

# Mục lục

<b>1</b>	<b>Các khái niệm cơ bản</b>	<b>2</b>
<b>2</b>	<b>Kiểm soát phản hồi trạng thái tối ưu</b>	<b>3</b>
2.1	Nguyên tắc tối ưu . . . . .	3
2.2	Học thuyết Hamilton-Jacobi-Bellman . . . . .	5
2.2.1	Điều kiện đủ cho giải pháp tối ưu . . . . .	5
2.2.2	Luận điểm hợp lý về lý thuyết HJB . . . . .	7
2.2.3	Vấn đề điều chỉnh LQ . . . . .	8
2.2.4	Trường hợp bất biến theo thời gian với không gian vô hạn . . . . .	10
2.3	Xấp xỉ điều khiển tối ưu . . . . .	13
2.3.1	Ký hiệu . . . . .	14
2.3.2	Phương pháp Lukes . . . . .	15

# Chương 1

## Các khái niệm cơ bản

## Chương 2

# Kiểm soát phản hồi trạng thái tối ưu

### 2.1 Nguyên tắc tối ưu

Xem xét vấn đề kiểm soát tối ưu sau đây với thời gian cuối cố định  $t_b$ :

Tìm một điều khiển chấp nhận được  $u: [t_a, t_b] \rightarrow \Omega \subseteq \mathbb{R}^m$  với các ràng buộc

$$\begin{aligned}x(t_a) &= x_a \\ \dot{x}(t) &= f(x(t), u(t), t) \quad \text{với } t \in [t_a, t_b] \\ x(t_b) &\in S \subseteq \mathbb{R}^n\end{aligned}$$

sao cho hàm mục tiêu đạt giá trị nhỏ nhất.

$$J(u) = K(x(t_b)) + \int_{t_a}^{t_b} L(x(t), u(t), t) dt$$

Giả sử rằng chúng ta đã tìm thấy giải pháp tối ưu toàn cầu duy nhất với quỹ đạo điều khiển tối ưu  $u_0: [t_a, t_b] \rightarrow \mathbb{R}^m$  và quỹ đạo trạng thái tối ưu tương ứng  $x_0: [t_a, t_b] \rightarrow \mathbb{R}^n$  thỏa mãn  $x_0(t_a) = x_a$  và  $x_0(t_b) \in S$ .

Bây giờ, chọn một thời gian tùy ý  $\tau \in (t_a, t_b)$  và chia đôi vấn đề kiểm soát tối ưu ban đầu thành một vấn đề kiểm soát tối ưu trước đó trong khoảng thời gian  $[t_a, \tau]$  và một vấn đề tối ưu thành công trong khoảng thời gian  $[\tau, t_b]$ .

Vấn đề kiểm soát tối ưu tiền đề là:

Tìm một điều khiển được chấp nhận  $u: [t_a, \tau] \rightarrow \Omega$ , sao cho hệ thống động

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

được chuyển từ trạng thái ban đầu

$$x(t_a) = x_a$$

đến trạng thái cố định cuối cùng

$$x(\tau) = x^0(\tau)$$

tại thời điểm cuối cùng cố định  $\tau$  và như vậy có hàm chi phí

$$J(u) = \int_{t_a}^{\tau} L(x(t), u(t), t) dt$$

là nhỏ nhất.

Vấn đề kiểm soát tối ưu thành công là:

Tìm một điều khiển được chấp nhận  $u: [\tau, t_b] \rightarrow \Omega$ , sao cho hệ thống động

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

được chuyển từ trạng thái ban đầu nhất định

$$x(\tau) = x^0(\tau)$$

đến trạng thái cuối cùng

$$x(t_b) \in S$$

tại thời gian cuối cùng cố định  $t_b$  và như vậy hàm chi phí

$$J(u) = K(x(t_b)) + \int_{\tau}^{t_b} L(x(t), u(t), t) dt$$

là nhỏ nhất.

### **Định lý: Nguyên tắc tối ưu**

- 1) Giải pháp tối ưu cho bài toán điều khiển tối ưu thành công trùng khớp với phần thành công của giải pháp tối ưu của bài toán ban đầu.
- 2) Giải pháp tối ưu cho vấn đề kiểm soát tối ưu tiên đề trùng khớp với phần tiên đề của giải pháp tối ưu cho vấn đề ban đầu.

### **Chứng minh**

- 1) Mặt khác, kết hợp giải pháp tối ưu của vấn đề kiểm soát tối ưu thành công với phần tiên đề của giải pháp cho vấn đề kiểm soát tối ưu ban đầu sẽ mang lại giải pháp tốt hơn cho vấn đề sau.

- 2) Mặt khác, kết hợp giải pháp tối ưu của vấn đề kiểm soát tối ưu tiên đề với phần thành công của giải pháp cho vấn đề kiểm soát tối ưu ban đầu sẽ mang lại giải pháp tốt hơn cho vấn đề sau.

Về mặt khái niệm, chúng ta có thể giải quyết vấn đề điều khiển tối ưu thành công cho bất kỳ trạng thái ban đầu tùy ý  $x \in \mathbb{R}^n$  tại thời điểm ban đầu  $\tau$ , thay vì chỉ cho giá trị cố định  $x_0(\tau)$ . Hơn nữa, chúng ta có thể lặp lại quá trình này trong một thời gian ban đầu tùy ý  $t \in [t_a, t_b]$ , thay vì chỉ cho giá trị được chọn  $\tau$  ban đầu. Chỉ tập trung vào giá trị tối ưu của hàm chi phí trong tất cả các trường hợp này mang lại gọi là hàm chi phí tối ưu

$$T(x, t) = \min_{u(\cdot)} \left\{ K(x(t_b)) + \int_t^{t_b} L(x(t), u(t), t) dt \mid x(t) = x \right\}$$

### Bổ đề

- 3) Giải pháp tối ưu cho vấn đề kiểm soát tối ưu tiên đề với một trạng thái cuối cùng tự do tại thời điểm cuối cùng cố định  $\tau$  và với giá trị hàm

$$J = T(x(\tau), \tau) + \int_{t_a}^{\tau} L(x(t), u(t), t) dt$$

trùng với phần tiên đề của giải pháp tối ưu cho vấn đề kiểm soát tối ưu ban đầu.

- 4) Vector chi phí tối ưu  $\lambda^0(\tau)$  tương ứng với độ dốc của hàm chi phí tối ưu để đi (cost-to-gofunction), tức là,

$$\lambda^0(\tau) = \nabla_x T(x^0(\tau), \tau) \text{ với tất cả } \tau \in [t_a, t_b]$$

với điều kiện  $T(x, \tau)$  là liên tục khác biệt đối với  $x$  tại  $x^0(\tau)$

## 2.2 Học thuyết Hamilton-Jacobi-Bellman

### 2.2.1 Điều kiện đủ cho giải pháp tối ưu

Xem xét công thức thông thường của một vấn đề kiểm soát tối ưu với trạng thái cuối cùng không xác định tại thời điểm cuối cùng cố định:

Tìm một điều khiển liên tục từng phần  $u : [t_a, t_b] \rightarrow \Omega$  sao cho hệ động lực

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

được chuyển từ trạng thái ban đầu  $x(t_a) = x_a$  sang trạng thái cuối cùng tùy ý tại thời điểm cuối cùng cố định và sao cho hàm chi phí

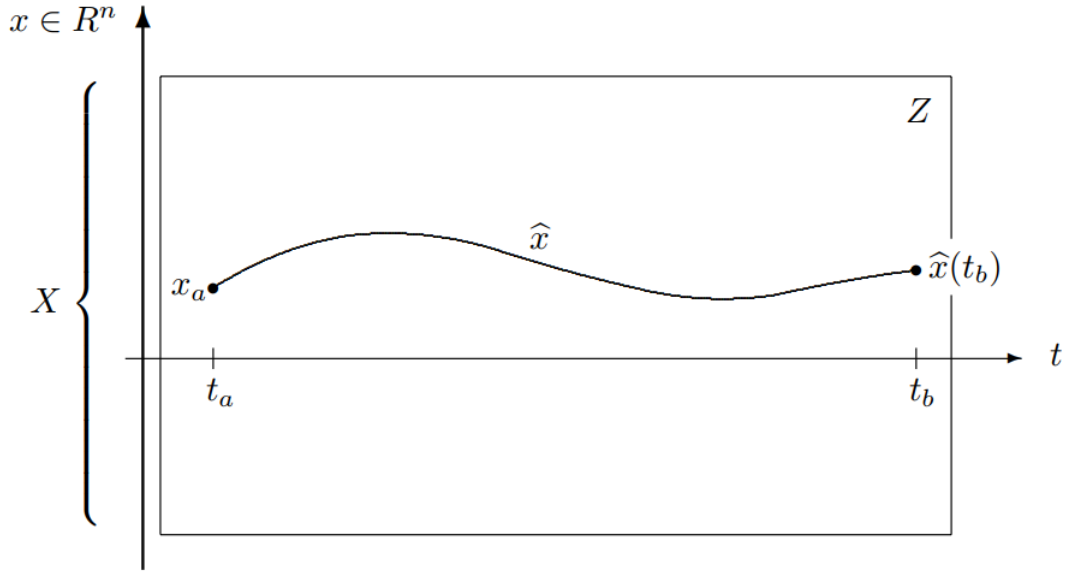
$$J(u) = K(x(t_b)) + \int_{t_a}^{t_b} L(x(t), u(t), t) dt$$

đạt min.

Xây dựng hàm Halminton với  $\lambda_0^0 = 1$

$$H(x, u, \lambda, t) = L(x, u, t) + \lambda^T f(x, u, t)$$

Xét tập hợp  $n + 1$  chiều  $Z = X \times [a, b] \subseteq \mathbb{R}^n \times \mathbb{R}$ , trong đó  $X$  là tập con (hy vọng rất lớn) của không gian trạng thái  $\mathbb{R}^n$  với bên trong không trống và  $[a, b]$  là tập hợp con của trục thời gian chứa ít nhất khoảng  $[t_a, t_b]$ , như trong Hình 2.1.



Hình 2.1: Ví dụ quỹ đạo trạng thái  $\hat{x}(\cdot)$  không rời  $X$

Chúng ta hãy xem xét các điều khiển được chấp nhận tùy ý  $u : [t_a, t_b] \rightarrow \Omega$  tạo ra các quỹ đạo trạng thái tương ứng  $\hat{x} : [t_a, t_b] \rightarrow \Omega$  bắt đầu từ  $x(t_a) = x_a$ . Chúng ta chủ yếu quan tâm đến các quỹ đạo không rời khỏi tập  $Z$ , tức là thỏa mãn  $x(t) \in X$  cho tất cả  $t \in [t_a, t_b]$ .

Với các giả thuyết sau đây, các điều kiện đủ cho sự tối ưu toàn cục của một giải pháp cho một vấn đề điều khiển tối ưu có thể được nêu trong Định lý Hamilton-Bellman-Jacobi dưới đây.

### Giải thuyết

- Đặt  $u : [t_a, t_b] \rightarrow \Omega$  là một điều khiển được chấp nhận tạo ra quỹ đạo trạng thái  $\hat{x} : [t_a, t_b] \rightarrow \mathbb{R}^n$  với  $\hat{x}(t_a) = x_a$  và  $\hat{x}(\cdot) \in Z$ .
- Với tất cả  $(x, t) \in Z$  và tất cả  $\lambda \in \mathbb{R}^n$ , để hàm Hamilton  $H(x, \omega, \lambda, t) = L(x, \omega, t) + \lambda^T f(x, \omega, t)$  có một giá trị nhỏ nhất toàn cục với  $\omega \in \Omega$  tại

$$\omega = \tilde{u}(x, \lambda, t) \in \Omega$$

- c) Đặt  $T(x, t) : Z \rightarrow \mathbb{R}$  là một hàm phân biệt liên tục thỏa mãn phương trình vi phân riêng phần Hamilton-Jacobi-Bellman

$$\frac{\partial T(x, t)}{\partial t} + H \left[ x, \tilde{u}(x, \nabla_x T(x, t), t), \nabla_x T(x, t), t \right] = 0$$

với điều kiện biên

$$T(x, t_b) = K(x)$$

với tất cả  $(x, t_b) \in Z$ .

Nhận xét:

- Hàm  $\hat{u}$  được gọi là điều khiển H thu nhỏ.
- Khigiar thuyết b được thỏa mãn, hàm Hamilton H được gọi là bình thường.

### Định lý Hamilton-Jacobi-Bellman

Nếu các giả thuyết a, b, và c được thỏa mãn và nếu quỹ đạo điều khiển  $\hat{u}(\cdot)$  và quỹ đạo trạng thái  $\hat{x}(\cdot)$  được tạo bởi  $\hat{u}(\cdot)$  có liên quan thông qua

$$\hat{u}(t) = \tilde{u}(\hat{x}(t), \nabla_x T(\hat{x}(t), t), t),$$

thì giải pháp  $\hat{u}, \hat{x}$  là tối ưu đối với tất cả các quỹ đạo trạng thái  $x$  được tạo bởi một quỹ đạo kiểm soát được chấp nhận  $u$ , không rời  $X$ . Hơn nữa,  $T(x, t)$  là hàm chi phí tối ưu để đi (cost-to-go function).

### Bổ đề

Nếu  $Z = \mathbb{R}^n \times [t_a, t_b]$  thì giải pháp  $\hat{u}, \hat{x}$  là tối ưu toàn cục.

### 2.2.2 Luận điểm hợp lý về lý thuyết HJB

Chúng ta có các sự kiện sau đây:

- 1) Nếu hàm Hamilton H bình thường, chúng ta có duy nhất sau H-tối thiểu hóa điều khiển tối ưu:

$$u^o(t) = \tilde{u}(x^o(t), \lambda^o(t), t)$$

- 2) Hàm chi phí tối ưu  $T(x, t)$  rõ ràng phải thỏa mãn điều kiện biên

$$T(x, t_b) = K(x)$$

vì tại thời điểm cuối cùng  $t_b$ , giá trị hàm giá chỉ bảo gồm trạng thái cuối cùng  $K(x)$ .

- 3) Nguyên lý tối ưu đã chỉ ra rằng chi phí tối ưu  $\lambda^o(t)$  tương ứng với độ dốc của hàm chi phí tối ưu,

$$\lambda^o(t) = \nabla_x T(x^o(t), t),$$

bất cứ đâu  $T(x^o(t), t)$  liên tục khác biệt đối với  $x$  tại  $x = x^o(t)$ .

- 4) Cùng với một quỹ đạo được chấp nhận tùy ý  $u(\cdot), x(\cdot)$ , tương ứng hàm chi phí tối ưu

$$J(x(t), t) = K(x(t_b)) + \int_t^{t_b} L(x(t), u(t), t) dt$$

tiến hóa theo phương trình vi phân sau:

$$\frac{dJ}{dt} = \frac{\partial J}{\partial x} \dot{x} + \frac{\partial J}{\partial t} = \lambda^T f(x, u, t) + \frac{\partial J}{\partial t} = -L(x, u, t)$$

Vì thế,

$$\frac{\partial J}{\partial t} = -\lambda^T f(x, u, t) - L(x, u, t) = -H(x, u, \lambda, t)$$

Điều này tương ứng với phương trình vi phân từng phần của hàm chi phí tối ưu  $T(x, t)$ , ngoại trừ luật điều khiển tối ưu chưa được áp dụng vào.

### 2.2.3 Vấn đề điều chỉnh LQ

*Báo cáo vấn đề điều khiển tối ưu*

Tìm một luật điều khiển phản hồi trạng thái tối ưu  $u: \mathbb{R}^n \times [t_a, t_b] \rightarrow \mathbb{R}^m$ , sao cho hệ động lực tuyến tính

$$\dot{x}(t) = A(t)x(t) + B(t)u(t)$$

được chuyển từ trạng thái ban đầu đã cho  $x(t_a) = x_a$  sang trạng thái cuối cùng tùy ý tại thời gian cuối cùng cố định  $t_b$  sao cho hàm chi phí bậc hai

$$J(u) = \frac{1}{2} x^T(t_b) F x(t_b) + \int_{t_a}^{t_b} \left( \frac{1}{2} x^T(t) Q(t) x(t) + x^T(t) N(t) u(t) + \frac{1}{2} u^T(t) R(t) u(t) \right) dt$$

đạt min, trong đó  $R(t)$  là ma trận đối xứng xác định dương và  $F, Q$  và  $\begin{bmatrix} Q(t) & N(t) \\ N^T(t) & R(t) \end{bmatrix}$  là đối xứng và nửa xác định dương.

*Phân tích vấn đề*

Xây dựng hàm Hamilton

$$H = \frac{1}{2} x^T Q x + x^T N u + \frac{1}{2} u^T R u + \lambda^T A x + \lambda^T B u$$



có điều khiển H-tối thiểu sau đây:

$$u = -R^{-1}[B^T\lambda + N^Tx] = -R^{-1}[B^T\nabla_x T + N^Tx]$$

Do đó, phương trình vi phân từng phần Hamilton-Jacobi-Bellman là

$$\begin{aligned} 0 &= \frac{\partial T}{\partial t} + H(x, \tilde{u}(x, \nabla_x T, t), \nabla_x T, t) \\ &= \frac{\partial T}{\partial t} + \frac{1}{2} \left( x^T Q x - x^T N R^{-1} N^T x - \nabla_x T^T B R^{-1} B^T \nabla_x T \right. \\ &\quad \left. + \nabla_x T^T [A - B R^{-1} N^T] x + x^T [A - B R^{-1} N^T]^T \nabla_x T \right) \end{aligned}$$

với điều kiện biên

$$T(x, t_b) = \frac{1}{2} x^T F x.$$

Rõ ràng, một giả định cho hàm chi phí tối ưu  $T(x, t)$  là bậc hai trong  $x$  nên hoạt động. Điều này dẫn đến một phương trình vi phân từng phần trong đó tất cả các số hạng là bậc hai theo  $x$ . (...)

$$T(x, t) = \frac{1}{2} x^T K(t) x$$

dẫn đến

$$\nabla_x T(x, t) = K(t)x \text{ và } \frac{\partial T(x, t)}{\partial t} = \frac{1}{2} x^T \dot{K}(t) x$$

Dạng cuối cùng sau đây của phương trình vi phân từng phần Hamilton-Jacobi-Bellman thu được:

$$\begin{aligned} &\frac{1}{2} x^T \left( \dot{K}(t) + Q - N R^{-1} N^T - K(t) B R^{-1} B^T K(t) \right. \\ &\quad \left. + K(t) [A - B R^{-1} N^T] + [A - B R^{-1} N^T]^T K(t) \right) x = 0 \\ &\quad T(x, t_b) = \frac{1}{2} x^T F x \end{aligned}$$

Do đó, chúng tôi nhận được luật kiểm soát phản hồi trạng thái tối ưu sau:

$$u(t) = -R^{-1}(t)[B^T(t)K(t) + N^T(t)]x(t)$$

trong đó ma trận đối xứng và (nếu) xác định dương phải được tính toán trước bằng cách giải phương trình vi phân Riccati ma trận

$$\begin{aligned} \dot{K}(t) &= -[A(t) - B(t)R^{-1}(t)N^T(t)]^T K(t) - K(t)[A(t) - B(t)R^{-1}(t)N^T(t)] \\ &\quad - K(t)B(t)R^{-1}(t)B^T(t)K(t) - Q(t) + N(t)R^{-1}(t)N^T(t) \end{aligned}$$

với điều kiện biên

$$K(t_b) = F$$

### 2.2.4 Trường hợp bất biến theo thời gian với không gian vô hạn

Trong phần này, các vấn đề điều khiển tối ưu bất biến theo thời gian với vector điều khiển không giới hạn trạng thái  $u(t) \in \mathbb{R}^m$ , trạng thái tự do cuối cùng  $x(t_b)$  tại thời điểm cuối cùng vô hạn  $t_b = \infty$  được xem xét.

Phát biểu chung nhất của vấn đề kiểm soát tối ưu này là:

Tìm một điều khiển liên tục từng phần  $u : [0, \infty) \rightarrow \mathbb{R}^m$ , sao cho hệ thống động hoàn toàn có thể điều khiển

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

được chuyển từ trạng thái ban đầu nhất định

$$x(0) = x_a$$

đến trạng thái cuối cùng tùy ý  $x(\infty) \in \mathbb{R}^n$  tại thời điểm cuối cùng vô hạn và hàm chi phí

$$J(u) = \int_0^\infty L(x(t), u(t)) dt$$

là nhỏ nhất và đạt được một giá trị tối ưu hữu hạn.

Để có một vấn đề được đặt ra, các biến của vấn đề nên được chọn theo cách sao cho trạng thái cân bằng tĩnh cố định ở  $x = 0$  và có thể đạt được bằng điều khiển biến mất không đối xứng  $u(t) \rightarrow 0$  như  $t \rightarrow \infty$ . Do đó,  $f(0, 0) = 0$  là bắt buộc. Hơn nữa, chọn tích phân  $L$  của hàm chi phí với  $L(0, 0) = 0$  và sao cho nó lồi hoàn toàn trong cả  $x$  và  $u$  và sao cho  $L(x, u)$  tăng mà không bị ràng buộc, với tất cả  $(x, u)$  trong đó  $x$ , hoặc  $u$ , hoặc cả  $x$  và  $u$  đi đến vô cùng theo bất kỳ hướng nào trong không gian trạng thái  $\mathbb{R}^n$  hoặc không gian điều khiển  $\mathbb{R}^m$ , tương ứng. Tất nhiên, chúng ta giả sử rằng cả  $f$  và  $L$  ít nhất một lần liên tục khác nhau đối với  $x$  và  $u$ .

Rõ ràng, trong trường hợp bất biến theo thời gian với không gian vô hạn, hàm chi phí tối ưu  $T(x, t)$  là bất biến theo thời gian, tức là,

$$T(x, t) \equiv T(x),$$

bởi vì giải pháp tối ưu là thay đổi bất biến (?). Không quan trọng hệ thống bắt đầu với trạng thái ban đầu  $x_a$  tại thời điểm ban đầu 0 hay với cùng trạng thái ban đầu  $x_a$  tại một số thời điểm bắt đầu  $t_a \neq 0$ .

Do đó, phương trình vi phân từng phần Hamilton-Jacobi-Bellman

$$\frac{\partial T(x, t)}{\partial t} + H\left[x, \tilde{u}\left(x, \nabla_x T(x, t), t\right), \nabla_x T(x, t), t\right] = 0$$

suy biến thành phương trình vi phân từng phần

$$H\left[x, \tilde{u}\left(x, \nabla_x T(x)\right), \nabla_x T(x)\right] = 0$$

và mất điều kiện biên cũ  $T(x, t_b) = K(x)$ .

Trong trường hợp đặc biệt của hệ động lực bậc 1 ( $n = 1$ ), đây là một phương trình vi phân thông thường có thể được tích hợp bằng điều kiện biên  $J(0) = 0$ .

Đối với các hệ thống động có bậc cao hơn ( $n \geq 2$ ), có sẵn các kỹ thuật giải quyết vấn đề thay thế sau:

- a) Chọn một hàm xác định dương  $K(x)$  tùy ý với  $K(0) = 0$  thỏa mãn điều kiện tăng thông thường. Tích hợp phương trình vi phân từng phần Hamilton-Jacobi-Bellman trên  $\mathbb{R}^n \times (-\infty, t_b]$  bằng cách sử dụng điều kiện biên  $T(x, t_b) = K(x)$ . Giải pháp  $T(x, t)$  hội tụ theo thời gian mong muốn chi phí tối ưu để thực hiện hàm  $T(x)$  là  $t \rightarrow \infty$ ,

$$T(x) = \lim_{t \rightarrow -\infty} T(x, t).$$

- b) Giải hai phương trình

$$\begin{aligned} H(x, u, \lambda) &= 0 \\ \nabla_u H(x, u, \lambda) &= 0 \end{aligned}$$

để tìm ra luật kiểm soát phản hồi trạng thái tối ưu mong muốn  $u^o(x)$  mà không tính toán hàm chi phí tối ưu  $T(x)$ . Vì cả hai phương trình này đều tuyến tính trong giá trị  $\lambda$ , nên có một cơ hội tốt mà có thể được loại bỏ mà không cần tính toán  $\lambda = \nabla_x T(x)$  một cách rõ ràng. Điều này dẫn đến một hình thức ngầm định của luật kiểm soát phản hồi trạng thái tối ưu.

### Ví dụ

Giả sử rằng chúng ta đã giải quyết vấn đề điều chỉnh LQ sau đây cho một hệ thống động không ổn định:

$$\begin{aligned} \dot{x}(t) &= ax(t) + bu(t) \text{ với } a > 0, b \neq 0 \\ x(0) &= x_a \\ J(u) &= \frac{1}{2} \int_0^\infty (qx^2(t) + u^2(t)) dt \text{ với } q > 0 \end{aligned}$$

Kết quả là bộ điều khiển phản hồi trạng thái tuyến tính

$$u(t) = -gx(t)$$

với

$$g = bk = \frac{a + a\sqrt{1 + \frac{b^2q}{a^2}}}{b}$$

Bây giờ, chúng ta muốn thay thế bộ điều khiển tuyến tính này bằng một bộ điều khiển phi tuyến cho các giá trị lớn của  $|x|$ , tức là, trong đó cho thấy một hành vi bão hòa phù hợp, nhưng mà vẫn giữ được "độ cứng" của bộ điều khiển LQ cho các tín hiệu nhỏ  $x$ . Lưu ý rằng, do sự bất ổn của nhà máy, điều khiển không được thấm vào một giá trị tối đa liên tục cho sự kiểm soát. Thay vào đó, chỉ có thể thấm vào một bộ điều khiển tuyến tính "mềm hơn" của dạng  $u = -g_\infty x$  cho  $|x|$  lớn với  $g_\infty > a/b$ .

Để đạt được mục tiêu này, hàm chi phí được sửa đổi như sau:

$$J(u) = \frac{1}{2} \int_0^\infty (qx^2(t) + u^2(t) + \beta u^4(t)) dt \text{ với } \beta > 0.$$

Theo **b**, ta có

$$\begin{aligned} H(x, u, \lambda) &= \frac{q}{2}x^2 + \frac{1}{2}u^2 + \frac{\beta}{2}u^4 + \lambda ax + \lambda bu = 0 \\ \frac{\partial H}{\partial u} &= u + 2\beta u^3 + \lambda b = 0. \end{aligned}$$

Loại bỏ  $\lambda$  mang lại luật kiểm soát phản hồi trạng thái tối ưu:

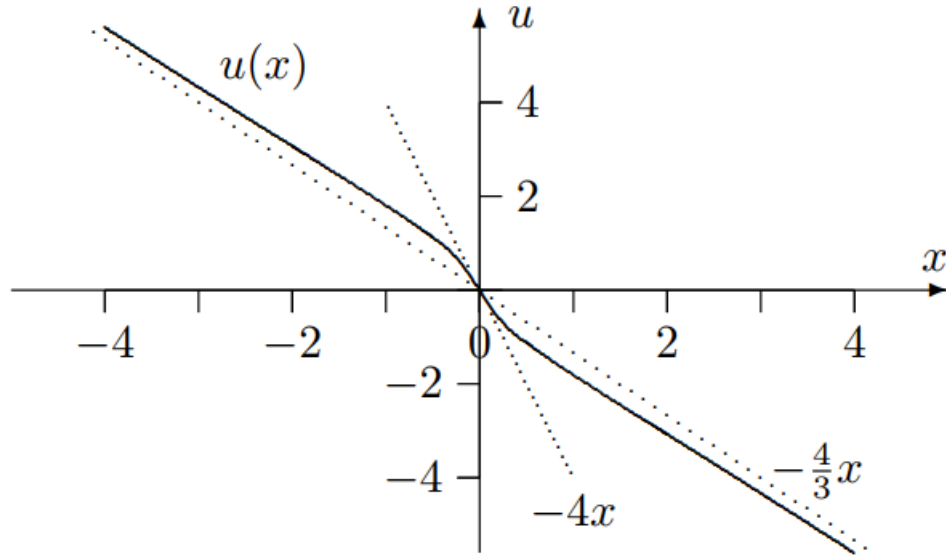
$$3\beta u^4 + \frac{2\beta a}{b}xu^3 + u^2 + \frac{2a}{b}xu - qx^2 = 0$$

Điều khiển phản hồi trạng thái tối ưu  $u(x)$  thu được bằng cách giải phương trình này:

$$u(x) = \arg \left\{ 3\beta u^4 + \frac{4\beta a}{b}xu^3 + u^2 + \frac{2a}{b}xu - qx^2 = 0 \left| \begin{array}{ll} u < -\frac{a}{b}x & \text{với } x > 0 \\ u = 0 & \text{với } x = 0 \\ u > -\frac{a}{b}x & \text{với } x < 0 \end{array} \right. \right\}$$

Đặc tính tín hiệu nhỏ giống với đặc tính của bộ điều chỉnh LQ vì các số hạng thứ tư  $u^4$  và  $xu^3$  không đáng kể. Ngược lại, đối với đặc tính tín hiệu lớn, các thuật ngữ bậc bốn chiếm ưu thế và các thuật ngữ bậc hai  $u^2$ ,  $xu$  và  $x^2$  là không đáng kể. Do đó, đặc tính tín hiệu lớn là

$$u \approx \frac{4a}{3b}x$$



Hình 2.2: Luật điều khiển phi tuyến tối ưu

Trong hình 2.2, luật điều khiển tối ưu phi tuyến được mô tả cho ví dụ trong đó  $a = 1, b = 1, q = 8$  và  $\beta = 1$  với bộ điều chỉnh LQ  $g = 4$  và mức tăng tín hiệu lớn  $g_\infty = -\frac{4}{3}$ .

Các đầu đọc nghiêng (the inclined reader is invited) được mời để xác minh rằng thay thế  $\beta u^4$  trong hàm chi phí  $\beta u^{2k}$  với  $k \geq 2$  dẫn đến kết quả là mức tăng trưởng lớn

$$u \approx -\frac{2ka}{(2k-1)b}x$$

## 2.3 Xấp xỉ điều khiển tối ưu

Trong hầu hết các trường hợp, không tìm thấy giải pháp phân tích nào của phương trình vi phân từng phần Hamilton-Jacobi-Bellman. Hơn nữa, việc giải quyết nó bằng số có thể cực kỳ cồng kềnh.

Do đó, một phương pháp được trình bày ở đây cho phép tìm các kết quả gần đúng cho luật kiểm soát phản hồi trạng thái cho một vấn đề kiểm soát tối ưu bất biến theo thời gian với không gian vô hạn.

Chúng ta hãy xem xét một vấn đề kiểm soát tối ưu bất biến theo thời gian với một không gian vô hạn cho một hệ động lực phi tuyến với hàm chi phí không bậc hai, được cấu trúc như sau:

Tìm một luật điều khiển phản hồi trạng thái tối ưu bất biến theo thời gian  $u: R^n \rightarrow R^m$ , sao cho hệ thống động

$$\dot{x}(t) = F(x(t), u(t)) = Ax(t) + Bu(t) + f(x(t), u(t))$$

được chuyển từ trạng thái ban đầu  $x(0) = x_a$  sang trạng thái cân bằng  $x = 0$  tại thời điểm cuối cùng vô hạn sao cho hàm chi phí

$$\begin{aligned} J(u) &= \int_0^\infty L(x(t), u(t)) dt \\ &= \int_0^\infty \left( \frac{1}{2} x^T(t) Q x(t) + x^T(t) N u(t) + \frac{1}{2} u^T(t) R u(t) + \ell(x(t), u(t)) \right) dt \end{aligned}$$

đạt giá trị nhỏ nhất. Trong vấn đề này, người ta cho rằng các điều kiện sau được thỏa mãn:

- $[A, B]$  cổ định
- $R > 0$
- $Q = C^T C \geq 0$
- $[A, C]$  có thể tìm được
- $\begin{bmatrix} Q & N \\ N^T & R \end{bmatrix} \geq 0$
- $f(x, u)$  chỉ chứa các bậc hai hoặc cao hơn trong  $x$  và/hoặc  $u$
- $\ell(x, u)$  chỉ chứa các bậc hai hoặc cao hơn trong  $x$  và/hoặc  $u$ .

### 2.3.1 Ký hiệu

#### Phép toán vi phân

Đối với ma trận Jacobi của các đạo hàm riêng của hàm có giá trị  $n$ -vector  $f$  đối với  $m$ -vector  $u$ , ký hiệu sau được sử dụng:

$$f_u = \frac{\partial f}{\partial u} = \begin{bmatrix} \frac{\partial f_1}{\partial u_1} & \cdots & \frac{\partial f_1}{\partial u_m} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial u_1} & \cdots & \frac{\partial f_n}{\partial u_m} \end{bmatrix}$$

Chú ý rằng kết quả của một vector hàng là đạo hàm vô hướng của hàm  $f$ .

### Sorting Powers (Phân loại theo lũy thừa (?))

Trong các hàm đa thức, chúng ta tổng hợp các thuật ngữ có cùng công suất (the same power) như sau:

$$\begin{aligned} f(x, u) &= f^{(2)}(x, u) + f^{(3)}(x, u) + f^{(4)}(x, u) \cdots \\ \ell(x, u) &= \ell^{(3)}(x, u) + \ell^{(4)}(x, u) + \ell^{(5)}(x, u) + \cdots \\ T(x) &= T^{(2)}(x) + T^{(3)}(x) + T^{(4)}(x) + \cdots \\ u^o(x) &= u^{o(1)}(x) + u^{o(2)}(x) + u^{o(3)}(x) + \cdots \end{aligned}$$

Ví dụ: Trong trường hợp đơn giản nhất của hàm vô hướng với các đối số vô hướng  $x$  và  $u$ , hàm  $\ell^{(3)}(x, u)$  có dạng tổng quát sau:

$$\ell^{(3)}(x, u) = \alpha x^3 + \beta x^2 u + \gamma x u^2 + \delta u^3$$

Đối với các đạo hàm của hàm  $f$  và  $\ell$ , các lũy thừa được sắp xếp theo cách tương tự, ví dụ:

$$\begin{aligned} f_u(x, u) &= f_u^{(1)}(x, u) + f_u^{(2)}(x, u) + f_u^{(3)}(x, u) + \cdots \\ \ell_u(x, u) &= \ell_u^{(2)}(x, u) + \ell_u^{(3)}(x, u) + \ell_u^{(4)}(x, u) + \cdots \end{aligned}$$

Lưu ý thực tế sau đây cho đạo hàm của hàm đối với đối số có giá trị vô hướng hoặc giá trị véc tơ:

$$\ell_u^{(k)}(x, u) = \frac{\partial \ell^{(k+1)}(x, u)}{\partial u}.$$

Trong ví dụ trước, chúng ta nhận được

$$\ell_u^{(2)}(x, u) = \frac{\partial \ell^{(3)}(x, u)}{\partial u} = \beta x^2 + 2\gamma x u + 3\delta u^2.$$

Từ bây giờ, chúng ta viết

$$T_x(x) = T_x^{[2]}(x) + T_x^{[3]}(x) + T_x^{[4]}(x) + \cdots$$

thay vì  $T_x^{(1)}(x) + T_x^{(2)}(x) + T_x^{(3)}(x) + \cdots$ . Nói chung, loại ký hiệu này sẽ được sử dụng trong phần tiếp theo.

### 2.3.2 Phương pháp Lukes

Trong bước xấp xỉ đầu tiên, điều khiển tuyến tính  $u^{o(1)}$  được xác định bằng cách giải bài toán điều chỉnh LQ cho hệ động lực tuyến tính và cho phần thuần phương trình bậc 2 của hàm chi phí.

Trong mỗi bước xấp xỉ bổ sung của phương pháp đệ quy Lukes, một số lũy

thừa bổ sung được thêm vào luật phản hồi  $u(x)$ , trong khi một số lũy thừa bổ sung của xấp xỉ  $f$  và  $A$  được xử lý.

Như đã trình bày trong 2.2.4, hai phương trình sau phải được giải gần đúng trong mỗi bước xấp xỉ:

$$H = T_x(x)F(x, u) + L(x, u) = 0 \quad (2.1)$$

$$H_u = T_x(x)F_u(x, u) + L_u(x, u) \quad (2.2)$$

Trong bài toán sắp tới, chúng ta có các phương trình sau:

$$\begin{aligned} H &= T_x(x)[Ax + Bu + f(x, u)] + \frac{1}{2}x^T Qx + x^T Nu \\ &\quad + \frac{1}{2}u^T Ru + \ell(x, u) = 0 \end{aligned} \quad (2.3)$$

$$H_u = T_x(x)(B + f_u(x, u)) + x^T N + u^T R + \ell_u(x, u) = 0 \quad (2.4)$$

Giải phương trình (2.4) với  $u^o$ :

$$u^{oT} = -[T_x(x)(B + f_u(x, u^o)) + x^T N + \ell_u(x, u^o)]R^{-1} \quad (2.5)$$

### 1<sup>st</sup> Approximation: LQ-Regulator

$$\begin{aligned} \dot{x}(t) &= Ax + Bu \\ J(u) &= \int_0^\infty \left( \frac{1}{2}x^T Qx + x^T Nu + \frac{1}{2}u^T Ru \right) dt \\ u^{o(1)} &= Gx \text{ với } G = -R^{-1}(B^T K + N^T), \end{aligned}$$

trong đó  $K$  là nghiệm duy nhất duy nhất của phương trình Riccati

$$[A - BR^{-1}N^T]^T K + K[A - BR^{-1}N^T] - KBR^{-1}B^T K + Q - NR^{-1}N^T = 0$$

Kết quả của hệ điều khiển tuyến tính được mô tả bởi phương trình vi phân

$$\dot{x}(t) = [A + BG]x(t) = A^o x(t)$$

và có giá trị hàm chi phí

$$T^{(2)}(x) = \frac{1}{2}x^T Kx \text{ với } T_x^{[2]}(x) = x^T K.$$

### 2<sup>nd</sup> Approximation

$$\begin{aligned} u^o(x) &= u^{o(1)}(x) + u^{0(2)}(x) \\ T_x(x) &= T_x^{[2]}(x) + T_x^{[3]}(x) \end{aligned}$$



**a)** Xác định  $T_x^{[3]}(x)$ :

Sử dụng (2.3):

$$\begin{aligned} 0 &= (T_x^{[2]} + T_x^{[3]})[Ax + B(u^{o(1)} + u^{o(2)}) + f(x, u^{o(1)} + u^{o(2)})] \\ &\quad + \frac{1}{2}x^T Qx + x^T N(u^{o(1)} + u^{o(2)}) + \frac{1}{2}(u^{o(1)} + u^{o(2)})^T R(u^{o(1)} + u^{o(2)}) \\ &\quad + \ell(x, u^{o(1)} + u^{o(2)}) \end{aligned}$$

Thuật ngữ khối: