开发者说丨Apollo控制算法之LQR

知识点

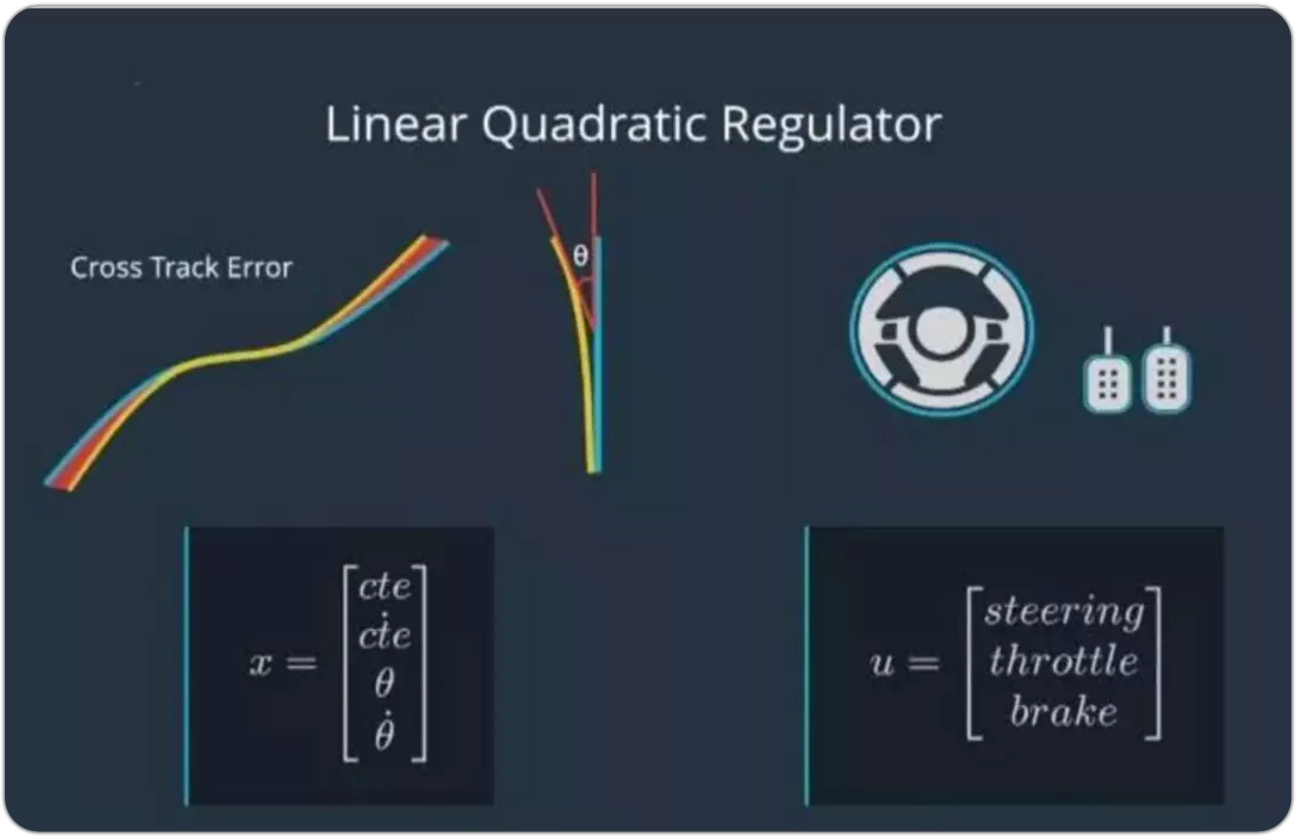
敲黑板，本文需要学习的知识点有

LQR 理论　**控制算法**
动态规划算法　**状态反馈控制**
最优控制量　**黎卡提方程**
等效闭环反馈矩阵

LQR 理论是现代控制理论中发展最早也最为成熟的一种状态空间设计法。特别可贵的是，LQR可得到状态线性反馈的最优控制规律，易于构成闭环最优控制。

LQR 最优设计是指设计出的状态反馈控制器 K 要使二次型目标函数 J 取最小值，而 K 由权矩阵 Q 与 R 唯一决定，故此 Q、R 的选择尤为重要。

而且 **Matlab 的应用**为 **LQR 理论仿真**提供了条件，更为我们实现稳、准、快的控制目标提供了方便。



Apollo使用LQR进行横向控制

线性二次调节器（Linear Quadratic Regulator 或LQR）是基于模型的控制器，它使用车辆的状态来使误差最小化。

Apollo 使用 LQR 进行横向控制。横向控制包含四个组件：

> **横向误差**
>
> **横向误差的变化率**
>
> **朝向误差**
>
> **朝向误差的变化率**

变化率与导数相同，我们用变量名上面的一个点来代表。

我们称这四个组件的集合为X，这个集合X捕获车辆的状态。除了状态之外，该车有三个控制输入：**转向**、**加速**和**制动**。我们将这个控制输入集合称为U。

本文由社区开发者——卜大鹏撰写，整理出**Apollo控制算法**之**LQR的推导过程**。

以下，ENJOY

我们考虑有如下离散线性系统：

$$x_{t+1} = Ax_t + Bu_t, x_0 = x^{init} \tag{1}$$

LQR的目标就是找到一组控制量 u_0, u_1 ,...使

x_0, x_1 ,...足够小，即系统达到稳定状态；

u_0, u_1 ,...足够小，即花费较小的控制代价。

为了达到上述效果，定义代价函数：

$$J=\sum_{\tau=0}^{N-1}(x_\tau^T Q x_\tau + u_\tau^T R u_\tau) + x_N^T Q_f x_N \tag{2}$$

其中x为**状态量**，u为**控制量**，Q为**状态权重矩阵**，R为**控制权重矩阵**， Q_f 为最终**状态权重矩阵，**N为到达最终状态的**控制序列数**。

根据代价函数，再定义$V_t(z)$表示为从t时刻的状态z开始，通过最优控制序列，获得的到结束后的最小代价。

$$V_t(z)=\min_{ut,\ldots,u_{N-1}}\sum_{\tau=t}^{N-1}(x_\tau^T Q x_\tau + u_\tau^T R u_\tau) + x_N^T Q_f x_N \tag{3}$$

根据公式3可得，当t=N时有：

$$V_N(z)=z^T Q_f z \tag{4}$$

即 V_N 只和最终状态有关。可以想到如果t时刻获取最优控制，上一时刻t-1也应该是最优控制，这样可以通过 V_N 倒推上一状态的控制来获取控制序列。

顺着这个思路，上一时刻与当前时刻的关系有：

$$V_t(z)=\min_w(z^T Q z + w^T R w + V_{t+1}(Az+Bw)) \tag{5}$$

其中公式6是当前时刻的代价，公式7是当前时刻到下一时刻到最终时刻的代价。

$$z^T Q z + w^T R w \tag{6}$$

$$V_{t+1}(Az+Bw) \tag{7}$$

至此，**求解最优控制序列转化为了动态规划问题。**

因为当前状态z与优化命题无关，公式5可以改写为：

$$V_t(z) = z^T Q z + \min_w (w^T R w + V_{t+1}(Az + Bw)) \tag{8}$$

根据公式4形式，假设t时刻的代价有如下表达形式：

$$V_t(z) = z^T P_t z \tag{9}$$

把 V_{t+1} 改写为与t时刻同样的形式，代入公式8后，有：

$$V_t(z) = z^T Q z + \min_w (w^T R w + (Az + Bw)^T P_{t+1} (Az + Bw)) \tag{10}$$

这样对于无约束凸优化问题，只需令其导数等于0即可求得最优解。对公式10中求最小控制量的矩阵求导得：

$$2w^T R + 2(Az + Bw)^T P_{t+1} B = 0 \tag{11}$$

由公式11求最优控制 w^* 有：

$$w^* = -(R + B^T P_{t+1} B)^{-1} B^T P_{t+1} A z \tag{12}$$

将该控制量带入公式8，有：

$$\begin{aligned} V_t(z) &= z^T Q z + w^{*T} R w^* + (Az + Bw^*)^T P_{t+1} (Az + Bw^*) && (13) \\ &= z^T (Q + A^T P_{t+1} A - A^T P_{t+1} B (R + B^T P_{t+1} B)^{-1} B^T P_{t+1} A) z && (14) \\ &= z^T P_t z && (15) \end{aligned}$$

公式15中 P_t 为：

$$P_t = Q + A^T P_{t+1} A - A^T P_{t+1} B (R + B^T P_{t+1} B)^{-1} B^T P_{t+1} A \qquad (16)$$

当t=N时有 $P_N = Q_f$ ，而

$$Q_f = Q_f^T \geq 0 \qquad (17)$$

在由公式16中P的推导表达式可以得到，任意时刻t有表达式

$$P_t = P_t^T \geq 0 \qquad (18)$$

至此，构造完满足凸优化条件的求解方程。

同时，当t相对于时域N很小时，P趋于稳定不变，有：

$$P_{ss} = Q + A^T P_{ss} A - A^T P_{ss} B (R + B^T P_{ss} B)^{-1} B^T P_{ss} A \qquad (19)$$

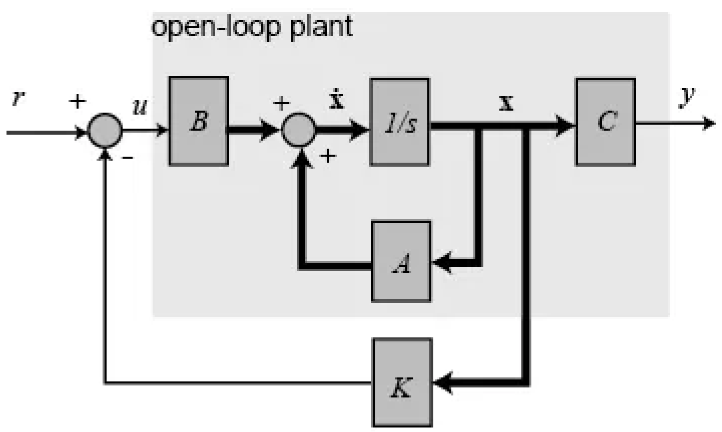
公式19被称为代数黎卡提方程，结合公式12，有：

$$u_t = K_{ss} x_t \qquad (20)$$

其中，

$$K_{ss} = -(R + B^T P_{ss} B)^{-1} B^T P_{ss} A z \qquad (21)$$

原控制系统框图可以表示为：



其中K即为通过迭代求解**黎卡提方程**得到的**等效闭环反馈矩阵**。

至此，**LQR算法求解过程**可以总结为：

> 1. 令P等于最终状态权重矩阵;
> 2. 迭代黎卡提方程求出新的P;
> 3. 当两次P的差值足够小时，计算反馈矩阵K;
> 4. 根据反馈矩阵K获取最优控制量u；

采用《Apollo控制算法之汽车动力学模型》一文中的模型，代入AD、BD、x_t，u_t到以上算法对应各项，因模型中的CD项为常量，在LQR求解完成后加入即可获得最终所需的系统控制量。

Apollo中还在LQR算出的控制量后加入前馈等操作，本文就不做详细介绍。

*参考文献：

Professor Stephen Boyd, Stanford University, Winter Quarter 2008-09 Linear Dynamical Systems