x_1 < x_0 - u_0, \ldots, 0 < x_n < x_{n-1} - u_{n-1} \} \\
\subset (0, \infty)^n,
\]

where \( u_m = f_m(x_m) \).

When \( n = 0 \), we have

\[
W_{x_0}[r_0] = x_0^2 + u_0^2.
\]

In the following, the sequence of states

\[
x \xrightarrow{u} y \xrightarrow{v} z \xrightarrow{w} a \xrightarrow{b} \ldots \xrightarrow{p} s \xrightarrow{q} t \rightarrow \ldots
\]

reads

\[
x_0 \xrightarrow{u_0} x_1 \xrightarrow{u_1} x_2 \xrightarrow{u_2} x_3 \xrightarrow{u_3} \ldots \xrightarrow{u_{n-2}} x_{n-1} \xrightarrow{u_{n-1}} x_n \rightarrow \ldots
\]
The first three weighted values are

\[ W_x [y^2 + v^2] = \int_C \frac{2(y^2 + v^2)}{y} dy \]
\[ W_x [z^2 + w^2] = \int\int_D \frac{2(z^2 + w^2)}{yz} dydz \]
\[ W_x [a^2 + b^2] = \int\int\int_E \frac{2(a^2 + b^2)}{yza} dydzda, \]

where

\[ C = \{ y \mid 0 < y < x - u(x) \} \subset (0, \infty) \]
\[ D = \{ (y, z) \mid 0 < y < x - u(x), 0 < z < y - v(y) \} \subset (0, \infty)^2 \]
\[ E = \{ (y, z, a) \mid 0 < y < x - u(x), 0 < z < y - v(y), 0 < a < z - w(z) \} \subset (0, \infty)^3 \]

We call

\[ W_{x_0} [x_n^2 + u_n^2] \quad \text{and} \quad \beta^n W_{x_0} [x_n^2 + u_n^2] \]

\( n \)-th weighted value and \( n \)-th discounted weighted value, respectively. The limit of series is called a total discounted weighted value. Thus the objective function (of \( x_0 \)) represents a total discounted weighted value by using policy \( \pi = \{ f_0, f_1, \ldots, f_{n-1}, \ldots \} \) from initial state \( x_0 \).

Then we consider the total discounted weighted value

\[ J(x_0; \pi) := W_{x_0} [r_0] + \beta W_{x_0} [r_1] + \cdots + \beta^n W_{x_0} [r_n] + \cdots \]

Thus our problem is to choose a policy which minimizes the discounted total weighted value. This is expressed as

\[ P(x_0) \quad \text{minimize} \quad J(x_0; \pi) \quad \text{subject to} \quad \pi \in \Pi. \]

### 6.1 Evaluation-optimization

First we evaluate any proportional policy \( \pi = f^\infty, f(x) = px \). The decision maker adopts a decision \( u = px \) on state \( x \) and the system will go to state \( y \) on open interval \( (0, x-u) = (0, (1-p)x) \) with the weight \( \frac{2}{y} \). Then we have inductively

\[ W_x [x_n^2 + u_n^2] = rq^n x^2. \]

**Lemma 6.1** A proportional policy \( \pi = f^\infty, f(x) = px \), yields the objective value

\[ \sum_{n=0}^{\infty} \beta^n W_{x_0} [x_n^2 + u_n^2] = \frac{r}{1 - \beta q} x_0^2, \]

where \( r = 1 + p^2 \), \( q = (1-p)^2 \).
Thus we have reached the same ratio minimization problem \((C_{\beta})\) as in deterministic dynamics. Therefore we have the optimal policy \(f^{\infty}\);

\[
\hat{f}(x) = \hat{p}x, \quad \hat{p} = \frac{\sqrt{4\beta^{2} + 1} - 1}{2\beta}
\]

and the value function

\[
v(x) = mx^2, \quad m = \frac{2\beta - 1 + \sqrt{4\beta^2 + 1}}{2\beta}.
\]

6.2 Dynamic programming

Let \(v(x_0)\) be the minimum value. Then the value function \(v : [0, \infty) \to \mathbb{R}^1\) satisfies the Bellman equation:

\[
v(x) = \min_{-\infty < u < \infty} \left[ x^2 + u^2 + \beta \int_{0}^{x-u} \frac{v(y)}{y} dy \right] v(0) = 0. \tag{7}
\]

This is also written as follows:

\[
v(x) = \min_{-\infty < u < \infty} \left[ x^2 + u^2 + \beta W_{x}^{u}[v] \right].
\]

We may assume that Eq.(7) has a quadratic form \(v(x) = vx^2\), where \(v \in \mathbb{R}^1\). We solve (7) as follows. Then we have

\[
\int_{0}^{x-u} 2 \frac{v(y)}{y} dy = \int_{0}^{x-u} 2vydy = v(x-u)^2.
\]

Eq.(7) is reduced to a minimum equation for scalar \(v\):

\[
v x^2 = \min_{-\infty < u < \infty} \left[ x^2 + u^2 + \beta v(x-u)^2 \right].
\]

Thus we have reached the same situation as in deterministic dynamics as was shown in (5).

**Lemma 6.2** The control process \((N)\) has the proportional optimal policy \(f^{\infty}, f(x) = px\), and the quadratic value function \(v(x) = vx^2\), where

\[
v = \frac{2\beta - 1 + \sqrt{4\beta^2 + 1}}{2\beta}, \quad p = \frac{\sqrt{4\beta^2 + 1} - 1}{2\beta}.
\]

Thus as in deterministic dynamics, we have the same result on Golden optimality:

**Theorem 6.1** The control process \((N)\) with unit discount rate \(\beta = 1\) has a Golden optimal policy \(f^{\infty}, f(x) = (\phi - 1)x\), and the Golden quadratic value function \(v(x) = \phi x^2\).
7 Golden Policies

Let us now discuss whether the desired optimal policy is Golden or not. Throughout three presections, we have obtained a common optimal solution. The optimal policy both for stochastic process and for non-deterministic process is identical with the optimal policy for the deterministic process. This is called *certainty equivalence principle*. The three control processes — (D), (S) and (N) — have a common proportional optimal policy

\[ f^\infty; f(x) = px \]

and the quadratic value function

\[ v(x) = \begin{cases} v x^2 & \text{for } (D), (N) \\ v x^2 + \rho & \text{for } (S), \end{cases} \]

where

\[ p = \frac{\sqrt{4\beta^2 + 1} - 1}{2\beta} = \frac{2\beta}{1 + \sqrt{4\beta^2 + 1}} \]
\[ v = \frac{2\beta - 1 + \sqrt{4\beta^2 + 1}}{2\beta}, \quad \rho = \frac{2\beta - 1 + \sqrt{4\beta^2 + 1}}{2(1 - \beta)} \]  

(8)

The rate \( p \) is determined by the coefficient \( v \):

\[ p = v - 1. \]

The proportional optimal policy \( f^\infty \) splits any interval \([0, x]\) into \([0, (1-p)x]\) and \([(1-p)x, x]\). We are interested in values of discount factor \( \beta \) which yields the two Golden sections. This asks us when \( 1 - p \) becomes \( \phi - 1 \) or \( 2 - \phi \).

Let us now consider both \( p \) and \( 1 - p \) as functions of \( \beta \). We take

\[ p(\beta) := \frac{2\beta}{1 + \sqrt{4\beta^2 + 1}}. \]

Then

\[ p'(\beta) = \frac{2}{\sqrt{4\beta^2 + 1} \left(1 + \sqrt{4\beta^2 + 1}\right)} > 0. \]

Thus \( p(\beta) \) is strictly increasing and

\[ 1 - p(0) = 1, \quad 1 - p(1) = 2 - \phi \approx 0.382 \]

This enables us to solve the equation

\[ 1 - p(\beta) = \begin{cases} 2 - \phi & \text{i.e. } p(\beta) = \phi - 1 \\ \phi - 1 & 2 - \phi. \end{cases} \]
This is reduced to
\[
\frac{2\beta}{1 + \sqrt{4\beta^2 + 1}} = \begin{cases} \phi^{-1} \\ \phi^{-2}. \end{cases}
\]
The equation has respective solutions
\[
\beta = \begin{cases} 
1 \\ \phi^2 \\ \phi^4 - 1 \\
\end{cases} = \frac{1}{\sqrt{5}}.
\]
We note that
\[
\frac{\phi^2}{\phi^4 - 1} = \frac{1}{(\phi^2 + 1)(\phi - 1)} = \frac{1}{2\phi - 1} = \frac{2\phi - 1}{5} = \frac{\sqrt{5}}{5} = \frac{1}{\sqrt{5}} \approx 0.4772
\]

7.1 Deterministic dynamics

The deterministic control process (D) has a discount factor $0 \leq \beta < \infty$.

7.1.1 Case $\beta = 1$

When $\beta = 1$, the quadratic coefficient $v$ is reduced to the Golden number
\[
v = \phi = \frac{1 + \sqrt{5}}{2} \approx 1.618
\]
and the proportional rate $p$ is reduced to
\[
p = \phi - 1 = \frac{\sqrt{5} - 1}{2} \approx 0.618
\]
Further the division of $[0, x]$ into $[0, \phi^{-2}x]$ and $[\phi^{-2}x, x]$ is Golden.

The Golden optimal policy $f^\infty, f(x) = px$, yields the optimal deterministic behavior as follows. A current state $x_n$ under the Golden optimal decision $u_n = px_n = \phi^{-1}x_n$ goes to a unique state $x_{n+1} = x_n - u_n = \phi^{-2}x_n$ into $R^1$. The next state is $x_{n+1} = \phi^{-2}x_n \approx 0.382x_n$ (Fig. 1). The dynamics is depicted as
\[
R^1 \ni x_n \xrightarrow{u_n = \phi^{-1}x_n} x_{n+1} = \phi^{-2}x_n \approx 0.382x_n \text{ uniquely.}
\]
Thus the Golden optimal dynamics says that next state becomes $x_{n+1} = \phi^{-2}x_n \approx 0.382x_n$. 
7.1.2 Case $\beta = \frac{1}{\sqrt{5}}$

We consider case $\beta = \frac{1}{\sqrt{5}} \approx 0.4772$. Then we have

$$v = 3 - \phi \approx 1.382, \quad p = \phi^{-2} \approx 0.382$$

The division of $[0, x]$ into $[0, \phi^{-1}x]$ and $[\phi^{-1}x, x]$ is Golden.

The Golden optimal policy $f^\infty, f(x) = px$, yields the optimal deterministic behavior as follows. A current state $x_n$ under the Golden optimal decision $u_n = \phi^{-2}x_n$ goes to a unique state $x_{n+1} = x_n - u_n = \phi^{-1}x_n \approx 0.618x_n$ (Fig. 2). The Golden optimal dynamics

$$R^1 \ni x_n \xrightarrow{\downarrow} u_n = \phi^{-2}x_n \rightarrow x_{n+1} = \phi^{-1}x_n \approx 0.618x_n$$

uniquely says that $x_{n+1} = \phi^{-1}x_n \approx 0.618x_n$.

7.2 Stochastic dynamics

The stochastic control process (S) has the discount factor restricted to $0 \leq \beta < 1$. We consider the Case $\beta = \frac{1}{\sqrt{5}}$ only.

7.2.1 Case $\beta = 1$

As we have shown in (6), the total noise is $\rho = \frac{2\beta - 1 + \sqrt{4\beta^2 + 1}}{2(1 - \beta)}$ for $0 \leq \beta < 1$. Thus it diverges to $\infty$ for $\beta = 1$.

7.2.2 Case $\beta = \frac{1}{\sqrt{5}}$

We have

$$v = 3 - \phi \approx 1.382, \quad p = \phi^{-2} \approx 0.382, \quad \rho = \frac{\sqrt{5}}{2} \approx 1.118$$

The state sequence $\{X_n\}_{0}^\infty$ defined by

$$X_{n+1} = X_n - pX_n - \epsilon_{n+1}, \quad X_0 = x_0$$

is stochastically Golden:

$$E[X_{n+1} \mid x_n] = \phi^{-1}x_n.$$ 

That is, the Golden optimal policy $f^\infty, f(x) = px$, yields the optimal stochastic behavior as follows. A current state $x_n$ under the Golden optimal decision $u_n = px_n = \phi^{-2}x_n$ goes
to $x_{n+1}$ on $R^1$ with transition probability $q(x_{n+1} \mid x_n, u_n) = \frac{1}{\sqrt{2\pi}} e^{-(x_{n+1} - \phi^{-1}x_n)^2/2}$. The next state (random variable) $x_{n+1}$ follows the normal distribution $N(\phi^{-1}x_n, 1)$. The mean is $\phi^{-1}x_n \approx 0.618x_n$ (see Fig. 2). The Golden optimal dynamics

$$R^1 \ni x_n \xrightarrow{u_n=\phi^{-2}x_n} x_{n+1} \text{ w.p. } q(x_{n+1} \mid x_n, u_n) \text{ for any } x_{n+1} \in R^1$$

says that current state goes down to $\phi^{-1}x_n \approx 0.618x_n$ on average.

### 7.3 Non-deterministic dynamics

The non-deterministic control process (N) has a discount factor $0 \leq \beta < \infty$.

#### 7.3.1 Case $\beta = 1$

When $\beta = 1$, it follows that

$$v = \phi \approx 1.618, \quad p = \phi^{-1} \approx 0.618$$

Further the division of $[0, x]$ into $[0, \phi^{-2}x]$ and $[\phi^{-2}x, x]$ is Golden optimal.

The Golden optimal policy $f^\infty, f(x) = px$, yields the optimal non-deterministic behavior as follows. A current state $x_n$ under the Golden optimal decision $u_n = px_n = \phi^{-1}x_n$ goes to $x_{n+1}$ on interval $(0, x_n - u_n) = (0, \phi^{-2}x_n)$ with transition weight $q(x_{n+1} \mid x_n, u_n) = 2/x_{n+1}$. The next state (non-deterministic variable) $x_{n+1}$ has the unbounded weight $2/x_{n+1}$ on $(0, \phi^{-2}x_n) \approx (0, 0.382x_n)$. The Golden optimal dynamics

$$(0 \infty) \ni x_n \xrightarrow{u_n=\phi^{-2}x_n} x_{n+1} \text{ w.w. } 2/x_{n+1} \text{ for any } x_{n+1} \in (0, \phi^{-2}x_n)$$

says that current state goes down on a shrunken interval $(0, \phi^{-2}x_n) \approx (0, 0.382x_n)$ with the Golden rate $\phi^{-2} \approx 0.382$ (see Fig. 1).

#### 7.3.2 Case $\beta = \frac{1}{\sqrt{5}}$

The case yields

$$v = 3 - \phi \approx 1.382, \quad p = \phi^{-2} \approx 0.382$$

The division of $[0, x]$ into $[0, \phi^{-1}x]$ and $[\phi^{-1}x, x]$ is Golden optimal.

The Golden optimal policy $f^\infty, f(x) = px$, yields the optimal non-deterministic behavior as follows. A current state $x_n$ under the Golden optimal decision $u_n = \phi^{-2}x_n$ goes to $x_{n+1}$ on interval $(0, x_n - u_n) = (0, \phi^{-1}x_n)$ with transition weight $q(x_{n+1} \mid x_n, u_n) = 2/x_{n+1}$. The non-deterministic $x_{n+1}$ has the unbounded weight $2/x_{n+1}$ on $(0, \phi^{-1}x_n) \approx (0, 0.618x_n)$. The Golden optimal dynamics

$$(0 \infty) \ni x_n \xrightarrow{u_n=\phi^{-2}x_n} x_{n+1} \text{ w.w. } 2/x_{n+1} \text{ for any } x_{n+1} \in (0, \phi^{-1}x_n)$$
says that next state goes down on a shrunken interval $(0, \phi^{-1}x_n) \approx (0, 0.618x_n)$ with the Golden rate $\phi^{-1} \approx 0.618$ (see Fig. 2).

Finally we have the following result.

**Theorem 7.1** For the discount rate $\beta = \frac{1}{\sqrt{5}}$, three processes (D), (S) and (N) have a common Golden optimal policy $g^\infty, g(x) = (2-\phi)x$. Then (D) and (N) have the quadratic value function $v(x) = (3-\phi)x^2$ and (S) has the quadratic value function $v(x) = (3-\phi)x^2 + \frac{\sqrt{5}}{2}$.

**References**


