2008. MPC for tracking piecewise constant references for constrained linear systems

用于跟踪约束线性系统的分段常数参考指令的 MPC

作者

Daniel Limon 是西班牙塞维利亚大学的正教授，主要研究方向为MPC， Tracking control。目前他正在从事非线性（非线性模型，非线性约束）MPC，以及learning-MPC。

Abstract

本文使用标准的 MPC 方法处理了带约束的（非方阵）线性系统的分段常数参考指令跟踪控制。控制器保证了**满足约束条件**，并且系统能够使系统状态渐近演化到任何**可达的稳定状态**。因此，所有的分段可达设定点能够被**无偏跟踪**。如果稳定的目标点不可达，控制器将引导系统到最近的可达稳定状态。

这些目标通过以下方法实现：

1. 添加人工稳态和输入作为决策变量
2. 使用修改的成本函数来惩罚从人工稳态到目标稳态的距离
3. 考虑基于跟踪不变集概念的增广终端约束

相比标准MPC该方法**有更大的吸引域**，但是由于额外的自由度也会导致最优性的缺失，不过这可以通过修改偏置价值项的权重来减少影响。

Introduction

MPC是可以考虑**状态和输入约束**的控制方法，这是通过预测系统的演化以及计算满足约束的可达序列得到的。这种计算方法可以通过在每个采样时间计算最优问题得到，并通过滚动优化(receding horizon policy)求解控制律。通过设置终端状态的惩罚以及额外的终端约束可以保证渐近稳定。举例如下：

给定状态空间方程以及约束：

$$x(t+1) = Ax(t) + Bu(t)$$
$$x(t) \in \mathcal{X}, u(t) \in \mathcal{U}, \forall t \geq 0$$

其中$\mathcal{X} \subseteq \mathbb{R}^n$,$\mathcal{U} \subseteq \mathbb{R}^m$，设定价值函数：

$$J(x(t)) = J(x(t), U_{t\to t+N|t}) = p(x_{t+N|t}) + \sum_{k=0}^{N-1} q(x_{t+k|t}, u_{t+k|t})$$

N 是滚动时域的长度，p 是终端代价， q 是阶段代价。要求解最优的控制序列U^* ,可以定义优化问题：

$$\begin{aligned} J_t^*(x(t)) = \min_{U_{t\to t+N|t}} \quad & J(x(t), U_{t\to t+N|t}) = p(x_{t+N|t}) + \sum_{k=0}^{N-1} q(x_{t+k|t}, u_{t+k|t}) \\ \text{subj. to} \quad & x_{t+k|t} \in \mathcal{X}, u_{t+k|t} \in \mathcal{U} \\ & x_{t+N|t} \in \mathcal{X}_f \end{aligned}$$

可以用递归方法求解这个问题，也可以用常规的优化方法**一个模型预测控制（MPC）的简单实现 - 知乎 (zhihu.com)**。

对于模型预测控制器**处理非零目标稳态**的实际应用，这些稳态通常由稳态目标优化器提供。如果设定的目标稳态不可达（由于约束），就可以通过重新计算终端集合来解决。由于算力有限，在线重新计算也许并不能及时解决问题。针对一组有限的目标稳态的控制器的离线设计可能是一种解决方案，但要跟踪的目标集是有限的。针对这种跟踪问题（可行性的丧失），文献中提出了几种解决方案。

在处理跟踪问题的现有结果中在存在约束的情况下，一种很牛的方法是so-called command governors（Gilbert，Kolmanovsky，&Tan，1994）;该技术基于对参考指令添加非线性低通滤波器，以保证系统向参考指令演变。这可以看作是添加一个人工参考（滤波器的输出），在每个采样时间计算该参考指令，以确保系统演变，收敛到所需的参考指令上。

本文提出了一种新的MPC，用于一般非方阵线性系统，用于跟踪允许演化中任何允许的目标稳态。主要成分是：（i）人工稳态和输入被视为决策变量，（ii）成本函数惩罚人工稳态与期望稳态之间的偏差，以及（iii）扩展稳定终端条件被认为是包括在成本函数中添加跟踪误差惩罚项，并在终端状态和人工稳态和输入中添加终端约束。

Problem description

系统：

$$\begin{aligned} x^+ &= Ax + Bu \\ y &= Cx + Du \end{aligned}$$

约束：

$$(x_k, u_z) \in Z = \{z \in \mathbb{R}^{n+m} : A_z z \leq b\}$$

控制目标是让系统状态 x 趋近于目标稳定值 $\hat{x}_s$ 。

Preliminary results

给定目标输出 y_t ，对应的稳定状态 $z_x = (x_s, u_s)$ 要满足以下方程：

$$\begin{cases} x^+ = x_s = Ax_s + Bu_s, \\ y_t = Cx_s + Du_s. \end{cases}$$

$$\Rightarrow \begin{cases} (A-I)x_s + Bu_s = 0, \\ y_t = Cx_s + Du_s. \end{cases}$$

$$\Rightarrow \begin{bmatrix} A-I & B & 0 \\ C & D & -I \end{bmatrix} \begin{bmatrix} x_s \\ u_s \\ y_t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{3}$$

- 假设1：(A, B) 可稳

根据假设1，方程 3 有非平凡解

那么将这个问题的解写成参数化的形式：

$$
\begin{aligned}
z_s &= M_\theta \theta \\
y_t &= N_\theta \theta
\end{aligned}
$$

这个方程解出来是一个集合，集合参数化表示，就是以上形式。解第一个方程可以得到 M_θ 和 θ 的关系（θ 是没有范围约束的，参考三维空间的二维解），；根据 M_θ 和第二个方程就可以得到 N_θ 。

写成参数化的形式能通过参数 θ 的允许范围表征稳定状态的子空间。

最后考虑约束：

$$
Z_s = \{(x_s, u_s) = M_\theta \theta : M_\theta \theta \in Z\}.
$$

那么容许集可以写为：$X_s = Proj_x(Z_s)$ 和 $U_s = Proj_u(Z_s)$

那么控制律可以改写为：

$$
\begin{aligned}
u = K(x - x_s) + u_s &= Kx - Kx_s + u_s \\
&= Kx + [-K \;\; I_m](x_s, u_s) \\
&= Kx + L\theta
\end{aligned}
$$

$L = [-K \;\; I_m]M_\theta$ 。

将系统状态增广为：$w = (x, \theta)$ ，那么得到闭环系统 $w^+ = A_w w$。

$$
A_w = \begin{bmatrix} A + BK & BL \\ 0 & I_{n\theta} \end{bmatrix}
$$

对应的约束 W_λ：

$$
W_\lambda = \{w = (x, \theta) : (x, Kx + L\theta) \in Z, M_\theta \theta \in \lambda Z\}
$$

λ 取值范围0到1，可以对应地修改稳态的范围。

最大的容许不变集：

$$
\mathcal{O}_\infty^w = \{w : A_w^i \in W_1, \forall i \geq 0\}
$$

该集合不一定有有限解，改用以下集合：

$$
\mathcal{O}_{\infty,\lambda}^w = \{w : A_w^i \in W_\lambda, \forall i \geq 0\}
$$

这一节做了两件事：

1. 将系统的稳态解参数化表示，通过对 θ 的约束可以表述稳态时系统状态和控制输入的范围。
2. 将系统状态 x 增广为 $w = (x, \theta)$ ，得到增广后的状态空间方程、约束方程以及容许不变集。

之所以要这么做，是因为本文提出的方法是基于将增加的稳态状态（用 θ 表示）和输入作为**决策变量**，以及对修改的价值函数和扩展的终端约束。

MPC for tracking

- 假设2：
    1. Q, R, T 正定
    2. 矩阵 $(A + BK)$ 是赫尔维兹矩阵（稳定）。
    3. $(A + BK)^T P(A + BK) - P = -Q(K^T RT)$ (和李雅普诺夫稳定性判据有点类似)
    4. X_f^w 是不变容许集

给出价值函数：

$$V_N(x,\hat{x}_s,u,\theta)=\sum_{i=0}^{N-1}||x(i)-x_s||_Q^2+||u(i)-u_s||_R^2+||x(N)-x_s||_P^2+||x_s-\hat{x}_s||_T^2$$

其中 $\hat{x}_s$ 是人工给定的稳定状态。这个cost function 和标准 MPC 相比，多了一项人工稳态和计算的目标稳态的偏差；此外，还增加了一个终端惩罚：终端状态和目标稳态之间的偏差（跟踪误差成本）。

- 因为 θ 是多解的，所以 θ 也是决策变量
- 通过增加人工稳态以及价值函数中的 $||x_s-\hat{x}_s||_T^2$，相当于让计算出的容许集中的 x_s 能够尽最大可能的逼近目标输出，防止出现给定的目标输出是非稳定的，导致 x_s 无解。

根据价值函数写出优化问题：

$$\begin{aligned} V_N^*(x,\hat{x}_s)=\min_{u,\theta} & V_N(x,\hat{x}_s,u,\theta) \\ s.t. \quad & x(0)=x, \\ & x(j+1)=Ax(j)+Bu(j), \\ & (x_s,u_s)\in M_\theta\theta \\ & (x(N),\theta)\in X_f^w \end{aligned}$$

直接求解即可得到反馈控制律。

人工稳态 $\hat{x}_s$ 的给定或许可以用 MP 逆求解最小二乘解。

到目前为止，该方法解决了提出的第一个问题：

- 如果给定的目标输出不可达，如何解决。

  在标准 MPC 中，如果目标稳态不可达，会导致优化问题无解，或者系统不稳定；通过本文提出的方法，如果给定的目标稳态不可达，则使系统状态最大逼近目标稳态。原因是目标稳态 x_s 通过 θ 决定，而 θ 是通过计算得到的，肯定能有解；标准方法是写死的。

第二个问题：

- 标准 MPC 的吸引域就是 $O_\infty(\hat{x}_s):A^i\hat{x}_s\in Z_x$ ，本文方法的吸引域是：$Proj_x\,O_{\infty,\lambda}^w$ 。

  很明显后者更大（比如 $x>1\ \ \&\ \ x+y>1$ 中 x 的范围）。（不知道对不对）