ANNALES DE L'I. H. P., SECTION C

I. CAPUZZO DOLCETTA
M. FALCONE

Discrete dynamic programming and viscosity solutions of the Bellman equation

Annales de l'I. H. P., section C, tome S6 (1989), p. 161-183
<http://www.numdam.org/item?id=AIHPC_1989__S6__161_0>

DISCRETE DYNAMIC PROGRAMMING AND VISCOSITY SOLUTIONS OF THE BELLMAN EQUATION

I. CAPUZZO DOLCETTA & M. FALCONE
Dipartimento di Matematica, Università di Roma "La Sapienza", P. Aldo Moro, 2, 00185 Roma, Italy

Abstract. This paper presents a technique for approximating the viscosity solution of the Bellman equation in deterministic control problems. This technique, based on discrete dynamic programming, leads to monotonically converging schemes and allows to prove a priori error estimates. Several computational algorithms leading to monotone convergence are reviewed and compared.

Key words. Bellman equation, dynamic programming, approximation schemes, viscosity solutions.

Introduction

The notion of viscosity solution of Hamilton-Jacobi equations recently introduced by Crandall and Lions [21] and developed by various authors (see Crandall-Ishii-Lions [19], Lions [40] and references therein) has proved to be an important tool in several applications and particularly in the dynamic programming approach to deterministic (see Lions [38], [39], Capuzzo Dolcetta-Evans [13], Bardi [1], Barles [4], Barles-Perthame [5],[6], Lions-Souganidis [42], Soner [50]) and stochastic (see Fleming [26] and references therein) optimal control problems.

The aim of this paper is to discuss some aspects of the theory of viscosity solutions related to approximation and computational methods in the framework of deterministic control theory. More precisely, our purpose here is to point out on a model problem how the solution of the classical discrete time dynamic programming functional equation (see Bellman [7]) can be regarded as a uniform approximation of the viscosity solution of the corresponding Bellman partial differential equation.

We will show how this PDE approach allow to establish general results on the rate of convergence of the approximate solutions. These results provide a wider theoretical basis to classical and more recent computational techniques for the value function and optimal feedbacks. The last section is completely devoted to review in this framework several methods leading to monotone convergence (namely, successive approximations, iteration in policy space and finite difference approximations).

Let us mention finally that most of these methods have been also applied to stochastic control problems. We refer the interested reader to Bellman-Dreyfus [8], Howard [29], and to the more recent works by Kushner [32], Kushner-Kleinman [33],[34], Lions-Mercier [41], Menaldi [46], Quadrat [49], Bensoussan-Rungaldier [10].

1. The infinite horizon problem: discrete Bellman equation and synthesis.

Let us suppose to observe the state $x(t) \in \mathbb{R}^n$ of some deterministic system evolving in time and to be able to affect its evolution by acting on some available controls. A standard model for the evolution can be expressed in terms of a system of controlled ordinary differential equations:

$$(1.1)\qquad \begin{cases} dx(t)/dt = b(x(t),a(t)), & t>0 \\ x(0) = x, & \end{cases}$$

where the control a is a measurable function of t taking values in a given compact subset A of $\mathbb{R}^m$.

The cost functional which measures the performance of the control a is taken of the form

$$(1.2) \qquad J(x,a) = \int_0^{\infty} f(x(t),a(t))\, e^{-\lambda t}\, dt$$

where f is a given function (the running cost) which we shall always assume to be bounded on $\mathbb{R}^n \times A$ and λ is a positive number (the discount factor).

The discounted infinite horizon problem is to determine, for any initial position $x \in \mathbb{R}^n$, the value

$$(P) \qquad v(x) = \operatorname{Inf}_{a(.) \in \mathcal{A}} J(x,a)\,, \qquad \mathcal{A} = \{\, a: [0,+\infty) \to A,\ a \text{ measurable}\}$$

and to identify a control a^* (depending on x) for which the infimum in (P) is attained, provided such a control exists (see Fleming-Rishel [27], Lee-Markus [37] as general references for optimal control problems and Carlson-Haurie [18] for the treatment of several examples of infinite horizon problems arising in the applications).

In this section we describe how discrete time dynamic programming methods (see Bellman [7], Bensoussan [9], Bertsekas-Shreve [11]) apply to problem (P). At this purpose, let h be a fixed positive number and assume that the evolution (1.1) is observed only at a sequence of instants $t_j = jh$, $j=0,1,2,....$ Assume as well that the dynamics b and the running cost f remain constant in any time interval $[t_j,t_{j+1}[$, namely,

$$(1.3) \qquad \begin{aligned} b(x(t),a(t)) &= b(x_j,a_j) \\ f(x(t),a(t)) &= f(x_j,a_j), \quad t \in [t_j,t_{j+1}[\end{aligned}$$

where

$$(1.4) \qquad a_j = a(t_j)$$

and $x_j = x(t_j)$ is given by the recursion

$$(1.5) \qquad \begin{aligned} x_0 &= x \\ x_{j+1} &= x_j + hb(x_j,a_j)\,, \qquad j = 0,1,2,... \end{aligned}$$

The total cost associated with the initial position x and the control a is given, in this discretized model, by the series

(1.6) $$J_h(x,a) = h\, \Sigma_{j=0,\infty}\ \beta^j\, f(x_j,a_j)$$

with $\beta = 1 - \lambda h$.

Let us introduce now the approximate value function v_h by setting

$$(P_h) \qquad v_h(x) = \operatorname{Inf}_{a \in \mathcal{A}_h} \{ J_h(x,a) \}$$

where $\mathcal{A}_h$ is the subset of $\mathcal{A}$ consisting of controls which are constant on each interval $[t_j,t_{j+1}[$.

The statement below comprises the well-known Bellman Dynamic Programming Principle for the problem under consideration:

Proposition 1.1 The following identity holds

$$(DPP_h) \qquad v_h(x) = \operatorname{Inf}_{a \in \mathcal{A}_h} \{ h\, \Sigma_{j=0,p-1}\ \beta^j f(x_j,a_j) + \beta^p v_h(x_p)\}$$

for all $x = x_0 \in \mathbb{R}^n$ *and* $p=1,2,\ldots$

Proof. Let $x = x_0$ be an arbitrary but fixed vector in $\mathbb{R}^n$. By (P_h), for any $\varepsilon > 0$ there exists $a^\varepsilon \in \mathcal{A}$ such that

$$v_h(x) + \varepsilon \geq J_h(x,a^\varepsilon).$$

Now it is easy to check that for any $p \geq 1$

$$J_h(x, a^\varepsilon) = h\Sigma_{j=0,p-1}\ \beta^j f(x^\varepsilon_j, a^\varepsilon_j) + \beta^p\, h\Sigma_{j=0,\infty}\ f(x^\varepsilon_{j+p}, a^\varepsilon_{j+p}).$$

The second term in the right hand member is actually $\beta^p J_h(x^\varepsilon_p, a^\varepsilon)$.
Hence,

$$v_h(x) + \varepsilon \geq h\Sigma_{j=0,p-1}\ \beta^j f(x^\varepsilon_j, a^\varepsilon_j) + \beta^p v_h(x^\varepsilon_p) \geq$$

$$\geq \operatorname{Inf}_{a \in \mathcal{A}_h} \{ h\Sigma_{j=0,p-1}\ \beta^j f(x_j, a_j) + \beta^p v_h(x_p) \}.$$

The reverse inequality is proved in a similar way. ■

The (DPP_h) yields the Bellman functional equation characterizing the value function v_h.

Proposition 1.2 Let us assume $| f(x,a) | \leq M$ *for all* $(x,a) \in \mathbb{R}^n \times A$. *Then the function* v_h *defined by* (P_h) *is the unique bounded solution of*

(B_h) $\qquad u(x) + \mathrm{Sup}_{a \in A} [-\beta u(x + hb(x,a)) - hf(x,a)] = 0 , \quad x \in \mathbb{R}^n;$

for $h \in [0,1/\lambda[$.

Proof. The boundedness of v_h is an easy consequence of the assumption on f . To check that v_h satisfies (B_h), take $p = 1$ in (DPP_h). This gives

$$v_h(x) + \mathrm{Sup}_{a \in A} [- \beta v_h(x + hb(x,a)) - hf(x,a)] \leq 0.$$

To prove the converse, take $p = 1$ in (DPP_h) and observe that for any $\varepsilon > 0$ ther exists a^ε such that

$$v_h(x) \geq hf(x,a^\varepsilon) + \beta v_h(x + hb(x,a^\varepsilon)) - \varepsilon$$

and therefore

$$v_h(x) + \mathrm{Sup}_{a \in A} [- \beta v_h(x + hb(x,a)) - hf(x,a)] \geq 0.$$

Finally, let u_1, u_2 be two bounded solutions of (B_h). It is immediate to deduce from (B_h) that

$$| u_1(x) - u_2(x) | \leq \beta \, \mathrm{Sup}_{x \in \mathbb{R}^n} | u_1(x) - u_2(x) |$$

and the uniqueness is proved. ■

On this characterization of the approximate value function v_h relies an algorithm for the synthesis of optimal feeback controls for the discrete time problem (P_h).

In order to describe it, let us assume that v_h is lower semicontinuous (this is the case if, for example, b is continuous and f is lower semicontinuous). Then, for any fixed $x \in \mathbb{R}^n$ the supremum in (B_h) is attained at some $a_h = a_h(x) \in A$. Define next

$$(1.7) \qquad x^*_0 = x, \qquad x^*_{j+1} = x^*_j + hb(x^*_j, a^*_h(x^*_j)), \qquad j=0,1,2,...$$

and set

$$(1.8) \qquad a^*_h(t) = a^*_h(x^*_j), \quad t \in [t_j, t_{j+1}[.$$

Therefore we have:

Proposition 1.3 Let $h \in [0,1/\lambda[$. *If* v_h *is lower semicontinuous and* A *is compact, then the piecewise constant control defined by (1.8) is optimal for* (P_h), *that is*

$$(1.9) \qquad v_h(x) = J_h(x, a^*_h), \text{ for all } x \in \mathbb{R}^n.$$

Proof. By the very definition of the mapping $x \to a^*_h(x)$ from equation (B_h) it follows that for any p

$$\beta^j \, [v_h(x^*_j) - \beta v_h(x^*_{j+1}) - hf(x^*_j, a^*_h(x^*_j))\,] = 0 .$$

This yields by addition

$$v_h(x) = \beta^{p+1} v_h(x^*_{p+1}) + h \, \Sigma_{j=0,p} \, \beta^j f(x^*_j, a^*_h(x^*_j))$$

and (1.9) follows by letting $p \to +\infty$. ■

2. Limiting behaviour as the time step vanishes.

In this section we shall investigate the behaviour of v_h and a^*_h as the time step h tends to zero. It is natural to expect that v_h and a^*_h should converge, respectively, to the value function v and to some optimal control (if it exists) for problem (P).

We report hereafter on some recent convergence results and error estimates. A crucial role in establishing these results is played by two key facts in the theory of viscosity solutions of Hamilton - Jacobi - Bellman equations. First, the observation due to P.L.Lions (see [38]) that the value v of problem (P), even when only continuous, satisfies nonetheless the corresponding Bellman equation

(B) $$\lambda u(x) + \operatorname{Sup}_{a\in A} [-b(x,a)\cdot Du(x) - f(x,a)] = 0 \ , \ x \in \mathbb{R}^n,$$

in the viscosity sense. Second, the uniqueness result of M.G.Crandall-P.L.Lions [21] for bounded uniformly continuous viscosity solutions of (B). Let us recall for the reader's convenience that a viscosity solution of (B) is a bounded uniformly continuous function u such that for each $\phi \in C^1(\mathbb{R}^n)$ the following holds:

(i) if $u-\phi$ attains a local maximum at x_0 then

$$\lambda\, u(x_0) + \operatorname{Sup}_{a \in A} [- b(x_0,a)\, D\phi(x_0) - f(x_0,a)] \leq 0$$

(ii) if $u-\phi$ attains a local minimum at x_0 then

$$\lambda\, u(x_0) + \operatorname{Sup}_{a \in A} [- b(x_0,a)\, D\phi(x_0) - f(x_0,a)] \geq 0.$$

Let us also point out that the use of this weak notion of solution of (B) allows to pass to the limit as $h \to 0$ in the nonlinear equations (B_h) under rather general conditions on b and f.

Let us state and sketch the proof of two results in this direction (see Capuzzo Dolcetta [12], Capuzzo Dolcetta-Ishii [14] and Loreti [43] for details and extensions).

Theorem 2.1 Let us assume that b *and* f *are continuous on* $\mathbb{R}^n \times A$ *and that for some constants* L, M *the following holds*

(2.1) $\quad \| b(x,a) - b(x',a) \| \leq L \mid x - x' \mid \ , \ \| b(x,a) \| \leq M$

(2.2) $\quad \mid f(x,a) - f(x',a) \mid \ \leq M \mid x - x' \mid \ , \ \mid f(x,a) \mid \ \leq M,$

for all x, x' *in* $\mathbb{R}^n$, $a \in A$, *and*

(2.3) $\quad \lambda > L.$

Then, $v_h \to v$ *locally uniformly on* $\mathbb{R}^n$ *as* $h \to 0^+$ *and*

(2.4) $\quad \text{Sup} \ (v(x) - v_h(x)) \ \leq \ Ch,$

for some constant $C > 0$.

Theorem 2.2 Under the assumptions of Theorem 2.1 and

(2.5) $\quad \| b(x+z,a) - 2b(x,a) + b(x-z,a) \| \ \leq \ M \ |z|^2$

(2.6) $\quad f(x+z,a) - 2f(x,a) + f(x-z,a) \ \leq \ M \ |z|^2,$

(2.7) $\quad \lambda > 2L$

the following estimate of the rate of convergence holds

(2.8) $\quad \text{Sup} \ (v_h(x) - v(x)) \ \leq \ Ch,$

for some constant $C > 0$.

Sketch of the proof of Theorem 2.1.
The first step is to establish the uniform estimates:

$$\text{Sup} \mid v_h(x) \mid \ \leq \ M/\lambda \quad , \ \text{Sup} \ (\mid v_h(x) - v_h(x') \mid / \mid x - x' \mid) \leq M/(\lambda - L),$$

for $h \in [0,1/\lambda[$. These are simple consequences of the fact that v_h satisfies (B_h). Then, by Ascoli-Arzelà theorem, there exists a Lipschitz continuous function u such that

$$v_h \to u \quad \text{locally uniformly in } \mathbb{R}^n$$

as $h \to 0^+$ (at least for a subsequence).

The next step is to show that u is a viscosity solution of (B). To do this, let $\phi \in C^1(\mathbb{R}^n)$, x_0 a local strict maximum point for $u - \phi$, S a closed ball centered at x_0, x_{0h} a maximum point for $v_h - \phi$ over S. Then (B_h) yields :

$$0 = v_h(x_{0h}) + \mathrm{Sup}_{a\in A} [- \beta h v_h(x_{0h} + hb(x_{0h},a)) - hf(x_{0h},a)] \geq$$

$$\geq \mathrm{Sup}_{a\in A} [\phi (x_{0h}) - \phi (x_{0h} + hb(x_{0h},a)) + \lambda h v_h(x_{0h} + hb(x_{0h},a)) - hf(x_{0h},a)].$$

Since $\phi \in C^1(\mathbb{R}^n)$, it follows that

$$0 \geq \mathrm{Sup}_{a\in A} [-D \phi (x_{0h}+\theta hb(x_{0h},a)\cdot b(x_{0h},a)+\lambda v_h(x_{0h}+hb(x_{0h},a)) - f(x_{0h},a)]$$

for some $\theta = \theta(h,a) \in [0,1]$. Hence, by uniform convergence,

$$0 \geq \lambda u(x_0) + \mathrm{Sup}_{a\in A} [- D \phi (x_0) \cdot b(x_0,a) - f(x_0,a)],$$

that is u is a viscosity subsolution of (B). A similar argument shows that u is also a viscosity supersolution of (B) and, by the uniqueness theorem of Crandall-Lions [21], $u \equiv v$.

In order to prove (2.4), observe that by definition,

$$v(x) - v_h(x) \leq \mathrm{Inf}_{a \in \mathcal{A}_h} J(x,a) - \mathrm{Inf}_{a \in \mathcal{A}_h} J_h(x,a)$$

$$\leq \mathrm{Sup}_{a \in \mathcal{A}_h} | J(x,a) - J_h(x,a) |$$

A simple computation shows that for any $a \in \mathcal{A}_h$

$$| J(x,a) - J_h(x,a) | \leq Kh + M \lambda (| 1-\theta | + h) \int_0^{+\infty} ((s+1) \max \{ e^{-\lambda s}; e^{-\lambda\theta s} \} ds ,$$

where $\theta = -\log(1-\lambda h)/\lambda h$. Since $\theta - 1$ is of order h as $h \to 0^+$, (2.4) follows. ■

Sketch of the proof of Theorem 2.2.

Let us consider the function Ψ defined by

$$\Psi (x,y) = v_h(x) - v(y) - |(x-y)/\varepsilon|^2 + \delta \, \xi(x,y) , \quad \delta,\varepsilon \in (0,1),$$

where ξ is an auxiliary function forcing Ψ to attain its maximum at some point $(x_0,y_0) \in \mathbb{R}^{2n}$. It is not difficult to show that $| x_0 - y_0 | \leq C\varepsilon^2$. Since v is the viscosity solution of (B), then

$$(2.9) \qquad \lambda v(y_0) + b(y_0,a^*) \cdot [- 2(x_0-y_0)/\varepsilon^2 - \delta D_y \xi (x_0,y_0)] - f(y_0,a^*) \geq 0$$

for some $a^* \in A$. The next step is to show that v_h satisfies the semiconcavity condition

$$(2.10) \qquad v_h(x+z) - 2v_h(x) + v_h(x-z) \leq C |z|^2 ,$$

for some $C > 0$. To prove this observe that v_h satisfies, by its very definition,

$$v_h(x+z)-2v_h(x) + v_h(x-z) \leq \operatorname{Sup}_{a \in \mathcal{A}_h} [J_h(x+z,a) - 2J_h(x,a) - J_h(x-z,a)] .$$

A rather technical computation based on (2.5), (2.6) shows (see [14] for details) that

$$(2.11) \qquad J_h(x+z,a) - 2J_h(x,a) + J_h(x-z,a) \leq e^{\theta\lambda h} \int_0^{\infty} \gamma (t) \, e^{-\lambda t} dt,$$

where

$$\gamma(t)=\min\{4M;M|z|^2[(1+Lh)2[t/h]+Mh(1+Lh)[t/h]-1(1+Lh)[t/h)-1/(1+Lh)-1]\}$$

$$\theta = -\log(1-Lh)/Lh.$$

Since

$$\gamma(t) \le \min \{ 4M \,;\, M\,|z|^2\,(1+M/lv)\,e^{2Lt} \},$$

with $v = \log(1+Lh)/Lh$, and $\lambda > 2L$, the integral in (2.11) can be estimated by $CM\,(1+M/vL)\,|z|^2$. This implies (2.10). As a consequence of the uniform semiconcavity (2.10) of v_h one has (see [14] for technical details)

$$v_h(x_0+x)-v_h(x_0)+(-2(x_0-y_0)/\varepsilon^2-\delta D\xi(x_0,y_0))\cdot x \le C\,|x|^2$$

for all $x\in \mathbb{R}^n$. The choice $x= hb(x_0,a^*)$ in the above yields

$$v_h(x_0+hb(x_0,a^*))-v_h(x_0)+(-2(x_0-y_0)/\varepsilon^2-\delta D\xi(x_0,y_0))\cdot hb(x_0,a^*)\le$$
$$\le C|hb(x_0,a^*)|^2.$$

From this and equation (B_h) evaluated at $x = x_0$ it follows that

$$v_h(x_0)-(1-\lambda h)[v_h(x_0)-hb(x_0,a^*))\cdot(-2(x_0-y_0)/\varepsilon^2+\delta D_x\xi(x_0,y_0))+$$
$$+C|b(x_0,a^*)|^2h^2] - hf(x_0,a^*) \le 0.$$

This inequality, combined with (2.9), yields

$$v_h(x_0)-v(y_0) \le C\,[|x_0-y_0|2/\varepsilon^2 + h|x_0-y_0|/\varepsilon^2 + |x_0-y_0| + h + \delta\,]$$

$$\le C(\varepsilon^2+h+\delta).$$

Choosing $\varepsilon = h^{1/2}$, from the above inequality it follows that

$$v_h(x) - v(x) \le Ch$$

and (2.8) is proved. ■

Remark 2.1. Under the sole assumptions of Theorem 2.1 it is possible to show that

$$\mathrm{Sup}\ (v_h(x) - v(x)) \leq Ch^{1/2} ;$$

for a proof we refer to [14] and also to the papers by Crandall - Lions [20] and Souganidis [51] where similar estimates are established for general Hamilton - Jacobi equations of the form

$$H\ (x,u(x),Du(x)) = 0,$$

including for example the Isaac's equation in differential games theory. The above Theorem 2.2 exploits, through the semiconcavity assumptions (2.5), (2.6), the special structure of the Bellman's equation (B) , namely the convexity of the mapping

$$p \to \mathrm{Sup}_{\, a\in A}\ [\ -b(x,a)\cdot p - f(x,a)\ ]$$

in order to obtain the sharper, and in fact optimal (see [14]), estimate (2.8).
Theorems 2.1 and 2.2 can be regarded as an extension of earlier results by J.Cullum [22], K.Malanowski [44] , M.M.Hrustalev [30] obtained via the Pontryagin maximum principle under rather restrictive convexity assumptions. The asymptotic behaviour of the discrete value function as the time step vanishes is studied in [16], [17] for other control problems such as the stopping time and the optimal switching problem, without use of the notion of viscosity solution. More recently the notion of discountinuous viscosity solutions (see [5]) has been applied to prove the convergence of a discretization scheme for the minimum time problem by Bardi-Falcone [2].

Remark 2. As a consequence of Theorem 2.1,2.2 and Proposition 1.3, the piecewise constant controls a^*_h (see §1) yield the value of problem (P) with any prescribed accuracy. Their limiting behaviour as $h \to 0^+$ is easily understood in the framework of relaxed controls in the sense of L.C.Young [53] (see also Lee-Markus [37], Warga [52] as general references on this subject). Namely, a^*_h converge weakly to an optimal control for the relaxed version of (P). The result relies on the fact that the Hamiltonian is left unchanged by relaxation (see Capuzzo Dolcetta -Ishii [10] for details and also to Mascolo-Migliaccio [45] for other relaxation results in optimal control).

3. Discretization in the state variable and computational methods.

In order to reduce equation (B_h) to a finite dimensional problem we need a discretization in the state variable x. To be able to build a grid in the state space, let us assume the existence of an open bounded subset Ω of $\mathbb{R}^n$ which is invariant for the dynamics (1.1). A triangulation of Ω into a finite number P of simplices S_j can be constructed (Falcone [23]) so that, the set $\Omega^k = \cup_{j=1,P} S_j$ is invariant with respect to the discretized trajectories, i.e.

(3.1) $$\exists h > 0 : \quad x+h\, b(x,a) \in \Omega^k , \qquad \forall (x,a) \in \Omega^k \times A.$$

where k=max diam(S_j). We denote x_i, i=1,...,N, a generic node of that grid and we replace (B_h) by the following system of N equations

(B_h^k) $$u(x_i) + \text{Sup}_{a \in A} [-\beta u(x_i + hb(x_i,a)) - hf(x_i,a)] = 0 , \quad i=1,\dots,N$$

which corresponds to the above outlined discretization in the space variable. We shall briefly refer in the sequel on some computational techniques available to solve this system of equations.

Remark 3.1. Notice that, due to the fact that we are dealing with the infinite horizon problem, the discretization in time (1.3), (1.4) and (1.5) is not sufficient in itself to compute an approximate solution of (B). Infact, even if in priciple the value of u(x) can be obtained by (B_h) from the previous knoledge of u at all points x+hb(x,a), i.e. the points which follow x in the discretization scheme (1.3), the classical backward scheme of the dynamic programming computational procedure (Larson [35]) cannot be applied since time varies in $[0,+\infty[$. That scheme is only suitable for finite horizon problems where a terminal condition is given.

Remark 3.2. Other numerical methods which can be applied directly to (B) do not lead to the system (B_h^k)(see also Remark 5). Let us just mention the finite element approach used by Gonzales-Rofman [28] to solve the Bellman equation related to a stopping time problem with continuous and impulsive controls. The procedure generates a monotone non-decreasing sequence and a convergence result is obtained by means of a discrete maximum principle (see also Menaldi-Rofman [47] where the convergence to the viscosity solution is proved).

Successive approximation (Bellman [7], Bellman-Dreyfus [8]).

Perhaps the most classical technique to compute a solution of (B_h^k) is to apply Picard method of successive approximations.
Starting from any assigned initial guess

$$(3.2) \qquad u(x_i)=u_i^0 , \qquad i=1,...,N$$

one can compute $u^0(x_i+hb(x_i,a))$ by a simple interpolation on the nodes of the grid and then define the following recursive sequence

$$(3.3) \qquad u^{n+1}(x_i) =T_h(u^n)=\mathrm{Inf}_{a\in A} [\beta u^n(x_i+hb(x_i,a))+hf(x_i,a)] , \quad i=1,...,N.$$

Since T_h is a contractive map (see Theorem 3.1 below), u^n will converge to a limit as h tends to 0.

In spite of its simplicity this technique has several computational disadvantages which do not compensate for the fact that convergence is assured for any initial choice u^0.

In order to illustrate this point, assume to have M admissible controls. Since the point $x_i+hb(x_i,a)$ could be anywhere in Ω^k, the computation of u^{n+1} would require to keep the N values of u^n at the nodes in the central memory and compare M different values to find the value of u^{n+1} at each node: this means that $N \times M$ comparisons are needed at each iteration. However, using state increment dynamic programming (see Larson [33]) it is possible to reduce considerably the huge demand for memory allocations. Larson' technique consists essentially in the introduction of a compatibility condition between the step in time h and the spacial step k . By defining $M_b = \max_{i,a} |b(x_i, a)|$, $h_1 = k/M_b$ and choosing h_2 such that the invariance condition (3.1) is verified , the choice of $h = \min\{h_1, h_2\}$ guarantees that the points $x_i + hb(x_i, a)$ belong to some simplex having x_i among its nodes. In that way only the values of u^n at neighbouring nodes are needed to compute u^{n+1} at x_i.

The second disadvantage is that the velocity of convergence of this algorithm becomes very poor when a great accuracy is requested, due to the fact that the contraction coefficient of T_h is $\beta = 1 - \lambda h$, which is close to 1 for h close to 0. We shall see in the section devoted to finite difference schemes an acceleration technique which permits to overcome this difficulty. This acceleration technique is essentially based on the monotonicity of T_h which guarantees a monotone convergence of u^n for an appropriate choice of u^0.

Approximation in policy space (Bellman-Dreyfus [8], Howard [29]).

A different computational procedure leading to monotone convergence is the Bellman-Howard approximation in policy space. The idea here is to fix an initial guess for the policy rather then for the value function (as in the successive approximation procedure). Assuming a_0 to be such an initial guess, we can compute the related return function by solving the equations

(3.4) $$u(x_i) = \beta u(x_i + hb(x_i, a_0)) + hf(x_i, a_0) \quad , \; i = 1, \dots, N$$

This can be done starting from an assigned $u^{0,0}$, and generating the recursive sequence

(3.5) $$u^{0,n+1} = \beta u^{0,n}(x_i + hb(x_i, a_0)) + hf(x_i, a_0)$$

which, under appropriate assumptions, will converge

$$(3.6) \qquad \lim_{n\to\infty} u^{0,n} = u^0 .$$

Then one looks for a new policy such that

$$(3.7) \qquad \beta u^0(x_i + hb(x_i,a_1)) - hf(x_i,a_1) =$$

$$= \operatorname{Inf}_{a\in A} [-\beta u^0(x_i + hb(x_i,a)) - hf(x_i,a)] , \quad i=1,\dots,N$$

and replaces a_0 with a_1 in (3.4) to compute a new return function u^1 by means of recursion.

Notice that

$$u^0(x_i) = \beta u^0(x_i + hb(x_i,a_0)) + hf(x_i,a_0) \geq$$

$$\geq \operatorname{Inf}_{a\in A} [-\beta u^0(x_i + hb(x_i,a)) - hf(x_i,a)] =$$

$$= \beta u^0(x_i + hb(x_i,a_1)) + hf(x_i,a_1) \geq u^1(x_i) \qquad , \quad i=1,\dots,N.$$

Continuing in this way, we obtain a monotone decreasing sequence u^n and the procedure stops when $a^{n+1} = a^n$. Since for any choice of the policy we obtain a return function which is greater or equal to the value function u, we have

$$(3.8) \qquad u^0(x_i) \geq u^1(x_i) \geq u^2(x_i) \geq \dots \geq u(x_i) \qquad , \quad i=1,\dots,N$$

so the approximation in policy space yields monotone convergence.

This procedure appear to have more appealing properties from the computational point of view. Infact, Kalaba [31] has shown that it is equivalent to the Newton-Kantorovich iteration procedure applied to the functional equation of dynamic programming and Puterman-Brumelle [48] have given sufficient conditions for the rate of convergence to be either superlinear or quadratic.

Remark 3.3. Assume to have M admissible controls. Even if this procedure does not require M comparisons to compute $u^{j,n}(x_i)$ in (3.5), they are requested in (3.7) . Then the cost of this procedure in terms of computational requirements is similar to the method of successive approximation and depends mainly from the velocity of convergence in (3.6). Larson' technique can also be applied to this scheme.

Finite difference approximation.

We look for a solution of $(B_h{}^k)$ in the following space of piecewise affine functions, $W = \{w \in C(\Omega^k)\colon Dw(x) = c_j \text{ in } S_j\}$. The following theorem holds (see Falcone [23])

Theorem 3.1 Under the same assumptions of Theorem 2.1, for any $h \in [0, 1/\lambda[$ *satisfying (3.1) there exists a unique solution* $v_h{}^k$ *of* $(B_h{}^k)$ *in* W *and the following estimate holds*

$$\text{(3.9)} \qquad \text{Sup } |v_h{}^k(x) - v(x)| \leq C_1 h^{1/2} + C_2 k \ .$$

Sketch of the proof.
Due to the invariance assumption (3.1) any point $x_i + hb(x_i, a)$ belongs to Ω^k and can be written as a convex combination of the nodes, with coefficients $\lambda_{ij}(a)$,. Since we look for a solution in W, equation (3.3) is equivalent to the following fixed point problem in $\mathbb{R}^N$

$$\text{(3.10)} \qquad U = T_h(U)$$

where $(T_h(U))_i \equiv \text{Inf}_{a \in A} \ [\beta \ \Lambda(a) \ U + h \ F(a)]_i$ and $\Lambda(a) = [\lambda_{ij}(a))]$ is a nonnegative matrix $N \times N$ such that the sum of the elements on each row is equal to 1, U and F are vectors of $\mathbb{R}^N$. By the properties of Λ follows that

$$\text{(3.11)} \qquad |(T_h(U) - T_h(U))_i| \leq \beta \max_i |(U-V)_i| \qquad \forall\, U, V \in \mathbb{R}^N$$

therefore T_h is a contracting operator in $\mathbb{R}^N$ and (3.10) has a unique solution U^*.

Estimate (3.9) can be derived combining the estimates on Sup $|v_h(x) - v(x)|$ obtained in the preceeding section with the lipschitz continuity of v_h and the fact that $U^* = (v_h(x_1), \ldots, v_h(x_N))$. ■

Even if the preceeding theorem establishes the convergence of the recursive sequence $U^{n+1} = T_h(U^n)$ starting from any initial guess U^0, special choices of U^0 are more appropriate for numerical purposes in order to obtain monotone convergence and to accelerate convergence. The acceleration technique is crucial to break the obstacle constituted by the fact the convergence is very slow whenever a great accuracy is requested, as we explained in the section devoted to successive approximation.
Let us consider the set of subsolutions for the fixed point problem (3.10), i.e. the set $\mathcal{U} = \{U: U \leq T_h(U)\}$. Since $U \leq V$ implies $T_h(U) \leq T_h(V)$, the recursive sequence U^n is monotone increasing and remains in $\mathcal{U}$ whenever U^0 is taken in $\mathcal{U}$. This set is closed and convex, by the definition of T_h, and U^* is its maximal element since it is the limit of any (monotone) sequence starting in $\mathcal{U}$. We can construct then the following

Accelerated algorithm

<u>Step 1</u>: take $\mathbf{U}^0$ in $\mathcal{U}$ and compute $\mathbf{U}^{1,0} = T_h(U^0)$;

<u>Step 2</u> : compute $\mathbf{U}^1 = \mathbf{U}^0 + \alpha\,(\mathbf{U}^{1,0} - \mathbf{U}^0)$ where

(3.12) $\alpha = \max\{\alpha \in R_+ : \mathbf{U}^0 + \alpha\,(\mathbf{U}^{1,0} - \mathbf{U}^0) \in \mathcal{U}\};$

<u>Step 3</u> : replace $\mathbf{U}^0$ with $\mathbf{U}^1$ and go back to Step 1.

The convergence of this algorithm to U^* is guaranteed by the convergence and monotonicity properties of U^n, since $U^n \leq \mathbf{U}^n \leq \mathbf{U}^{n+1} \leq U^*$.
The first interesting feature of this acceleration method is that its velocity of convergence is not strictly related to the magnitude of $\beta = 1 - \lambda h$, since T_h is used only to find a direction of displacement. Infact several

numerical tests have shown that a decrease of h does not always imply an increase in the number of iterations (for a detailed discussion of some numerical tests as well as for the practical implementation of the algorithm see Falcone [24] and [25]). The second appealing property is that, once a direction of displacement has been computed, the algorithm proceeds in Step 2 by a one-dimensional constrained optimization problem to recover the next approximation. Finally, it is quite interesting to compare the actual acceleration parameter α_n with 1, which is the parameter value corresponding to standard successive approximation, in order to have an empirical estimate of the velocity of convergence. Numerical tests have shown that $\alpha_n \approx h^{-1}$, i.e. $\alpha_n \approx 100$ when $h=10^{-2}$.

Remark 3.4. Assume to have M admissible controls. The cost for finding a direction of displacement is exactly the same of successive approximation and the search of U^1 at Step 2 would require a certain (a priori unknown) number of additional comparisons to check that the constraint in (3.12) is verified. Nonetheless, the total amount of operations needed in order to find a fixed point with the above acceleration technique is much lower since the total number of iterations is considerably reduced, e.g. from order 10^3 to order 10 when $h=10^{-2}$. The application of Larson technique to this scheme yields a particular structure of the matrix $\Lambda(a)$, namely it becomes a tridiagonal matrix for scalar control problems.

Remark 3.5. Other finite difference schemes have been considered to solve the Bellman equation (B). In particular Crandall-Lions [18] have studied some finite difference schemes which can be regarded as an adaptation of classical schemes for conservation laws, such as Lax-Friedrichs scheme. It is interesting to point out here that they obtain an estimate of order 1/2 and monotone convergence as well.

We conclude with section with some final comments and remarks related to the schemes that we have presented here. It is worthwhile to notice that these schemes can be adapted to treat other deterministic control problem such as finite horizon and optimal stopping control problems (infact, successive approximation and approximation in policy space were originally developped for those problems, see [7], [8], [35]). This can be done adding some control, say a', to A, and writing those problems as an infinite horizon problem by an appropriate definition of f(x,a'), b(x,a').

The second, and perhaps more important remark, is that all these schemes provide the approximate value function and, without any additional computation, the related approximate feedback law at least for all nodes of the grid. It can also be proved (see Falcone [23]) that such a feedback law is "quasi optimal" for the original continuous problem (P).

Due to the monotonicity, the acceleration method can be applied also to the approximation in policy space. In this case the fixed point will be the minimal element in the set of supersolutions and the sequence $\mathbf{U}^n$ will be monotone decreasing. Moreover, due the local quadratic convergence of this approximation, it should be quite interesting from a numerical point of view to combine this procedure with the finite difference approximation scheme. Infact, one can first look for a rough approximation of the optimal policy by applying that scheme and then use it as initial guess in (3.4).

References

[1] M. Bardi, A boundary value problem for the minimum time function, to appear on SIAM J. Control and Optimization.

[2] M. Bardi, M. Falcone, An approximation scheme for the minimum time function, preprint, july 1988

[3] G. Barles, Inéquations quasi-variationnelles du premier ordre et équations d'Hamilton-Jacobi,, C.R.Acad.Sc.Paris,t.296, 1983.

[4] G. Barles, Deterministic impulse control problems, SIAM J. Control and Optimization, vol.23, 3, 1985

[5] G. Barles, B.Perthame, Discontinuous solutions of deterministic optimal stopping time problems, to appear in Math. Methods and Num. Anal.

[6] G. Barles, B.Perthame, Exit time problems in optimal control and vanishing viscosity method, preprint

[7] R. Bellman, *Dynamic programming,* Princeton University Press, 1957

[8] R. Bellman, S. Dreyfus, *Applied Dynamic Programming,* Princeton University Press, 1962

[9] A. Bensoussan, Contrôle stochastique discret, Cahiers du Ceremade

[10] A. Bensoussan, W. Runggaldier, An approximation method for stochastic control problems with partial observation of the state- A method for constructing ε-optimal controls, Acta Applicandae Mathematicae, vol. 10, 2, 1987

[11] D.P. Bertsekas, S.E. Shreve, *Stochastic optimal control: the discrete time case,* Academic Press, New York, 1978.

[12] I. Capuzzo Dolcetta, On a discrete approximation of the Hamilton - Jacobi equation of dynamic programming, Appl.Math.and Optim. 10,1983

[13] I. Capuzzo Dolcetta,L.C. Evans, Optimal switching for ordinary differential equations, SIAM J. Control and Optimization, vol.22,1984

[14] I. Capuzzo Dolcetta, H. Ishii, Approximate solutions of the Bellman equation of deterministic control theory, Appl. Math.and Optim.11,1984.

[15] I. Capuzzo Dolcetta, P.L. Lions, Hamilton-Jacobi equations and state-constrained problems, to appear in Transactions of the AMS

[16] I. Capuzzo Dolcetta, M. Matzeu, On the dynamic programming inequalities associated with the optimal stopping problem in discrete and continuous time, Numer. Funct. Anal. Optim., 3 (4),1981.

[17] I. Capuzzo Dolcetta,M. Matzeu,J.L. Menaldi, On a system of first order quasi-variational inequalities connected with the optimal switching problem, System and Control Letters, Vol.3,No.2,1983.

[18] D.A. Carlson, A. Haurie, Infinite horizon optimal control, Lecture Notes in Economics and Mathematical Systems, 290, Springer Verlag,1987.

[19] M.G. Crandall, H. Ishii, P.L. Lions, Uniqueness of viscosity solutions of Hamilton-Jacobi equations revisited, J. Math. Soc. Japan, vol. 39, 4, 1987

[20] M.G. Crandall, P.L. Lions, Two approximations of solutions of Hamilton-Jacobi equations, Mathematics of Computation, vol.43, n.167, 1984

[21] M.G. Crandall, P.L. Lions, Viscosity solutions of Hamilton-Jacobi equations, Trans. A.M.S. 277, 1983.

[22] J. Cullum, An explicit procedure for discretizing continuous optimal control problems, J. Optimization Theory and Applications, 8(1),1971.

[23] M. Falcone, A numerical approach to the infinite horizon problem of deterministic control theory, Appl. Math. Optim. 15, 1987

[24] M. Falcone, Numerical solution of deterministic continuous control problems, Proceedings International Symposium on Numerical Analysis, Madrid, 1985

[25] M. Falcone, Approximate feedback controls for deterministic control problems, in preparation

[26] W.H. Fleming, Controlled Markov processes and viscosity solutions of nonlinear evolution equations, Lecture Notes, Scuola Normale Superiore Pisa, to appear

[27] W.H. Fleming, R. Rishel, *Deterministic and stochastic optimal control*, Springer Verlag, Berlin, 1975.

[28] R. Gonzales, E. Rofman, On deterministic control problems: an approximation procedure for the optimal cost, part I and II, SIAM J. Control and Optimization, vol. 23, n. 2, 1985

[29] R. A. Howard, *Dynamic programming and Markov processes*, Wiley, New York, 1960

[30] M.M. Hrustalev, Necessary and sufficient optimality conditions in the form of Bellman's equation, Soviet Math.Dokl. ,19 (5), 1978.

[31] R. Kalaba, On nonlinear differential equations, the maximum operation and monotone convergence, J. of Math. and Mech., vol 8, n. 4, 1959

[32] H.J. Kushner, *Probability methods for approximation in stochastic control and for elliptic equations*, Academic Press, New York, 1977

[33] H.J. Kushner, A.J. Kleinman, Accelerated procedures for the solution of discrete Markov control problems, IEEE Trans. Automatic control, AC-16, 1971

[34] H.J. Kushner, A.J. Kleinman, Mathematical programming and the control of Markov chains, Int. J. Control, 13, 1971

[35] R.E. Larson, *State increment dynamic programming* , American Elsevier, New York, 1967

[36] R.E. Larson, Korsak, A dynamic programming successive approximaton technique with convergence proofs, parts I end II, Automatica, 1969

[37] E.B. Lee,L. Markus, *Foundations of optimal control*, J.Wiley, New York, 1967.

[38] P.L. Lions, *Generalized solutions of Hamilton-Jacobi equations*, Pitman, London, 1982.

[39] P.L.Lions, Optimal control and viscosity solutions, in I. Capuzzo Dolcetta-W.H. Fleming- T. Zolezzi (Eds.), *Recent Mathematical Methods in Dynamic Programming*, Lecture Notes in Mathematics 1119, Springer Verlag Berlin, 1985.

[40] P.L. Lions, Recent progress on first order Hamilton-Jacobi equations, in *Directions in Partial Differential Equations*, Academic Press, 1987

[41] P.L. Lions, B. Mercier, Approximation numérique des équations de Hamilton-Jacobi-Bellman, R.A.I.R.O Analyse numérique, vol.14, n. 4, 1980

[42] P.L. Lions, P.E. Souganidis, Differential games, optimal control and directional derivatives of viscosity solutions of Bellman-Isaacs equations, to appear in SIAM J. Control and Optimization

[43] P.Loreti, Approssimazione di soluzioni viscosità dell'equazione di Bellman, Boll. U.M.I (6), 5-B ,1986.

[44] K.Malanowski, On convergence of finite difference approximations to optimal control problems for systems with control appearing linearly, Archivum Automatyki i Telemechaniki, 24, Zeszyt 2, 1979.

[45] E.Mascolo,L.Migliaccio, Relaxation methods in optimal control, to appear in J. Optim. Th. App.

[46] J.L. Menaldi, Probabilistic view of estimates for finite difference methods, to appear in Mathematicae Notae

[47] J.L. Menaldi, E. Rofman, An algorithm to compute the viscosity solution of the Hamilton-Jacobi-Bellman equation, in C.I. Byrnes, A. Lindquist (eds.), *Theory and Applications of Nonlinear Control Systems*, North-Holland, 1986

[48] M.L. Puterman, S.L. Brumelle, On the convergence of policy iteration in stationary dynamic programming, Math. of Operation Research, vol.4, n. 1, 1979

[49] J.P. Quadrat, Existence de solution et algorothme de resolutions numeriques de problemes stochastiques degenerées ou non, SIAM J. Control and Optimization, vol.18, 1980

[50] M.H.Soner Optimal control with state space constraint, SIAM J. Control and Optimization, vol.24, 3, 1986

[51] P.E.Souganidis, Approximation schemes for viscosity solutions of Hamilton-Jacobi equations, Journal of Differential Equations,57,1985.

[52] J.Warga,*Optimal control of differential and functional equations,* Academic Press,New York,1972.

[53] L.C.Young, *Lectures on the Calculus of Variations and Optimal Control Theory,* W.B.Saunders, Philadelfia, 1969.