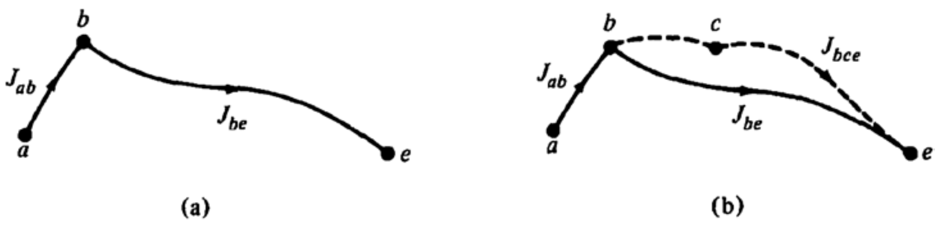
برنامه ریزی دینامیکی/پویا

Dynamic programming

اصل بهینه بلمن Bellman Principle of Optimality

هر تکه از مسیر بهینه، یک مسیر بهینه است.



بهترین مسیر:

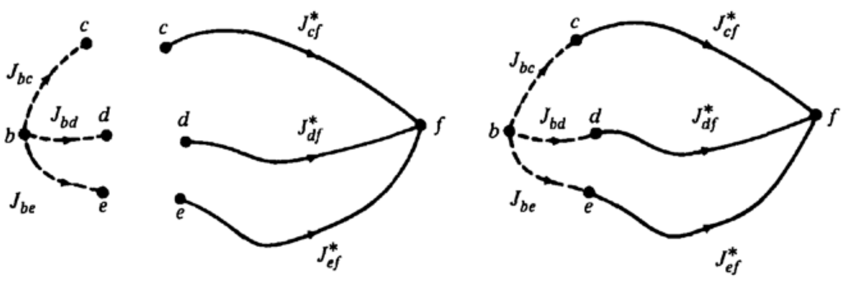
$$J^*_{ae} = J_{ab} + J_{be}$$

برای تمام مسیرهای دیگر از a به e باید $J \geq J^*_{ae}$

بنابراین:

$$J_{ab} + J_{bce} \geq J_{ab} + J_{be} = J^*_{ae} \quad \Rightarrow \quad J_{bce} \geq J_{be}$$

کاربرد اصل بلمن در تعیین مسیر بهینه



بین مسیرهای زیر بهینه سازی شود:

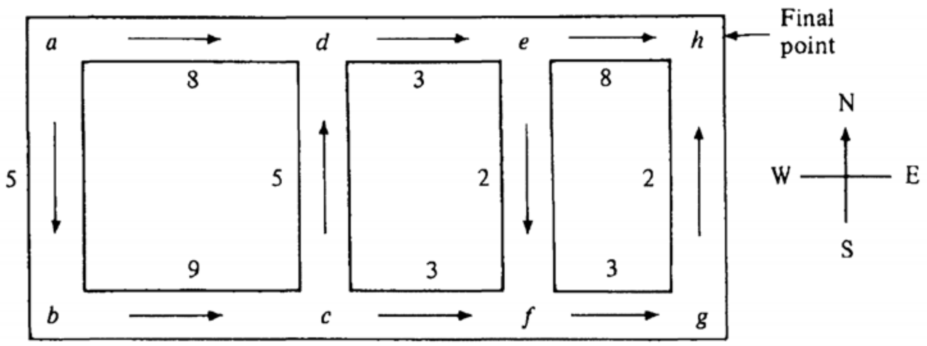
$$C^*_{bcf} = J_{bc} + J^*_{cf}$$

$$C^*_{bdf} = J_{bd} + J^*_{df}$$

$$C^*_{bef} = J_{be} + J^*_{ef}$$

یعنی بخشی از مسیر قبلا بهینه شده است و فقط یک تصمیم بهینه می شود.

مثال: انتخاب مسیر



نحوه کار: بین قدم به قدم بین هزینه های مقطعی:

$$C^*_{\alpha x_i h} = J_{\alpha x_i} + J^*_{x_i h}$$

بهینه سازی شود:

$$J^*_{\alpha h} = \min\{C^*_{\alpha x_1 h}, C^*_{\alpha x_2 h}, \ldots, C^*_{\alpha x_i h}, \ldots\}.$$

که در آن

- α is the current state (intersection).
- u_i is an allowable decision (control) elected at the state α. In this example i can assume one or more of the values 1, 2, 3, 4, corresponding to the headings N, E, S, W.
- x_i is the state (intersection) adjacent to α which is reached by application of u_i at α.
- h is the final state.
- $J_{\alpha x_i}$ is the cost to move from α to x_i.
- $J^*_{x_i h}$ is the *minimum cost* to reach the final state h from x_i.
- $C^*_{\alpha x_i h}$ is the minimum cost to go from α to h *via* x_i.
- $J^*_{\alpha h}$ is the minimum cost to go from α to h (by any allowable path).
- $u^*(\alpha)$ is the optimal decision (control) at α.

$C^*_{cdh} = J_{cd} + J^*_{dh} =$ minimum cost to reach h from c via d

$C^*_{cfh} = J_{cf} + J^*_{fh} =$ minimum cost to reach h from c via f.

$$\begin{aligned} J^*_{ch} &= \min\{C^*_{cdh}, C^*_{cfh}\} \\ &= \min\{15, 8\} \\ &= 8 \end{aligned}$$

جواب بهینه:

Current intersection	*Heading*	*Next intersection*	*Minimum cost from α to h via x_i*	*Minimum cost to reach h from α*	*Optimal heading at α*
α	u_i	x_i	$J_{\alpha x_i} + J^*_{x_i h} = C^*_{\alpha x_i h}$	$J^*_{\alpha h}$	$u^*(\alpha)$
g	N	*h*	2 + 0 = 2	2	N
f	E	*g*	3 + 2 = 5	5	E
e	E	*h*	8 + 0 = 8		
	S	*f*	2 + 5 = 7	7	S
d	E	*e*	3 + 7 = 10	10	E
c	N	*d*	5 + 10 = 15		
	E	*f*	3 + 5 = 8	8	E
b	E	*c*	9 + 8 = 17	17	E
a	E	*d*	8 + 10 = 18	18	E
	S	*b*	5 + 17 = 22		

مثال ۲: کنترل بهینه یک سیستم دینامیکی

معادله سیستم

$$\frac{d}{dt}[x(t)] = ax(t) + bu(t),$$

تابع هزینه

$$J = x^2(T) + \lambda \int_0^T u^2(t)\,dt,$$

نیاز است سیستم به یک مسیر با تصمیم گیری (کنترل) در میان مسیر تبدیل شود. زمان گسسته سازی می شود

$$\frac{x(t+\Delta t) - x(t)}{\Delta t} \approx ax(t) + bu(t),$$

$$x(t+\Delta t) = [1 + a\,\Delta t]x(t) + b\,\Delta t\,u(t).$$

$$x([k+1]\,\Delta t) = [1 + a\,\Delta t]x(k\,\Delta t) + b\,\Delta t\,u(k\,\Delta t); \qquad k = 0, 1, \ldots, N-1.$$

$$x(k+1) = [1 + a\,\Delta t]x(k) + b\,\Delta t\,u(k).$$

تابع هزینه:

$$J = x^2(N\,\Delta t) + \lambda\left[\int_0^{\Delta t} u^2(0)\,dt + \int_{\Delta t}^{2\,\Delta t} u^2(\Delta t)\,dt + \cdots + \int_{(N-1)\,\Delta t}^{N\,\Delta t} u^2([N-1]\,\Delta t)\,dt\right],$$

$$J = x^2(N) + \lambda\,\Delta t\left[u^2(0) + u^2(1) + \cdots + u^2(N-1)\right]$$
$$= x^2(N) + \lambda\,\Delta t \sum_{k=0}^{N-1} u^2(k).$$

نمونه برای دو مرحله:

$$a = 0,\ b = 1,\ \lambda = 2,\ T = 2, \Delta t = 1, N = 2$$

$$x(k+1) = x(k) + u(k); \qquad k = 0, 1$$

$$J = x^2(2) + 2u^2(0) + 2u^2(1)$$

$$0.0 \le x(k) \le 1.5; \qquad k = 0, 1, 2$$

$$-1.0 \le u(k) \le 1.0; \qquad k = 0, 1.$$

مقداردهی به متغیرهای حالت و تصمیم گیری (کنترل)Quantization

$$x(k) = 0.0, 0.5, 1.0, 1.5$$

$$u(k) = -1.0, -0.5, 0.0, 0.5, 1.0.$$

بهینه سازی برای قدم انتها

$$C_{12}\left(x(1), u(1)\right) = J_{22}\left(x(2)\right) + J_{12}\left(x(1), u(1)\right)$$
$$C_{12}^*\left(x(1), u(1)\right) = J_{22}^*\left(x(2)\right) + J_{12}\left(x(1), u(1)\right) = x^2(2) + 2u^2(1) = \left(x(1) + u(1)\right)^2 + 2u^2(1)$$

$$J_{12}^*\left(x(1)\right) = \min_{u(1)} C_{12}^*\left(x(1), u(1)\right)$$

Current state $x(1)$	*Control* $u(1)$	*Next state* $x(2) = x(1) + u(1)$	*Cost* $x^2(2) + 2u^2(1) = J_{12}(x(1), u(1))$		*Minimum cost* $J_{12}^*(x(1))$	*Optimal control applied at* $k = 1$ $u^*(x(1), 1)$
1.5	0.0	1.5	$(1.5)^2 + 2(0.0)^2 =$	2.25		
	−0.5	1.0	$(1.0)^2 + 2(-0.5)^2 =$	1.50	$J_{12}^*(1.5) = 1.50$	$u^*(1.5, 1) = -0.5$
	−1.0	0.5	$(0.5)^2 + 2(-1.0)^2 =$	2.25		
1.0	0.5	1.5	$(1.5)^2 + 2(0.5)^2 =$	2.75		
	0.0	1.0	$(1.0)^2 + 2(0.0)^2 =$	1.00		
	−0.5	0.5	$(0.5)^2 + 2(-0.5)^2 =$	0.75	$J_{12}^*(1.0) = 0.75$	$u^*(1.0, 1) = -0.5$
	−1.0	0.0	$(0.0)^2 + 2(-1.0)^2 =$	2.00		
0.5	1.0	1.5	$(1.5)^2 + 2(1.0)^2 =$	4.25		
	0.5	1.0	$(1.0)^2 + 2(0.5)^2 =$	1.50		
	0.0	0.5	$(0.5)^2 + 2(0.0)^2 =$	0.25	$J_{12}^*(0.5) = 0.25$	$u^*(0.5, 1) = 0.0$
	−0.5	0.0	$(0.0)^2 + 2(-0.5)^2 =$	0.50		
0.0	1.0	1.0	$(1.0)^2 + 2(1.0)^2 =$	3.00		
	0.5	0.5	$(0.5)^2 + 2(0.5)^2 =$	0.75		
	0.0	0.0	$(0.0)^2 + 2(0.0)^2 =$	0.00	$J_{12}^*(0.0) = 0.00$	$u^*(0.0, 1) = 0.0$

حل تحلیلی:

$$J_{12}^{*}\left(x(1)\right)=\min_{u(1)} C_{12}^{*}\left(x(1),u(1)\right)\longrightarrow$$

$$2\left(x(1)+u(1)\right)+4u(1)=0\Rightarrow u^{*}(x(1),1)=-\frac{1}{3}x(1)$$

$$J_{12}^{*}\left(x(1)\right)=\left(x(1)-\frac{1}{3}x(1)\right)^{2}+\frac{2}{9}x^{2}(1)=\frac{2}{3}x^{2}(1)$$

بهینه سازی برای قدم دوم از انتها (قدم اول از ابتدا)

$$C_{02}^{*}(x(0),u(0))=J_{01}(x(0),u(0))+J_{12}^{*}(x(1)),$$

$$J_{02}^{*}(x(0))=\min_{u(0)}\left[J_{01}(x(0),u(0))+J_{12}^{*}(x(1))\right],$$

- $C_{02}^{*}(x(0),u(0))$ is the minimum cost of operation over the last two stages for one quantized value of $x(0)$ *given a particular trial quantized value of* $u(0)$.
- $J_{01}(x(0),u(0))$ is the cost of operation in the interval $k=0$ to $k=1$ for specified quantized values of $x(0)$ and $u(0)$.
- $J_{12}^{*}(x(1))$ is the cost of the optimal last-stage trajectory which is a function of the state $x(1)$.
- $J_{02}^{*}(x(0))$ is the minimum cost of operation over the last two stages for a specified quantized value of $x(0)$.

Current state $x(0)$	*Control* $u(0)$	*Next state* $x(1)=x(0)+u(0)$	*Minimum cost over last two stages for trial value* $u(0)$: $J_{01}(x(0),u(0))+J_{12}^{*}(x(1))=2u^2(0)+J_{12}^{*}(x(1))=C_{02}^{*}(x(0),u(0))$	*Minimum cost over last two stages* $J_{02}^{*}(x(0))$	*Optimal control applied at* $k=0$ $u^{*}(x(0),0)$
1.5	0.0	1.5	$2(0.0)^2+1.50=1.50$		
	−0.5	1.0	$2(-0.5)^2+0.75=1.25$	$J_{02}^{*}(1.5)=1.25$	$u^{*}(1.5,0)=-0.5$
	−1.0	0.5	$2(-1.0)^2+0.25=2.25$		
1.0	0.5	1.5	$2(0.5)^2+1.50=2.00$		
	0.0	1.0	$2(0.0)^2+0.75=0.75$	$J_{02}^{*}(1.0)=\begin{Bmatrix}0.75\\0.75\end{Bmatrix}$	$u^{*}(1.0,0)=\begin{Bmatrix}0.0\\-0.5\end{Bmatrix}$
	−0.5	0.5	$2(-0.5)^2+0.25=0.75$		
	−1.0	0.0	$2(-1.0)^2+0.00=2.00$		
0.5	1.0	1.5	$2(1.0)^2+1.50=3.50$		
	0.5	1.0	$2(0.5)^2+0.75=1.25$		
	0.0	0.5	$2(0.0)^2+0.25=0.25$	$J_{02}^{*}(0.5)=0.25$	$u^{*}(0.5,0)=0.0$
	−0.5	0.0	$2(-0.5)^2+0.00=0.50$		
0.0	1.0	1.0	$2(1.0)^2+0.75=2.75$		
	0.5	0.5	$2(0.5)^2+0.25=0.75$		
	0.0	0.0	$2(0.0)^2+0.00=0.00$	$J_{02}^{*}(0.0)=0.00$	$u^{*}(0.0,0)=0.0$

حل تحلیلی:

$$C_{02}^*\left(x(0),u(0)\right)=J_{01}\left(x(0),u(0)\right)+J_{12}^*\left(x(1)\right)$$
$$=2u^2(0)+\frac{2}{3}x^2(1)=2u^2(0)+\frac{2}{3}\left(x(0)+u(0)\right)^2$$

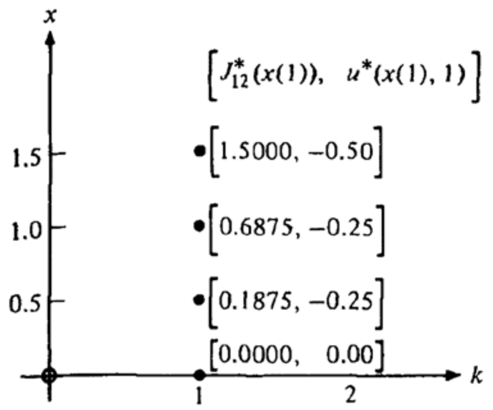
$$J_{02}^*\left(x(0)\right)=\min_{u(0)}C_{02}^*\left(x(0),u(0)\right)\longrightarrow$$
$$4u(0)+\frac{4}{3}\left(x(0)+u(0)\right)=0\Rightarrow u^*(x(0),0)=-\frac{4}{16}x(0)=-\frac{1}{4}x(0)$$
$$J_{02}^*\left(x(0)\right)=\frac{1}{8}x^2(0)+\frac{2}{3}\left(x(0)-\frac{1}{4}x(0)\right)^2=\frac{4}{8}x^2(0)=\frac{1}{2}x^2(0)$$

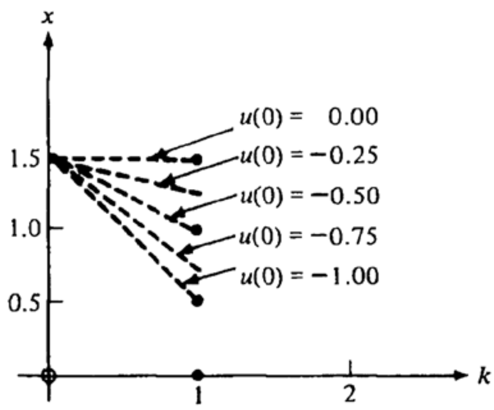
اگر هرچند مرحله دیگر بود:

$$C_{kN}^*(x(k),u(k))=J_{k,k+1}(x(k),u(k))+J_{k+1,N}^*(x(k+1)),$$
$$J_{kN}^*(x(k))=\min_{u(k)}\left[C_{kN}^*(x(k),u(k))\right].$$

میانیابی

اگر دقیقا در نقاط x(k) داده بوجود نیامد باید برای $u^*(x(k),k)$ و $J_{kN}^*\left(x(k)\right)$ میانیابی کرد





Current state $x(0)$	Control $u(0)$	Next state $x(1)=x(0)+u(0)$	Minimum cost over last two stages for trial value $u(0)$: $J_{01}(x(0),u(0))+J_{12}^*(x(1))=2u^2(0)+J_{12}^*(x(1))=C_{02}^*(x(0),u(0))$	Minimum cost over last two stages $J_{02}^*(x(0))$	Optimal control applied at $k=0$ $u^*(x(0),0)$
1.50	0.00	1.50	$2(0.00)^2+1.50000=1.50000$		
	−0.25	1.25	$2(-0.25)^2+1.09375=1.21875$		
	−0.50	1.00	$2(-0.50)^2+0.68750=1.18750$	$J_{02}^*(1.5)=1.18750$	$u^*(1.5,0)=-0.50$
	−0.75	0.75	$2(-0.75)^2+0.43750=1.56250$		
	−1.00	0.50	$2(-1.00)^2+0.18750=2.18750$		

$$J_{12}^*(1.25)=0.68750+\tfrac{1}{2}[1.50000-0.68750]=1.09375$$

$$J_{12}^*(0.75)=0.18750+\tfrac{1}{2}[0.68750-0.18750]=0.43750$$

کنترل بهینه با برنامه ریزی دینامیکی برای سیستم گسسته

اگر سیستم پیوسته بود:

$$\dot{\mathbf{x}}(t) = \mathbf{a}(\mathbf{x}(t), \mathbf{u}(t)).$$

$$J = h(\mathbf{x}(t_f)) + \int_0^{t_f} g(\mathbf{x}(t), \mathbf{u}(t))\, dt,$$

با هر روش انتگرال گیری عددی می توان به سیستم گسسته تبدیل کرد.

مثلا مرتبه ۱، برای سیستم

$$\frac{\mathbf{x}(t + \Delta t) - \mathbf{x}(t)}{\Delta t} \approx \mathbf{a}(\mathbf{x}(t), \mathbf{u}(t))$$

$$\mathbf{x}(t + \Delta t) = \mathbf{x}(t) + \Delta t\, \mathbf{a}(\mathbf{x}(t), \mathbf{u}(t)).$$

$$\mathbf{x}(k + 1) = \mathbf{x}(k) + \Delta t\, \mathbf{a}(\mathbf{x}(k), \mathbf{u}(k)),$$

$$\mathbf{x}(k + 1) \triangleq \mathbf{a}_D(\mathbf{x}(k), \mathbf{u}(k)).$$

و تابع هزینه:

$$J = h(\mathbf{x}(N\, \Delta t)) + \int_0^{\Delta t} g\, dt + \int_{\Delta t}^{2\,\Delta t} g\, dt + \cdots + \int_{(N-1)\,\Delta t}^{N\,\Delta t} g\, dt,$$

$$J \approx h(\mathbf{x}(N)) + \Delta t \sum_{k=0}^{N-1} g(\mathbf{x}(k), \mathbf{u}(k)),$$

$$J = h(\mathbf{x}(N)) + \sum_{k=0}^{N-1} g_D(\mathbf{x}(k), \mathbf{u}(k)).$$

هزینه قدم آخر:

$$J_{NN}(\mathbf{x}(N)) \triangleq h(\mathbf{x}(N));$$

هزینه یک قدم مونده به آخر:

$$\begin{aligned} J_{N-1,\,N}(\mathbf{x}(N-1), \mathbf{u}(N-1)) &\triangleq g_D(\mathbf{x}(N-1), \mathbf{u}(N-1)) + h(\mathbf{x}(N)) \\ &= g_D(\mathbf{x}(N-1), \mathbf{u}(N-1)) + J_{NN}(\mathbf{x}(N)) \\ &= g_D(\mathbf{x}(N-1), \mathbf{u}(N-1)) \\ &\quad + J_{NN}(\mathbf{a}_D(\mathbf{x}(N-1), \mathbf{u}(N-1))) \end{aligned}$$

بهینه سازی:

$$J^*_{N-1,\,N}(\mathbf{x}(N-1)) \triangleq \min_{\mathbf{u}(N-1)} \{g_D(\mathbf{x}(N-1), \mathbf{u}(N-1)) + J_{NN}(\mathbf{a}_D(\mathbf{x}(N-1), \mathbf{u}(N-1)))\}$$

یک بهینه سازی غیر خطی (عملا فقط بهینه سازی m پارامتری است و اگر مثل جداول عمل شود می توان گفت روش grid search یا به عبارتی brute-force search است)

به همین ترتیب برای هزینه دو قدم مونده به آخر:

$$
\begin{aligned}
&J_{N-2,N}(\mathbf{x}(N-2),\mathbf{u}(N-2),\mathbf{u}(N-1)) \\
&\quad = g_D(\mathbf{x}(N-2),\mathbf{u}(N-2)) + g_D(\mathbf{x}(N-1),\mathbf{u}(N-1)) + h(\mathbf{x}(N)) \\
&\quad = g_D(\mathbf{x}(N-2),\mathbf{u}(N-2)) + J_{N-1,N}(\mathbf{x}(N-1),\mathbf{u}(N-1)),
\end{aligned}
$$

بهینه سازی (اگر قرار بود دو قدم مسیر با هم بهینه سازی شود)

$$
J^*_{N-2,N}(\mathbf{x}(N-2)) \triangleq \min_{\mathbf{u}(N-2),\,\mathbf{u}(N-1)} \{g_D(\mathbf{x}(N-2),\mathbf{u}(N-2)) + J_{N-1,N}(\mathbf{x}(N-1),\mathbf{u}(N-1))\}
$$

بهینه سازی (در برنامه ریزی دینامیکی که قرار است فقط یک قدم از مسیر بهینه سازی شود)

$$
\begin{aligned}
J^*_{N-2,N}(\mathbf{x}(N-2)) &= \min_{\mathbf{u}(N-2)} \{g_D(\mathbf{x}(N-2),\mathbf{u}(N-2)) + J^*_{N-1,N}(\mathbf{x}(N-1))\}. \\
&= \min_{\mathbf{u}(N-2)} \{g_D(\mathbf{x}(N-2),\mathbf{u}(N-2)) + J^*_{N-1,N}(\mathbf{a}_D(\mathbf{x}(N-2),\mathbf{u}(N-2)))\}
\end{aligned}
$$

به همین ترتیب برای سایر مراحل:

بهینه سازی برای کل مسیر

$$
J^*_{N-K,N}(\mathbf{x}(N-K)) = \min_{\mathbf{u}(N-K),\,\mathbf{u}(N-K+1),\,\ldots,\,\mathbf{u}(N-1)} \left\{h(\mathbf{x}(N)) + \sum_{k=N-K}^{N-1} g_D(\mathbf{x}(k),\mathbf{u}(k))\right\}
$$

بهینه سازی یک قدم (برنامه ریزی دینامیکی)

$$
\begin{aligned}
J^*_{N-K,N}(\mathbf{x}(N-K)) = \min_{\mathbf{u}(N-K)} &\{g_D(\mathbf{x}(N-K),\mathbf{u}(N-K)) \\
&+ J^*_{N-(K-1),N}(\mathbf{a}_D(\mathbf{x}(N-K),\mathbf{u}(N-K)))\}
\end{aligned}
$$

که برای $k=1,2,\ldots,N$ قابل استفاده است. برای شروع:

$$
J^*_{NN}(\mathbf{x}(N)) = h(\mathbf{x}(N))
$$

مزایا و معایب برنامه ریزی دینامیکی

مزایا:

- بهینه مطلق (Global/Absolute Optimum) را می دهد (عملا تمام فضا را جستجو می کند)
- اعمال قیود مسیر و کنترل ساده است
- فرم بسته از جواب می دهد

$$\vec{u}^{*} = \vec{u}^{*}(\vec{x}(t), t)$$

- نسبت به اینکه تمام حالت ها بررسی شود کم هزینه تر است

برای سیستم مرتبه ۱ با ۱۰ عدد quantization برای x و 4 تا برای u

Number of stages in the process N	*Number of calculations required by dynamic programming*	*Number of calculations required by direct enumeration*	*Number of calculations required by direct enumeration (assuming 50% of state values admissible and distinct)*
1	40	40	40
2	80	200	120
3	120	840	280
4	160	3,400	600
5	200	13,640	1,240
6	240	54,600	2,520
L	$40L$	$\sum_{k=1}^{L} [10 \cdot 4^k]$	$\sum_{k=1}^{L} [20 \cdot 2^k]$

ایراد بزرگ:

- هنوز هم مشکل ابعادی دارد (Curse of dimensioanlity):
  - محاسبات اولیه (امکان دارد off-line انجام شود)
  - نگهداری جدول محاسبات
    - نیاز به فضای بزرگ داده
    - بزرگ کردن فواصل quantization و استفاده از میانیابی

کنترل بهینه سیستم خطی گسسته با تابع هزینه درجه ۲

Discrete Linear Quadratic Regulator Problem (Discrete LQR)

سیستم خطی

$$\mathbf{x}(k+1) = \mathbf{A}(k)\mathbf{x}(k) + \mathbf{B}(k)\mathbf{u}(k).$$

تابع هزینه درجه ۲

$$J = \tfrac{1}{2}\mathbf{x}^T(N)\mathbf{H}\mathbf{x}(N) + \tfrac{1}{2}\sum_{k=0}^{N-1}[\mathbf{x}^T(k)\mathbf{Q}(k)\mathbf{x}(k) + \mathbf{u}^T(k)\mathbf{R}(k)\mathbf{u}(k)],$$

که

$\mathbf{H}$ and $\mathbf{Q}(k)$ are real symmetric positive semi-definite $n \times n$ matrices.
$\mathbf{R}(k)$ is a real symmetric positive definite $m \times m$ matrix.
N is a fixed integer greater than 0.

هزینه مرحله نهایی:

$$J_{NN}(\mathbf{x}(N)) = \tfrac{1}{2}\mathbf{x}^T(N)\mathbf{H}\mathbf{x}(N) = J^*_{NN}(\mathbf{x}(N)) \triangleq \tfrac{1}{2}\mathbf{x}^T(N)\mathbf{P}(0)\mathbf{x}(N)$$

هزینه یک قدم مانده به آخر

$$\begin{aligned} J_{N-1,N}(\mathbf{x}(N-1), \mathbf{u}(N-1)) &= \tfrac{1}{2}\mathbf{x}^T(N-1)\mathbf{Q}\mathbf{x}(N-1) \\ &+ \tfrac{1}{2}\mathbf{u}^T(N-1)\mathbf{R}\mathbf{u}(N-1) + \tfrac{1}{2}\mathbf{x}^T(N)\mathbf{P}(0)\mathbf{x}(N), \end{aligned}$$

که باید بهینه شود

$$\begin{aligned} J^*_{N-1,N}(\mathbf{x}(N-1)) &\triangleq \min_{\mathbf{u}(N-1)} \{J_{N-1,N}(\mathbf{x}(N-1), \mathbf{u}(N-1))\} \\ &= \min_{\mathbf{u}(N-1)} \{\tfrac{1}{2}\mathbf{x}^T(N-1)\mathbf{Q}\mathbf{x}(N-1) + \tfrac{1}{2}\mathbf{u}^T(N-1)\mathbf{R}\mathbf{u}(N-1) \\ &\quad + \tfrac{1}{2}[\mathbf{A}\mathbf{x}(N-1) + \mathbf{B}\mathbf{u}(N-1)]^T\mathbf{P}(0)[\mathbf{A}\mathbf{x}(N-1) + \mathbf{B}\mathbf{u}(N-1)]\} \end{aligned}$$

برای بهینه سازی:

$$\begin{bmatrix} \dfrac{\partial J_{N-1,N}}{\partial u_1(N-1)} \\ \dfrac{\partial J_{N-1,N}}{\partial u_2(N-1)} \\ \vdots \\ \dfrac{\partial J_{N-1,N}}{\partial u_m(N-1)} \end{bmatrix} \triangleq \frac{\partial J_{N-1,N}}{\partial \mathbf{u}(N-1)} = \mathbf{0}.$$

$$\mathbf{Ru}(N-1) + \mathbf{B}^T\mathbf{P}(0)[\mathbf{Ax}(N-1) + \mathbf{Bu}(N-1)] = \mathbf{0}$$

$$\begin{aligned}\mathbf{u}^*(N-1) &= -[\mathbf{R} + \mathbf{B}^T\mathbf{P}(0)\mathbf{B}]^{-1}\mathbf{B}^T\mathbf{P}(0)\mathbf{Ax}(N-1) \\ &\triangleq \mathbf{F}(N-1)\mathbf{x}(N-1)\end{aligned}$$

مشابه فیدبک خطی از x شد

این u حتما مینیمم می کند چون مشتق دوم (هشن) مثبت معین است:

$$\frac{\partial^2 J_{N-1,N}}{\partial \mathbf{u}^2(N-1)} = \mathbf{R} + \mathbf{B}^T\mathbf{P}(0)\mathbf{B}$$

مقدار $J^*_{N-1,N}$

$$\begin{aligned}J^*_{N-1,N}(\mathbf{x}(N-1)) &= \tfrac{1}{2}\mathbf{x}^T(N-1)\{[\mathbf{A} + \mathbf{BF}(N-1)]^T\mathbf{P}(0)[\mathbf{A} + \mathbf{BF}(N-1)] \\ &\quad + \mathbf{F}^T(N-1)\mathbf{RF}(N-1) + \mathbf{Q}\}\mathbf{x}(N-1) \\ &\triangleq \tfrac{1}{2}\mathbf{x}^T(N-1)\mathbf{P}(1)\mathbf{x}(N-1)\end{aligned}$$

چون شکل آن شبیه $J^*_{N,N}$ شد، همان مسیر قبلی طی شود، نتیجه زیر حاصل می شود

$$\begin{aligned}\mathbf{u}^*(N-2) &= -[\mathbf{R} + \mathbf{B}^T\mathbf{P}(1)\mathbf{B}]^{-1}\mathbf{B}^T\mathbf{P}(1)\mathbf{Ax}(N-2) \\ &\triangleq \mathbf{F}(N-2)\mathbf{x}(N-2)\end{aligned}$$

$$\begin{aligned}J^*_{N-2,N}(\mathbf{x}(N-2)) &= \tfrac{1}{2}\mathbf{x}^T(N-2)\{[\mathbf{A} + \mathbf{BF}(N-2)]^T\mathbf{P}(1)[\mathbf{A} + \mathbf{BF}(N-2)] \\ &\quad + \mathbf{F}^T(N-2)\mathbf{RF}(N-2) + \mathbf{Q}\}\mathbf{x}(N-2) \\ &\triangleq \tfrac{1}{2}\mathbf{x}^T(N-2)\mathbf{P}(2)\mathbf{x}(N-2)\end{aligned}$$

و با استقرا

$$\begin{aligned}\mathbf{u}^*(N-K) &= -[\mathbf{R} + \mathbf{B}^T\mathbf{P}(K-1)\mathbf{B}]^{-1}\mathbf{B}^T\mathbf{P}(K-1)\mathbf{Ax}(N-K) \\ &\triangleq \mathbf{F}(N-K)\mathbf{x}(N-K)\end{aligned}$$

$$\begin{aligned}J^*_{N-K,N}(\mathbf{x}(N-K)) &= \tfrac{1}{2}\mathbf{x}^T(N-K)\{[\mathbf{A} + \mathbf{BF}(N-K)]^T\mathbf{P}(K-1)[\mathbf{A} + \mathbf{BF}(N-K)] \\ &\quad + \mathbf{F}^T(N-K)\mathbf{RF}(N-K) + \mathbf{Q}\}\mathbf{x}(N-K) \\ &\triangleq \tfrac{1}{2}\mathbf{x}^T(N-K)\mathbf{P}(K)\mathbf{x}(N-K).\end{aligned}$$

اگر ماتریس های تابع زمان باشند:

$$\begin{aligned}\mathbf{u}^*(N-K) &= -[\mathbf{R}(N-K) + \mathbf{B}^T(N-K)\mathbf{P}(K-1)\mathbf{B}(N-K)]^{-1} \\ &\quad \times \mathbf{B}^T(N-K)\mathbf{P}(K-1)\mathbf{A}(N-K)\mathbf{x}(N-K) \\ &\triangleq \mathbf{F}(N-K)\mathbf{x}(N-K)\end{aligned}$$

$$\begin{aligned} J^*_{N-K,N}(\mathbf{x}(N-K)) = &\tfrac{1}{2}\mathbf{x}^T(N-K)\{[\mathbf{A}(N-K) \\ &+\mathbf{B}(N-K)\mathbf{F}(N-K)]^T \\ &\times\mathbf{P}(K-1)[\mathbf{A}(N-K)+\mathbf{B}(N-K)\mathbf{F}(N-K)] \\ &+\mathbf{F}^T(N-K)\mathbf{R}(N-K)\mathbf{F}(N-K) \\ &+\mathbf{Q}(N-K)\}\mathbf{x}(N-K) \\ \triangleq &\tfrac{1}{2}\mathbf{x}^T(N-K)\mathbf{P}(K)\mathbf{x}(N-K) \end{aligned}$$

هزینه کل:

$$J^*_{0,N}(\mathbf{x}_0) = \tfrac{1}{2}\mathbf{x}_0^T\mathbf{P}(N)\mathbf{x}_0$$

برای محاسبه فیدبک:

$$\begin{aligned} \mathbf{F}(N-K) = &-[\mathbf{R}(N-K) + \mathbf{B}^T(N-K)\mathbf{P}(K-1)\mathbf{B}(N-K)]^{-1} \\ &\times \mathbf{B}^T(N-K)\mathbf{P}(K-1)\mathbf{A}(N-K) \end{aligned}$$

$$\begin{aligned} \mathbf{P}(K) = &[\mathbf{A}(N-K) + \mathbf{B}(N-K)\mathbf{F}(N-K)]^T\mathbf{P}(K-1) \\ &\times [\mathbf{A}(N-K) + \mathbf{B}(N-K)\mathbf{F}(N-K)] \\ &+ \mathbf{F}^T(N-K)\mathbf{R}(N-K)\mathbf{F}(N-K) + \mathbf{Q}(N-K) \end{aligned}$$

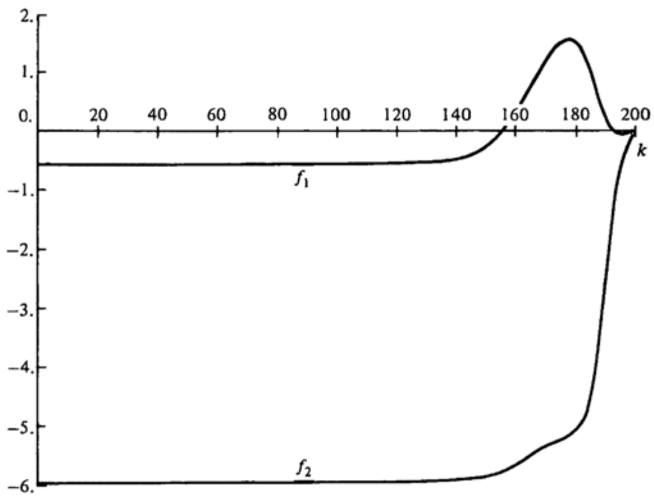
با $\mathbf{P}(0) = \mathbf{H}$

برای سادگی، با تعریف:

$$\mathbf{V}(N-K) \triangleq \mathbf{A}(N-K) + \mathbf{B}(N-K)\mathbf{F}(N-K)$$

$$\begin{aligned} \mathbf{P}(K) = &\mathbf{V}^T(N-K)\mathbf{P}(K-1)\mathbf{V}(N-K) \\ &+ \mathbf{F}^T(N-K)\mathbf{R}(N-K)\mathbf{F}(N-K) + \mathbf{Q}(N-K) \end{aligned}$$

مشابه قبل، برای سیستم ثابت با زمان، ماتریس فیدبک به مقدار خاصی همگرا می شود



معادلات همیلتون-ژاکوبی-بلمن

Hamilton-Jacobi-Bellman Equation

کنترل بهینه با برنامه ریزی دینامیکی برای سیستم پیوسته غیرخطی

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

$$J = h(\mathbf{x}(t_f), t_f) + \int_{t_0}^{t_f} g(\mathbf{x}(\tau), \mathbf{u}(\tau), \tau)\, d\tau$$

طبق اصل بلمن:

$$J(\mathbf{x}(t), t, \underset{t \le \tau \le t_f}{\mathbf{u}(\tau)}) = h(\mathbf{x}(t_f), t_f) + \int_{t}^{t_f} g(\mathbf{x}(\tau), \mathbf{u}(\tau), \tau)\, d\tau$$

اگر به صورت بهینه سازی کلی دیده شود (مانند grid search برای کل u ها)

$$\begin{aligned} J^*(\mathbf{x}(t), t) &= \min_{\substack{\mathbf{u}(\tau) \\ t \le \tau \le t_f}} \left\{ \int_t^{t_f} g(\mathbf{x}(\tau), \mathbf{u}(\tau), \tau)\, d\tau + h(\mathbf{x}(t_f), t_f) \right\} \\ &= \min_{\substack{\mathbf{u}(\tau) \\ t \le \tau \le t_f}} \left\{ \int_t^{t+\Delta t} g\, d\tau + \int_{t+\Delta t}^{t_f} g\, d\tau + h(\mathbf{x}(t_f), t_f) \right\} \end{aligned}$$

ولی هدف نوشتن به فرم برنامه ریزی دینامیکی است:

$$\begin{aligned} J^*(\mathbf{x}(t), t) = &\min_{\substack{\mathbf{u}(\tau) \\ t \le \tau \le t+\Delta t}} \left\{ \int_t^{t+\Delta t} g\, d\tau + J^*(\mathbf{x}(t+\Delta t), t+\Delta t) \right\} \\ = &\min_{\substack{\mathbf{u}(\tau) \\ t \le \tau \le t+\Delta t}} \left\{ \int_t^{t+\Delta t} g\, d\tau + J^*(\mathbf{x}(t), t) + \left[\frac{\partial J^*}{\partial t}(\mathbf{x}(t), t) \right] \Delta t \right. \\ &+ \left[\frac{\partial J^*}{\partial \mathbf{x}}(\mathbf{x}(t), t) \right]^T \left[\mathbf{x}(t+\Delta t) - \mathbf{x}(t) \right] \\ &\left. + \text{ terms of higher order} \right\} \end{aligned}$$

برای Δt دیفرانسیلی (بسیار کوچک)

$$\begin{aligned} J^*(\mathbf{x}(t), t) = \min_{\mathbf{u}(t)} \{ &g(\mathbf{x}(t), \mathbf{u}(t), t)\, \Delta t + J^*(\mathbf{x}(t), t) \\ &+ J_t^*(\mathbf{x}(t), t)\, \Delta t + J_{\mathbf{x}}^{*T}(\mathbf{x}(t), t)[\mathbf{a}(\mathbf{x}(t), \mathbf{u}(t), t)]\, \Delta t \\ &+ o(\Delta t) \},\dagger \end{aligned}$$

طبق تعریف:

$$J_{\mathbf{x}}^* \triangleq \frac{\partial J^*}{\partial \mathbf{x}} = \left[\frac{\partial J^*}{\partial x_1} \quad \frac{\partial J^*}{\partial x_2} \quad \cdots \quad \frac{\partial J^*}{\partial x_n} \right]^T \quad \text{and} \quad J_t^* \triangleq \frac{\partial J^*}{\partial t}$$

ساده سازی روابط:

$$0 = J_t^*(\mathbf{x}(t), t)\,\Delta t + \min_{\mathbf{u}(t)} \{g(\mathbf{x}(t), \mathbf{u}(t), t)\,\Delta t + J_{\mathbf{x}}^{*T}(\mathbf{x}(t), t)[\mathbf{a}(\mathbf{x}(t), \mathbf{u}(t), t)]\,\Delta t + o(\Delta t)\}$$

نهایتا می شود معادلات PDE (Partial Differential Equation) زیر

$$0 = J_t^*(\mathbf{x}(t), t) + \min_{\mathbf{u}(t)} \{g(\mathbf{x}(t), \mathbf{u}(t), t) + J_{\mathbf{x}}^{*T}(\mathbf{x}(t), t)[\mathbf{a}(\mathbf{x}(t), \mathbf{u}(t), t)]\}$$

به همراه شرط مرزی:

$$J^*(\mathbf{x}(t_f), t_f) = h(\mathbf{x}(t_f), t_f)$$

با تعریف همیلتونین (Hamiltonian)

$$\mathscr{H}(\mathbf{x}(t), \mathbf{u}(t), J_{\mathbf{x}}^*, t) \triangleq g(\mathbf{x}(t), \mathbf{u}(t), t) + J_{\mathbf{x}}^{*T}(\mathbf{x}(t), t)[\mathbf{a}(\mathbf{x}(t), \mathbf{u}(t), t)]$$

$$\mathscr{H}(\mathbf{x}(t), \mathbf{u}^*(\mathbf{x}(t), J_{\mathbf{x}}^*, t), J_{\mathbf{x}}^*, t) = \min_{\mathbf{u}(t)} \mathscr{H}(\mathbf{x}(t), \mathbf{u}(t), J_{\mathbf{x}}^*, t),$$

معادله Hamilton-Jacobi که چون معادله بلمن برای سیستم های پیوسته است به عنوان معادله Hamilton-Jacobi-Bellman شناخته می شود:

$$0 = J_t^*(\mathbf{x}(t), t) + \mathscr{H}(\mathbf{x}(t), \mathbf{u}^*(\mathbf{x}(t), J_{\mathbf{x}}^*, t), J_{\mathbf{x}}^*, t)$$

مثال:

سیستم:

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

تابع هزینه:

$$J = \tfrac{1}{4}x^2(T) + \int_0^T \tfrac{1}{4}u^2(t)\,dt$$

همیلتونین:

$$\mathscr{H}(x(t), u(t), J_x^*, t) = \tfrac{1}{4}u^2(t) + J_x^*[x(t) + u(t)]$$

بهینه کردن همیلتونین، شرط لازم

$$\frac{\partial \mathscr{H}}{\partial u} = \frac{1}{2}u(t) + J_x^*(x(t), t) = 0 \qquad \Rightarrow \qquad u^*(t) = -2J_x^*(x(t), t)$$

شرط کافی

$$\frac{\partial^2 \mathscr{H}}{\partial u^2} = \frac{1}{2} > 0$$

معادله HJB

$$0 = J_t^* + \tfrac{1}{4}[-2J_x^*]^2 + [J_x^*]x(t) - 2[J_x^*]^2$$
$$= J_t^* - [J_x^*]^2 + [J_x^*]x(t)$$

با شرط مرزی

$$J^*(x(T), T) = \tfrac{1}{4}x^2(T)$$

فرض می شود که شکل جواب اینچنین است (با دانستن شکل جواب های مشابه و ...)

$$J^*(x(t), t) = \tfrac{1}{2}K(t)x^2(t)$$

بعد از جایگذاری:

$$J_x^*(x(t), t) = K(t)x(t)$$

$$u^*(t) = -2K(t)x(t)$$

$$J_t^*(x(t), t) = \tfrac{1}{2}\dot{K}(t)x^2(t)$$

معادله HJB

$$0 = \tfrac{1}{2}\dot{K}(t)x^2(t) - K^2(t)x^2(t) + K(t)x^2(t) \quad \Rightarrow \quad \tfrac{1}{2}\dot{K}(t) - K^2(t) + K(t) = 0$$

شرط مرزی:

$$K(T) = \tfrac{1}{2}$$

جواب:

$$K(t) = \frac{\epsilon^{(T-t)}}{\epsilon^{(T-t)} + \epsilon^{-(T-t)}}$$

مساله LQR سیستم پیوسته

Continuous Linear Quadratic Regulator Problem (LQR)

سیستم زیر:

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

$$J = \tfrac{1}{2}\mathbf{x}^T(t_f)\mathbf{H}\mathbf{x}(t_f) + \int_{t_0}^{t_f} \tfrac{1}{2}[\mathbf{x}^T(t)\mathbf{Q}(t)\mathbf{x}(t) + \mathbf{u}^T(t)\mathbf{R}(t)\mathbf{u}(t)]\, dt$$

همیلتونین:

$$\mathscr{H}(\mathbf{x}(t), \mathbf{u}(t), J_{\mathbf{x}}^*, t) = \tfrac{1}{2}\mathbf{x}^T(t)\mathbf{Q}(t)\mathbf{x}(t) + \tfrac{1}{2}\mathbf{u}^T(t)\mathbf{R}(t)\mathbf{u}(t) + J_{\mathbf{x}}^{*T}(\mathbf{x}(t), t)\cdot[\mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)]$$

بهینه سازی همیلتونین، شرط لازم

$$\frac{\partial \mathscr{H}}{\partial \mathbf{u}}(\mathbf{x}(t), \mathbf{u}(t), J_{\mathbf{x}}^*, t) = \mathbf{R}(t)\mathbf{u}(t) + \mathbf{B}^T(t)J_{\mathbf{x}}^*(\mathbf{x}(t), t) = \mathbf{0}$$

$$\mathbf{u}^*(t) = -\mathbf{R}^{-1}(t)\mathbf{B}^T(t)J_{\mathbf{x}}^*(\mathbf{x}(t), t)$$

شرط کافی:

$$\frac{\partial^2 \mathscr{H}}{\partial \mathbf{u}^2} = \mathbf{R}(t)$$

جایگذاری در همیلتونین:

$$\begin{aligned}\mathscr{H}(\mathbf{x}(t), \mathbf{u}^*(t), J_{\mathbf{x}}^*, t) &= \tfrac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} + \tfrac{1}{2}J_{\mathbf{x}}^{*T}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T J_{\mathbf{x}}^* \\ &\quad + J_{\mathbf{x}}^{*T}\mathbf{A}\mathbf{x} - J_{\mathbf{x}}^{*T}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T J_{\mathbf{x}}^* \\ &= \tfrac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} - \tfrac{1}{2}J_{\mathbf{x}}^{*T}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T J_{\mathbf{x}}^* + J_{\mathbf{x}}^{*T}\mathbf{A}\mathbf{x}\end{aligned}$$

معادلع HJB

$$0 = J_t^* + \tfrac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} - \tfrac{1}{2}J_{\mathbf{x}}^{*T}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T J_{\mathbf{x}}^* + J_{\mathbf{x}}^{*T}\mathbf{A}\mathbf{x}$$

با شرط مرزی

$$J^*(\mathbf{x}(t_f), t_f) = \tfrac{1}{2}\mathbf{x}^T(t_f)\mathbf{H}\mathbf{x}(t_f)$$

فرض می شود جواب به فرم زیر باشد (یه نوع جداسازی متغیرهای)

$$J^*(\mathbf{x}(t), t) = \tfrac{1}{2}\mathbf{x}^T(t)\mathbf{K}(t)\mathbf{x}(t)$$

قراردادن در HJB

$$0 = \tfrac{1}{2}\mathbf{x}^T\dot{\mathbf{K}}\mathbf{x} + \tfrac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} - \tfrac{1}{2}\mathbf{x}^T\mathbf{K}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T\mathbf{K}\mathbf{x} + \mathbf{x}^T\mathbf{K}\mathbf{A}\mathbf{x}$$

ماتریس ترکیب زیر است (جمع بخش متقارن و غیرمتقارن) که با توجه به تقارن بقیه ماتریس ها، فقط بخش متقارن می ماند

$$\mathbf{KA} = \tfrac{1}{2}[\mathbf{KA} + (\mathbf{KA})^T] + \tfrac{1}{2}[\mathbf{KA} - (\mathbf{KA})^T]$$

نهایتا می شود معادله ریکاتی و شرط مرزی زیر:

$$\mathbf{0} = \dot{\mathbf{K}}(t) + \mathbf{Q}(t) - \mathbf{K}(t)\mathbf{B}(t)\mathbf{R}^{-1}(t)\mathbf{B}^T(t)\mathbf{K}(t) + \mathbf{K}(t)\mathbf{A}(t) + \mathbf{A}^T(t)\mathbf{K}(t)$$

$$\mathbf{K}(t_f) = \mathbf{H}$$

نهایتا سیگنال کنترل می شود

$$\mathbf{u}^*(t) = -\mathbf{R}^{-1}(t)\mathbf{B}^T(t)\mathbf{K}(t)\mathbf{x}(t)$$