Solution of Hamilton Jacobi Bellman Equations

C. L. Navasca
and
A. J. Krener [1]
Department of Mathematics
University of California
Davis, CA 95616-8633

Abstract

We present a new method for the numerical solution of the Hamilton Jacobi Bellman PDE that arises in an infinite time optimal control problem. The method can be of higher order to reduce "the curse of dimensionality". It proceeds in two stages. First the HJB PDE is solved in a neighborhod of the origin using the power series method of Al'brecht. From a boundary point of this neighborhood, an extremal trajectory is computed backward in time using the Pontryagin Maximum Principle. Then ordinary differential equations are developed for the higher partial derivatives of the solution along the extremal. These are solved yielding a power series for the approximate solution in a neighborhood of the extremal. This is repeated for other extremals and these approximate solutions are fitted together by transferring them to a rectangular grid using splines.

Key words: Infinite horizon optimal control, Numerical soluton of Hamilton Jacobi Bellman PDE.

1 Introduction

The Hamilton Jacobi Bellman (HJB) Partial Differential Equation and related equations such as Hamilton Jacobi Isaacs (HJI) equation arise in many control problems. Perhaps the simplest is the infinite horizon optimal control problem of minimizing the cost

$$\int_t^\infty l(x,u)\ dt \tag{1.1}$$

subject to the dynamics

$$\dot{x} = f(x,u) \tag{1.2}$$

and initial condition

$$x(t) = x^0. \tag{1.3}$$

The state vector x is an n dimensional column vector, the control u is an m dimensional column vector and the dynamics $f(x,u)$ and Lagrangian $l(x,u)$ are assumed to be sufficiently smooth.

If the minimum exists and is a smooth function $\pi(x^0)$ of the initial condition then it satisfies the HJB PDE

$$\min_u \left\{ \frac{\partial \pi}{\partial x}(x) f(x,u) + l(x,u) \right\} = 0 \tag{1.4}$$

and the optimal control $\kappa(x)$ satisfies

$$\kappa(x) = \arg\min_u \left\{ \frac{\partial \pi}{\partial x}(x) f(x,u) + l(x,u) \right\} = 0 \tag{1.5}$$

These are conveniently expressed in terms of the Hamiltonian

$$H(p,x,u) = pf(x,u) + l(x,u) \tag{1.6}$$

where the argument p is an n dimensional row vector. The HJB PDE becomes

$$0 = \min_u H(\frac{\partial \pi}{\partial x}(x), x, u) \tag{1.7}$$

$$\kappa(x) = \arg\min_u H(\frac{\partial \pi}{\partial x}(x), x, u) \tag{1.8}$$

If the minimum exists but is not a smooth function then the HJB PDE must be interpreted in the viscosity sense, see [2], [3] and [4] for details.

If the Hamiltonian $H(p,x,u)$ is strictly convex in u for all $p,\ x$ then (1.4, 1.5) become

$$\frac{\partial \pi}{\partial x}(x) f(x,\kappa(x)) + l(x,\kappa(x)) = 0 \tag{1.9}$$

and

$$\frac{\partial \pi}{\partial x}(x) \frac{\partial f}{\partial u}(x,\kappa(x)) + \frac{\partial l}{\partial u}(x,\kappa(x)) = 0 \tag{1.10}$$

In the nonlinear case even if a smooth solution exists, the solution of the HJB PDE is quite difficult. A standard approach is to discretize the optimal control problem (1.1, 1.2) in space and time and solve the corresponding nonlinear program, see [10] and the appendix by Falcone in [2]. Other methods for solving the HJB

[1] Corresponding author. Research supported in part by AFOSR-49620-95-1-0409 and by NSF 9970998.

PDE are similar to those for conservation laws [12] and marching methods [14]. All these methods suffer from the "curse of dimensionality", the computation grows exponentially in the dimensions $n,\ m$ of the state and control $x,\ u$. There are problems of academic interest where $n \leq 2$ and $m = 1$, but in most real world problems n is at least 6 and m is at least 3. Since the number of dimensions is so substantial, one needs a method that uses only a few discretizations in each dimension, a higher order method. But the solutions of HJB PDEs have discontinuities and hence higher order methods may fail. We propose a higher order approach that locally approximates the solution by a smooth solution and then pieces these solutions together in appropriate fashion by computing in characteristic directions to mitigate the effect of discontinuities. The local regions are chosen to be large so as to mitigate the curse of dimensionality.

2 Al'brecht's Method: Power Series Expansion Around Zero

Al'brecht [1] solved the HJB PDE locally around zero by expanding the problem in a power series,

$$\begin{aligned} f(x,u) &= Ax + Bu + f^{[2]}(x,u) \\ &\quad + f^{[3]}(x,u) + \ldots & (2.11) \\ l(x,u) &= \frac{1}{2}\left(x'Qx + 2x'Su + u'Ru\right) \\ &\quad + l^{[3]}(x,u) + l^{[4]}(x,u) + \ldots & (2.12) \\ \pi(x) &= \frac{1}{2}x'Px + \pi^{[3]}(x) \\ &\quad + \pi^{[4]}(x) + \ldots & (2.13) \\ \kappa(x) &= Kx + \kappa^{[2]}(x) + \kappa^{[3]}(x) + \ldots & (2.14) \end{aligned}$$

where $\cdot^{[d]}$ denotes a homogeneous polynomial of degree d. He plugged these into the HJB PDE (1.9, 1.10) and equated terms of like degree to obtain a sequence of algebraic equations for the unknowns.

The first level is the pair of equations obtained by collecting the quadratic terms of (1.9) and the linear terms of (1.10). We denote the d^{th} level as the pair of equations garnered from the $[d+1]^{th}$ degree terms of (1.9) and the d^{th} degree terms of (1.10).

Al'brecht's Algebraic Equations

First Level:

$$\begin{aligned} 0 &= A'P + PA + Q - \\ &\quad (PB+S)R^{-1}(PB+S)' & (2.15) \\ K &= -R^{-1}(PB+S)' & (2.16) \end{aligned}$$

The quadratic terms of the HJB PDE reduce to the familiar Riccati equation (2.15) and the linear optimal feedback (2.16).

We assume $A,\ B$ is stabilizable and $Q,\ A$ is detectable then the Riccati equation has a unique positive definite solution P and the linear feedback locally exponentially stabilizes the closed loop system. Moreover the optimal quadratic cost is a local Lyapunov function for the closed loop system.

The d^{th} Level

Suppose we have solved through the $d-1^{th}$ level. It is convenient to incorporate this solution into the dynamics and the cost. Let $\kappa^{k]}(x) = Kx + \kappa^{[2]}(x) + \kappa^{[3]}(x) + \ldots + \kappa^{[k]}(x)$ and define

$$\begin{aligned} \bar{f}(x,u) &= f(x, \kappa^{d-1]}(x) + u) & (2.17) \\ \bar{l}(x,u) &= l(x, \kappa^{d-1]}(x) + u) & (2.18) \end{aligned}$$

These have power series expansions through terms of degree d and $d+1$ of the form

$$\begin{aligned} \bar{f}(x,u) &= (A+BK)x + Bu + \bar{f}^{[2]}(x,u) + \ldots \bar{f}^{[d]}(x,u) + \ldots \\ \bar{l}(x,u) &= \frac{1}{2}\left(x'Qx + 2x'SKx + x'K'RKx\right) \\ &\quad + x'Su + u'Ru + \\ &\quad \bar{l}^{[3]}(x,u) + \ldots + \bar{l}^{[d+1]}(x,u) + \ldots \end{aligned}$$

We plug these into the HJB PDE (1.9, 1.10) and find that they are satisfied through the $d-1$ level and don't involve u. Since $u = \kappa^{[d]}(x) + \ldots$, the d level equations are

$$\begin{aligned} 0 &= \frac{\partial \pi^{[d+1]}}{\partial x}(x)(A+BK)x & (2.19) \\ &\quad + \sum_{i=2}^{d-1} \frac{\partial \pi^{[d+2-i]}}{\partial x}(x)\bar{f}^{[i]}(x,0) + x'PBu \\ &\quad + \bar{l}^{[d+1]}(x,0) + x'Su + \frac{1}{2}x'K'Ru \end{aligned}$$

and

$$\begin{aligned} 0 &= \frac{\partial \pi^{[d+1]}}{\partial x}(x)B + \sum_{i=2}^{d} \frac{\partial \pi^{[d+2-i]}}{\partial x}\frac{\partial \bar{f}^{[i]}}{\partial u}(x,0) \\ &\quad + \frac{\partial \bar{l}^{[d+1]}}{\partial u}(x,0) + u'R & (2.20) \end{aligned}$$

Because of (2.16), u drops out of the first equation which becomes

$$\begin{aligned} 0 &= \frac{\partial \pi^{[d+1]}}{\partial x}(x)(A+BK)x & (2.21) \\ &\quad + \sum_{i=2}^{d-1} \frac{\partial \pi^{[d+2-i]}}{\partial x}(x)\bar{f}^{[i]}(x,0) \\ &\quad + \bar{l}^{[d+1]}(x,0). \end{aligned}$$

After this has been solved for $\pi^{[d+1]}(x)$, we can solve the second for $\kappa^{[d]}(x)$,

$$\begin{aligned}\kappa^{[d]}(x) &= -R^{-1}\left(\frac{\partial \pi^{[d+1]}}{\partial x}(x)B\right. \qquad (2.22)\\ &\left. + \sum_{i=2}^{d} \frac{\partial \pi^{[d+2-i]}}{\partial x}\frac{\partial \bar{f}^{[i]}}{\partial u}(x,0) + \frac{\partial \bar{l}^{[d+1]}}{\partial u}(x,0)\right)\end{aligned}$$

These equations admit a unique solution up to the smoothness of f and l since $A + BK$ has all its eigenvalues in the left half plane. If f and l are real analytic, the power series converges to the solution of the HJB PDE locally around $x = 0$ [11], [9]. The higher degree equations are linear, of Sylvester type, and hence are easily solvable. There is a MATLAB package to solve them to arbitrary degree [8].

The weakness of the Al'brecht approach is that the power series frequently does not converge quickly on a large neighborhood of 0. Increasing the degree d does not necessarily lead to a larger region where the power series accurately approximates the true solution. This is a familiar problem with polynomial approximations to functions based on Taylor series.

We take an alternative approach. Recall that the true optimal cost is a Lyapunov function for the dynamics under the true optimal feedback. Following Al'brecht we have an approximation of these. We compute the time derivative of the approximate cost following the true dynamics using the approximate feedback to determine a sublevel set of the approximate cost on which it is an acceptable Lyapunov function and the approximate feedback is stabilizing. We emphasize the stabilizing property of the control law rather than its optimality because usually optimality is only a tool to find a stabilizing feedback. Typically the goal is to find a control law to stabilize the system and the optimal control problem is formulated as a way of finding one.

Let $\pi^{d+1]}$ and $\kappa^{d]}$ denote the approximate cost and feedback to degrees $d+1$ and d respectively. We assume $l(x, u) \geq 0$. We find the largest sublevel set

$$\pi^{d+1]}(x) \leq c \qquad (2.23)$$

such that

$$\frac{\partial \pi^{d+1]}}{\partial x}(x) f(x, \kappa^{d]}(x)) \leq -(1-\epsilon) l(x, \kappa^{[d]}(x)) \qquad (2.24)$$

The parameter ϵ controls the rate of exponential stability of the closed loop system.

3 Power Series Expansions Along the Optimal Trajectories

We specialize to problems with dynamics that is affine in u and with a Lagrangian that is quadratic in u,

$$\begin{aligned} f(x,u) &= g_0(x) + g_1(x)u & (3.25)\\ l(x,u) &= l_0(x) + l_1(x)u + u' l_2(x) u & (3.26)\end{aligned}$$

where $l_2(x)$ is an invertible $m \times m$ matrix for all x.

Assume that we have solved the HJB PDE locally around 0 and have choosen a sublevel set of value c subject to the condition (2.23, 2.24). We generate an optimal trajectory emanating from the level set and construct a moving power series solution along the trajectory.

From a point x^0 on the level set $\pi^{d+1]}(x) = c$, we numerically compute backward in time the optimal trajectory $\xi(t)$ that satisfies $\xi(0) = x^0$. This trajectory and its corresponding optimal control $\mu(t)$ and costate $p(t)$ satisfy the Pontryagin Maximum Principle

$$\begin{aligned} \dot{\xi} &= \frac{\partial H}{\partial p}(p, \xi, \mu) & (3.27)\\ \dot{p} &= -\frac{\partial H}{\partial x}(p, \xi, \mu) & (3.28)\\ \mu &= \arg\min_u H(p, \xi, u) & (3.29)\end{aligned}$$

with the terminal conditions

$$\begin{aligned} \xi(0) &= x^0 & (3.30)\\ p(0) &= \frac{\partial \pi}{\partial x}(x^0) \end{aligned}$$

Since π is not known exactly we replace the latter by

$$p(0) = \frac{\partial \pi^{d+1]}}{\partial x}(x^0) \qquad (3.31)$$

We introduce some notation. Let $\alpha = (\alpha_1, \ldots, \alpha_n)$ be a multi-index of nonnegative integers and $|\alpha| = \sum_i \alpha_i$. Let $\beta = (\beta_1, \ldots, \beta_n)$. We say $\beta \leq \alpha$ if $\beta_i \leq \alpha_i$, $i = 1, \ldots, n$ and $\beta < \alpha$ if $\beta \leq \alpha$ and for at least one i, $\beta_i < \alpha_i$. Let $\mathbf{0} = (0, \ldots, 0)$.

Define the differential operator

$$D^\alpha = (\frac{\partial}{\partial x_1})^{\alpha_1} \ldots (\frac{\partial}{\partial x_n})^{\alpha_n},$$

the multifactorial

$$\alpha! = \alpha_1! \ldots \alpha_n!,$$

the monomial

$$x^\alpha = x_1^{\alpha_1} \ldots x_n^{\alpha_n},$$

and the coefficient

$$C(\alpha, \beta) = \begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix} \cdots \begin{pmatrix} \alpha_n \\ \beta_n \end{pmatrix}$$

where $\beta \le \alpha$.

We derive a system of ordinary differential equations for

$$D^\alpha \pi(\xi(t)), \qquad D^\alpha \kappa(\xi(t))$$

for $t \le 0$. We already know that if α is the i^{th} unit vector

$$\begin{aligned}
D^{\mathbf{0}}\pi(\xi(t)) &= \pi(\xi(t)) = \pi(x^0) + \int_t^0 l(\xi(\tau), \mu(\tau))\, d\tau \\
D^{\mathbf{0}}\kappa(\xi(t)) &= \mu(t) \\
D^\alpha \pi(\xi(t)) &= \frac{\partial \pi}{\partial x_i}(\xi(t)) = p(t).
\end{aligned}$$

Assume that we have derived ODEs for $D^\beta\pi(\xi(t))$ and algebraic equations for $D^\beta\kappa(\xi(t))$ for $\beta < \alpha$. We apply D^α to (1.9) to obtain

$$\begin{aligned}
0 &= \frac{\partial}{\partial x}(D^\alpha \pi(x)) f(x, \kappa(x)) \qquad (3.32)\\
&\quad + \sum_{\mathbf{0}<\beta\le\alpha} C(\alpha,\beta)(D^{\alpha-\beta}\frac{\partial \pi}{\partial x}(x)) D^\beta f(x,\kappa(x)) \\
&\quad + D^\alpha l(x, \kappa(x)).
\end{aligned}$$

This yields an ODE for $D^\alpha\pi(\xi(t))$ because

$$\frac{d}{dt} D^\alpha \pi(\xi(t)) = \left(\frac{\partial}{\partial x} D^\alpha \pi(\xi(t))\right) f(\xi(t), \kappa(\xi(t))) \qquad (3.33)$$

We apply D^α to (1.10) to obtain

$$\begin{aligned}
0 &= \sum_{\mathbf{0}\le\beta\le\alpha} C(\alpha,\beta)(D^{\alpha-\beta}\frac{\partial \pi}{\partial x}(x)) D^\beta g_1(x) \\
&\quad + D^\alpha l_1(x) + (D^\alpha\kappa(x))' l_2(x) \\
&\quad \sum_{\mathbf{0}<\beta\le\alpha} C(\alpha,\beta)(D^{\alpha-\beta}\kappa(x))' D^\beta l_2(x) \qquad (3.34)
\end{aligned}$$

Notice that this equation (3.34) only contains $D^\alpha\kappa(x)$ in one term multiplied by an invetible matrix so we can express $D^\alpha\kappa(x)$ as a function of $D^\beta\kappa(x)$ for $\beta < \alpha$ and $D^\gamma\pi(x)$. Notice also that the right side of the ODE (3.32) depends on terms of the form $D^\gamma\pi(\xi(t))$ for $\gamma > \alpha$ so the system is not triangular, we cannot solve term by term. The right side also depends on terms containing $D^\beta\kappa(\xi(t))$ for $\beta \le \alpha$ but these cause no problem because of (3.34).

We fix a level d and solve the above for all $|\alpha| \le d$ assuming that $D^\alpha\pi(\xi(t)) = 0$ for $|\alpha| > d$ to close the system of equations. The result is a system of coupled ODEs which cannot be solved term by term as in Al'brecht's method. They are solved backward in time from $t = 0$ with final conditions coming from the Al'brecht solution,

$$\begin{aligned}
D^\alpha\pi(\xi(0)) &= D^\alpha \pi^{d+1]}(x^0) \qquad (3.35)\\
D^\alpha\kappa(\xi(0)) &= D^\alpha \kappa^{d]}(x^0) \qquad (3.36)
\end{aligned}$$

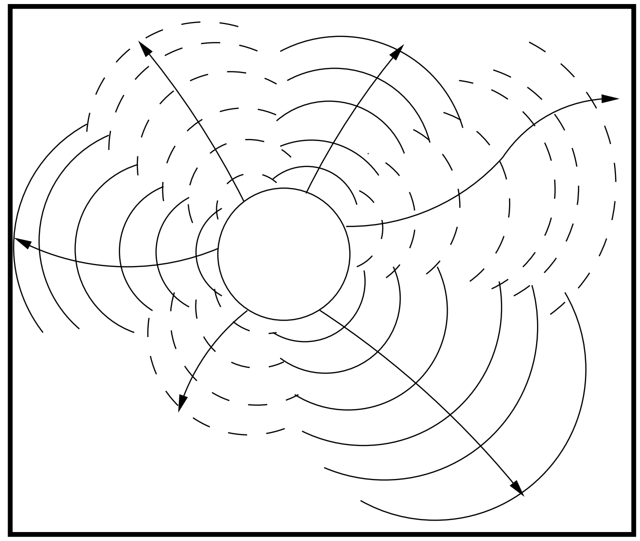

Figure 1: Solutions around the characteristic curves generated by the optimal trajectories from the level set $\pi^{d+1]}(x) \le c$

Although we cannot solve these ODE term by term as with Al'brecht method, they are explicit equations for the derivatives as opposed to the implicit linear equations (2.21).

In this way we obtain the approximations

$$\begin{aligned}
\pi(x) &\approx \sum_{|\alpha|\le d} \frac{1}{\alpha!} D^\alpha \pi(\xi(t))(x - \xi(t))^\alpha \qquad (3.37)\\
\kappa(x) &\approx \sum_{|\alpha|\le d} \frac{1}{\alpha!} D^\alpha \kappa(\xi(t))(x - \xi(t))^\alpha \qquad (3.38)
\end{aligned}$$

where x is close to $\xi(t)$. If $f_i(\xi(\tau), \kappa(\xi(\tau))) \ne 0,\ \tau \le 0$ then we can choose t such that $\xi_i(t) = x_i$.

The process is repeated at other points on the level set $\pi^{d+1]}(x) = c$ and the solutions fitted together to obtain the desired approximation to π and κ. See Fig. 1.

The boundaries between adjacent approximations are found by equating the values of π because the lower approximation is the optimal one. There is the problem of how to store the approximate solutions and their regions of validity. We propose transferring the approximate solutions to a rectangular grid of x space using splines of degree d similar to the method of Prager [13].

In most applications $d = 1$ or $d = 3$ so one would use piecewise linear approximations or cubic splines. The data would be transferred to the grid using a variation on the fast marching methods of Sethian [14].

4 Conclusion

We have sketched a numerical procedure for solving infinite horizon optimal control problems. The MATLAB code exists to solve the first stage of the procedure [8] and we are developing code for the second stage.

References

[1] E. G. Al'brecht, *On the optimal stabilization of nonlinear systems*, PMM-J. Appl. Math. Mech., 25:1254-1266, 1961.

[2] M. Bardi and I. Capuzzo-Dolcetta *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*, Birkhäuser, Boston, 1997.

[3] L. C. Evans, *Partial Differential Equations*. American Mathematical Society, Providence, 1998.

[4] W. H. Fleming and H. M. Soner, *Controlled Markov Processes and Viscosity Solutions*. Springer-Verlag, New York, 1992.

[5] A. J. Krener. *The construction of optimal linear and nonlinear regulators*, in A. Isidori and T.J. Tarn, editors, *Systems, Models and Feedback: Theory and Applications*, Birkhauser, Boston, 1992, 301–322.

[6] A. J. Krener. *Optimal model matching controllers for linear and nonlinear systems*, in M. Fliess, editor, *Nonlinear Control System Design 1992*, Pergamon Press, Oxford, 1993, 209–214.

[7] A. J. Krener. *Necessary and sufficient conditions for nonlinear worst case (H-infinity) control and estimation*, summary and electronic publication, J. Mathematical Systems, Estimation, and Control, 4:485-488, 1994, full manuscript in J. Mathematical Systems, Estimation, and Control, 7:81-106, 1997.

[8] A. J. Krener. *Nonlinear Systems Toolbox V. 1.0, 1997*, MATLAB based toolbox available by ftp from scad.utdallas.edu

[9] A. J. Krener. *The existence of optimal regulators*, Proc. of 1998 CDC, Tampa, FL, 3081–3086.

[10] H. J. Kushner and P. G. Dupuis, *Numerical Methods for Stochastic Control Problems in Continuous Time*, Springer-Verlag, New York, 1992.

[11] D. L. Lukes. *Optimal regulation of nonlinear dynamical systems*, SIAM J. Contr., 7:75–100, 1969.

[12] S. Osher and C. W. Shu. *High-order Essentially Nonoscillatory Schemes for Hamilton Jacobi Equations*, SIAM J. Numerical Analysis, 28:907-922, 1991.

[13] W. Prager. *Numerical Computation of the optimal feedback law for nonlinear infinite horizon control problems*, CALCOLO, 37:97-123.

[14] J. A. Sethian, *Level Set Methods and Fast Marching Methods*. Cambridge University Press, 1999.