

ĐIỀU KHIỂN TỐI ƯU TRỰC TUYẾN CHO CÁC HỆ PHI TUYẾN LIÊN TỤC

Vũ Văn Tú

Khoa Điện - Cơ

Email: tuvv@dhhp.edu.vn

Nguyễn Thị Thu Hiền

Khoa Điện-Cơ

Email: hienntt85@dhhp.edu.vn

Ngày nhận bài: 18/3/2019

Ngày PB đánh giá: 27/4/2019

Ngày duyệt đăng: 29/4/2019

TÓM TẮT

Bài báo này trình bày vấn đề của luật điều khiển tối ưu thích nghi cho các hệ phi tuyến với nhiều đầu vào và các tham số chưa biết. Mục tiêu chính là tìm ra luật điều khiển tối ưu thích nghi dựa trên phương pháp quy hoạch động thích nghi (Adaptive Dynamic Programming (ADP)), trong đó qui hoạch động thích nghi online (Online Adaptive Dynamic Programming (OADP)) là giải thuật được phân tích và thiết kế và nó có thể ổn định hệ thống vòng kín. Bên cạnh đó, các đặc tính hội tụ của thuật toán đề xuất được chỉ ra. Các phân tích lý thuyết và kết quả mô phỏng chứng minh sự hiệu quả của thuật toán đề xuất.

Từ khóa: Xấp xỉ quy hoạch động, điều khiển tối ưu thích nghi, qui hoạch động thích nghi online.

ONLINE OPTIMAL CONTROL OF CONTINUOUS-TIME AFFINE NONLINEAR SYSTEMS

ABSTRACT

This paper presents the problem of adaptive optimal control law for nonlinear systems with input disturbance and unknown parameters. The main objective is to find an adaptive optimal control law based on the adaptive dynamic programming (ADP) method, in which Online Adaptive Dynamic Programming (OADP) is an analyzed and designed algorithm and it is able to stabilize the closed-loop system. Besides, the convergence properties of proposed algorithm is pointed out. The theoretical analysis and simulation results demonstrate the performance of the proposed algorithm.

Keywords: Adaptive dynamic programming (ADP), adaptive optimal control law, Online Adaptive Dynamic Programming (OADP).

1. ĐẶT VẤN ĐỀ

Vấn đề của điều khiển tối ưu cho những hệ phi tuyến động liên tục được nghiên cứu sâu trong vài thập kỷ. Một thách thức cốt lõi của vấn đề điều khiển tối ưu phi tuyến là nó thường rơi vào giải quyết phương trình toán HJB. Người ta biết rằng phương trình HJB thực ra là một phương trình vi phân từng phần, rất khó khăn hoặc không thể nào giải quyết được bởi phương pháp phân tích. Để giải quyết vấn đề này, chúng ta phát triển một chương trình xấp xỉ điều khiển tối ưu cho những hệ phi tuyến liên tục affine sử dụng phương pháp thích nghi quy hoạch động.

2. NỘI DUNG NGHIÊN CỨU

2.1. Hệ chuyển mạch phi tuyến có nhiều đầu vào

Xét hệ chuyển mạch có nhiều đầu vào được xác định bởi phương trình sau:

$$\dot{x} = f_i(x) + g_i(x)(u + \Delta(x, t)) \quad (1)$$

Trong đó $x(t) \in \Omega_x \subset R^n$ là vector trạng thái và đo được, $u(t) \in \Omega_u \subset R^m$ là tín hiệu điều khiển. Hàm $\sigma: [0, +\infty) \rightarrow \{1, 2, \dots, M\}$ là hàm chọn chế độ của hệ, đã biết theo thời gian, liên tục trên từng khoảng, và M là số lượng hệ con. $f_i(x)$ là các vector hàm chưa biết và thỏa mãn $f_i(0) = 0$. $g_i(x)$ là các vector hàm đã biết và thỏa mãn: $G_{\min} \leq \|g_i(x)\| \leq G_{\max}$. $\Delta(x, t)$ là nhiễu, bất định hệ thống không xác định được.

Giả sử 1: Tồn tại hàm đã biết thỏa mãn:

Với hệ (1), ta xét hàm mục tiêu sau:

$$J(x(t), u(t)) = \int_t^\infty r(x(\tau), u(\tau)) d\tau \quad (2)$$

Trong đó $r(x(\tau), u(\tau)) = x^T Q x + u^T R u$ và các ma trận $Q = Q^T > 0$ $R = R^T > 0$

Mục tiêu là tìm ra bộ điều khiển u và xác định được hàm chặn trên $K(u)$ thỏa mãn $K(u) \leq M$, để hệ (1) là ổn định bền và hàm mục tiêu (2) phải thỏa mãn:

$$J(x, u) \leq K(u)$$

Định nghĩa: Hàm $K(u)$ là chặn trên của hàm mục tiêu (1) nên nó có tên là hàm đảm bảo chi phí. Hơn thế nữa, tín hiệu điều khiển u^* với $u^* = \arg \min_{u \in \Omega_u} K(u)$ sẽ được gọi là luật điều khiển tối ưu chi phí. (Theo nguyên lý tối ưu Bellman)

Xét hệ không nhiễu sau:

$$\dot{x} = f_i(x) + g_i(x)u \quad (3)$$

Và xét hàm mục tiêu của (3) được định nghĩa như sau [1]:

$$J_1(x(t), u(t)) = \int_t^\infty [r(x(\tau), u(\tau)) + \lambda(\rho(x))^2] d\tau \quad (4)$$

Theo vavoudakis, K.G and Lewis, F.L (2010), hàm mục tiêu của hệ thống khác mô hình toán của hệ phi tuyến. Ở phần dưới tác giả chứng minh $J_1(x(t), u(t))$ là hàm mục tiêu của hệ phi tuyến có nhiễu (1) và từ đó chứng minh u^* là tín hiệu điều khiển của hệ phi tuyến có nhiễu (1).

Ta chứng minh rằng $J_1(x(t), u(t))$ với $\lambda \geq \|R\|$ sẽ là một trong những hàm đảm bảo chi phí của hệ (1).

Từ việc định nghĩa $J(x(t), u(t))$ trong (2) và $J_1(x(t), u(t))$ trong (4), ta có thể dễ dàng thấy ngay:

$$J(x(t), u(t)) \leq J_1(x(t), u(t)) \quad (5)$$

Ta thấy rằng trong biểu thức của $J_1(x(t), u(t))$ có thêm thành phần $\lambda(\rho(x))^2$ là một hàm dương ($\lambda \geq \|R\| > 0 \Rightarrow \lambda(\rho(x))^2 > 0$) do đó

$$J(x(t), u(t)) = \int_t^\infty (r(x(\tau), u(\tau))) d\tau \leq J_1(x(t), u(t)) = \int_t^\infty [r(x(\tau), u(\tau)) + \lambda(\rho(x))^2] d\tau$$

Đặt $V^*(t) = \min_{u \in \Omega_u} J_1(x(t), u(t))$, ta có:

$$V^*(t) = \min_{u \in \Omega_u} \int_t^\infty [r(x(\tau), u(\tau)) + \lambda(\rho(x))^2] d\tau \quad (6)$$

$$V^*(t) = \min_{u \in \Omega_u} \int_t^{t+\Delta t} [r(x(\tau), u(\tau)) + \lambda(\rho(x))^2] d\tau \quad (7)$$

$$+ \min_{u \in \Omega_u} \int_{t+\Delta t}^\infty [r(x(\tau), u(\tau)) + \lambda(\rho(x))^2] d\tau$$

$$V^*(t) = \min_{u \in \Omega_u} \int_t^{t+\Delta t} [r(x(\tau), u(\tau)) + \lambda(\rho(x))^2] d\tau + V^*(t + \Delta t) \quad (8)$$

$$\min_{u \in \Omega_u} \left\{ \frac{1}{\Delta t} \int_t^{t+\Delta t} [r(x(\tau), u(\tau)) + \lambda(\rho(x))^2] d\tau + \frac{V^*(t + \Delta t) - V^*(t)}{\Delta t} \right\} = 0 \quad (9)$$

Theo định nghĩa của đạo hàm của hàm số và theo nguyên lý Bellman thì hàm mục tiêu là giá trị nhỏ nhất của hàm chi phí và để tối ưu thì giá trị nhỏ nhất của hàm mục tiêu phải bằng không.

Cho $\Delta t \rightarrow 0^+$ ta có:

$$\min_{u \in \Omega_u} \left\{ \frac{1}{\Delta t} \int_t^{t+\Delta t} [r(x(\tau), u(\tau)) + \lambda(\rho(x))^2] d\tau + \frac{V^*(t + \Delta t) - V^*(t)}{\Delta t} \right\} = 0 \quad (10)$$

$$\min_{u \in \Omega_u} [r(x(t), u(t)) + \lambda \rho^2(x) + (\nabla V^*)^T (f_1(x) + g_1(x)u)] = 0 \quad (11)$$

Kí hiệu: $\nabla \bar{h} = \frac{\partial \bar{h}}{\partial x}$ là đúng cho mọi hàm $\bar{h}$ trong bài báo này.

$$\text{Đặt: } H(x, u, V^*) = r(x(t), u(t)) + \lambda \rho^2(x) + (\nabla V^*)^T (f_1(x) + g_1(x)u) \quad (12)$$

$$\text{Ta có thể thấy ngay: } H(x, u^*, V^*) = \min_{u \in \Omega_u} H(x, u, V^*) = 0 \quad (13)$$

$$\left. \frac{\partial H(x, u, V^*)}{\partial u} \right|_{u=u^*} = 0 \Rightarrow u^* = -\frac{1}{2} R^{-1} (g_i(x))^T \nabla V^* \quad (14)$$

Theo lý thuyết quy hoạch động lúc này u^* là tín hiệu điều khiển của hệ phi tuyến có nhiễu đầu vào.

Định lý 1 [1],[2]: Bằng việc sử dụng tín hiệu điều khiển $u(x) = -\frac{1}{2} R^{-1} (g_i(x))^T \nabla V$ và chọn

$$V(t) = \int_t^\infty [r(x(\tau), u(\tau)) + \lambda (\rho(x))^2] d\tau$$

là hàm Lyapunov, trong đó $\lambda \geq \|R\|$ sẽ đảm bảo hệ (1) ổn định.

Chứng minh:

Đạo hàm của V được xác định theo công thức:

$$\dot{V}(t) = (\nabla V)^T (f_i(x) + g_i(x)(u + \Delta(x, t)))$$

Bằng việc sử dụng $u(x) = -\frac{1}{2} R^{-1} (g_i(x))^T \nabla V^*$, ta có thể suy ra:

$$\begin{aligned} \dot{V}(t) &= -r(x(t), u(t)) - \lambda \rho^2(x) + (g_i(x))^T \nabla V^* \Delta(x, t) \\ &= -x^T Qx - u^T Ru - \lambda \rho^2(x) - 2u^T R \Delta(x, t) \end{aligned} \quad (15)$$

$$= -x^T Qx - \lambda \rho^2(x) - (u^T + \Delta(x, t)^T) R (u + \Delta(x, t)) + \Delta(x, t)^T R \Delta(x, t) \quad (16)$$

$$= -x^T Qx - (\lambda \rho^2(x) - \Delta(x, t)^T R \Delta(x, t) - (u + \Delta(x, t))^T R (u + \Delta(x, t))) \quad (17)$$

Với giả sử 1 và $\lambda \geq \|R\|$ ta có:

$$\begin{aligned} \dot{V}(t) &\leq -x^T Qx - (\lambda \rho^2(x) - \|R\| \|\Delta(x, t)\|^2) \leq -x^T Qx - (\lambda - \|R\|) \rho^2(x) \\ \dot{V}(t) &\leq -x^T Qx \end{aligned} \quad (18)$$

Do đó, hệ (1) là ổn định bền.

Rất khó để giải trực tiếp phương trình HJB (13), do đó, hàm mục tiêu V^* của hệ (3) có thể được xấp xỉ bằng mạng RBF trên tập Ω_x :

$$V^* = w^T \sigma(x) + \varepsilon(x) \quad (19)$$

Theo lý thuyết toán học

Xét hàm khả vi liên tục $f(x)$ được xấp xỉ hàm bằng vector hàm nơron (NN) truyền thẳng một lớp theo công thức:

$$f(x) = W^T \varphi(x) + \varepsilon(x)$$

Trong đó W là trọng số NN lý tưởng, $\varphi(x): \mathbb{R}^n \rightarrow \mathbb{R}^n$; $\varphi(x)$ là vector hàm tác động, n_h là số lượng nút ở lớp ẩn và $\varepsilon(x)$ là sai số xấp xỉ NN. Khi đó, nếu hệ $\{\varphi_i(x): i = \overline{(1, \infty)}\}$ là vector cơ sở độc lập hoàn toàn thì $f(x)$ và đạo hàm của nó:

$\frac{\partial f(x)}{\partial x} = \left(\frac{\partial \varphi(x)}{\partial x} \right)^T W + \frac{\partial \varepsilon(x)}{\partial x}$ sẽ được xấp xỉ đều. Ngoài ra, sai số ε và đạo hàm của nó sẽ hội tụ đều về 0 khi $n_h \rightarrow \infty$. Hơn nữa, nếu n_h hữu hạn thì ε và $\frac{\partial \varepsilon}{\partial x}$ sẽ bị chặn bởi các hằng số dương trong tập đóng.

Còn việc luyện mạng nơ ron và tìm ra thông số tác giả lập trình trên phần mềm Matlab.

Trong đó $\sigma(x): \mathbb{R}^n \rightarrow \mathbb{R}^N$; $\sigma(0) = 0$ là (NN), N là số neural ẩn, và $\varepsilon(x)$ là sai số xấp xỉ do NN, $w \in \mathbb{R}^N$ là vector hàm trọng lượng. $\sigma(x)$ có thể được chọn sao cho với $N \rightarrow \infty$, ta có: $\varepsilon(x) \rightarrow 0$ và $\nabla \varepsilon(x) \rightarrow 0$, do đó việc lựa chọn N hữu hạn, ta có thể giả sử:

$$\text{Giả sử 2: } \|\varepsilon(x)\| \leq \varepsilon_{\max}; \|\nabla \varepsilon(x)\| \leq \nabla \varepsilon_{\max}; \nabla \sigma_{\min} \leq \|\nabla \sigma(x)\| \leq \nabla \sigma_{\max}; \|w\| \leq w_{\max}$$

Bằng việc kết hợp (13), (14) ta có thể suy ra:

$$H(x, u^*, V^*) = r(x(t), u^*(t)) + \lambda \rho^2(x) + (\nabla V^*)^T (f_i(x) + g_i(x)u^*) = 0 \quad (20)$$

$$= x^T Qx + \lambda \rho^2(x) + (\nabla V^*)^T f_i(x) - \frac{1}{4} (\nabla V^*)^T g_i(x) R^{-1} g_i(x)^T (\nabla V^*) = 0 \quad (21)$$

$$\text{Phương trình (19) có thể dẫn tới: } \nabla V^* = (\nabla \sigma(x))^T w + \nabla \varepsilon(x) \quad (22)$$

Sử dụng mạng NN vào phương trình HJB trên ta có:

$$e_{NN} = x^T Qx + \lambda \rho^2(x) + (\nabla \sigma(x)^T w)^T f_i(x) \quad (23)$$

$$\begin{aligned} & - \frac{1}{4} (\nabla \sigma(x)^T w)^T g_i(x) R^{-1} g_i(x)^T (\nabla \sigma(x)^T w) \\ e_{NN} &= x^T Qx + \lambda \rho^2(x) + w^T \nabla \sigma(x) f_i(x) \\ & - \frac{1}{4} w^T \nabla \sigma(x) g_i(x) R^{-1} g_i(x)^T \nabla \sigma(x)^T w \end{aligned} \quad (24)$$

Sai số của hệ NN gây ra có thể được viết lại:

$$e_{NN} = -\nabla \varepsilon(x)^T (f_i(x) + g_i(x)u^*) + \frac{1}{4} \nabla \varepsilon(x)^T g_i(x) R^{-1} g_i(x)^T \nabla \varepsilon(x) \quad (25)$$

Với $N \rightarrow \infty$, e_{NN} có thể hội tụ về không. Nên với số N hữu hạn, e_{NN} sẽ nằm trong một tập compact, do đó ta có thể giả sử $e_{NN} \leq e_{\max}$.

Sử dụng mạng NN để có những xấp xỉ sau:

$$\hat{V} = \hat{w}^T \sigma(x) = \sigma(x)^T \hat{w}; \hat{u} = -\frac{1}{2} R^{-1} (g_i(x))^T \nabla \hat{V} \quad (26)$$

Khi đó, sai số xấp xỉ sẽ được sinh ra [3], [4]:

$$e_{HJB} = r(x(t), \hat{u}(t)) + \lambda \rho^2(x) + \left(\nabla \sigma(x)^T \hat{w} \right)^T (f_i(x) + g_i(x) \hat{u}) \quad (27)$$

$$e_{HJB} = x^T Q x + \lambda \rho^2(x) + \hat{w}^T \nabla \sigma(x) f_i(x) - \frac{1}{4} \hat{w}^T \nabla \sigma(x) g_i(x) R^{-1} g_i(x)^T \nabla \sigma(x)^T \hat{w} \quad (28)$$

$$\text{Vector trọng lượng được cập nhật: } \dot{\hat{w}} = -\alpha \frac{\partial E}{\partial \hat{w}} \quad (29)$$

$$\text{với } E = \frac{1}{2} e_{HJB}^T e_{HJB}$$

Định lý 2 [4]: Bằng cách sử dụng luật điều khiển (26) và vector trọng lượng được học theo mạng NN và cập nhật theo luật (29), thì sai lệch vector trọng lượng $\tilde{w} = w - \hat{w}$ và trạng thái $x(t)$ của hệ kín sẽ là ổn định UUB.

Chứng minh:

$$\text{Đặt: } \tilde{w} = w - \hat{w} \Rightarrow \dot{\tilde{w}} = -\dot{\hat{w}}$$

Xét hàm Lyapunov sau:

$$V(t) = V_1(t) + V_2(t); V_1(t) = \frac{1}{2\alpha} \tilde{w}(t)^T \tilde{w}(t); V_2(t) = V^*$$

Tác giả chọn theo bài báo đã ghi trong danh mục tài liệu tham khảo.

Đạo hàm hàm $V_1(t)$, ta thu được:

$$\dot{V}_1(t) = \frac{1}{\alpha} \tilde{w}(t)^T \dot{\tilde{w}}(t) = -\frac{1}{\alpha} \tilde{w}(t)^T \dot{\hat{w}}(t) = \tilde{w}(t)^T \frac{\partial E}{\partial \hat{w}}$$

$$\dot{V}_1 = \tilde{w}^T e_{HJB} \nabla \sigma(x) (f_i(x) + g_i(x) \hat{u})$$

Kết hợp (14) và (26) ta có:

$$\begin{aligned} \hat{u} - u^* &= -\frac{1}{2} R^{-1} (g_i(x))^T (\nabla \hat{V} - \nabla V^*) \\ &= -\frac{1}{2} R^{-1} (g_i(x))^T (\nabla \sigma(x)^T \hat{w} - (\nabla \sigma(x))^T w - \nabla \varepsilon(x)) \\ &= \frac{1}{2} R^{-1} (g_i(x))^T ((\nabla \sigma(x))^T \tilde{w} + \nabla \varepsilon(x)) \end{aligned}$$

Mặt khác ta có:

$$\begin{aligned} \nabla \sigma(x) (f_i(x) + g_i(x) \hat{u}) &= \nabla \sigma(x) (f_i(x) + g_i(x) u^*) + \nabla \sigma(x) g_i(x) (\hat{u} - u^*) \\ &= \nabla \sigma(x) \left(f_i(x) + g_i(x) u^* + \frac{1}{2} g_i(x) R^{-1} g_i(x)^T (\nabla \sigma(x)^T \tilde{w} + \nabla \varepsilon(x)) \right) \end{aligned}$$

Từ (24) và (28) suy ra:

$$e_{HJB} - e_{NN} = -\left(\nabla \sigma(x)^T \tilde{w}\right)^T f_i(x) \quad (29)$$

$$+ \frac{1}{4} \left(\begin{aligned} & \left((w^T - \hat{w}^T) \nabla \sigma(x) g_i(x) R^{-1} g_i(x)^T \nabla \sigma(x)^T w \right) \\ & + \hat{w}^T \nabla \sigma(x) g_i(x) R^{-1} g_i(x)^T \nabla \sigma(x)^T (w - \hat{w}) \end{aligned} \right) \quad (30)$$

$$e_{HJB} - e_{NN} = \left(\left(\nabla \sigma(x)^T \hat{w} \right)^T f_i(x) - \left(\nabla \sigma(x)^T w \right)^T f_i(x) \right)$$

$$+ \frac{1}{4} \left(\begin{aligned} & w^T \nabla \sigma(x) g_i(x) R^{-1} g_i(x)^T \nabla \sigma(x)^T w \\ & - \hat{w}^T \nabla \sigma(x) g_i(x) R^{-1} g_i(x)^T \nabla \sigma(x)^T \hat{w} \end{aligned} \right)$$

$$+ \frac{1}{2} \left(\nabla \sigma(x)^T \hat{w} \right)^T g_i(x) R^{-1} \left(g_i(x) \right)^T \left(\left(\nabla \sigma(x) \right)^T \tilde{w} + \nabla \varepsilon(x) \right)$$

$$e_{HJB} - e_{NN} = -\tilde{w}^T \nabla \sigma(x) f_i(x) \quad (31)$$

$$+ \frac{1}{4} \left(\begin{aligned} & \tilde{w}^T \nabla \sigma(x) g_i(x) R^{-1} g_i(x)^T \nabla \sigma(x)^T w \\ & + \hat{w}^T \nabla \sigma(x) g_i(x) R^{-1} g_i(x)^T \nabla \sigma(x)^T \tilde{w} \end{aligned} \right)$$

$$e_{HJB} - e_{NN} = -\tilde{w}^T \nabla \sigma(x) f_i(x) \quad (32)$$

$$+ \frac{1}{2} \tilde{w}^T \nabla \sigma(x) g_i(x) R^{-1} g_i(x)^T \nabla \sigma(x)^T w$$

$$- \frac{1}{4} \tilde{w}^T \nabla \sigma(x) g_i(x) R^{-1} g_i(x)^T \nabla \sigma(x)^T \tilde{w}$$

Vì $u^* = -\frac{1}{2} R^{-1} \left(g_i(x) \right)^T \left(\left(\nabla \sigma(x) \right)^T w + \nabla \varepsilon(x) \right)$ nên ta có:

$$e_{HJB} - e_{NN} = -\tilde{w}^T \nabla \sigma(x) f_i(x) + \tilde{w}^T \nabla \sigma(x) g_i(x) \left(-u^* - \frac{1}{2} R^{-1} g_i(x)^T \nabla \varepsilon(x) \right) \quad (33)$$

$$- \frac{1}{4} \tilde{w}^T \nabla \sigma(x) g_i(x) R^{-1} g_i(x)^T \nabla \sigma(x)^T \tilde{w}$$

$$e_{HJB} = e_{NN} - \tilde{w}^T \nabla \sigma(x) \left(f_i(x) + g_i u^* \right) - \frac{1}{2} \tilde{w}^T \nabla \sigma(x) g_i(x) R^{-1} g_i(x)^T \nabla \varepsilon(x) \quad (34)$$

$$- \frac{1}{4} \tilde{w}^T \nabla \sigma(x) g_i(x) R^{-1} g_i(x)^T \nabla \sigma(x)^T \tilde{w}$$

Giả sử 3: $\|f_i(x) + g_i(x) u^*\| \leq \mu_{\max}$

Kí hiệu: $\mu_i = f_i(x) + g_i(x) u^*$; $G_i = g_i(x) R^{-1} g_i(x)^T$; $\nabla \sigma = \nabla \sigma(x)$; $\nabla \varepsilon = \nabla \varepsilon(x)$

Từ phương trình (34) ta có:

$$\dot{V}_1(t) = -\tilde{w}^T \begin{pmatrix} -e_{NN} + \tilde{w}^T \nabla \sigma \mu_i \\ + \frac{1}{2} \tilde{w}^T \nabla \sigma G_i \nabla \varepsilon \\ + \frac{1}{4} \tilde{w}^T \nabla \sigma G_i \nabla \sigma^T \tilde{w} \end{pmatrix} \nabla \sigma(x) \begin{pmatrix} \mu_i + \\ \frac{1}{2} G_i (\nabla \sigma^T \tilde{w} + \nabla \varepsilon) \end{pmatrix} \quad (35)$$

$$\dot{V}_1(t) = -\frac{1}{8} (\tilde{w}^T \nabla \sigma G_i \nabla \sigma^T \tilde{w})^2 \quad (36)$$

$$- \left[\frac{1}{2} (\tilde{w}^T \nabla \sigma \mu_i) + \frac{1}{4} (\tilde{w}^T \nabla \sigma G_i \nabla \varepsilon) + \frac{1}{4} (\tilde{w}^T \nabla \sigma \mu_i) + \frac{1}{2} e_{NN} \right] (\tilde{w}^T \nabla \sigma G_i \nabla \sigma^T \tilde{w})$$

$$- (\tilde{w}^T \nabla \sigma \mu_i)^2 + \left[e_{NN} - \frac{1}{2} (\tilde{w}^T \nabla \sigma G_i \nabla \varepsilon) \right] (\tilde{w}^T \nabla \sigma \mu_i)$$

$$\dot{V}_1(t) = -\frac{1}{8} (\tilde{w}^T \nabla \sigma G_i \nabla \sigma^T \tilde{w})^2 - \left[\frac{3}{4} (\tilde{w}^T \nabla \sigma \mu_i) + \frac{1}{4} (\tilde{w}^T \nabla \sigma G_i \nabla \varepsilon) + \frac{1}{2} e_{NN} \right] (\tilde{w}^T \nabla \sigma G_i \nabla \sigma^T \tilde{w}) \quad (37)$$

$$- (\tilde{w}^T \nabla \sigma \mu_i)^2 + \left[e_{NN} - \frac{1}{2} (\tilde{w}^T \nabla \sigma G_i \nabla \varepsilon) \right] (\tilde{w}^T \nabla \sigma \mu_i)$$

Đặt:

$$A = \tilde{w}^T \nabla \sigma G_i \nabla \sigma^T \tilde{w}; B = \frac{3}{4} (\tilde{w}^T \nabla \sigma \mu_i) + \frac{1}{4} (\tilde{w}^T \nabla \sigma G_i \nabla \varepsilon) + \frac{1}{2} e_{NN};$$

$$C = \tilde{w}^T \nabla \sigma \mu_i; D = e_{NN} - \frac{1}{2} (\tilde{w}^T \nabla \sigma G_i \nabla \varepsilon)$$

Dễ dàng nhận thấy:

$$A + 4B \geq \|\tilde{w}\|^2 (G_{\min})^2 \lambda_{\min}(R^{-1}) (\nabla \sigma_{\min})^2 \quad (38)$$

$$- \|\tilde{w}\| \left(3 \nabla \sigma_{\max} \mu_{\max} + \nabla \sigma_{\max} (G_{\max})^2 \lambda_{\max}(R^{-1}) \nabla \varepsilon_{\max} \right) - 2e_{\max} \quad (39)$$

$$\dot{V}_1(t) = -\frac{1}{8} A^2 - BA - C^2 - DC = -\frac{1}{8} (A + 4B)^2 + 2B^2 - \left(C + \frac{D}{2} \right)^2 + \frac{D^2}{4}$$

$$\dot{V}_1(t) \leq -\frac{1}{8} \left[(A + 4B)^2 - (16B^2 + 2D^2) \right]$$

$$16B^2 + 2D^2 \leq \left(\|\tilde{w}\| \left(3 \nabla \sigma_{\max} \mu_{\max} + \nabla \sigma_{\max} (G_{\max})^2 \lambda_{\max}(R^{-1}) \nabla \varepsilon_{\max} \right) + 2e_{\max} \right)^2$$

$$+ 2 \left(e_{\max} + \frac{1}{2} \|\tilde{w}\| \nabla \sigma_{\max} \nabla \sigma_{\max} (G_{\max})^2 \lambda_{\max}(R^{-1}) \nabla \varepsilon_{\max} \right)^2 \quad (40)$$

Từ (39), (40), ta có bất đẳng thức $(A + 4B)^2 - (16B^2 + 2D^2) \geq \pi_1$

Với $\pi_1 > 0$ có thể được suy ra khi $\|\tilde{w}\|$ đủ lớn, bởi vì hệ số bậc cao nhất là

$\left((G_{\min})^2 \lambda_{\min}(R^{-1})(\nabla \sigma_{\min})^2\right)^2 > 0$. Do đó, ta hoàn toàn có thể tìm được số thực $\mathcal{G}_1$ sao cho

$\forall \|\tilde{w}\| > \mathcal{G}_1$ ta có: $(A+4B)^2 - (16B^2 + 2D^2) \geq \pi_1$, từ (40) ta có: $\dot{V}_1(t) \leq -\pi_1$. Tương tự với

$V_2(t)$, từ (21) ta tính đạo hàm của nó:

$$\begin{aligned} \dot{V}_2 &= (\nabla V^*)^T (f_i + g_i(\hat{u} + \Delta)) \\ &= -(x^T Q x + \lambda \rho^2(x)) + \frac{1}{4} (\nabla V^*)^T g_i(x) R^{-1} g_i(x)^T (\nabla V^*) \end{aligned} \quad (41)$$

$$\begin{aligned} &+ (\nabla V^*)^T g_i \left(-\frac{1}{2} R^{-1} (g_i(x))^T \nabla \sigma(x)^T \hat{w} + \Delta \right) \\ &= -(x^T Q x + \lambda \rho^2(x)) - \frac{1}{4} (\nabla V^*)^T g_i R^{-1} g_i^T (\nabla V^*) \end{aligned} \quad (42)$$

$$+ \frac{1}{2} (\nabla V^*)^T g_i R^{-1} g_i^T (\nabla \sigma(x)^T \tilde{w} + \nabla \varepsilon(x)) + (\nabla V^*)^T g_i \Delta$$

Để thuận tiện ta giả sử $\rho(x) = \varpi \|x\|$. Từ (46) ta có:

$$\dot{V}_2 \leq -(\lambda_{\min}(Q) + \lambda \varpi) \|x\|^2 + \theta^2 \quad (43)$$

2.2. Mô phỏng hệ chuyển mạch phi tuyến có nhiều đầu vào

Tác giả lấy hệ dưới theo ví dụ của một bài báo đã xây dựng không phải đối tượng thực nên về quá trình vật lý hệ thống là chưa thể hiện hết.

Phần lớn những bài báo trên tạp chí lớn của nhiều Nhà nghiên cứu về thuật toán ADP (Adaptive Dynamis Programming) và để chứng minh thuật toán thì chỉ cần chọn lựa ví dụ một hệ tuân theo mô hình chung của hệ phi tuyến với nhiều đầu vào như ở dưới tác giả chọn, như tài liệu tham khảo [1]. Ở đây tác giả giả định nhiều đầu vào bị chặn. Sau khi mô phỏng sự đáp ứng của tín hiệu là chính xác khi tham chiếu với thuật toán. Do đó kết luận của bài báo tác giả kết luận hoàn toàn chính xác. Với kết quả thuật toán này tác giả đã áp dụng cho mô hình con lắc ngược và đã được đăng trong một hội thảo chuyên ngành.

Xét hệ chuyển mạch sau:

$$\begin{cases} \dot{x} = \begin{bmatrix} -x_1 + x_2 \\ -0.2x_2 - \sin x_1 \cos x_3 \\ x_1 x_2 - \sin x_3 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix} u + \Delta_1(x, t) \\ \dot{x} = \begin{bmatrix} -x_1 + x_2^2 x_3 \\ -x_2 - \sin(x_1 \cos x_3) \\ \sin x_2 - x_3 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix} u + \Delta_2(x, t) \end{cases}$$

Trong đó,

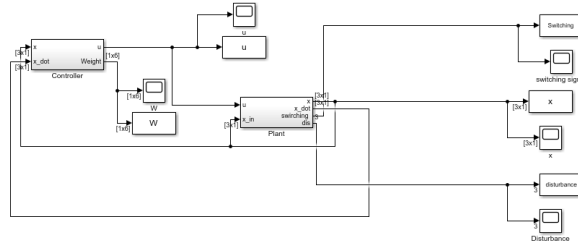
$$x_0 = [5; -3; 2]^T; \Delta_i(x, t) = \frac{1}{10} \begin{bmatrix} d_1 x_1 \\ d_2 x_2 \\ d_3 x_3 \end{bmatrix}; \|\Delta_i\| \leq \frac{1}{10} \|x\|$$

Với, $d_1; d_2; d_3$ được chọn là `rand()` trong matlab.

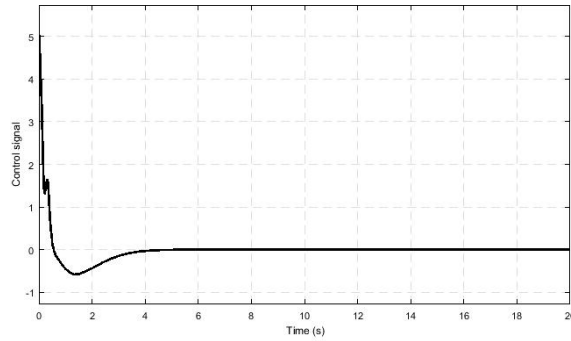
Ta chọn các tham số của bộ điều khiển:

$$R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \alpha = 0.01; Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}; \lambda = 5$$

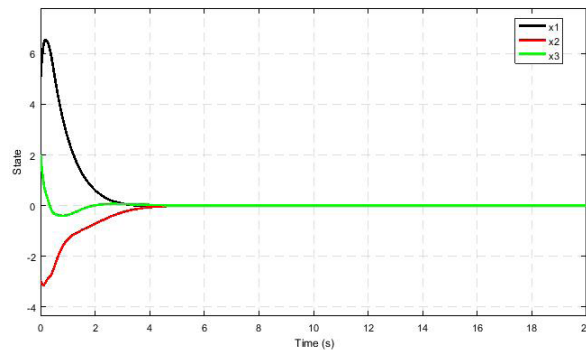
Khi đó ta có kết quả mô phỏng sau:



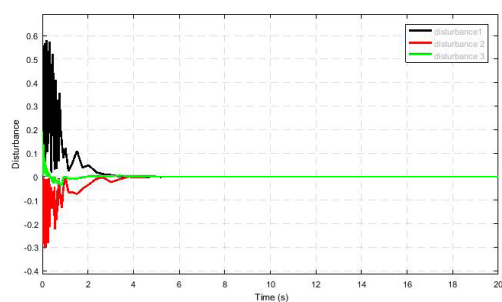
Hình 1. Sơ đồ cấu trúc điều khiển của hệ thống



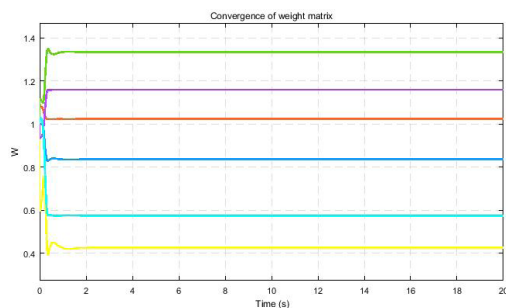
Hình 2. Tín hiệu điều khiển của hệ thống



Hình 3. Trạng thái của hệ thống



Hình 4. Nhiễu của hệ thống



Hình 5. Sự hội tụ của ma trận trọng lượng

3. KẾT LUẬN

Bài báo đã trình bày về lý thuyết và ứng dụng của thuật toán ADP trong điều khiển hệ phi tuyến. Tác giả đã thiết kế thuật toán và mô phỏng ví dụ đơn giản trên phần mềm Matlab. Để áp dụng vào một đối tượng có mô hình phức tạp nhiều bất định cần thêm những nghiên cứu nữa từ tác giả. Đối tượng điều khiển ở đây có thể phát triển với hệ phức tạp, quy mô lớn, đặc biệt đã có những nghiên cứu bước đầu đối với hệ multi agency.

TÀI LIỆU THAM KHẢO

1. Vamvoudakis, K.G. and Lewis, F.L. (2010), "Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem," *Automatica*, 46, pp.878-888.
2. Vamvoudakis, K.G. (2011), *Online learning algorithms for differential dynamic games and optimal control*, Ph.D. Thesis, Univ. Texas at Arlington.
3. Frank Lewis, Derong Liu (2012) *Reinforcement learning and approximate dynamic programming for feedback control*, IEEE Press, Wiley,
4. Derong Liu, Qinglai Wei, Ding Wang, Xiong Yang (2016), *Adaptive dynamic programming with applications in optimal control*, Springer,
5. Nguyễn Doãn Phước (2001) *Lý thuyết điều khiển tuyến tính*, NXB KH& KT.
6. Nguyễn Phùng Quang (2006), *MATLAB và Simulink dành cho kỹ sư điều khiển tự động*, Nhà xuất bản Khoa học và Kỹ thuật.