Available online at www.sciencedirect.com
SCIENCE DIRECT®

Computers and Chemical Engineering 28 (2004) 2845–2849

Computers & Chemical Engineering

www.elsevier.com/locate/compchemeng

Short note

Stochastic maximum principle for optimal control under uncertainty

Vicente Rico-Ramirez[a], Urmila M. Diwekar[b,*]

[a] *Instituto Tecnologico de Celaya, Ave. Tecnologico y Garcia Cubas SN Celaya, Gto., CP 38010, Mexico*
[b] *Departments of Bio, Chemical, and Industrial Engineering, Center for Uncertain Systems: Tools for Optimization and Management (CUSTOM), Institute of Environmental Science and Policy, University of Illinois at Chicago, 851 S. Morgan St. (M/C 063), Chicago, IL 60607, USA*

Received 17 December 2003; received in revised form 3 August 2004; accepted 3 August 2004
Available online 22 September 2004

Abstract

Optimal control problems involve the difficult task of determining time-varying profiles through dynamic optimization. Such problems become even more complex in practical situations where handling time dependent uncertainties becomes an important issue. Approaches to stochastic optimal control problems have been reported in the finance literature and are based on real option theory, combining Ito's Lemma and the dynamic programming formulation. This paper describes a new approach to stochastic optimal control problems in which the stochastic dynamic programming formulation is converted into a stochastic maximum principle formulation. An application of such method has been reported by Rico-Ramirez et al. (*Computers and Chemical Engineering, 2003, 27, 1867*) but no details of the derivation were provided. The main significance of this approach is that the solution to the partial differential equations involved in the dynamic programming formulation is avoided. The classical isoperimetric problem illustrates this approach.

PACS: 02.30.W; 05.40

Keywords: Stochastic optimal control; Stochastic maximum principle; Ito's Lemma

1. Introduction

Optimal control problems in engineering have received considerable attention in the literature. In general, solution to these problems involves finding the time-dependent profiles of the control variables so as to optimize a particular performance index. The dynamic nature of the decision variables makes these problems much more difficult to solve compared to normal optimization where the decision variables are scalar. In general mathematical methods to solve these problems involve calculus of variations, the maximum principle and the dynamic programming technique. Nonlinear programming (NLP) techniques can also be used to solve this problem provided the system of differential equations is converted to nonlinear algebraic equations. For details of these methods, please see the work by Diwekar (2003).

Calculus of variations considers the entire path of the function and optimizes the integral through the minimization of the functional by vanishing the first derivative, resulting in second-order differential equations that can be difficult to solve. Other two approaches keep the first-order differential system by using transformation.

In the maximum principle, the objective function is reformulated as a linear function in terms of final values of state variables and the values of a vector of constants (linear Mayer form). However, this maximum principle transformation needs to include additional variables and corresponding first-order differential equations, referred to as adjoint variables and adjoint equations, respectively.

Dynamic programming formulation results in a first-order system of partial differential equations (the Hamilton–Jacobi–Bellman, HJB equations) that may not be easy to solve. However, this dynamic programming method provides the basis for stochastic optimal control problems and is used in this paper to derive the stochastic maximum principle.

* Corresponding author. Tel.: +1 312 3553277; fax: +1 312 9965921.
E-mail address: urmila@uic.edu (Urmila M. Diwekar).

0098-1354/$ – see front matter
doi:10.1016/j.compchemeng.2004.08.001

2. The stochastic maximum principle

Although the mathematics of dynamic programming looks different from the maximum principle formulation, in most cases they lead to the same results. As a matter of fact, in this section we first show that, starting from the dynamic programming optimality conditions (HJB equations), the derivation of the adjoint equations of the maximum principle can be achieved. This is not surprising and has been reported elsewhere (see for instance (Diwekar, 1995, 2003)) for the deterministic case. Financial literature reports an extension of the HJB equation for the stochastic case (Dixit & Pyndick, 1994; Merton & Samuelson, 1990; Thompson & Sethi, 1994) but an equivalent maximum principle is not reported. In this paper, we use the mathematical equivalence between dynamic programming and the maximum principle to extend the maximum principle to the stochastic case. Recently, we presented an application of this method (Diwekar, 1995; Rico-Ramirez, Morel, & Diwekar, 2003) but no details of the derivation were provided. Those references also the present advantages of the maximum principle.

The main aspect of the derivations consists on obtaining the expressions for the *adjoint equations*. The adjoint equations provide the dynamics of the *adjoint variables* in the maximum principle. For the deterministic case, it is shown that the adjoint variables in the maximum principle are equivalent to the derivatives of the objective function with respect to the state variables of the dynamic programming approach. It is possible to show such an equivalence for the stochastic case also and that provides the basis of our reformulation.

In the remaining of this section, we provide the steps involved in the derivation of the adjoint equations used in the maximum principle approach to the solution of stochastic optimal control problems. This derivation is achieved by using the optimality conditions established for the dynamic programming approach. First, the deterministic case will be considered and then an analogous analysis for the stochastic problem will be described. Notice that, seeking simplicity, the derivations correspond to a case where just one state variable, x, is present in the formulation (no vectorial representation is used); nevertheless, the interpretation of the derivation can be extended to the situation of x being a vector of state variables.

2.1. Deterministic case

Consider the definition of an optimal control problem:

$$\max_\theta L = \int_0^T k(x,\theta)\,\mathrm{d}t$$

Subject to:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(x,\theta) = f$$

and the dynamic programming optimality conditions given by

$$0 = \max_\theta \left[k + \frac{\partial L}{\partial t} + \frac{\partial L}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} \right] \tag{1}$$

Eq. (1) is the HJB equation and can be rewritten as

$$0 = \max_\theta [k + L_t + L_x f] \tag{2}$$

$$0 = \max_\theta [L_t + H]$$

where H is the Hamiltonian function defined by

$$H = k + L_x f \tag{3}$$

Taking derivative of Eq. (2) with respect to x:

$$L_{xt} + k_x + L_{xx} f + f_x L_x = 0$$

$$L_{xt} + L_{xx} f = -k_x - f_x L_x \tag{4}$$

and using chain rule:

$$\frac{\mathrm{d}L_x}{\mathrm{d}t} = \frac{\partial L_x}{\partial t} + \frac{\partial L_x}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t}$$

$$\frac{\mathrm{d}L_x}{\mathrm{d}t} = L_{xt} + L_{xx} f \tag{5}$$

Substituting (4) in (5):

$$\frac{\mathrm{d}L_x}{\mathrm{d}t} = -k_x - f_x L_x \tag{6}$$

Finally, if we consider that, for the case of the maximum principle formulation, the problem is represented in linear Mayer form, then $k = 0$ and Eqs. (3) and (6) become:

$$H = L_x f \tag{7}$$

$$\frac{\mathrm{d}L_x}{\mathrm{d}t} = -f_x L_x \tag{8}$$

Rewriting Eqs. (7) and (8) in terms of variables, μ, where μ are equivalent to the derivatives of the objective function L with respect to the state variables x in dynamic programming, results in following adjoint equations for maximum principle:

$$\frac{\mathrm{d}\mu}{\mathrm{d}t} = -f_x \mu \tag{9}$$

and, for n-dimensional vectors:

$$\frac{\mathrm{d}\mu_i}{\mathrm{d}t} = -\sum_{j=1}^{n} \mu_j \frac{\partial f_j}{\partial x_i}$$

2.2. Stochastic case

Once again, we will use a scalar representation although the extension to vectorial analysis can be made. Consider the stochastic optimal control problem given by

$$\max_\theta L = E\left[\int_0^T k(x,\theta)\,\mathrm{d}t \right]$$

where E is the expectation operator.

Subject to an Ito process (Diwekar, 1995):

$$\mathrm{d}x = f\mathrm{d}t + g\mathrm{d}z \tag{10}$$

where dz is a stochastic Wiener process, and f and g are both given functions.

The corresponding optimality conditions are now:

$$0 = \max_\theta \left[k + \frac{1}{\mathrm{d}t} E(\mathrm{d}L) \right]$$

and the corresponding HJB equation becomes:

$$0 = \max_\theta \left[k + \frac{\partial L}{\partial t} + \frac{\partial L}{\partial x} f + \frac{1}{2} g^2 \frac{\partial^2 L}{(\partial x)^2} \right]$$

$$= \max_\theta \left[k + L_t + L_x f + \frac{1}{2} g^2 L_{xx} \right] \tag{11}$$

The Hamiltonian function, H, for the stochastic case is given below:

$$H = k + L_x f + \tfrac{1}{2} g^2 L_{xx}$$

Note the second-order term coming from the consideration of the state variable being an Ito process (Ito's Lemma). We then follow the same procedure as in the deterministic case. Taking derivative of Eq. (11) with respect to x:

$$L_{xt} + k_x + L_{xx} f + f_x L_x + \tfrac{1}{2} g^2 L_{xxx} + \tfrac{1}{2} (g^2)_x L_{xx} = 0 \tag{12}$$

and, therefore,

$$L_{xt} + L_{xx} f + \tfrac{1}{2} g^2 L_{xxx} = -k_x - f_x L_x - \tfrac{1}{2} (g^2)_x L_{xx} \tag{13}$$

Also, using chain rule and considering second-order contributions of the derivatives with respect to x (Ito's Lemma) results in:

$$\mathrm{d}L_x = \frac{\partial L_x}{\partial t} \mathrm{d}t + \frac{\partial L_x}{\partial x} \frac{\mathrm{d}x}{\mathrm{d}t} \mathrm{d}t + \frac{1}{2} \frac{\partial^2 L_x}{(\partial x)^2} \mathrm{d}x^2$$

Since, because of Ito's Lemma, $E[\mathrm{d}(x^2)] = g^2\, \mathrm{d}t$, the previous equation reduces to:

$$\frac{\mathrm{d}L_x}{\mathrm{d}t} = \frac{\partial L_x}{\partial t} + \frac{\partial L_x}{\partial x} \frac{\mathrm{d}x}{\mathrm{d}t} + \frac{1}{2} \frac{\partial^2 L_x}{(\partial x)^2} g^2$$

$$\frac{\mathrm{d}L_x}{\mathrm{d}t} = L_{xt} + L_{xx} f + \tfrac{1}{2} L_{xxx} g^2 \tag{14}$$

Substituting (13) in (14):

$$\frac{\mathrm{d}L_x}{\mathrm{d}t} = -k_x - f_x L_x - \frac{1}{2} (g^2)_x L_{xx}$$

Finally, if the problem is stated in Mayer linear form, $k = 0$, and then:

$$\frac{\mathrm{d}L_x}{\mathrm{d}t} = -f_x L_x - \frac{1}{2} (g^2)_x L_{xx} \tag{15}$$

Eq. (15) provides the expression for the dynamics of the adjoint variables in the stochastic case. By comparing Eqs. (15)–(8), one can observe the second-order term included in Eq. (15) as a result of the state variable being a stochastic variable behaving as an Ito process.

2.2.1. *Representing second derivatives contributions using adjoint equations*

The use of Eq. (15) implies that the calculation of L_{xx} is also needed. In order to obtain an expression for the dynamics of L_{xx}, a derivation similar to the one presented in the previous subsection must be used. The resulting equation will also be named as adjoint equation. Let us start by deriving again Eq. (12) with respect to x:

$$L_{xxt} + k_{xx} + L_{xx} f_x + L_{xxx} f + L_{xx} f_x + L_x f_{xx} + \tfrac{1}{2} g^2 L_{xxxx} + \tfrac{1}{2} (g^2)_x L_{xxx} + \tfrac{1}{2} (g^2)_x L_{xxx} + \tfrac{1}{2} (g^2)_{xx} L_{xx} = 0 \tag{16}$$

and, therefore,

$$L_{xxt} + L_{xxx} f + \tfrac{1}{2} g^2 L_{xxxx} = -k_{xx} - 2L_{xx} f_x - L_x f_{xx} - (g^2)_x L_{xxx} - \tfrac{1}{2} (g^2)_{xx} L_{xx} \tag{17}$$

Using chain rule and considering second-order contributions in the derivatives with respect to x (Ito's Lemma):

$$\mathrm{d}L_{xx} = \frac{\partial L_{xx}}{\partial t} \mathrm{d}t + \frac{\partial L_{xx}}{\partial x} \frac{\mathrm{d}x}{\mathrm{d}t} \mathrm{d}t + \frac{1}{2} \frac{\partial^2 L_{xx}}{(\partial x)^2} \mathrm{d}x^2 \tag{18}$$

By using Ito's Lemma:

$$\frac{\mathrm{d}L_{xx}}{\mathrm{d}t} = \frac{\partial L_{xx}}{\partial t} + \frac{\partial L_{xx}}{\partial x} \frac{\mathrm{d}x}{\mathrm{d}t} + \frac{1}{2} \frac{\partial^2 L_{xx}}{(\partial x)^2} g^2$$

$$\frac{\mathrm{d}L_{xx}}{\mathrm{d}t} = L_{xxt} + L_{xxx} f + \frac{1}{2} L_{xxxx} g^2 \tag{19}$$

Substituting Eq. (17) in Eq. (19):

$$\frac{\mathrm{d}L_{xx}}{\mathrm{d}t} = -k_{xx} - 2L_{xx} f_x - L_x f_{xx} - (g^2)_x L_{xxx} - \frac{1}{2} (g^2)_{xx} L_{xx}$$

Finally, if $k = 0$ and we neglect contribution of third order (assumption also consistent with Ito's Lemma):

$$\frac{\mathrm{d}L_{xx}}{\mathrm{d}t} = -2L_{xx} f_x - L_x f_{xx} - \frac{1}{2} (g^2)_{xx} L_{xx} \tag{20}$$

Equating adjoint variable, μ, to the first derivatives of the objective function L with respect to the state variables x and ω as the second derivatives of the objective function with respect to the state variables, Eqs. (15) and (20) can be rewritten as

$$\frac{d\mu}{dt} = -f_x\mu - \frac{1}{2}(g^2)_x\omega \tag{21}$$

$$\frac{d\omega}{dt} = -2\omega f_x - \mu f_{xx} - \frac{1}{2}(g^2)_{xx}\omega \tag{22}$$

2.2.2. A summary for the stochastic case

Summarizing the results for the stochastic case, the Hamiltonian function and the adjoint equations to be solved in the stochastic maximum principle formulation are ($k = 0$):

$$H = \mu f + \frac{1}{2}g^2\omega$$

$$\frac{d\mu}{dt} = -f_x\mu - \frac{1}{2}(g^2)_x\omega \quad \mu(T) = c$$

$$\frac{d\omega}{dt} = -2\omega f_x - \mu f_{xx} - \frac{1}{2}(g^2)_{xx}\omega \quad \omega(T) = 0$$

Notice that the resulting problem is a two-point boundary value problem. However, it is possible to circumvent the two point boundary value problem as explained in Diwekar (1995, 1992).

3. The isoperimetric problem

We consider here the historic isoperimetric problem of Queen Dido (Diwekar, 2003). Queen Dido's (1000 BC) problem was to find maximum area that can be covered by a rope (curve) whose length (perimeter) is fixed (λ). Using kinematics this problem can be written as follows:

$$\text{maximize}_{u_t} J = \int_0^T x_1(t)\,dt \quad \text{area}$$

subject to

$$dx_1 = u_t\,dt \quad x_1(0) = 0.0 \quad \text{kinematic constraint}$$

$$\frac{dx_2}{dt} = \sqrt{1+u_t^2} \quad x_2(0) = 0.0\ x_2(T) = \lambda = 16.0$$

perimeter constraint

We define a new state variable in order to obtain the linear Mayer form of the objective function:

$$x_3(t) = \int_0^t x_1(t)\,dt$$

And the problem can then be written as

$$\max_{u_t} L = x_3(T) \tag{23}$$

Subject to differential equations for x_1, x_2, and x_3. As in Example 7.5 in Ref. Diwekar (2003), assume that the vertical displacement variable x_1 is stochastic and follows a Brownian motion (instance of an Ito process). This results in the following differential equations:

$$dx_1 = u_t dt + \sigma dz \quad x_1(0) = 0.0 \tag{24}$$

where $dz = \sqrt{dt}$ and $\sigma = 0.5$

$$\frac{dx_2}{dt} = \sqrt{1+u_t^2} \quad x_2(0) = 0.0\ x_2(T) = \lambda = 16.0 \tag{25}$$

$$\frac{dx_3}{dt} = x_1(t) \quad x_3(0) = 0.0 \tag{26}$$

The optimality conditions for this problem are

$$0 = \max_\theta \left[k + \frac{\partial L}{\partial t} + \frac{\partial L}{\partial x}f + \frac{1}{2}g^2\frac{\partial^2 L}{(\partial x)^2}\right],$$

$$0 = \max_{u_t}\left[\frac{\partial L}{\partial t} + u_t\frac{\partial L}{\partial x_1(t)} + \sqrt{1+u_t^2}\frac{\partial L}{\partial x_2(t)} + x_1(t)\frac{\partial L}{\partial x_3(t)} + \frac{\sigma^2}{2}\frac{\partial^2 L}{\partial x_1(t)^2}\right]$$

or

$$\max_{u_t}\left[L_t + L_{x_1}u_t + L_{x_2}\sqrt{1+u_t^2} + L_{x_3}x_1(t) + \frac{L_{x_1x_1}\sigma^2}{2}\right] = 0$$

Expressing in terms of the adjoint variables results in

$$\max_{u_t}\left[L_t + \mu_1 u_t + \mu_2\sqrt{1+u_t^2} + \mu_3 x_1(t) + \frac{\omega\sigma^2}{2}\right] = 0 \tag{27}$$

The adjoint equations are, therefore:

$$\frac{d\mu_1}{dt} = -1 \Longrightarrow \mu_1 = -t + c_1 \tag{28}$$

$$\frac{d\mu_2}{dt} = 0 \Longrightarrow \mu_2 = c_2 \tag{29}$$

$$\frac{d\mu_3}{dt} = 0 \quad \mu_3(T) = 1 \Longrightarrow \mu_3 = 1 \tag{30}$$

$$\frac{d\omega}{dt} = 0 \quad \omega(T) = 0 \Longrightarrow \omega = 0 \tag{31}$$

Finally, maximizing Eq. (27) with respect to u_t leads to:

$$\mu_1 + \frac{u_t}{\sqrt{1+u_t^2}}\mu_2 = 0 \tag{32}$$

Therefore, the velocity parameter u_t follows the path given by

$$u_t = \frac{t - c_1}{\sqrt{c_2 - (t-c_1)^2}} \tag{33}$$

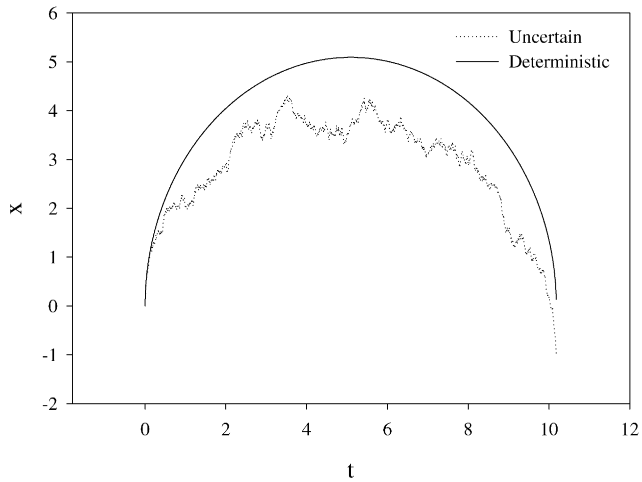

Fig. 1. Deterministic and stochastic path of variable x_1.

Solution to the deterministic version of the isoperimetric problem has been reported in Diwekar (2003). By comparing such solution to Eq. (33), it would seem that the deterministic solution and the stochastic solution are the same. However, stochasticity is embedded in the differential equation for x_1 given by Eq. (24). This is also obvious when one simulates a single instant of stochasticity (Fig. 1) by choosing a normal random process with a mean of zero and variance σ represented by the parameter ϵ in the following form of Eq. (34):

$$\mathrm{d}x_1 = u_t\,\mathrm{d}t + \sigma\epsilon\sqrt{\mathrm{d}t} \quad x_1(0) = 0.0 \tag{34}$$

It can be seen that, although the stochastic solution follows a circular path, the expected area obtained in the stochastic case is smaller than the area obtained in the deterministic case for the same perimeter.

4. Conclusions

This paper presented the stochastic maximum principle for optimal control under uncertainty. This paper presents the theoretical basis for the approach and illustrates it using a simple iso-perimetric example problem. This approach to stochastic optimal control is advantageous as it avoids the solution to partial differential equations resulting from stochastic dynamic programming. The usefulness of this approach for batch and bio processing to handle time dependent thermodynamic uncertainties inherent in such processes has recently been illustrated in Rico-Ramirez et al. (2003) and in Ulas and Diwekar (2004). This approach has a great potential to solve problems in various areas such as ecosystem dynamics (Duggempudi & Diwekar, 2003), financial modeling, and economics.

Acknowledgments

We want to thank the financial support given by NSF/CONACYT "Consejo Nacional de Ciencia y Tecnologia".

References

Dixit, A. K., & Pindyck, R. S. (1994). *Investment Under Uncertainty*. Princeton, NJ, USA: Princeton University Press.

Diwekar, U. M. (1992). Unified approach to solving optimal design control problems in batch distillation. *AIChE Journal*, *38*, 1551.

Diwekar, U. M. (1995). *Batch Distillation. Simulation, Optimal Design and Control*. Washington, DC, USA: Taylor and Francis.

Diwekar, U. M. (2003). *Introduction to Applied Optimization*. Norwell, MA: Kluwer Academic Publishers.

Duggempudi, J., & Diwekar, U. (2003). Sustainability, Options Theory, and Off-line Quality Control. *Proceedings of Topical Conference on Sustainability and Life Cycle, AIChE Annual Meeting, SanFrancisco, CA*.

Merton, R. C., & Samuelson, P. A. (1990). *Continuous-Time Finance*, Cambridge, MA, USA: Blackwell Publishing.

Rico-Ramirez, V., Morel, B., & Diwekar, U. M. (2003). Real option theory from finance to batch distillation. *Comp. Chem. Eng.*, *27*, 1867.

Thompson, G. L., & Sethi, S. P. (1994). *Optimal Control Theory*, Boston, USA: Martinus Nijhoff Publishing.

Ulas, S., & Diwekar, U. (2003). Optimal Control of Batch Separations for Bio-processing: A Real Options Theory Approach. *Proceedings of Topical Conference on Systems Engineering in Pharmaceutical Processes and Biotechnology, AIChE Annual Meeting, San Francisco, CA*.

Ulas S., & Diwekar, U. (2004). Thermodynamic Uncertainties in Batch Processing and Optimal Control. *Comp. Chem. Eng.*, *28*(11), 2245–2258.