PMM U.S.S.R.,Vol.45,pp.179-183 0021-8928/82/2 0179 $7.50/0

UDC 62-50

ASYMPTOTIC BEHAVIOR OF THE BELLMAN FUNCTION IN A STOCHASTIC OPTIMAL CONTROL PROBLEM*

A.I. OVSEEVICH

A quasi-optimal control giving a near-optimal result under small power resources is found by studying the asymptotic behavior of the Bellman function of the optimal control problem for the motion of a system perturbed by white noise with limited control power resources. The determination of the asymptotic behavior of the Bellman function and of the quasi-optimal control calls for solving the partial differential equations introduced in /1/. A converging iteration process is found for solving them.

1. Statement of the problem. Let the motion of a point in an n-dimensional space R^n be described by the stochastic differential equation

$$dx_t = u_t dt + dw_t, \quad x_s = x \tag{1.1}$$

where w_t is the n-dimensional Brownian motion. It is assumed that control u_t depends upon the trajectory x_τ up to the instant t and is such that ($|u_t|$ is the Euclidean norm of vector u_t)

$$\int_s^T |u_t|^2 dt \leqslant Q^2 \tag{1.2}$$

In addition, let $F(x)$ be some fixed function. It is required to minimize the mean $E_{x,s}^u(F(x_T))$ of the random quantity $F(x_T)$ under the condition $x_s = x$. Such a setting can be treated as a model optimal motion control problem with limited control power, while the quantity Q can be interpreted as the initial reserve of the power resources.

This and similar problems were analyzed in /1,2/. Below we study the asymptotic behavior of the Bellman function $S(x, t, Q) = \inf E_{x,t}^u(F(x_T))$ as $Q \to 0$. To be precise, we introduce the asymptotic representation

$$S(x, t, Q) = S_0(x, t) + QS_1(x, t) + O(Q^2)$$

and the partial differential equations satisfied by functions S_0 and S_1, as well as a converging iteration process for solving these equations. The equations have the following form:

$$\frac{\partial S_0}{\partial t} + \frac{1}{2}\Delta S_0 = 0, \quad S_0(x, T) = F(x) \tag{1.3}$$

$$\frac{\partial S_1}{\partial t} + \frac{1}{2}\Delta S_1 + \min\left((u, \operatorname{grad} S_0) - \frac{|u|^2}{2} S_1\right) = 0, \quad S_1(x, T) = 0 \tag{1.4}$$

where Δ and grad are, respectively, the Laplacian and the gradient in R^n; in what follows these equations are called the zero- and first-approximation equations, respectively. Further, with respect to S_0 and S_1 we construct a quasi-optimal control u such that

$$S = E^u(F(x_T)) + O(Q^2)$$

The main idea of the paper is that the function $S_0 + QS_1$, just as S, is the Bellman function of some stochastic extremal problem, and the sets of admissible controls in both problems coincide, while the functionals differ by a quantity of the order of Q^2. Henceforth, for simplicity of notation we take it that $n = 1$, i.e., x_t, u_t, w_t are scalars.

2. Refinement of the problem statement. The set $U(Q)$ of admissible controls with resource Q consists of functionals $u_t = u(t, x_{\tau\leqslant t})$ which measurably depend on the trajectory $x_\tau, \tau \leqslant t$ of the random process x_t (thus, u_t is a nonanticipative functional in the terminology of /3/) and for which

*Prikl.Matem.Mekhan.,45,No.2,249-255,1981

$$\int\limits_{1}^{T} u_t^2\,dt \leqslant Q^2$$

with probability one. Concerning the risk function $F(x)$ we assume that it is infinitely smooth with exponentially bounded derivatives, i.e., any derivative is estimated from above by (const) $\exp(\lambda|x|)$, where $\lambda = \text{const}$. We write this assumption as $F \in C^\infty(\exp)$. We define the functionals

$$I(u) = E_{x,s}^u(F(x_T)), \qquad K(u) = E_{x,s}\Big(F(w_T)\Big(1 + \int\limits_s^T u_t\,dw_t\Big)\Big)$$

The Brownian motion w_t in the definition of $K(u)$ is assumed to issue from x at instant s, while the control $u_t = u(t, w_{\tau\leqslant t})$ is computed along the trajectory w_τ.

Theorem 1. $|I(u) - K(u)| \leqslant CQ^2$ for $u \in U(Q)$ and sufficiently small Q, where $C = C(x, s)$ is a function of exponential growth (i.e., $C(x,s) \leqslant$ (const) $\exp(\lambda|x|)$ for s lying in a bounded interval).

Proof. The idea is to apply the Cameron—Martin formula (see /3—5/)

$$I(u) = E(F(w_T)\exp X), \quad X = \int\limits_s^T u_t\,dw_t - \frac{1}{2}\int\limits_s^T u_t^2\,dt \tag{2.1}$$

to expand the exponential into a series and to retain in it only terms of first order in Q, which constitute the functional $K(u)$. Formula (2.1) is true for $u \in U(Q)$ since by virtue of Theorems 6.1 and 4.13 of /4/ it suffices that

$$E\Big(\exp\frac{1}{2}\int\limits_s^T u_t^2\,dt\Big) < \infty$$

but this follows from (1.2).

Thus, a part of the program projected at the end of section 1 has been realized. It remains to solve the problem on the minimum for the simplifies functional $K(u)$. We begin by reducing $K(u)$ to a more convenient form.

3. Transformation of functional $K(u)$. We define function S_0 by the formula $S_0(x, s) = E_{x,s}(F(w_T))$, where the Brownian motion w_t starts from x at instant s. Then, as is well known (see /5/, for example), S_0 is a smooth solution of the zero-approximation Eq.(1.3). We set $\varphi = \partial S_0/\partial x$. Using the Itô formula (see /3/) as applied to $S_0(w_t, t)$, with due regard to (1.3) we can show that

$$E\Big(F(w_T)\int\limits_s^T u_t\,dw_t\Big) = E\int\limits_s^T \varphi(w_t, t)\,u_t\,dt \tag{3.1}$$

We denote the right-hand side of (3.1) by $J(u)$. Then

$$K(u) = S_0(x, s) + J(u)$$

For the complete realization of the program marked at the end of section 1 it remains to show that the minimum of $J(u)$ with respect to $u \in U(Q)$ exists and that $\min J(u) = QS_1(x, s)$, where S_1 satisfies the first-approximation Eq.(1.4).

4. Regularization of the extremal problem. We complement the random process $x_t = w_t$ by the component

$$Q_t = \Big(Q^2 - \int\limits_s^t u_\tau^2\,d\tau\Big)^{1/2}$$

(see /1,6,7/) and we obtain the system

$$dx_t = dw_t,\ x_s = x, \qquad dQ_t = -(u_t^2/(2Q_t))\,dt, \quad Q_s = Q$$

We denote $\inf J(u)$ over $u \in U(Q)$ by $S_+(x, s, Q)$. Then it is natural to expect that S_+ satisfies the Bellman equation

$$\frac{\partial S_+}{\partial t} + \frac{1}{2}\frac{\partial^2 S_+}{\partial x^2} + \min\Big(u\varphi - \frac{u^2}{2Q}\frac{\partial S_+}{\partial Q}\Big) = 0 \tag{4.1}$$

$S_+ = 0$ when $Q = 0$ or when $t = T$. Setting $S_+(x, s, Q) = QS_1(x, S)$, for S_1 we obtain the first-approximation equation

$$\frac{\partial S_1}{\partial t} + \frac{1}{2}\frac{\partial^2 S_1}{\partial x^2} + \min\left(u\varphi - \frac{u^2}{2} S_1\right) = 0, \quad S_1(x, T) = 0. \tag{4.2}$$

We note that $\min(u\varphi - (u^2/2)s_1) = \varphi^2/(2S_1)$ when $S_1 < 0$. Together with the boundary condition $S_1(x, T) = 0$ this shows that Eq.(4.2) is singular when $t = T$. Therefore, we begin with the solution of the regularized problem

$$\frac{\partial \Sigma_\varepsilon}{\partial t} + \frac{1}{2}\frac{\partial^2 \Sigma_\varepsilon}{\partial x^2} + \min\left(u\varphi - \frac{u^2}{2}\Sigma_\varepsilon\right) = 0, \quad \Sigma_\varepsilon(x, T) = -\varepsilon, \quad \varepsilon > 0 \tag{4.3}$$

which is connected with the minimization of the functional $J_\varepsilon(u) = J(u) - \varepsilon E(Q_T)$. The solution of the regularized problem (4.3) is obtained by using the standard Bellman interation process.

5. The Bellman method. We set $\Phi_0 = -\varepsilon$, $u_0 = 0$. If Φ_n has been defined, then we set $u_{n+1} = \varphi/\Phi_n$ (since $\varphi u_{n+1} - (u_{n+1}^2/2)\Phi_n = \min(\varphi u - (u^2/2)\Phi_n)$) and we determine Φ_{n+1} as the solution of the Cauchy problem

$$\frac{\partial \Phi_{n+1}}{\partial t} + \frac{1}{2}\frac{\partial^2 \Phi_{n+1}}{\partial x^2} + u_{n+1}\varphi - \frac{u_{u+1}^2}{2}\Phi_{n+1} = 0, \quad \Phi_{n+1}(x, T) = -\varepsilon \tag{5.1}$$

We note that the function $\varphi = \partial S_0/\partial x$ and $S_0 = E(F(W_T))$ are infinitely smooth with exponentially bounded derivatives because F has these properties (in the notation of Sect.2, $\varphi \in C^\infty(\exp)$). The following lemma ensures the possibility of an unrestricted continuation of the iterations.

Lemma 1. Let $u, \psi \in C^\infty(\exp), \psi \leqslant 0$. Then the Cauchy problem

$$\frac{\partial \Phi}{\partial t} + \frac{1}{2}\frac{\partial^2 \Phi}{\partial x^2} + \psi - \frac{u^2}{2}\Phi = 0, \quad \Phi(x, T) = -\varepsilon$$

has the unique solution $\Phi \in C^\infty(\exp)$, while $\Phi \leqslant -\varepsilon$.

Lemma 2. The sequence Φ_n decreases monotonically and converges to the function $\Sigma_\varepsilon \in C^\infty(\exp)$ i.e., to the unique solution of the problem:

$$\frac{\partial \Sigma_\varepsilon}{\partial t} + \frac{1}{2}\frac{\partial^2 \Sigma_\varepsilon}{\partial x^2} + \frac{1}{2}\frac{\varphi^2}{\Sigma_\varepsilon} = 0, \quad \Sigma_\varepsilon(x, T) = -\varepsilon \tag{5.2}$$

Proof. We set $W = \Phi_{n+1} - \Phi_n$. Then

$$W(x, T) = 0, \quad \frac{\partial W}{\partial t} + \frac{1}{2}\frac{\partial^2 W}{\partial x^2} - \frac{u_{n+1}^2}{2}W \geqslant 0.$$

The inequality follows from the fact that

$$\frac{\partial \Phi_n}{\partial t} + \frac{1}{2}\frac{\partial^2 \Phi_n}{\partial x^2} + u_{n+1}\varphi - \frac{u_{n+1}^2}{2}\Phi_n \leqslant \frac{\partial \Phi_n}{\partial t} + \frac{1}{2}\frac{\partial^2 \Phi_n}{\partial x^2} + u_n\varphi - \frac{u_n^2}{2}\Phi_n$$

because of the special choice of u_{n+1}. Therefore, the maximum principle yields $W \leqslant 0$, as required. To prove the convergence of Φ_n it is sufficient to establish the lower bound for Φ_n. We now set $\Sigma_\varepsilon^+(x, t, Q) = Q\Sigma_\varepsilon(x, t)$ and show that Σ_ε^+ possesses a certain "quasi-optimal property".

Lemma 3. 1) $\Sigma_\varepsilon^+(x, s, Q) = \min J_\varepsilon(u)$, where functional $J_\varepsilon(u) = J(u) - \varepsilon E(Q_T)$ has been defined in Sect.4 and the minimum is taken over $u \in U(Q)$.

2) The minimum in 1) is achieved by the control

$$u_t^\varepsilon = (\varphi/\Sigma_\varepsilon)(w_t, t)\, Q_t^\varepsilon$$

$$Q_t^\varepsilon = Q\exp\left(-\frac{1}{2}\int_s^t \left(\frac{\varphi}{\Sigma_\varepsilon}\right)^2 (w_\tau, \tau)\, d\tau\right)$$

6. Solution of the first-approximation equation. The next theorem completes the program marked as the end of Sect.1.

Theorem 2. 1) Functions Σ_ε increase monotonically as $\varepsilon \downarrow 0$ and the limit function $S_1 = \lim \Sigma_\varepsilon$ is a generalized solution of the Cauchy problem:

$$\frac{\partial S_1}{\partial t} + \frac{1}{2}\frac{\partial^2 S_1}{\partial x^2} + \min\left(u\varphi - \frac{u^2}{2} S_1\right) = 0, \quad S_1(x, T) = 0 \tag{6.1}$$

2) Let $S_+(x, t, Q) = QS_1(x, t)$. Then $S_+(x, t, Q) = \min J(u)$, where functional $J(u)$ was defined in Sect.3 and the minimum is taken over $u \in U(Q)$.

3) the minimum in 2) is achieved by the control

$$u_t = Q_t(\varphi/S_1)(w_t, t)$$

Proof. 1) Let $\varepsilon_1 < \varepsilon_2$, $V_1 = \Sigma_{\varepsilon_1}$, $V_2 = \Sigma_{\varepsilon_2}$, $W = V_1 - V_2$. Then

$$\frac{\partial W}{\partial t} + \frac{1}{2}\frac{\partial^2 W}{\partial x^2} - \frac{1}{2}\frac{\varphi^2}{V_1 V_2} W = 0, \quad W(x, T) \geqslant 0$$

Hence by the maximum principle we obtain $W \geqslant 0$, as required. To prove 1) it is necessary to establish certain **a priori** estimates for Σ_ε, uniform with respect to ε. Let $\rho \in C^\infty$, $\rho(x) = (\text{const})\,|x|$ for $|x| \geqslant 1$. Then, multiplying (5.2) by $\Sigma_\varepsilon e^{-\rho}$, integrating by parts and applying Gronwall's lemma, we obtain

$$\int_s^T\int \left|\frac{\partial \Sigma_\varepsilon}{\partial x}\right|^2 e^{-\rho}\,dx\,dt + \int |\Sigma_\varepsilon(x, s)|^2 e^{-\rho}\,dx \leqslant C\left(\int_s^T\int |\varphi|^2 e^{-\rho}\,dx\,dt + 1\right) \tag{6.2}$$

where C is independent of ε (the summand 1 in the right-hand side of (6.2) arises because $\Sigma_\varepsilon(x, T) \neq 0$). Multiplying (5.2) by $e^{-\rho}$, integrating by parts and using (6.2), we obtain

$$\int_s^T\int \frac{\varphi^2}{|\Sigma_\varepsilon|} e^{-\rho}\,dx\,dt \leqslant C\left(\int_s^T\int \varphi^2 e^{-\rho}\,dx\,dt + 1\right)^{1/2} \tag{6.3}$$

where again C is independent on ε. Estimates (6.2) and (6.3) permit a passage to the limit as $\varepsilon \to 0$ in the equality

$$\int_s^T\int \Sigma_\varepsilon\left(-\frac{\partial f}{\partial t}\right)dx\,dt + \frac{1}{2}\int_s^T\int \Sigma_\varepsilon \frac{\partial^2 f}{\partial x^2}\,dx\,dt + \frac{1}{2}\int_s^T \frac{\varphi^2}{\Sigma_\varepsilon} f\,dx\,dt = \int \Sigma_\varepsilon(x, s) f(x, s)\,dx - \varepsilon\int f(x, T)\,dx$$

where f is a smooth finite function, and to get that $S_1 = \lim \Sigma_\varepsilon$ is a generalized solution of Eq.(6.1). Assertion 1) has been proved. The theorem's assertion 2) follows at once from Lemma 3. The proof of assertion 3) reduces to justifying the possibility of passing to the limit as $\varepsilon \to 0$ in the equality $\Sigma_\varepsilon^+ = J_\varepsilon(u^\varepsilon)$ from Lemma 3. We recall that

$$u_t^\varepsilon = (\varphi/\Sigma_\varepsilon)(w_t, t)\,Q_t^\varepsilon \quad \text{where} \quad Q_t^\varepsilon = Q\exp\left(-\frac{1}{2}\int_s^t \left(\frac{\varphi}{\Sigma_\varepsilon}\right)^2 (w_\tau, \tau)\,d\tau\right)$$

We need to show that u_t^ε converges to $u(w_t, t, Q_t)$ as $\varepsilon \to 0$ and that the functions $\varphi(w_t, t)\,u_t^\varepsilon$ are uniformly integrable with respect to the product of the Lebesgue measure of interval $[s, T]$ by the Wiener measure. To prove the uniform integrability it is enough to estimate

$$E\int_s^T |\varphi(w_t, t)\,u_t^\varepsilon|^2\,dt$$

from above uniformly with respect to ε. This expression can be estimated through

$$E\left((\sup|\varphi(w_t, t)|^2)\int_s^T (u_t^\varepsilon)^2\,dt\right) \leqslant Q^2 E(\sup|\varphi(w_t, t)|^2), \quad s \leqslant t \leqslant T$$

since $u^\varepsilon \in U(Q)$. Further, φ has an exponential growth and, therefore, it remains to note that for any λ

$$E(\sup\exp(\lambda\,|w_t|)) < \infty, \; s \leqslant t \leqslant T$$

7. Example. We consider the controlled motion of a rigid body undergoing random perturbations around a fixed point O. The equation of motion is

$$dM_t = ([M_t, \omega_t] + u_t)dt + dw_t, \; M_s = x$$

where $M = J\omega$ is the kinetic moment vector relative to O in the body, ω is the angular velocity vector in the body, J is the inertia tensor, u is the control, w_t is the three-dimensional

Brownian motion. We are required to minimize the mathematical expectation $E^u_{x,s}(|M_T|^2)$ of the square of the moment's absolute value at a fixed instant T under constraint of form (1.2) on the control. By the Itô lemma it follows that the scalar $r_t = |M_t|$ satisfies the equation

$$dr_t = \left(\left(u_t, \frac{M_t}{|M_t|}\right) + \frac{2}{r_t}\right)dt + d\xi_t, \quad r_s = |x| \tag{7.1}$$

where ξ_t is a scalar Wiener process. By this same Itô lemma we get that if the vector process x_t satisfies the equation

$$dx_t = u_t\,dt + dw_t, \; x_s = x$$

then the process $|x_t|$ satisfies an equation analogous to (7.1) and the optimization problems for $E^u_{x,s}(|M_T|^2)$ and $E^u_{x,s}(|x_T|^2)$ under constraint (1.2) are equivalent. Therefore, from Theorem 2 we obtain the control quasi-optimal under small values of resource Q, given in the form of the synthesis

$$u_t = u(t, M_t, Q_t) = Q_t \frac{2M_t}{S_1(t, |M_t|)}$$

where $S_1(t, r)$ is a nonpositive solution of the equation

$$\frac{\partial S_1}{\partial t} + \frac{1}{2}\left(\frac{\partial^2 S_1}{\partial r^2} + \frac{2}{r}\frac{\partial S_1}{\partial r}\right) + \frac{2r^2}{S_1} = 0, \quad S_1(T, r) = 0$$

Function S_1 can be written in selfsimilar variables as $S_1(t, r) = r^2 f(\xi)$, where $\xi = r(T-t)^{-1/2}$, while $f(\xi)$ satisfies the equation

$$\tfrac{1}{2} f''(\xi)\xi^2 + f'(\xi)(\tfrac{1}{2}\xi^3 + 3\xi) + 3f(\xi) + 2/f(\xi) = 0$$

Methods for the numerical solution of such equations are given in /6/.

The author thanks F.L. Chernous'ko for posing the problem.

REFERENCES

1. CHERNOUS'KO F.L., Self-similar solutions of the Bellman equation for optimal correction of random disturbances. PMM Vol.35, No.2, 1971.
2. CHERNOUS'KO F.L., and KOLMANOVSKII V.B., Optimal Control Under Random Perturbations. Moscow, NAUKA, 1978.
3. STROOCK D.W. and VARADHAN S.R.S., Diffusion processes with continuous coefficients. I, II. Comm. Pure Appl. Math., Vol.22, No.2, 3, 1969.
4. LIPTSER R.Sh. and SHIRIEV A.N., Statistics of Random Processes. Moscow, NAUKA, 1973.
5. McKEAN Jr., H.P., Stochastic Integrals. New York, Academic Press, Inc., 1969.
6. KRASOVSKII N.N. and TRET'IAKOV V.E., On a problem of contact of motions. Dokl. Akad. Nauk, SSSR, Vol.173, No.2. 1967.
7. USHAKOV V.N. Extremal strategies in differential games with integral constraints. PMM Vol. 36, No.1, 1972.

Translated by N.H.C.