六. 动态规划
(Dynamic Programming)

6.5 连续系统动态规划

(1) 连续系统最优性原理

考虑 n 阶连续系统状态方程

$$\dot{x}(t) = f[x(t), u(t), t] \qquad (6\text{-}5\text{-}1)$$

和初始状态

$$x(t_0) = x_0 \qquad (6\text{-}5\text{-}2)$$

其中 $x(t) \in R^n, u(t) \in R^m$ ，

要确定控制函数 $u(t)$，使性能指标

$$J = \int_{t_0}^{t_f} L\left[x(t), u(t), t\right] dt \qquad (6\text{-}5\text{-}4)$$

达极小值（假定 t_f 给定，$x(t_f)$ 自由）。

假定 $u^*(t), x^*(t)$ 都已找到，代入(6-5-4)中即为最优性能指标 J^*。显然与 t_0 和 $x(t_0)$ 有关，即 J^* 是 t_0 和 $x(t_0)$ 的函数，记为

$$J^*[x(t_0), t_0] \triangleq J[x^*(t), u^*(t), t] = \int_{t_0}^{t_f} L\left[x^*(t), u^*(t), t\right] dt$$

$$= \min_{u[t_0, t_f]} \int_{t_0}^{t_f} L\left[x(t), u(t), t\right] dt \qquad (6\text{-}5\text{-}5)$$

与离散动态规划类似，连续动态规划也存在如下最优性原理。

连续系统最优性原理：

初始状态为 $x(t_0)$ 的最优控制策略 $u^*[t_0,t_f]$ 后面的一部分 $u^*[t_1,t_f]$（$t_1 > t_0$）仍是最优控制策略，其初始状态是在区间 $[t_0,t_1]$ 上用最优控制策略 $u^*[t_0,t_1]$，由系统状态方程 $\dot{x}(t) = f[x(t),u(t)]$ 和初态 $x(t_0) = x_0$ 所得到的 $x(t_1)$。

这与离散情况下的最优性原理本质上一致，证明也完全类似。

(2) 连续系统动态规划——H-J-B方程

基本方程的推导：

由(6-5-5)式定义，对于$[t_0, t_f]$上任意t，依据最优性原理，最优性能指标为

$$
\begin{aligned}
J^*[x(t),t] &= \min_{u[t,t_f]} \int_t^{t_f} L[x(\tau),u(\tau),\tau]d\tau \\
&= \min_{u[t,t_f]} \left\{ \int_t^{t+\Delta t} L[x(\tau),u(\tau),\tau]d\tau + \int_{t+\Delta t}^{t_f} L[x(\tau),u(\tau),\tau]d\tau \right\} \\
&= \min_{u[t,t+\Delta t]} \left\{ \min_{u[t+\Delta t,t_f]} \left\{ \int_t^{t+\Delta t} L[x(\tau),u(\tau),\tau]d\tau + \int_{t+\Delta t}^{t_f} L[x(\tau),u(\tau),\tau]d\tau \right\} \right\} \\
&= \min_{u[t,t+\Delta t]} \left\{ \int_t^{t+\Delta t} L[x(\tau),u(\tau),\tau]d\tau + \min_{u[t+\Delta t,t_f]} \int_{t+\Delta t}^{t_f} L[x(\tau),u(\tau),\tau]d\tau \right\}
\end{aligned}
\tag{6-5-8}
$$

又依据(6-5-5)式定义，有

$$
J^*[x(t+\Delta t),t+\Delta t] = \min_{u[t+\Delta t,t_f]} \int_{t+\Delta t}^{t_f} L[x(\tau),u(\tau),\tau]d\tau \tag{6-5-9}
$$

将(6-5-9)代入(6-5-8)得

$$J^*[x(t),t] = \min_{u[t,t+\Delta t]} \left\{ \int_t^{t+\Delta t} L\big[x(\tau),u(\tau),\tau\big]d\tau + J^*[x(t+\Delta t),t+\Delta t] \right\} \tag{6-5-10}$$

假定$J^*[x(t),t]$存在并连续可微，上式右端第二项利用Taloy公式展开为

$$J^*[x(t+\Delta t),t+\Delta t] = J^*[x(t),t] + \left[\frac{\partial J^*[x(t),t]}{\partial x(t)}\right]^{\mathrm{T}} \frac{dx(t)}{dt}\Delta t + \frac{\partial J^*[x(t),t]}{\partial t}\Delta t + O(\Delta t^2) \tag{6-5-11}$$

其中$\dfrac{\partial J^*[x(t),t]}{\partial x(t)}$为$J^*[x(t),t]$对$x(t)$的梯度。上式右端第一项可以写成

$$\int_t^{t+\Delta t} L\big[x(\tau),u(\tau),\tau\big]d\tau = L[x(t+\theta\Delta t),u(t+\theta\Delta t),t+\theta\Delta t]\Delta t \tag{6-5-12}$$

其中$0<\theta<1$。将(6-5-11)、(6-5-12)式代入(6-5-10)式有

$$J^*[x(t),t] = \min_{u[t,t+\Delta t]} \Big\{ L[x(t+\theta\Delta t),u(t+\theta\Delta t),t+\theta\Delta t]\Delta t + J^*[x(t),t] + \left[\frac{\partial J^*[x(t),t]}{\partial x(t)}\right]^{\mathrm{T}} \frac{dx(t)}{dt}\Delta t + \frac{\partial J^*[x(t),t]}{\partial t}\Delta t + O(\Delta t^2) \Big\} \tag{6-5-13}$$

消去$J^*[x(t),t]$，并考虑$\dfrac{\partial J^*[x(t),t]}{\partial t}$不是$u(t)$的函数，可移出求最小的括号外，除以$\Delta t$，并令$\Delta t \to 0$，则由上式可求得

$$-\frac{\partial J^*[x(t),t]}{\partial t} = \min_{u(t)}\left\{ L[x(t),u(t),t] + \left[\frac{\partial J^*[x(t),t]}{\partial x(t)}\right]^{\mathrm{T}} f[x(t),u(t),t] \right\} \quad \text{(6-5-14)}$$

即为连续动态规划的基本方程。

令
$$\frac{\partial J^*[x(t),t]}{\partial x(t)} = \lambda(t) \quad \text{(6-5-15)}$$

(6-5-14)式即可表示为

$$\frac{\partial J^*[x(t),t]}{\partial t} + H[x(t),u(t),\lambda(t),t] = 0 \quad \text{(6-5-16)}$$

的形式，其中$H = L + \lambda^T f$。

在(6-5-14)式中，f、L是$x(t)$、$u(t)$、t的已知函数，$J^*[x(t),t]$是关于$x(t)$和t的未知函数，所以，该方程是关于$J^*[x(t),t]$的一阶偏微分方程；又因要求右边函数对于$u(t)$的全局最小值，所以是偏微和函数方程的混合物。

考虑该式的求解：

先求右端括号内函数对$\boldsymbol{u}(t)$的全局最小值，从而确定$\boldsymbol{u}^*(t)$，一般可认为该$\boldsymbol{u}^*(t)$是$\boldsymbol{x}(t),\dfrac{\partial J^*}{\partial \boldsymbol{x}}$和$t$的函数，记为

$$\bar{\boldsymbol{u}} = \boldsymbol{u}^*\left[\boldsymbol{x}(t),\frac{\partial J^*}{\partial \boldsymbol{x}},t\right] \tag{6-5-17}$$

代入(6-5-14)，有

$$-\frac{\partial J^*}{\partial t} = L\left[\boldsymbol{x}(t),\bar{\boldsymbol{u}}[\boldsymbol{x}(t),\frac{\partial J^*}{\partial \boldsymbol{x}},t],t\right] + \left[\frac{\partial J^*}{\partial \boldsymbol{x}}\right]^{\mathrm{T}} \boldsymbol{f}\left[\boldsymbol{x}(t),\bar{\boldsymbol{u}}[\boldsymbol{x}(t),\frac{\partial J^*}{\partial \boldsymbol{x}},t],t\right] \tag{6-5-18}$$

该式是最优性能指标函数$J^*[\boldsymbol{x}(t),t]$以$\boldsymbol{x}(t)$、t为自变量的一阶偏微分方程，称为**Hamilton-Jacobi-Bellman方程(H-J-B方程)**。

方程(6-5-18)有无穷多解，还需边界条件。由(6-5-5)知：

$$J^*\left[\boldsymbol{x}(t_f),t_f\right] = \min_{\boldsymbol{u}(t_f,t_f)} \int_{t_f}^{t_f} L[\boldsymbol{x}(t),\boldsymbol{u}(t),t] = 0 \tag{6-5-19}$$

则**H-J-B**方程**(6-5-18)**和边界条件**(6-5-19)**即为问题**(6-5-1) - (6-5-4)**的最优控制和最优性能指标函数取最优值的充分条件。即，满足可微性条件的$J^*[x(t),t]$若满足**(6-5-18)**和**(6-5-19)**，则一定是问题**(6-5-1) - (6-5-4)**的最优性能指标函数。

关于**H-J-B**方程的求解：

H-J-B方程为一阶偏微分方程，通常难求解析解，只能求其数值解；

与极大值原理不同，**H-J-B**方程代表的是系统性能指标达最小值的充分条件，而不是必要或充要条件；

能利用极大值原理求解的最优控制问题，不一定能列出**H-J-B**方程。因许多实际工程问题，最优性能指标函数的可微性经常不能保证。