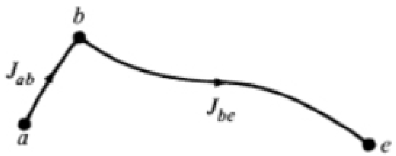
Optimal Control Theory

Dynamic Programming

Principle of Optimality

- Optimal path for a multistage decision process



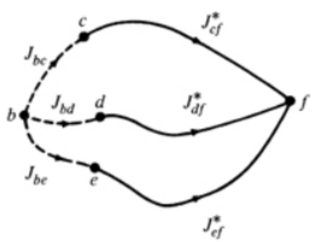
$$J^*_{ae} = J_{ab} + J_{be}$$

If a-b-e is the optimal path from a to e, then b-e is the optimal path from b to e.

➢ Bellman's principle of optimality (1962)

> An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.

(ex) Candidates for optimal paths from b to f



$$C^*_{bcf} = J_{bc} + J^*_{cf}$$
$$C^*_{bdf} = J_{bd} + J^*_{df}$$
$$C^*_{bef} = J_{be} + J^*_{ef}$$

⇨ Optimal decision at point b

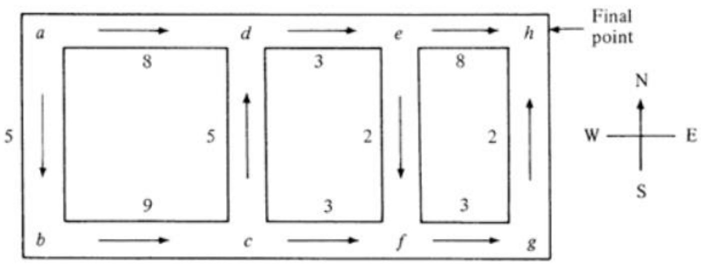
$$\min_{c,d,e} \{C^*_{bcf}, C^*_{bdf}, C^*_{bef}\}$$

➢ Dynamic Programming

Computational procedure which extends the above decision-making concept to sequences of decisions

Principle of Optimality

- Routing Problem



(notation)

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.

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

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

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

$$u^*(c) = f$$

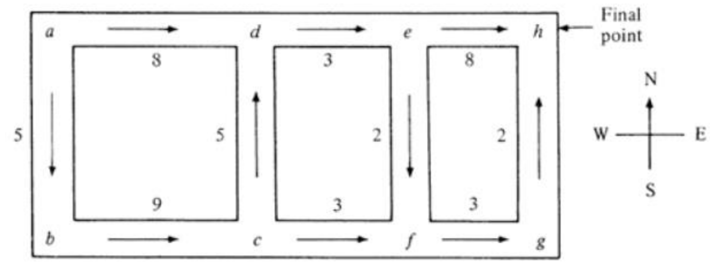
$$\begin{aligned} J^*_{fh} &= J_{fg} + J^*_{gh} \\ &= 3 + 2 \\ &= 5. \end{aligned}$$

$$J^*_{eh} = \min\{J_{eh}, [J_{ef} + J^*_{fh}]\}$$

Calculating backward from h

Dynamic Programming

- Routing Problem



Current intersection α	*Heading* u_i	*Next intersection* x_i	*Minimum cost from* α *to h via* x_i $J_{\alpha x_i} + J^*_{x_i h} = C^*_{\alpha x_i h}$	*Minimum cost to reach h from* α $J^*_{\alpha h}$	*Optimal heading at* α $u^*(\alpha)$
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	$5 + 17 = 22$		

Dynamic Programming

- Optimal Control Problem (scalar)
  - Differential equation

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

$$0.0 \le x(t) \le 1.5 \qquad -1.0 \le u(t) \le 1.0$$

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

  - Difference equation

$$\frac{x(t+\Delta t) - x(t)}{\Delta t} \approx ax(t) + bu(t) \quad \Rightarrow \quad x(t+\Delta t) = [1 + a\,\Delta t]x(t) + b\,\Delta t\,u(t)$$

$$t = 0, \Delta t, \ldots, (N-1)\,\Delta t \qquad t = k\,\Delta t,$$

$$x(k) \equiv x(k\Delta t)$$

$$\Rightarrow \quad \boxed{x(k+1) = [1 + a\,\Delta t]x(k) + b\,\Delta t\,u(k)}$$

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

$$\Rightarrow \quad \boxed{\begin{aligned} J &= x^2(N) + \lambda\,\Delta t[u^2(0) + u^2(1) + \cdots + u^2(N-1)] \\ &= x^2(N) + \lambda\,\Delta t \sum_{k=0}^{N-1} u^2(k). \end{aligned}}$$

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

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

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

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

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

Dynamic Programming

- Optimal Control Problem

$$x(k+1) = x(k) + u(k); \qquad k = 0, 1$$
$$J = x^2(2) + 2u^2(0) + 2u^2(1)$$
$$0.0 \le x(k) \le 1.5; \qquad k = 0, 1, 2$$
$$-1.0 \le u(k) \le 1.0; \qquad k = 0, 1.$$

  – Quantization: $x(k) = 0.0, 0.5, 1.0, 1.5, u(k) = -1.0, -0.5, 0.0, 0.5, 1.0$

(Principle of optimality)

$$C^*_{kN}(x(k), u(k)) = J_{k,k+1}(x(k), u(k)) + J^*_{k+1,N}(x(k+1)),$$
$$J^*_{kN}(x(k)) = \min_{u(k)} [C^*_{kN}(x(k), u(k))].$$

(Backward computation)

$$k = N-1, N-2, \cdots, 2, 1$$

Table 3-2 COSTS OF OPERATION OVER THE LAST STAGE

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	$(0.0)^2 + 2(0.0)^2 = 0.00$	$J^*_{12}(0.0) = 0.00$	$u^*(0.0, 1) = 0.0$

Table 3-3 COSTS OF OPERATION OVER THE LAST TWO STAGES

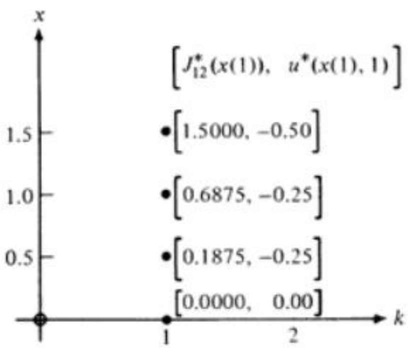
Current state $x(0)$	Control $u(0)$	Next state $x(1) = x(0) + u(0)$	Minimum cost over last two stages for trial value $u(0)$ $J_{01}(x(0), u(0)) + J^*_{12}(x(1)) =$ $2u^2(0) + J^*_{12}(x(1)) = C^*_{02}(x(0), u(0))$	Minimum cost over last two stages $J^*_{02}(x(0))$	Optimal control applied at $k = 0$ $u^*(x(0), 0)$
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	$2(0.0)^2 + 0.00 = 0.00$	$J^*_{02}(0.0) = 0.00$	$u^*(0.0, 0) = 0.0$

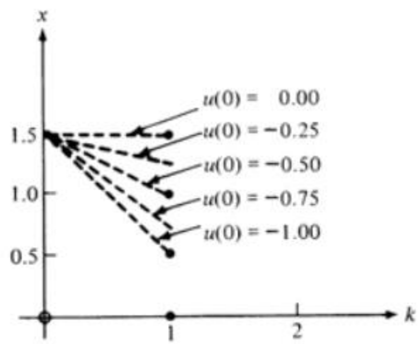
If $x(0) = 1.5$,
=> $u^*(0) = -0.5, u^*(1) = -0.5,\ J^* = 1.25$

Dynamic Programming

- Interpolation

(ex) quantization $x(k) = 0.0, 0.5, 1.0, 1.5,$ $u(k) = -1.0, -0.75, -0.5, -0.25, 0.0, 0.25, 0.5, 0.75, 1.0$





fine grid => increased memory

- Linear interpolation

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

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

Table 3-4 COSTS OF OPERATION OVER THE LAST TWO STAGES FOR $x(0) = 1.50$

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	$2(-1.00)^2 + 0.18750 = 2.18750$		

Dynamic Programming

(Q) **3-4.** The discrete approximation to a nonlinear continuously operating system is given by

$$x(k+1) = x(k) - 0.4x^2(k) + u(k).$$

The state and control values are constrained by

$$0.0 \leq x(k) \leq 1.0$$
$$-0.4 \leq u(k) \leq 0.4.$$

Quantize the state into the levels 0, 0.5, 1, and the control into the levels −0.4, −0.2, 0, 0.2, 0.4. The performance measure to be minimized is

$$J = 4\,|x(2)| + \sum_{k=0}^{1} |u(k)|.$$

(a) Use dynamic programming with linear interpolation to complete the tables shown below.

$x(0)$	$J^*_{0,2}(x(0))$	$u^*(x(0), 0)$
0.0		
0.5		
1.0		

$x(1)$	$J^*_{1,2}(x(1))$	$u^*(x(1), 1)$
0.0		
0.5		
1.0		

(b) From the results of part (a) find the optimal control sequence $\{u^*(0), u^*(1)\}$ and the minimum cost if the initial state is 1.0.

Dynamic Programming

- Optimal Control Problem (vector)
  - State equation

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

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

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

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

  - Performance measure

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

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

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

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

  - Optimal control law

$$\boxed{\mathbf{u}^*(\mathbf{x}(0), 0),\ \mathbf{u}^*(\mathbf{x}(1), 1),\ \ldots,\ \mathbf{u}^*(\mathbf{x}(N - 1), N - 1)}$$

Dynamic Programming

- Recurrence relation

$$\boxed{\begin{aligned} \mathbf{x}(k+1) &\triangleq \mathbf{a}_D(\mathbf{x}(k), \mathbf{u}(k)). \\ J &= h(\mathbf{x}(N)) + \sum_{k=0}^{N-1} g_D(\mathbf{x}(k), \mathbf{u}(k)) \end{aligned}}$$

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

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

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

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

(stage *N-3*) $$J^*_{N-3,N}(\mathbf{x}(N-3)) = \min_{\mathbf{u}(N-3)} \{g_D(\mathbf{x}(N-3), \mathbf{u}(N-3)) + J^*_{N-2,N}(\mathbf{a}_D(\mathbf{x}(N-3), \mathbf{u}(N-3)))\}$$

(stage *N-k*) $$\boxed{J^*_{N-K,N}(\mathbf{x}(N-K)) = \min_{\mathbf{u}(N-K)} \{g_D(\mathbf{x}(N-K), \mathbf{u}(N-K)) + J^*_{N-(K-1),N}(\mathbf{a}_D(\mathbf{x}(N-K), \mathbf{u}(N-K)))\}}$$

Dynamic Programming

- Computational Procedure
  - Recurrence equation

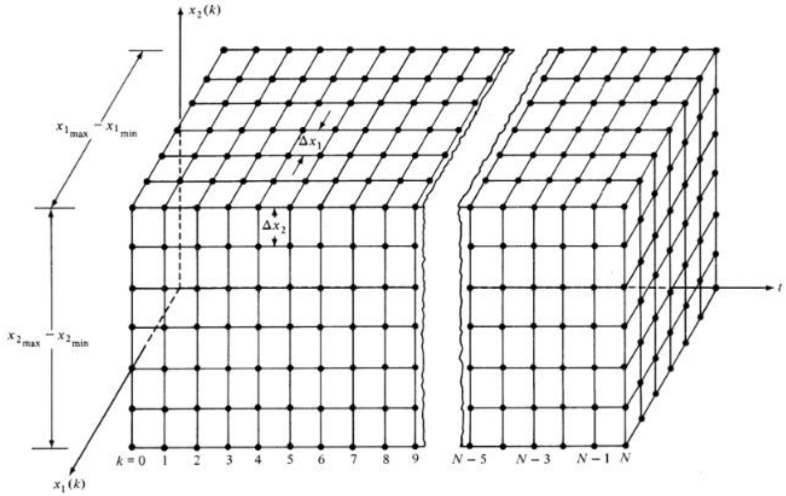
$$J^*_{N-K,N}(\mathbf{x}(N-K)) = \min_{\mathbf{u}(N-K)} \{g_D(\mathbf{x}(N-K), \mathbf{u}(N-K)) + J^*_{N-(K-1),N}(\mathbf{a}_D(\mathbf{x}(N-K), \mathbf{u}(N-K)))\}$$

$$K = 1 \rightarrow 2 \rightarrow \cdots \rightarrow N$$

  - Optimal control

$\mathbf{u}^*(\mathbf{x}(N-K), N-K)$ $\quad K = 1 \rightarrow 2 \rightarrow \cdots \rightarrow N$ $\quad$ Stage (t) = $N-1 \rightarrow \cdots \rightarrow 1 \rightarrow 0$

  - Grid of admissible state values



- Characteristics
  1) global (absolute) minimum
  2) curse of dimensionality

     for high-dimensional system, the storage locations (memory) become prohibitive
  3) offline control law / infinite horizon

Dynamic Programming

- Discrete Linear Regulator Problems – Analytical Results
  - Linear time-varying system

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

  - Performance measure: regulator

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

$\mathbf{H}$ 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.

  - Recurrence equation

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

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

$$= \tfrac{1}{2}\mathbf{x}^T(N-1)\mathbf{Q}\mathbf{x}(N-1) + \tfrac{1}{2}\mathbf{u}^T(N-1)\mathbf{R}\mathbf{u}(N-1) + \tfrac{1}{2}[\mathbf{A}\mathbf{x}(N-1) + \mathbf{B}\mathbf{u}(N-1)]^T\mathbf{P}(0)[\mathbf{A}\mathbf{x}(N-1) + \mathbf{B}\mathbf{u}(N-1)]$$

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

$$\leftrightarrow \quad \frac{\partial J_{N-1,N}}{\partial \mathbf{u}(N-1)} = \mathbf{0} \quad \text{and} \quad \frac{\partial^2 J_{N-1,N}}{\partial \mathbf{u}^2(N-1)} > 0 \text{ (positive definite)}$$

$\because J$ is convex function

$$\leftrightarrow \quad \frac{\partial J_{N-1,N}}{\partial \mathbf{u}(N-1)} = \mathbf{R}\mathbf{u}(N-1) + \mathbf{B}^T\mathbf{P}(0)[\mathbf{A}\mathbf{x}(N-1) + \mathbf{B}\mathbf{u}(N-1)] = \mathbf{0}.$$

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

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

$$J^*_{N-1,N}(\mathbf{x}(N-1)) = \tfrac{1}{2}\mathbf{x}^T(N-1)\{[\mathbf{A} + \mathbf{B}\mathbf{F}(N-1)]^T\mathbf{P}(0)[\mathbf{A} + \mathbf{B}\mathbf{F}(N-1)] + \mathbf{F}^T(N-1)\mathbf{R}\mathbf{F}(N-1) + \mathbf{Q}\}\mathbf{x}(N-1)$$

$$\triangleq \tfrac{1}{2}\mathbf{x}^T(N-1)\mathbf{P}(1)\mathbf{x}(N-1)$$

Dynamic Programming

- Discrete Linear Regulator Problems – Analytical Results
  - Recurrence equation (continued)

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

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

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

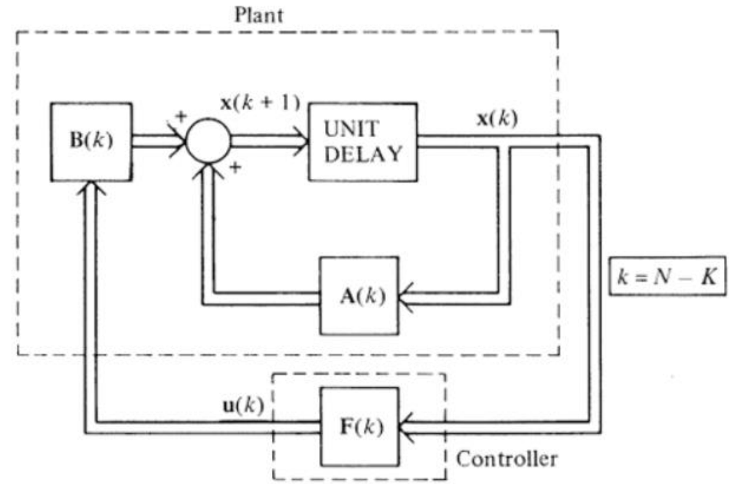
$$\begin{aligned}J^*_{N-K,N}(\mathbf{x}(N-K)) &= \tfrac{1}{2}\mathbf{x}^T(N-K)\{[\mathbf{A} + \mathbf{B}\mathbf{F}(N-K)]^T\mathbf{P}(K-1)[\mathbf{A} + \mathbf{B}\mathbf{F}(N-K)] + \mathbf{F}^T(N-K)\mathbf{R}\mathbf{F}(N-K) + \mathbf{Q}\}\mathbf{x}(N-K)\\ &\triangleq \tfrac{1}{2}\mathbf{x}^T(N-K)\mathbf{P}(K)\mathbf{x}(N-K)\end{aligned}$$

Dynamic Programming

- Linear time-varying system

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

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



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

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

$$
\begin{aligned}
\mathbf{P}(K) &= [\mathbf{A}(N-K)+\mathbf{B}(N-K)\mathbf{F}(N-K)]^T\mathbf{P}(K-1) \\
&\quad\times[\mathbf{A}(N-K)+\mathbf{B}(N-K)\mathbf{F}(N-K)] \\
&\quad+\mathbf{F}^T(N-K)\mathbf{R}(N-K)\mathbf{F}(N-K)+\mathbf{Q}(N-K) \qquad \mathbf{P}(0)=\mathbf{H}
\end{aligned}
$$

$\mathbf{P}(0) \to \mathbf{F}(N-1) \to \mathbf{P}(1) \to \mathbf{F}(N-2) \to \cdots \to \mathbf{P}(N) \to \mathbf{F}(0)$

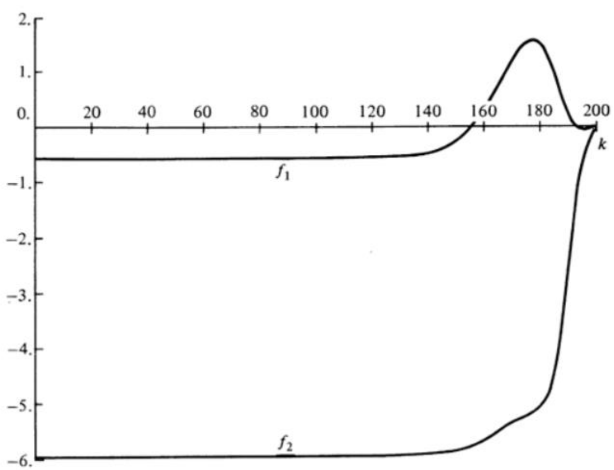
Dynamic Programming

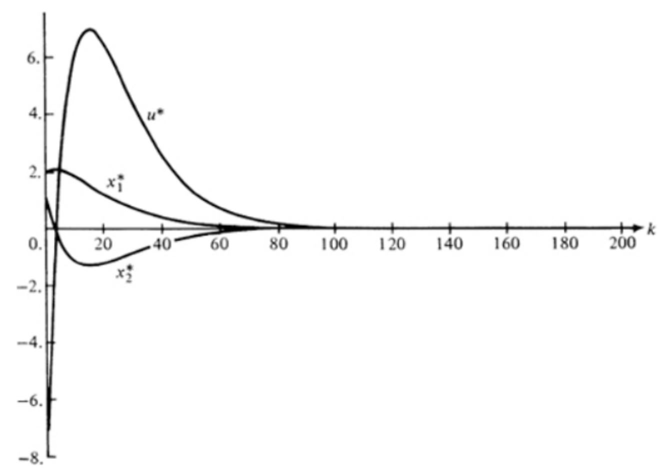
(Ex)

$$\mathbf{x}(k+1) = \begin{bmatrix} 0.9974 & 0.0539 \\ -0.1078 & 1.1591 \end{bmatrix} \mathbf{x}(k) + \begin{bmatrix} 0.0013 \\ 0.0539 \end{bmatrix} u(k)$$

$$J = \frac{1}{2} \sum_{k=0}^{N-1} \left[0.25 x_1^2(k) + 0.05 x_2^2(k) + 0.05 u^2(k) \right] \quad \rightarrow \quad \mathbf{H} = \mathbf{0}, \quad \mathbf{Q} = \begin{bmatrix} 0.25 & 0.00 \\ 0.00 & 0.05 \end{bmatrix}, \text{ and } R = 0.05$$

$N = 200$





Hamilton-Jacobi-Bellman Equation

- Continuous-Time System

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

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

  - Cost minimization

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

$$= \min_{\substack{\mathbf{u}(\tau)\\ t \le \tau \le t_f}} \left\{ \int_t^{t+\Delta t} g\, d\tau + \underline{\int_{t+\Delta t}^{t_f} g\, d\tau + h(\mathbf{x}(t_f), t_f)} \right\}$$

(subdividing interval)

$$= \min_{\substack{\mathbf{u}(\tau)\\ t \le \tau \le t+\Delta t}} \left\{ \int_t^{t+\Delta t} g\, d\tau + \underline{J^*(\mathbf{x}(t+\Delta t), t+\Delta t)} \right\}$$

(principle of optimality)

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

(Taylor series)

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

(approximation for small Δt)

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

$\Rightarrow$ $$\boxed{0 = J_t^*(\mathbf{x}(t), t) + \min_{\mathbf{u}(t)} \left\{ g(\mathbf{x}(t), \mathbf{u}(t), t) + J_{\mathbf{x}}^{*T}(\mathbf{x}(t), t)[\mathbf{a}(\mathbf{x}(t), \mathbf{u}(t), t)] \right\}} \quad \boxed{J^*(\mathbf{x}(t_f), t_f) = h(\mathbf{x}(t_f), t_f)}$$

Hamilton-Jacobi-Bellman Equation

- Continuous-Time System (continued)
  - Hamiltonian $\mathcal{H}$

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

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

  - Hamilton-Jacobi-Bellman (HJB) equation

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

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

(cf. discrete time case)

$$J_{N-K,N}^*(\mathbf{x}(N-K)) = \min_{\mathbf{u}(N-K)} \{g_D(\mathbf{x}(N-K), \mathbf{u}(N-K)) + J_{N-(K-1),N}^*(\mathbf{a}_D(\mathbf{x}(N-K), \mathbf{u}(N-K)))\}$$

Hamilton-Jacobi-Bellman Equation

(Ex) A first order system $\dot{x}(t) = x(t) + u(t)$

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

- Hamiltonian $\mathscr{H}(\mathbf{x}(t), \mathbf{u}(t), J_{\mathbf{x}}^*, t) \triangleq g(\mathbf{x}(t), \mathbf{u}(t), t) + J_{\mathbf{x}}^{*T}(\mathbf{x}(t), t)[\mathbf{a}(\mathbf{x}(t), \mathbf{u}(t), t)]$ $\quad g = \tfrac{1}{4}u^2(t) \quad a = x(t) + u(t)$

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

- Minimization of Hamiltonian

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

$$\Rightarrow u^*(t) = -2J_x^*(x(t), t) \quad (\bigstar)$$

- HJB equation

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

$$(\bigstar) \Rightarrow 0 = J_t^* + \tfrac{1}{4}[-2J_x^*]^2 + [J_x^*]x(t) - 2[J_x^*]^2 \qquad J^*(x(T), T) = \tfrac{1}{4}x^2(T)$$

$$= J_t^* - [J_x^*]^2 + [J_x^*]x(t). \quad (\bigstar\bigstar)$$

Assume $J^*(x(t), t) = \tfrac{1}{2}K(t)x^2(t)$ then $J_x^*(x(t), t) = K(t)x(t)$ $\quad J_t^*(x(t), t) = \tfrac{1}{2}\dot{K}(t)x^2(t)$

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

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

- Optimal control

$$\boxed{u^*(t) = -2K(t)x(t)}$$

Hamilton-Jacobi-Bellman Equation

(Q) The first-order linear system

$$\dot{x}(t) = -10x(t) + u(t)$$

is to be controlled to minimize the performance measure

$$J = \tfrac{1}{2}x^2(0.04) + \int_0^{0.04} [\tfrac{1}{4}x^2(t) + \tfrac{1}{2}u^2(t)]\, dt.$$

The admissible state and control values are not constrained by any boundaries. Find the optimal control law by using the Hamilton-Jacobi-Bellman equation.

Hamilton-Jacobi-Bellman Equation

- Continuous Linear Regulator

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

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

  - Hamiltonian

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

  - Minimization of Hamiltonian

$$\frac{\partial \mathscr{H}}{\partial \mathbf{u}}(\mathbf{x}(t), \mathbf{u}(t), J_{\mathbf{x}}^*, t) = \mathbf{R}(t)\mathbf{u}(t) + \mathbf{B}^T(t)J_{\mathbf{x}}^*(\mathbf{x}(t), t) = \mathbf{0}. \quad \rightarrow \quad \mathbf{u}^*(t) = -\mathbf{R}^{-1}(t)\mathbf{B}^T(t)J_{\mathbf{x}}^*(\mathbf{x}(t), t)$$

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

  - HJB equation

$$0 = J_t^* + \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} - \frac{1}{2}J_{\mathbf{x}}^{*T}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T J_{\mathbf{x}}^* + J_{\mathbf{x}}^{*T}\mathbf{A}\mathbf{x} \qquad J^*(\mathbf{x}(t_f), t_f) = \frac{1}{2}\mathbf{x}^T(t_f)\mathbf{H}\mathbf{x}(t_f)$$

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

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

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

$$\Rightarrow \quad \boxed{\begin{array}{l} \mathbf{0} = \dot{\mathbf{K}}(t) + \mathbf{Q}(t) - \mathbf{K}(t)\mathbf{B}(t)\mathbf{R}^{-1}(t)\mathbf{B}^T(t)\mathbf{K}(t) + \mathbf{K}(t)\mathbf{A}(t) + \mathbf{A}^T(t)\mathbf{K}(t) \\ \mathbf{K}(t_f) = \mathbf{H} \end{array}}$$

Ricatti equation

  - Optimal control

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

Hamilton-Jacobi-Bellman Equation

(Q) (Simulation)

It is desired to determine the control law that causes the plant

$$\dot{x}_1(t) = x_2(t)$$
$$\dot{x}_2(t) = -x_1(t) - 2x_2(t) + u(t)$$

to minimize the performance measure

$$J = 10x_1^2(T) + \tfrac{1}{2}\int_0^T [x_1^2(t) + 2x_2^2(t) + u^2(t)]\, dt.$$

The final time T is 10, and the states and control are not constrained by any boundaries. Find the optimal control law by

(a) Integrating the Riccati equation (3.12-14) with an integration interval of 0.02.

$$\boxed{\begin{aligned}\mathbf{0} = \dot{\mathbf{K}}(t) + \mathbf{Q}(t) - \mathbf{K}(t)\mathbf{B}(t)\mathbf{R}^{-1}(t)\mathbf{B}^T(t)\mathbf{K}(t) \\ + \mathbf{K}(t)\mathbf{A}(t) + \mathbf{A}^T(t)\mathbf{K}(t),\end{aligned}} \quad (3.12\text{-}14) \qquad \boxed{\mathbf{K}(t_f) = \mathbf{H}.}$$

(b) Plot the optimal control $u^*(t), x_1^*(t), x_2^*(t)$,

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