根据具体情况选取足够长的末端状态，或者调整代价权重系数，使系统更快地达到稳态。

4.3.3 动态规划的递归关系——连续系统

上面两个小节讨论了离散系统动态规划的递归关系，本小节将讨论连续系统的递归关系。这两者的推导思路非常相似。考虑如下一般形式的不含约束的连续系统

$$\dot{\boldsymbol{x}}_{(t)}=f(\boldsymbol{x}_{(t)},\boldsymbol{u}_{(t)},t) \tag{4.3.27}$$

其中 $\dot{\boldsymbol{x}}_{(t)}=\frac{\mathrm{d}\boldsymbol{x}_{(t)}}{\mathrm{d}t}$。从任意的中间时刻 t 到末端时间 t_f 的系统的性能指标定义为

$$J_{t\to t_f}(\boldsymbol{x}_{(t)},t,\boldsymbol{u}_{(\tau)})=h(\boldsymbol{x}_{(t_f)},t_f)+\int_t^{t_f}g(\boldsymbol{x}_{(\tau)},\boldsymbol{u}_{(\tau)},\tau)\mathrm{d}\tau,\quad t\leqslant\tau\leqslant t_f \tag{4.3.28}$$

$J_{t\to t_f}(\boldsymbol{x}_{(t)},t,\boldsymbol{u}_{(\tau)})$是状态 $\boldsymbol{x}_{(t)}$、时间 t 和控制量 $\boldsymbol{u}_{(\tau)}$ 的函数。$h()$和 $g()$函数分别为末端代价和运行代价。最优控制的目标是找到 $\tau\in[t,t_f]$区间内的实时最优控制量 $\boldsymbol{u}^*_{(\tau)}$ 使性能指标值最小，最优代价函数可以写成

$$J^*_{t\to t_f}(\boldsymbol{x}_{(t)},t)=\min_{\boldsymbol{u}_{(\tau)}}\left\{h(\boldsymbol{x}_{(t_f)},t_f)+\int_t^{t_f}g(\boldsymbol{x}_{(\tau)},\boldsymbol{u}_{(\tau)},\tau)\mathrm{d}\tau\right\},\quad t\leqslant\tau\leqslant t_f \tag{4.3.29}$$

使用逆向分级的思路，可以将时间间隔$[t,t_f]$分成两段$[t,t+\Delta t]$和$[t+\Delta t,t_f]$，式(4.3.29)变成

$$\begin{aligned}J^*_{t\to t_f}(\boldsymbol{x}_{(t)},t)=\min_{\boldsymbol{u}_{(\tau)}}\Big\{&h(\boldsymbol{x}_{(t_f)},t_f)+\int_{t+\Delta t}^{t_f}g(\boldsymbol{x}_{(\tau)},\boldsymbol{u}_{(\tau)},\tau)\mathrm{d}\tau+\\&\int_t^{t+\Delta t}g(\boldsymbol{x}_{(\tau)},\boldsymbol{u}_{(\tau)},\tau)\mathrm{d}\tau\Big\},\quad t\leqslant\tau\leqslant t_f\end{aligned} \tag{4.3.30}$$

参考式(4.3.28)，其中

$$h(\boldsymbol{x}_{(t_f)},t_f)+\int_{t+\Delta t}^{t_f}g(\boldsymbol{x}_{(\tau)},\boldsymbol{u}_{(\tau)},\tau)\mathrm{d}\tau=J_{t+\Delta t\to t_f}(\boldsymbol{x}_{(t)},t,\boldsymbol{u}_{(\tau)}),\quad t+\Delta t\leqslant\tau\leqslant t_f \tag{4.3.31}$$

$J_{t+\Delta t\to t_f}(\boldsymbol{x}_{(t)},t,\boldsymbol{u}_{(\tau)})$代表了系统从 $t+\Delta t$ 到末端时间 t_f 的性能指标，其最优代价为 $J^*_{t+\Delta t\to t_f}(\boldsymbol{x}_{(t+\Delta t)},t+\Delta t)$，对应的控制量为 $\boldsymbol{u}^*_{(\tau)}$，请注意这里的控制区间为 $\tau\in[t+\Delta t,t_f]$。根据贝尔曼最优化理论，式(4.3.30)一定可以写成

$$J^*_{t\to t_f}(\boldsymbol{x}_{(t)},t)=\min_{\boldsymbol{u}_{(\tau)}}\left\{J^*_{t+\Delta t\to t_f}(\boldsymbol{x}_{(t+\Delta t)},t+\Delta t)+\int_t^{t+\Delta t}g(\boldsymbol{x}_{(\tau)},\boldsymbol{u}_{(\tau)},\tau)\mathrm{d}\tau\right\},\quad t\leqslant\tau\leqslant t+\Delta t \tag{4.3.32}$$

它包含了两个部分：区间 $\tau\in[t+\Delta t,t_f]$的剩余最优代价 $J^*_{t+\Delta t\to t_f}(\boldsymbol{x}_{(t+\Delta t)},t+\Delta t)$；区间$[t,t+\Delta t]$的代价。

> 到这一步为止，推导过程与离散型动态规划类似。但是在连续系统中，我们无法像离散系统式(4.3.5)那样直接代入状态空间方程以消除 $\boldsymbol{x}_{(t+\Delta t)}$，因为连续系统的状态空间方程式(4.3.27)是由微分方程表示的。因此使用 $\boldsymbol{u}_{(t)}$ 表示 $J^*_{t\to t_f}(\boldsymbol{x}_{(t)},t)$需要一些技巧。