

随机最优控制理论

Stochastic Optimal Control Theory

蔡远利 教授

2026

西安交通大学 自动化学院

内容大纲

最优性原理

确定性最优控制基础

随机动态规划技术

连续时间 LQG 问题

离散时间 LQG 问题

非线性随机系统次优控制

最优性原理

最优性原理

- 美国理查德·贝尔曼 (Richard Bellman) 教授 1957 年提出

最优性原理

对于一个最优策略，在任何状态下，无论过去的历史如何，做出的选择都应该从当前状态到最终状态的最优选择。

确定性最优控制基础

连续时间系统模型

- 系统描述:

$$\dot{\mathbf{x}} = \mathbf{g}[\mathbf{x}(t), \mathbf{u}(t), t], \quad \mathbf{x}(t_0) = \mathbf{x}_0 \quad (1)$$

- 性能指标:

$$J = \varphi[\mathbf{x}(t_f), t_f] + \int_{t_0}^{t_f} L[\mathbf{x}(t), \mathbf{u}(t), t] dt \quad (2)$$

- 三种核心方法: 动态规划、变分法、极小值原理

哈密顿-雅可比-贝尔曼 (HJB) 方程

- 定义最优指标:

$$J^*[\mathbf{x}(t), t] \triangleq \min_{u(\tau), t \leq \tau \leq t_f} J[\mathbf{x}(t), t] \quad (3)$$

- Hamilton 函数:

$$H[\mathbf{x}, \mathbf{u}, J_x^*, t] = L[\mathbf{x}, \mathbf{u}, t] + J_x^{*T} \mathbf{g}[\mathbf{x}, \mathbf{u}, t] \quad (4)$$

- HJB 方程:

$$-J_t^* = \min_{u(t)} H[\mathbf{x}, \mathbf{u}, J_x^*, t] \quad (5)$$

- 边界条件:

$$J^*[\mathbf{x}(t_f), t_f] = \varphi[\mathbf{x}(t_f), t_f] \quad (6)$$

欧拉-拉格朗日方程

- Hamilton 函数:

$$H[\mathbf{x}, u, \boldsymbol{\lambda}, t] = L[\mathbf{x}, u, t] + \boldsymbol{\lambda}^T(t) \mathbf{g}(\mathbf{x}, u, t) \quad (7)$$

- 协态方程:

$$\dot{\boldsymbol{\lambda}}(t) = -\frac{\partial H}{\partial \mathbf{x}}, \quad \boldsymbol{\lambda}(t_f) = -\frac{\partial \varphi}{\partial \mathbf{x}(t_f)} \quad (8)$$

- 控制方程 (无约束时):

$$\frac{\partial H}{\partial u} = \mathbf{0} \quad (9)$$

以上为最优解的必要条件。充分条件需 $\delta^2 J > 0$ 。

庞特里亚金极小值原理

当 $u(t)$ 存在容许集 $\mathcal{U}$ (约束集) 时, 边界处 $\partial H / \partial u = 0$ 不再成立。

$$u^*(t) = \arg \min_{u(t) \in \mathcal{U}} H[\mathbf{x}(t), u(t), \boldsymbol{\lambda}(t), t] \quad (10)$$

$$\dot{\boldsymbol{\lambda}}(t) = -\frac{\partial H}{\partial \mathbf{x}}, \quad \boldsymbol{\lambda}(t_f) = -\frac{\partial \varphi}{\partial \mathbf{x}(t_f)} \quad (11)$$

- 本质: 一个两点边值问题 (TPBVP)

连续时间 LQR 问题

系统与性能指标:

$$\dot{\mathbf{x}} = A(t)\mathbf{x} + B(t)\mathbf{u}, \quad \mathbf{x}(t_0) = \mathbf{x}_0 \quad (12)$$

$$J = \frac{1}{2}\mathbf{x}^T(t_f)S_f\mathbf{x}(t_f) + \frac{1}{2}\int_{t_0}^{t_f} (\mathbf{x}^T Q \mathbf{x} + \mathbf{u}^T R \mathbf{u}) dt \quad (13)$$

其中 $S_f \geq 0$, $Q \geq 0$, $R > 0$ 为对称加权矩阵。

LQR 最优解

- 最优控制:

$$\mathbf{u}^*(t) = -R^{-1}B^T P(t)\mathbf{x}(t) \quad (14)$$

- 矩阵 Riccati 微分方程:

$$\dot{P} + PA + A^T P - PBR^{-1}B^T P + Q = 0, \quad P(t_f) = S_f \quad (15)$$

- 最优指标:

$$J^*[\mathbf{x}_0, t_0] = \frac{1}{2}\mathbf{x}_0^T P(t_0)\mathbf{x}_0 \quad (16)$$

应用：拦截与交会——问题描述

简化运动模型（以初始视线为参考基准）：

$$\dot{v} = a(t), \quad \dot{y} = v \quad (17)$$

$$J = \frac{1}{2} (c_1 v^2 + c_2 y^2)_{t=t_f} + \frac{1}{2} \int_{t_0}^{t_f} a^2 dt \quad (18)$$

其中 y 为横向偏差， v 为横向速度， a 为横向加速度控制量； c_1, c_2 为终端加权系数。

写为 LQR 标准形式：

$$\mathbf{x} = \begin{bmatrix} v \\ y \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{S}_f = \begin{bmatrix} c_1 & 0 \\ 0 & c_2 \end{bmatrix}, \quad Q = 0, \quad R = 1$$

应用：拦截与交会——Riccati 方程求解

Riccati 方程 $\dot{P} + PA + A^T P - PBB^T P = 0$, $P(t_f) = S_f$ 。设 $P = \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix}$, 得

$$\dot{p}_{11} = p_{11}^2 - 2p_{12}, \quad p_{11}(t_f) = c_1$$

$$\dot{p}_{12} = p_{11}p_{12} - p_{22}, \quad p_{12}(t_f) = 0$$

$$\dot{p}_{22} = p_{12}^2, \quad p_{22}(t_f) = c_2$$

由 $u^* = -B^T P x$ 得 $a^* = -p_{11}v - p_{12}y$ 。记 $t_{go} = t_f - t$, 解为

$$\Lambda_v = p_{11} = \frac{c_2^{-1} + c_1^{-1}t_{go}^2 + t_{go}^3/3}{\Delta}, \quad \Lambda_y = p_{12} = \frac{c_1^{-1}t_{go} + t_{go}^2/2}{\Delta}$$

$$\Delta = (c_2^{-1} + t_{go}^3/3)(c_1^{-1} + t_{go}) - t_{go}^4/4$$

应用：拦截与交会——拦截问题

$c_1 \rightarrow 0$ (对终端速度不作要求), 且 $c_2 \rightarrow \infty$ ($y(t_f) \rightarrow 0$)

$$a(t) = -\frac{3}{t_{go}}v(t) - \frac{3}{t_{go}^2}y(t) \quad (19)$$

在小角近似下, $y \approx r\sigma$ 、 $v \approx V\dot{\sigma}$, 则

$$a(t) \approx -3V\ddot{\sigma} \quad (20)$$

此即拦截问题中的**比例导引** (PNG), 最优导航系数 $N = 3$; 工程中通常取 $3 \sim 5$ 。

应用：拦截与交会——交会问题

$c_1, c_2 \rightarrow \infty$ ($v(t_f) \rightarrow 0, y(t_f) \rightarrow 0$, 即交会问题)

$$a(t) = -\frac{4}{t_{go}}v(t) - \frac{6}{t_{go}^2}y(t) \quad (21)$$

也可写为

$$a(t) = -V\left(4\dot{\sigma} + \frac{2\sigma}{t_{go}}\right) \quad (22)$$

此为交会问题的变形比例导引。

- 可将标量 v, y, a 替换为向量，得到三维拦截律

微分对策——问题描述

- 传统最优控制 = **单边**最优控制；**微分对策** = 相互对抗的**双边最优控制**（零和微分对策）

空间追逃博弈：追逐方 p 与逃逸方 e

$$\dot{\mathbf{v}}_p = \mathbf{g}_p + \mathbf{a}_p, \quad \dot{\mathbf{r}}_p = \mathbf{v}_p, \quad \dot{\mathbf{v}}_e = \mathbf{g}_e + \mathbf{a}_e, \quad \dot{\mathbf{r}}_e = \mathbf{v}_e$$

距离远小于地心距，引力差可忽略。性能指标（min-max 问题）：

$$J = \frac{b}{2} \|\mathbf{r}_p - \mathbf{r}_e\|_{t_f}^2 + \frac{1}{2} \int_0^{t_f} (\mathbf{c}_p^{-1} \mathbf{a}_p^T \mathbf{a}_p - \mathbf{c}_e^{-1} \mathbf{a}_e^T \mathbf{a}_e) dt \quad (23)$$

- 追逐方 $\mathbf{a}_p$ ：**极小化** J ，逃逸方 $\mathbf{a}_e$ ：**极大化** J

微分对策——求解思路

根据 LQR 理论, 构造相对运动状态 $\mathbf{x} = \mathbf{r}_p - \mathbf{r}_e$, $\mathbf{v} = \mathbf{v}_p - \mathbf{v}_e$:

$$\dot{\mathbf{x}} = \mathbf{v}, \quad \dot{\mathbf{v}} = \mathbf{a}_p - \mathbf{a}_e$$

$$J = \frac{b}{2} \|\mathbf{x}(t_f)\|^2 + \frac{1}{2} \int (c_p^{-1} \mathbf{a}_p^T \mathbf{a}_p - c_e^{-1} \mathbf{a}_e^T \mathbf{a}_e) dt$$

应用极小值原理, 构造 Hamilton 函数:

$$H = \frac{1}{2} (c_p^{-1} \mathbf{a}_p^T \mathbf{a}_p - c_e^{-1} \mathbf{a}_e^T \mathbf{a}_e) + \boldsymbol{\lambda}_x^T \mathbf{v} + \boldsymbol{\lambda}_v^T (\mathbf{a}_p - \mathbf{a}_e) \quad (24)$$

微分对策——最优性条件

最优性条件:

$$\partial H / \partial \mathbf{a}_p = 0 \Rightarrow c_p^{-1} \mathbf{a}_p + \boldsymbol{\lambda}_v = 0$$

$$\partial H / \partial \mathbf{a}_e = 0 \Rightarrow -c_e^{-1} \mathbf{a}_e - \boldsymbol{\lambda}_v = 0$$

由于 $\mathbf{a}_p$ 和 $\mathbf{a}_e$ 分别出现在 $\pm \boldsymbol{\lambda}_v$ 项中, 可得双方最优策略关系:

$$\mathbf{a}_p = -c_p \boldsymbol{\lambda}_v, \quad \mathbf{a}_e = c_e \boldsymbol{\lambda}_v \quad (25)$$

即追逐方与逃逸方的最优控制方向相反、幅值受各自能量加权系数调节。

微分对策——最优策略

求解协态方程并与终端条件联立，可得双方最优控制策略：

$$a_p = \frac{-c_p(t_f - t)[r_p - r_e + (v_p - v_e)(t_f - t)]}{1/b + (c_p - c_e)(t_f - t)^3/3} \quad (26)$$

$$a_e = \frac{c_e}{c_p} a_p \quad (27)$$

当 $b \rightarrow \infty$ (精确拦截) 时简化为

$$a_p = \frac{-3[r_p - r_e + (v_p - v_e)(t_f - t)]}{(1 - c_e/c_p)(t_f - t)^2} \quad (28)$$

$$a_p^\perp = \frac{3}{1 - c_e/c_p} V \dot{\sigma} \quad (29)$$

微分对策——有效导引系数

侧向最优控制：

$$a_p^\perp = \frac{3}{1 - c_e/c_p} V \dot{\sigma} \quad (30)$$

定义**有效导引系数（有效导航比）**

$$N \triangleq \frac{3}{1 - c_e/c_p} \quad (31)$$

特例分析：

- $c_e = 0$ （目标不机动）： $N = 3$ ，退化为经典比例导引
- $c_e/c_p = 2/5$ ： $N = 5$ ；工程中通常取 $N = 3 \sim 5$ ，与经验一致

结论：微分对策理论从双边最优控制的角度解释了比例导引中导航系数 $3 \sim 5$ 的取值依据。

离散时间最优控制

系统模型:

$$\mathbf{x}_{k+1} = \mathbf{g}(\mathbf{x}_k, \mathbf{u}_k), \quad k = 0, \dots, N-1 \quad (32)$$

性能指标:

$$J_0(\mathbf{x}_0) = \varphi(\mathbf{x}_N, N) + \sum_{k=0}^{N-1} L_k(\mathbf{x}_k, \mathbf{u}_k) \quad (33)$$

贝尔曼方程:

$$J_k^*(\mathbf{x}_k) = \min_{\mathbf{u}_k} [L_k(\mathbf{x}_k, \mathbf{u}_k) + J_{k+1}^*(\mathbf{x}_{k+1})] \quad (34)$$

边界条件: $J_N^*(\mathbf{x}_N) = \varphi(\mathbf{x}_N, N)$

离散 LQR 问题

系统与指标:

$$\mathbf{x}_{k+1} = A_k \mathbf{x}_k + B_k \mathbf{u}_k, \quad \mathbf{x}_0 \text{ 给定} \quad (35)$$

$$J_0 = \frac{1}{2} \mathbf{x}_N^T S_N \mathbf{x}_N + \frac{1}{2} \sum_{k=0}^{N-1} (\mathbf{x}_k^T Q_k \mathbf{x}_k + \mathbf{u}_k^T R_k \mathbf{u}_k) \quad (36)$$

· 最优控制:

$$\mathbf{u}_k^* = -K_k \mathbf{x}_k, \quad K_k = (R_k + B_k^T P_{k+1} B_k)^{-1} B_k^T P_{k+1} A_k \quad (37)$$

· 离散 Riccati 方程:

$$P_k = Q_k + K_k^T R_k K_k + (A_k + B_k K_k)^T P_{k+1} (A_k + B_k K_k), \quad P_N = S_N \quad (38)$$

数值解法分类

间接法

- 极小值原理 $\rightarrow$ TPBVP
- 打靶法、拟线性化、配置法
- 缺点：协态初值难猜、收敛域小

直接法

- 泛函优化 $\rightarrow$ 数学规划
- 梯度法、伪谱法
- 不需求解协态方程

智能优化方法：遗传算法、模拟退火、粒子群、蚁群等

随机动态规划技术

离散时间随机系统

系统描述：

$$\mathbf{x}_{k+1} = \mathbf{g}_k(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{G}_k \mathbf{w}_k \quad (39)$$

其中 $\mathbf{w}_k$ 为白噪声序列。完全状态信息假设。

期望代价：

$$J_i(\mathbf{x}_i) = E[J_i(\mathbf{x}_i)] \quad (40)$$

$$\bar{J}_i(\mathbf{x}_i) = E[J_i(\mathbf{x}_i) | \mathbf{x}_i] \quad (41)$$

随机贝尔曼方程

根据最优性原理，可得完整状态信息下的贝尔曼方程：

$$\bar{J}_k^*(\mathbf{x}_k) = \min_{u_k} E_{\mathbf{x}_{k+1}} \{L_k(\mathbf{x}_k, u_k) + \bar{J}_{k+1}^*(\mathbf{x}_{k+1})\} \quad (42)$$

边界条件：

$$\bar{J}_N^*(\mathbf{x}_N) = \phi(N, \mathbf{x}_N) \quad (43)$$

与确定性问题的区别：期望运算 $E_{\mathbf{x}_{k+1}} \{\cdot\}$

连续时间随机系统

标量随机系统:

$$\dot{x} = a(x, u, t) + w, \quad w(t) \sim \mathcal{N}(0, \sigma^2) \quad (44)$$

期望代价:

$$j(t_0) = E\{J[x(t_0), t_0]\} \quad (45)$$

随机 HJB 方程:

$$-\frac{\partial \bar{J}^*}{\partial t} = \min_{u(t)} \left[L + \frac{\partial \bar{J}^*}{\partial x} a + \frac{1}{2} \frac{\partial^2 \bar{J}^*}{\partial x^2} \sigma^2 \right] \quad (46)$$

向量随机系统的 HJB 方程

对于向量系统 $\dot{x} = A(x, u, t) + Gw$, $w \sim \mathcal{N}(0, Q')$:

$$-\frac{\partial \bar{J}^*}{\partial t} = \min_{u(t)} \left[L + \left(\frac{\partial \bar{J}^*}{\partial x} \right)^T A + \frac{1}{2} \text{tr} \left(\frac{\partial^2 \bar{J}^*}{\partial x^2} G Q' G^T \right) \right] \quad (47)$$

- 完整论证需引入伊藤随机微分理论
- 求解 HJB 方程非常困难, 可借助机器学习方法

连续时间 LQG 问题

问题描述

系统:

$$\dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{u} + G\mathbf{w}, \quad \mathbf{w}(t) \sim \mathcal{N}(0, Q') \quad (48)$$

初始状态:

$$\mathbf{x}(t_0) \sim \mathcal{N}(\bar{\mathbf{x}}_0, P_0)$$

性能指标:

$$J = \frac{1}{2}\mathbf{x}^T(T)S_T\mathbf{x}(T) + \frac{1}{2} \int_{t_0}^T (\mathbf{x}^T Q \mathbf{x} + \mathbf{u}^T R \mathbf{u}) dt \quad (49)$$

目标: 求解 $\mathbf{u}^*(t)$ 最小化期望代价 $j(t_0) = E[J]$

由随机 HJB 方程导出柯尔莫哥洛夫方程：

$$-\bar{J}_t^* = \frac{1}{2} \mathbf{x}^T Q \mathbf{x} - \frac{1}{2} (\bar{J}_x^*)^T B R^{-1} B^T \bar{J}_x^* + (\bar{J}_x^*)^T A \mathbf{x} + \frac{1}{2} \text{tr}(\bar{J}_{xx}^* G Q' G^T) \quad (50)$$

试凑解 $\bar{J}^* = \frac{1}{2} [\mathbf{x}^T S \mathbf{x} + \text{tr} \int_t^T S G Q' G^T d\tau]$ ：

$$-\dot{S} = A^T S + S A - S B R^{-1} B^T S + Q, \quad S(T) = S_T \quad (51)$$

最优控制：

$$\mathbf{u}^*(t) = -R^{-1} B^T S \mathbf{x}(t) = -K(t) \mathbf{x}(t) \quad (52)$$

最优性能评估

闭环系统:

$$\dot{x} = (A - BK)x + Gw \quad (53)$$

均方状态满足 Lyapunov 方程:

$$\dot{X} = (A - BK)X + X(A - BK)^T + GQ'G^T \quad (54)$$

$$X(t_0) = P_0 + \bar{x}_0 \bar{x}_0^T \quad (55)$$

控制信号的均方值:

$$E[uu^T] = KXK^T \quad (56)$$

非完全状态信息与分离原理

量测方程:

$$z = Hx + v, \quad v(t) \sim \mathcal{N}(0, R') \quad (57)$$

分离原理 (确定性等价原理)

非完全状态信息下 LQG 问题的最优解 = 确定性 LQR 最优控制 + 卡尔曼滤波器, 即

$$u^*(t) = -K(t)\hat{x}(t) \quad (58)$$

其中 $\hat{x}(t)$ 为卡尔曼滤波给出的最优状态估计。

分离原理的实现：卡尔曼滤波器

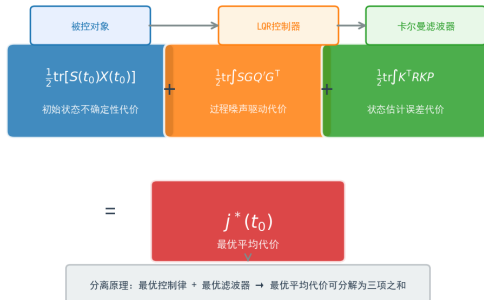
$$\dot{\hat{\mathbf{x}}} = A\hat{\mathbf{x}} + B\mathbf{u} + L(\mathbf{z} - H\hat{\mathbf{x}}), \quad \hat{\mathbf{x}}(0) = \bar{\mathbf{x}}_0 \quad (59)$$

$$L = PH^T R'^{-1} \quad (60)$$

$$\dot{P} = AP + PA^T + GQ'G^T - PH^T R'^{-1}HP \quad (61)$$

分离原理的实现：最优平均代价

$$\begin{aligned} j^*(t_0) &= \frac{1}{2} \text{tr}[S(t_0)X(t_0)] \\ &+ \frac{1}{2} \text{tr} \int_{t_0}^{t_f} SGQ'G^T dt \\ &+ \frac{1}{2} \text{tr} \int_{t_0}^{t_f} K^T R K P dt \end{aligned}$$



LQG 调节器结构

- 闭环系统构成：
 - 被控对象: $\dot{x} = Ax + Bu + Gw$
 - 卡尔曼滤波器 (状态观测器)
 - 状态反馈控制律 $u = -K\hat{x}$
- 主方程:

$$\dot{\hat{x}} = (A - LH)\hat{x} + Bu + Lz \quad (62)$$

$$u = -K\hat{x} \quad (63)$$

- 分离原理的关键: 估计误差 $\tilde{x} = x - \hat{x}$ 独立于控制

实例：导弹滚动姿态调节器

状态变量： $\mathbf{x} = [\delta, \omega, \phi]^T$

$$\dot{\delta} = u, \quad \dot{\omega} = -\frac{1}{\tau}\omega + \frac{q}{\tau}\delta + n, \quad \dot{\phi} = \omega$$

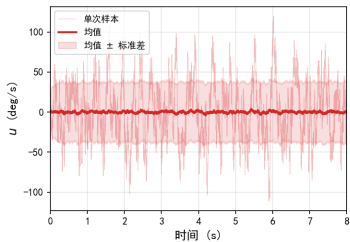
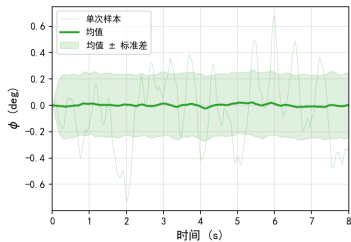
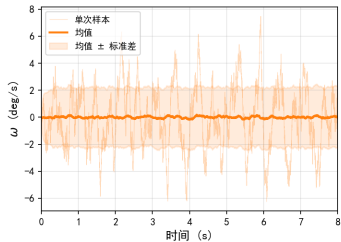
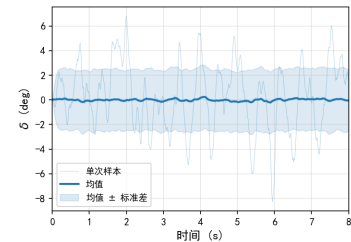
性能指标：

$$J = \lim_{t_f \rightarrow \infty} \frac{1}{2} \int \left(\frac{\phi^2}{\phi_0^2} + \frac{\delta^2}{\delta_0^2} + \frac{u^2}{u_0^2} \right) dt$$

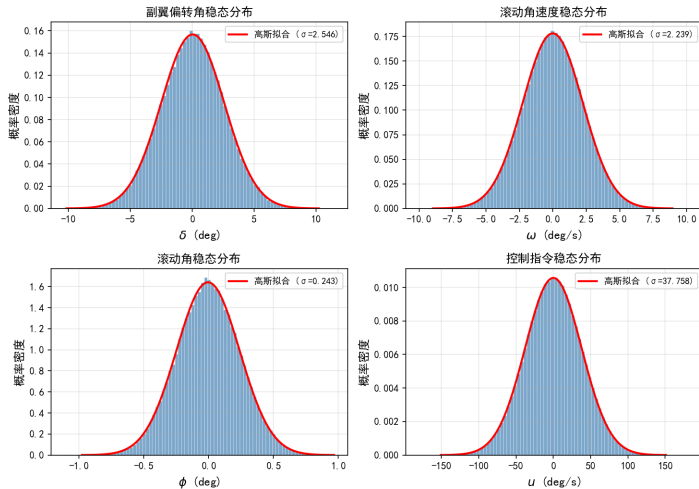
求解结果 ($\phi_0 = \pi/180$, $\delta_0 = \pi/12$, $u_0 = \pi$):

$$K^T = [27, 29, 180], \quad \sqrt{E(\delta^2)} = 2.50^\circ, \quad \sqrt{E(\phi^2)} = 0.24^\circ$$

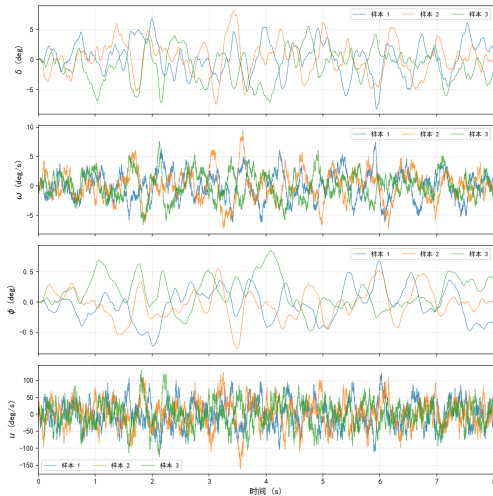
仿真结果：时域响应



仿真结果：稳态分布



仿真结果：典型样本时域响应



实例：飞机横向自动驾驶仪

五阶横向扰动运动方程：

- 状态： β (侧滑角), r (偏航角速率), p (滚转角速率), ϕ (滚转角), ψ (偏航角)
- 控制： δ_r (方向舵), δ_a (副翼)

稳态最优反馈：

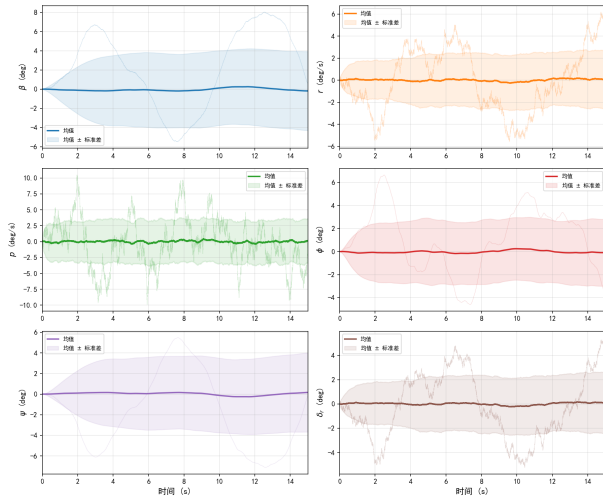
$$\delta_r = -0.317\beta - 1.01r - 0.069p - 0.076\phi - 0.551\psi$$

$$\delta_a = 0.177\beta + 0.388r + 0.737p + 1.03\phi + 0.834\psi$$

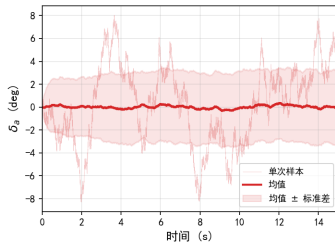
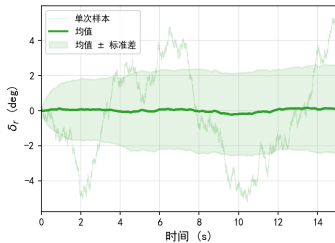
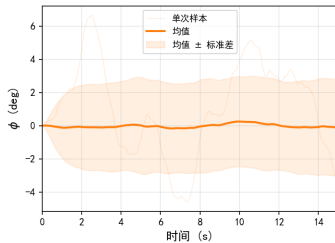
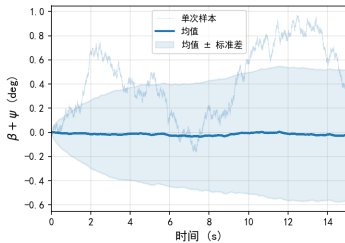
性能评估：

$$\sqrt{E(\beta + \psi)^2} = 0.60^\circ, \quad \sqrt{E(\phi^2)} = 2.87^\circ, \quad \sqrt{E(\delta_r^2)} = 2.48^\circ, \quad \sqrt{E(\delta_a^2)} = 3.30^\circ$$

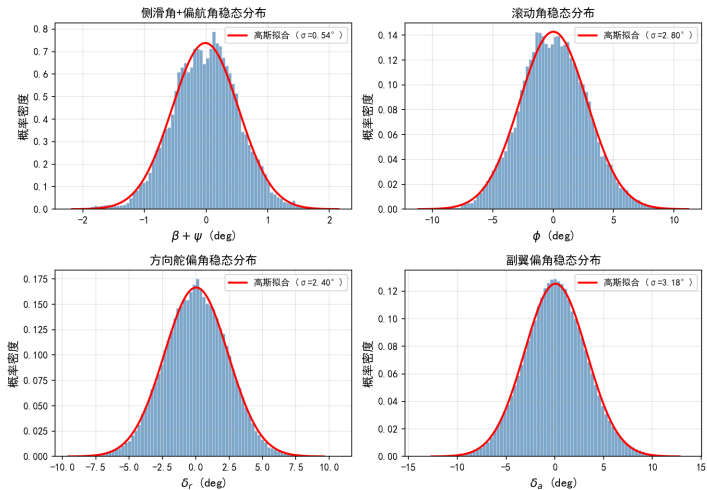
仿真结果：状态时域响应



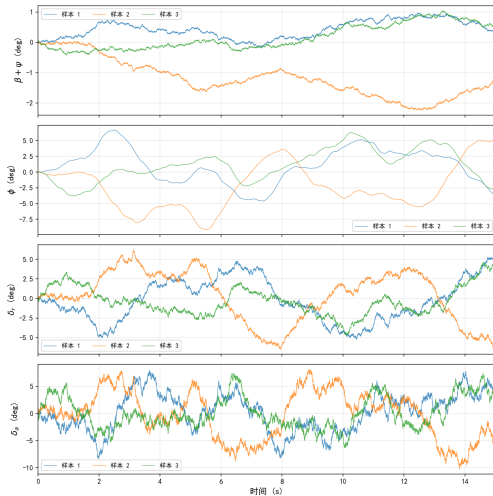
仿真结果：关键输出变量



仿真结果：稳态分布



仿真结果：典型样本时域响应



离散时间 LQG 问题

离散 LQG 完全状态信息

系统与指标:

$$\mathbf{x}_{k+1} = F\mathbf{x}_k + B\mathbf{u}_k + G\mathbf{w}_k, \quad \mathbf{w}_k \sim \mathcal{N}(0, Q') \quad (64)$$

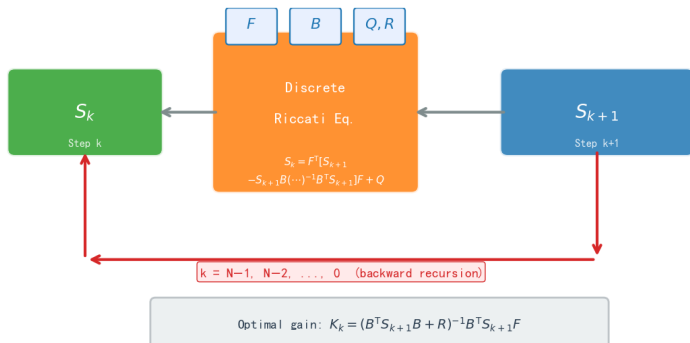
$$J_i = \frac{1}{2}\mathbf{x}_N^T S_N \mathbf{x}_N + \frac{1}{2} \sum_{k=i}^{N-1} (\mathbf{x}_k^T Q \mathbf{x}_k + \mathbf{u}_k^T R \mathbf{u}_k) \quad (65)$$

最优控制:

$$\mathbf{u}_k^* = -K_k \mathbf{x}_k, \quad K_k = (B^T S_{k+1} B + R)^{-1} B^T S_{k+1} F \quad (66)$$

离散 LQG: Riccati 方程

$$S_k = F^T [S_{k+1} - S_{k+1} B (B^T S_{k+1} B + R)^{-1} B^T S_{k+1}] F + Q \quad (67)$$



离散 LQG 非完全状态信息

量测: $\mathbf{z}_k = H\mathbf{x}_k + \mathbf{v}_k$, $\mathbf{v}_k \sim \mathcal{N}(0, R')$

分离原理同样成立:

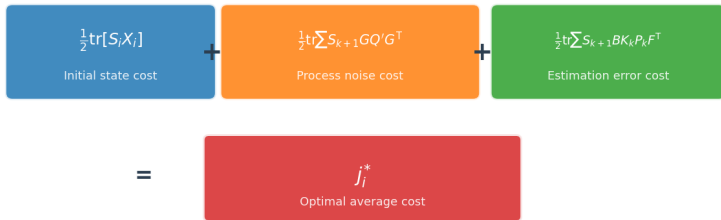
$$\mathbf{u}_k = -K_k \hat{\mathbf{x}}_k \tag{68}$$

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k|k-1} + L_k(\mathbf{z}_k - H\hat{\mathbf{x}}_{k|k-1}) \tag{69}$$

$$\hat{\mathbf{x}}_{k+1|k} = F\hat{\mathbf{x}}_k + B\mathbf{u}_k \tag{70}$$

离散 LQG: 最优平均代价

$$j_i^* = \frac{1}{2} \text{tr} \left[S_i X_i + \sum_{k=i}^{N-1} (S_{k+1} G Q' G^T + S_{k+1} B K_k P_k F^T) \right] \quad (71)$$



非线性随机系统次优控制

摄动法

DP 理论上可解非线性随机最优控制，但存在“维数灾难”。

摄动法步骤：

1. 忽略不确定性，设计确定性最优标称轨线 $\bar{x}(t)$ 、 $\bar{u}(t)$
2. 在标称轨线附近线性化，设计 LQR 摄动控制 $\delta u^* = K_\delta \delta x$
3. 设计摄动系统最优滤波器，获得 $\delta \hat{x}$
4. 实际控制： $u(t) = \bar{u}(t) + K_\delta(t) \delta \hat{x}(t)$

优点：简单实用，工程中广泛应用

摄动法算例

非线性系统:

$$\dot{x} = -x + ax^3 + u + w, \quad y = x + v$$

设计控制使 $y(t)$ 跟踪设定值 y_d 。

解: 平衡状态 $\bar{x} = y_d$, 标称控制 $\bar{u} = y_d - ay_d^3$ 。

线性化摄动系统:

$$\delta\dot{x} = (-1 + 3ay_d^2)\delta x + \delta u + w$$

对摄动系统应用 LQG 理论得 δu^* , 实际控制 $u = \bar{u} + \delta u^*$ 。

强迫分离法

直接应用 LQG 分离原理的思想（最优性无理论保证，但工程可行）：

1. 忽略不确定性，设计确定性最优控制 $\hat{u}(t) = \hat{u}[x(t)]$
2. 设计非线性滤波器（EKF/UKF/CKF），获得状态估计 $\hat{x}(t)$
3. 实际控制： $\hat{u}(t) = \hat{u}[\hat{x}(t)]$

注：与人工智能、机器学习结合是值得关注的方向发展。

本章小结

小结

核心理论与方法

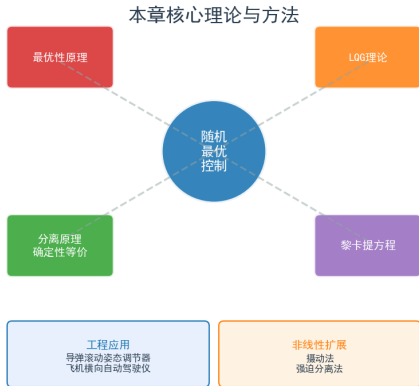
- **最优性原理**——动态规划的理论基石
- **LQG 理论**——线性二次型最优解
- **分离原理**——估计与控制可独立设计
- **黎卡提方程**——LQR/LQG 核心工具

工程应用

- 导弹滚动姿态调节器
- 飞机横向自动驾驶仪

非线性扩展

- 摄动法 / 强迫分离法



当前发展趋势

- 基于**强化学习**的非线性随机最优控制
- **模型预测控制** (MPC) 与随机优化的深度融合
- **数据驱动**控制——利用在线数据更新模型
- **鲁棒 LQG** 与**风险敏感**控制
- **多智能体**随机博弈——集群分布式控制
- 与**深度学习**结合求解高维 HJB 方程

关键挑战

- 高维状态/控制空间的维数灾难
- 模型未知时的探索-利用权衡
- 安全关键系统的实时性与鲁棒性

