

Hamilton-Jacobi-Bellman 方程式の可解性に関する研究

A Study on the Solvability of Hamilton-Jacobi-Bellman Equation

80122827 中山慶一 (Nakayama Keiichi) Supervisor 志水清孝 (Shimizu Kiyotaka)

1. はじめに

本論文ではクリティカルな (critical) 非線形制御システムの安定化制御について主に考える。非線形制御システムがクリティカルであるとは、その線形化システムが少なくとも1つ以上の不可制御な虚軸上の固有値をもち、それ以外の固有値はすべて可安定である場合を意味する。クリティカルな非線形制御システムに滑らかなフィードバック制御則を施すとき、その閉ループ系には中心多様体と呼ばれる局所不変多様体が存在する。本論文では中心多様体に収束する過渡状態を最適化するようなフィードバック制御則を提案しており、従来の制御則¹⁾と比較して中心多様体の吸引領域を拡大することが確認され、その有効性が示された。最適フィードバック制御則は Hamilton-Jacobi-Bellman 偏微分方程式 (以下、H-J 方程式) の解である値関数を用いて表現される。H-J 方程式の解析解を求めることは一般に困難である。本論文ではニューラルネットを用いた H-J 方程式の近似解法を提案する。Riccati 方程式の準正定解を利用することにより、原点における値関数の0次~2次の境界条件²⁾を、任意の結合重み行列としきい値に対して成立させている。提案手法によって近似解の値関数への収束性と学習効率を改善することができた。

2. 非線形最適制御問題と H-J 方程式

次のような時不変のアフィン非線形システムに対する非線形最適制御問題を考える。

$$\min_u \int_0^\infty \{q(x(t)) + u(t)^T R u(t)\} dt \quad (1a)$$

$$\text{subj. to } \dot{x} = f(x) + G(x)u, \quad x(0) = x_0 \quad (1b)$$

ここで $x(t) \in R^n$ は状態ベクトル, $u(t) \in R^r$ は制御入力ベクトルである。そして、次の仮定をおく。

[仮定 1] $f: R^n \rightarrow R^n$ は C^1 級で、かつ $f(0) = 0$ である。

[仮定 2] $q(0) = 0, q(x) \geq 0, R > 0$ とする。

問題 (1) は次の定常 H-J 方程式に帰着する。

$$\begin{aligned} 0 &= \min_u H(x, u, V_x(x)) \\ &= \min_u \{q(x) + u^T R u + V_x(x)(f(x) + G(x)u)\} \end{aligned} \quad (2)$$

ここで、 $H(x, u, V_x(x))$ は Hamilton 関数である。 $V(x)$ は値関数であり、次のように定義される。

$$\begin{aligned} V(x) &= \min_u \left\{ \int_t^\infty (q(x(t)) + u(t)^T R u(t)) dt \right. \\ &\quad \left. \left| \dot{x} = f(x) + G(x)u, x(t) = x \right. \right\} \end{aligned} \quad (3)$$

値関数は境界条件 $V(0) = 0$ を満たす。また、(2) を満たす u に対して次の最適性必要条件がなりたつ。

$$\nabla_u H(x, u, V_x(x)) = 2Ru + G(x)^T V_x(x)^T = 0 \quad (4)$$

これを u について解くと次のようになる。

$$u(x) = -(2R)^{-1} G(x)^T V_x(x)^T \quad (5)$$

これを (2) へ代入すると、次の H-J 方程式が得られる。

$$q(x) + V_x(x)f(x) - \frac{1}{4} V_x(x)G(x)R^{-1}G(x)^T V_x(x)^T = 0 \quad (6)$$

[定理 1]²⁾ 値関数 $V(x)$ の原点におけるヘッセ行列 $P \triangleq \nabla^2 V(0)$ は次の Riccati 方程式の準正定解である。

$$PA + A^T P - PB(2R)^{-1}B^T P + 2Q = 0 \quad (7)$$

ただし、 $A = f_x(0), B = G(0), 2Q = \nabla^2 q(0)$ とする。

3. ニューラルネットによる H-J 方程式の近似解法

次のような三層ニューラルネットを用いて近似を行なう。

$$z = W_1 x + \theta \quad (8a)$$

$$y = W_2(\sigma(z) - \sigma(\theta)) \quad (8b)$$

ここで、 $y \in R^n$ はニューラルネットの出力ベクトル、 $z \in R^q$ は内部状態、 $W_1 \in R^{q \times n}, W_2 \in R^{n \times q}$ は結合重み行列、 $\theta \in R^q$ はしきい値であり、 q は中間層ニューロンの数を表す。 $\sigma: R^q \rightarrow R^q$ はシグモイド関数であり、双曲関数 $\sigma_i(z_i) = \tanh z_i$ ($i = 1, \dots, q$) を用いる。ニューラルネットに対して定理 1 で述べた $\nabla^2 V(0) = P$ の関係を利用する。 $P \geq 0$ なので $P = S^T S$ を満たすような平方根 $S \in R^{n \times n}$ が存在する。 S を用いて $w \in R^n$ を次のように定義する。

$$w(x) = y + (S - y_x(0))x \quad (9)$$

ここで、 $y_x(0) = W_2 \nabla \sigma(\theta) W_1$ と計算される。そして、値関数 $V(x)$ を w を用いて次のように近似する。

$$V^N(x) = \frac{1}{2} w(x)^T w(x) \geq 0 \quad (10)$$

このとき、任意の W_1, W_2, θ に対して $V^N(0) = 0, \nabla V^N(0) = 0, \nabla^2 V^N(0) = P$ がなりたつ。H-J 方程式 (6) を満たす値関数のニューラルネットによる学習問題は次のようになる。

$$\min_{W_1, W_2, \theta} E[W_1, W_2, \theta] = \sum_{p=1}^P |e(x^p)|^2 \quad (11a)$$

$$\text{subj. to } (8a), (8b), (9), (10)$$

$$\begin{aligned} e(x^p) &= q(x^p) + V_x^N(x^p) f_0(x^p) \\ &\quad - \frac{1}{4} V_x^N(x^p) G(x^p) R^{-1} G(x^p)^T V_x^N(x^p)^T \end{aligned} \quad (11b)$$

ただし x^p は状態空間 R^n の部分集合 Ω を離散化して得られる集合 $\Delta \triangleq \{x^p \mid x^p \in \Omega, p = 1, 2, \dots, P\}$ の要素である。評価関数 (11a) の W_1, W_2, θ に関する勾配を計算すれば、学習是最急降下法により実行される。(5) に

$$V_x^N(x)^T = \{S^T + W_1^T (\nabla \sigma(z) - \nabla \sigma(\theta)) W_2^T\} w \quad (12)$$

を代入することにより最適フィードバック制御則は、学習の結果得られる W_1, W_2, θ を用いて次のように与えられる。

$$u_N(x) = -\frac{1}{2} R^{-1} G(x)^T \{S^T + W_1^T (\nabla \sigma(z) - \nabla \sigma(\theta)) W_2^T\} w \quad (13)$$

4. 中心多様体に収束する過渡状態の最適化

(1b) がクリティカルであるとき、

$$T^{-1}AT = \begin{bmatrix} A^\kappa & O \\ O & A^c \end{bmatrix}, \quad T^{-1}B = \begin{bmatrix} B^\kappa \\ B^c \end{bmatrix} \quad (14)$$

において、 $\{A^\kappa, B^\kappa\}$ は可安定であり、 A^c の固有値はすべて虚軸上に存在し、 $\{A^c, B^c\}$ は不可制御であるような正則行列 T が存在する。(1b) を $x = T \begin{bmatrix} y \\ z \end{bmatrix}$ ($y \in R^s, z \in R^c$) と座標変換したシステムを次のように表すことにする。

$$\dot{y} = f^s(y, z) + G^s(y, z)u \quad (15a)$$

$$\dot{z} = f^c(y, z) + G^c(y, z)u \quad (15b)$$

ここでは $B^c = G^c(0, 0) = O$ を仮定する。このとき、次の条件を満たすフィードバック制御則 $u = \alpha(y, z)$ を求める。

(P1) $\dot{y} = f^s(y, 0) + G^s(y, 0)\alpha(y, 0)$ を $y = 0$ まわりで線形化したシステムの平衡点が漸近安定である。

(P2) 閉ループ系における中心多様体上のダイナミクスの平衡点 $z = 0$ が漸近安定である。

[定理 2] (P1), (P2) を満たすような $u = \alpha(y, z)$ が存在するための必要十分条件は、(17) の平衡点が漸近安定となるような (16) の解 $\pi(z), c(z)$ が存在することである。

$$\frac{\partial \pi(z)}{\partial z} \{f^c(\pi(z), z) + G^c(\pi(z), z)c(z)\} = f^s(\pi(z), z) + G^s(\pi(z), z)c(z) \quad (16)$$

$$\pi(0) = 0, \quad c(0) = 0$$

$$\dot{z} = f^c(\pi(z), z) + G^c(\pi(z), z)c(z) \quad (17)$$

ここでは (17) が漸近安定となるような (16) の解 $\pi(z), c(z)$ が求まっているとする。(P1), (P2) を満たすような $u = \alpha(y, z)$ として、中心多様体に収束する過渡状態を最適にするものを考える。制御入力 $v \in R^r$ を次のように定義して入力変換する。

$$v = u - c(z) \quad (18)$$

そして、次のような最適制御問題を考える。

$$\min_v \int_0^\infty \{(y - \pi(z))^T Q_\xi (y - \pi(z)) + v^T R v\} dt \quad (19a)$$

$$\text{subj. to } \dot{y} = f^s(y, z) + G^s(y, z)\{v + c(z)\}, \quad y(0) = y_0 \quad (19b)$$

$$\dot{z} = f^c(y, z) + G^c(y, z)\{v + c(z)\}, \quad z(0) = z_0 \quad (19c)$$

ただし $Q_\xi > 0, R > 0$ とする。この問題の値関数 $V(y, z)$ の原点におけるヘッセ行列 $\nabla^2 V(0, 0)$ は、次の Riccati 方程式

$$PA^\kappa + (A^\kappa)^T P - PB^\kappa(2R)^{-1}(B^\kappa)^T P + 2Q_\xi = O \quad (20)$$

の安定化解 P^- を用いて、次のように求められる。

$$\nabla^2 V(0, 0) = \begin{bmatrix} P^- & -P^- \nabla \pi(0)^T \\ -\nabla \pi(0) P^- & \nabla \pi(0) P^- \nabla \pi(0)^T \end{bmatrix} \quad (21)$$

上式を用いると、3節の近似解法を適用することができる。(5) より最適フィードバック制御則は次のようになる。

$$v(y, z) = -(2R)^{-1} \begin{bmatrix} G^s(y, z) \\ G^c(y, z) \end{bmatrix}^T \nabla V(y, z) \quad (22)$$

上式と (18) より、求めるフィードバック制御則は

$$u(y, z) = -(2R)^{-1} \begin{bmatrix} G^s(y, z) \\ G^c(y, z) \end{bmatrix}^T \nabla V(y, z) + c(z) \quad (23)$$

となる。上式が (P1), (P2) を満たすことは容易に確認できる。従来の制御則¹⁾ は、(23) の右辺第 1 項を $(y - \pi(z), z)$ に関して線形化したものであり、次のようになる。

$$u(y, z) = -(2R)^{-1}(B^\kappa)^T P^- \{y - \pi(z)\} + c(z) \quad (24)$$

5. シミュレーション結果

制御対象として次のような非線形システムを考える。

$$\dot{y} = -y + (1 - y^2)u, \quad y(0) = y_0 \quad (25a)$$

$$\dot{z} = y^3, \quad z(0) = z_0 \quad (25b)$$

(17) が漸近安定となるような (16) の解の 1 つには、 $\pi(z) = -7z, c(z) = -7z$ が 2 次の近似解として考えられる。(19a) において $Q_\xi = 1, R = 1$ と設定し、次のような最適制御問題を考える。

$$\min_v \int_0^\infty \{(y + 7z)^2 + v^2\} dt \quad (26a)$$

$$\text{subj. to } \dot{y} = -y + (1 - y^2)(v - 7z), \quad y(0) = y_0 \quad (26b)$$

$$\dot{z} = y^3, \quad z(0) = z_0 \quad (26c)$$

この問題の値関数 $V(y, z)$ の原点におけるヘッセ行列は、(21) より次のように計算される。

$$\nabla^2 V(0, 0) = \begin{bmatrix} 0.82843 & 5.7990 \\ 5.7990 & 40.593 \end{bmatrix} \quad (27)$$

最適フィードバック制御則は (23) より次のように計算される。

$$u(y, z) = -0.5(1 - y^2) \frac{\partial V}{\partial y}(y, z) - 7z \quad (28)$$

$(y_0, z_0) = (1.1, 1.5)$ のときの制御結果を下に示す。

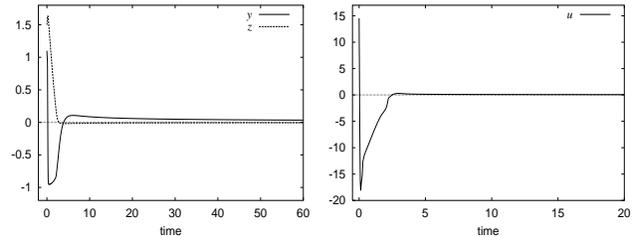


Fig.1 最適フィードバック制御，左： y, z ，右： u

6. 結論

従来の制御則を用いると、Fig.1 の初期状態から出発する解軌道は発散する。したがって (28) によって中心多様体の吸引領域を拡大することができた。また、クリティカルな非線形制御システムに対して H-J 方程式の安定化解を近似することができた。なお、4節の議論は出力レギュレーション問題に対しても有効であることが確認されている。

参考文献

- 1) A. Isidori, C. I. Byrnes: Output Regulation of Nonlinear Systems, IEEE Transactions on Automatic Control, Vol.35, No.2, 131/140 (1990)
- 2) A. van der Schaft: L_2 Gain and Passivity Techniques in Nonlinear Control, Lecture Notes in Control and Information Science 218, Springer-Verlag (1996)