基于值迭代的未知线性系统输出调节问题的数据驱动输出型方法

Haoyan Lin 和 Jie Huang，IEEE 终身会士

arXiv:2601.02748v1 [math.OC] 6 Jan 2026

摘要——在无扰动这一特殊情形下，研究者已采用基于状态和基于输出的内模方法，研究未知线性系统的输出调节问题。本文进一步通过值迭代，采用数据驱动输出型方法研究未知线性系统的输出调节问题。为此，我们首先提出一种新颖的输出反馈控制律，无需显式依赖观测器增益即可解决输出调节问题。随后，我们证明，为给定被控对象设计输出反馈控制律的数据驱动方法，可归结为对定义良好的增广辅助系统设计状态反馈控制律的数据驱动方法。因此，我们提出了一种系统的数据驱动方法，通过值迭代解决未知线性系统的输出调节问题。最后，我们在LQR意义下建立了数据驱动状态反馈控制律与数据驱动输出反馈控制律之间的关系。

索引词——自适应动态规划、数据驱动控制、线性输出调节问题、输出反馈。

I. 引言

作为一种数据驱动技术，强化学习（RL）也称为自适应动态规划（ADP），已被发展用于近似求解具有部分或完全未知动态特性的系统的各种最优控制问题 [1], [10], [14], [18], [20]。文献 [19]首次研究了系统矩阵未知的线性系统的线性二次调节器（LQR）问题，通过迭代更新控制策略来求解，这种方法称为策略迭代（PI）方法。文献 [8] 进一步研究了系统矩阵和输入矩阵均未知的线性系统的LQR问题。然而，由于PI方法要求初始稳定化反馈增益，而在系统未知时该增益可能难以获得，文献 [2] 提出了一种值迭代（VI）方法，通过迭代更新黎卡提方程的解，从而无需以初始稳定化反馈增益作为起点。上述所有文献均假设系统完整状态可用。当只能获得系统输出时，文献 [24] 考虑了输入矩阵已知的输出反馈线性二次调节问题，假设系统可通过静态输出反馈实现可稳定。文献 [15] 进一步通过引入折扣代价函数，研究了完全未知线性系统的输出反馈线性二次调节问题，其中折扣因子的上界由系统模型决定，以确保稳定性。为避免依赖折扣代价函数，文献 [16] 和 [17] 分别提出了基于PI和基于VI的输出反馈自适应动态规划算法，这些算法涉及构造无模型观测器。文献 [16] 和 [17] 中的方法尤其具有吸引力，因为它们保留了广为人知的分离原理。然而，这些方法要求参数化矩阵具有满行秩。由于该矩阵未知，验证其秩颇具挑战。最近，文献 [3]证明，当被控对象可控时，该矩阵具有满行秩，但这一结果不适用于被控对象仅可稳定的情况。更近来，文献[12] 提出了一种新方法，利用文献 [17]中的无模型观测器，将输出反馈线性二次调节问题成功转换为辅助线性系统的状态反馈线性二次调节问题。与文献 [17],相比，文献 [12] 中的方法消除了参数化矩阵的满行秩条件，并降低了计算成本，因为辅助系统的输入矩阵是已知的。

一种更一般的控制问题涉及对一类参考输入进行渐近跟踪，并抑制一类扰动。当参考输入和扰动均由一个称为外系统的线性自治系统生成时，该问题称为输出调节问题。该问题比单纯的稳定化更复杂，且已得到广泛研究（例如，见 [4], [5], [7]）。解决该问题主要有两种方法：前馈控制和内模原理。后者将原系统的输出调节问题转化为一个增广系统的稳定化问题，该增广系统由被控对象和动态补偿器（内模）组成。这种方法的优点是不需要求解调节器方程——一组线性矩阵方程。近年来，许多研究致力于利用数据驱动方法解决未知或部分未知系统的线性输出调节问题（LORP）。文献[6] 采用前馈控制方法，研究了系统矩阵、输入矩阵和扰动矩阵未知时的问题，该方法需要求解调节器方程。随后， [13] 考虑了单输入

本研究获得香港特别行政区研究资助局资助，项目编号为14203924。作者任职于香港中文大学机械与自动化工程系，香港（电子邮箱：hylin@mae.cuhk.edu.hk；jhuang@mae.cuhk.edu.hk。通讯作者：Jie Huang。）

单输出系统的LORP，并采用内模方法，允许输出矩阵未知。文献[22] 对 [13] 的方法进行了改进，以降低数据存储需求。近期， [11] 采用内模方法研究了多输入多输出系统的问题，并提出了一种新的学习算法，进一步降低计算成本并放宽可解条件。上述文献均假设拥有全状态信息。近期研究则着重于通过输出反馈控制解决未知系统的LORP [3], [23]。具体而言， [3] 研究了LORP的一种无扰动特殊情形，并采用基于PI方法的输出反馈控制律。此外， [23] 采用变分不等式方法研究了一般情形。然而， [23] 中的控制律依赖于外系统状态，因此并非真正的输出反馈。 [3] 和 [23] 均采用了 [16], 中的参数化观测器，因此要求参数化矩阵满行秩——而只有在被控对象可控时，才能保证这一条件成立。本文提出一种新颖的无模型输出反馈控制律来解决LORP。本文的主要贡献总结如下：

1）我们首先提出一种新颖的输出反馈控制律，无需显式依赖观测器增益即可求解输出调节问题。该控制律使我们能够设计一种数据驱动的输出反馈方法来求解输出调节问题。2）我们定义了一个增广辅助系统，并证明针对该系统求解状态反馈镇定问题，即可得到原系统输出调节问题的输出反馈解。3）我们的方法不要求参数化矩阵具有满行秩，因此适用于被控对象仅可稳定的情形。4）利用增广辅助系统的输入矩阵已知这一事实，我们提出了一种未知参数显著更少的数据驱动方法。这不仅大幅降低了计算成本，还放宽了可解条件。5）我们建立了增广辅助系统的数据驱动状态反馈LQR解与增广系统的数据驱动输出反馈LQR解之间的关系，并表明前者按照定理4所阐述的意义渐近收敛于后者。

本文其余部分安排如下。第二节回顾通过值迭代求解未知线性系统线性二次调节问题的状态反馈和输出反馈方法。第三节提出一种新颖的输出反馈控制律，用于求解输出调节问题，且不显式依赖观测器增益。第四节给出主要结果，包括一种系统的数据驱动输出反馈控制方法，该方法通过值迭代求解输出未知线性系统的调节问题。第五节进一步探讨在线性二次调节问题背景下，数据驱动状态反馈控制律与输出反馈控制律之间的关系。第六节通过两个仿真示例展示所提方法的有效性。最后，第七节总结全文。

符号说明 本文中， $\mathbb{R}$、$\mathbb{N}$和 $\mathbb{N}_+$分别表示实数集、非负整数集和正整数集。 $\mathcal{P}^n$ 表示所有$n \times n$ 实对称半正定矩阵的集合。 $||\cdot||$ 表示向量的欧几里得范数和矩阵的诱导范数。对于 $b = [b_1, b_2, \cdots, b_n]^T \in \mathbb{R}^n$，vecv($b$) $= [b_1^2, b_1b_2, \cdots, b_1b_n, b_2^2, b_2b_3, \cdots, b_{n-1}b_n, b_n^2]^T \in \mathbb{R}^{\frac{n(n+1)}{2}}$。

对于对称矩阵 $P = [p_{ij}] \in \mathbb{R}^{n\times n}$，vecs($P$) $= [p_{11}, 2p_{12}, \cdots, 2p_{1n}, p_{22}, 2p_{23}, \cdots, 2p_{n-1,n}, p_{nn}]^T \in \mathbb{R}^{\frac{n(n+1)}{2}}$ $v \in \mathbb{R}^n$ $|v|_P = v^TPv$。

对于 ，。对于列向量 a_i、$i = 1$、$\cdots$、s，col($a_1, \cdots, a_s$) $= [a_1^T, \cdots, a_s^T]^T$；并且，若 $A = (a_1, \cdots, a_s)$，则vec(A) $=$ col($a_1, \cdots, a_s$)。对于 $A \in \mathbb{R}^{n\times n}$， $\lambda(A)$表示由 A的所有特征值组成的集合。“blockdiag”表示分块对角矩阵算子。 I_n 表示维数为 n的单位矩阵。对于半正定矩阵 $Q \in \mathcal{P}^n$，$\sqrt{Q}$ 表示其半正定平方根。

II. 预备知识与问题表述

本节首先分别基于 [2] 和 [12], ，回顾用于求解无模型LQR问题的基于状态的VI方法和基于输出的VI方法。

A. 在未知 A、 B 和 C [2]的情况下求解LQR问题的基于状态的值迭代方法

考虑如下线性系统：

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx \end{aligned} \tag{1}$$

其中， $x \in \mathbb{R}^n$ 是系统状态， $u \in \mathbb{R}^m$ 是输入，$y \in \mathbb{R}^p$ 是测量输出。我们作出如下假设：

假设1。(A, B)是可稳定的。

假设2。对偶 (A, C) 可观测。

系统 (1) 的线性二次调节问题是求一个控制律 $u = Kx$，使代价 $\int_0^\infty (y^TQ_yy + u^TRu)d\tau$ 最小化，其中 $Q_y = {Q_y}^T \geq 0, R = R^T > 0$，且 ($A, \sqrt{Q_y}C$) 可观测。

令 $Q = C^TQ_yC$。根据假设1，以下代数黎卡提方程

$$A^TP^* + P^*A + Q - P^*BR^{-1}B^TP^* = 0 \tag{2}$$

存在唯一的正定解 $P^* > 0$。那么，系统 (1) 的线性二次调节问题的解为 $u^* = K^*x$，其中 $K^* = -R^{-1}B^TP^*$。

当未知时， [2] 提出了以下迭代方法，用于近似求解代数黎卡提方程 (2)。

为方便起见，对于任意向量 $a \in \mathbb{R}^n$、 $b \in \mathbb{R}^m$ 和任意整数 $s \in \mathbb{N}_+$，定义

$$\begin{aligned}\delta_a =&[\text{vecv}(a(t_1)) - \text{vecv}(a(t_0)), \cdots, \\ &\text{vecv}(a(t_s)) - \text{vecv}(a(t_{s-1}))]^T \\ I_{aa} =&[\int_{t_0}^{t_1} \text{vecv}(a)d\tau, \int_{t_1}^{t_2} \text{vecv}(a)d\tau, \cdots, \int_{t_{s-1}}^{t_s} \text{vecv}(a)d\tau]^T \\ I_{ab} =&[\int_{t_0}^{t_1} a \otimes Rbd\tau, \int_{t_1}^{t_2} a \otimes Rbd\tau, \cdots, \int_{t_{s-1}}^{t_s} a \otimes Rbd\tau]^T\end{aligned} \tag{3}$$

令 $H_k = A^T P_k + P_k A, K_k = -R^{-1}B^T P_k$。那么，将 $d(x^T(t)P_k x(t))\,dt$ 沿着方程 (1) 的解积分，得到

$$|x(t+\delta t)|_{P_k} - |x(t)|_{P_k} = \int_t^{t+\delta t} [|x|_{H_k} - 2u^T RK_k x]d\tau \tag{4}$$

(3) 和 (4) 可得

$$\Psi \begin{bmatrix}\text{vecs}(H_k) \\ \text{vec}(K_k)\end{bmatrix} = \Phi_k \tag{5}$$

其中 $\Psi = \begin{bmatrix}I_{xx} & -2I_{xu}\end{bmatrix}$ 且 $\Phi_k = \delta_x \text{vecs}(P_k)$。

(5) 的可解性由以下引理 [2]保证。

引理1. 若满足以下条件，则矩阵 Ψ 满列秩

$$\text{rank}([I_{xx}, I_{xu}]) = \frac{n(n+1)}{2} + mn \tag{6}$$

令 ϵ_k 为满足以下条件的时间步长序列

$$\epsilon_k > 0, \sum_{k=0}^{\infty} \epsilon_k = \infty, \sum_{k=0}^{\infty} \epsilon_k^2 < \infty \tag{7}$$

$\{B_j\}_{j=0}^{\infty}$ 是 $\mathcal{P}^n$ 中满足条件的有界子集集合

$$B_j \subseteq B_{j+1},\ j \in \mathbb{N},\ \lim_{j\to\infty} B_j = \mathcal{P}^n \tag{8}$$

并令 $\varepsilon > 0$ 为一个较小的实数，用于确定收敛判据。然后，可以通过算法1迭代求得式（2）的解。

Algorithm 1 Model-free VI Algorithm for Solving (2) [2]

Choose $P_0 = (P_0)^T > 0$. $k, j \leftarrow 0$.
loop
  Substitute P_k into (5) to solve H_k and K_k.
  $$\tilde{P}_{k+1} \leftarrow P_k + \epsilon_k(H_k + Q - (K_k)^T RK_k) \tag{9}$$
  if $\tilde{P}_{k+1} \notin B_j$ then
    $P_{k+1} \leftarrow P_0$. $j \leftarrow j + 1$.
  else if $||\tilde{P}_{k+1} - P_k||/\epsilon_k < \varepsilon$ then
    return P_k and K_k as approximations to P^* and K^*, respectively.
  else
    $P_{k+1} \leftarrow \tilde{P}_{k+1}$
  end if
  $k \leftarrow k + 1$
end loop

B. 在不知道 A、 B 和 C [12]的情况下求解线性二次型调节问题的基于输出的值迭代方法

[2] 中的方法假设被控对象的完整状态可用于反馈。在大多数应用中，无法获得完整状态。文献 [17] 进一步研究了仅输出 y 可用于反馈时的线性二次调节问题，并提出了一种用于估计状态的参数化方法。文献 [12] 进一步改进了 [17] 中的方法，将基于输出的线性二次调节问题转化为辅助系统的基于状态的线性二次调节问题。概括 [17] 和 [12], 中的方法，首先注意到，根据假设2，存在一个观测器增益 L ，使得 $A - LC$ 是赫尔维茨的。于是，以下Luenberger观测器:

$$\begin{aligned}\dot{\hat{x}} =&A\hat{x} + Bu + L(y - C\hat{x}) \\ =&(A - LC)\hat{x} + Bu + Ly\end{aligned} \tag{10}$$

将使状态$\hat{x}$ 指数地趋于 x 。

令$\Lambda(s) = \det(sI_n - A + LC) = s^n + \alpha_{n-1}s^{n-1} + \cdots + \alpha_1 s + \alpha_0$ $sI_n - A + LC^{-1} = \frac{D_{n-1}s^{n-1}+D_{n-2}s^{n-2}+\cdots+D_1 s+D_0}{\Lambda(s)}$。则 $D_i \in \mathbb{R}^{n\times n}$ $i = 0\ 1\ \cdots\ n-1$ $\zeta^i\ t \in \mathbb{R}^n$,,,。令ζ由以下系统描述:

$$\dot{\zeta}^i(t) = \mathcal{A}\zeta^i(t) + bw_i(t), \quad i = 1, \cdots, m+p \tag{11}$$

其中 $w_i(t) = u_i(t)$, $i = 1, \cdots, m$, $w_i(t) = y_{i-m}(t)$, $i = m+1, \cdots, m+p$, 并且

$$\mathcal{A} = \begin{bmatrix}0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -\alpha_0 & -\alpha_1 & -\alpha_2 & \cdots & -\alpha_{n-1}\end{bmatrix},\ b = \begin{bmatrix}0 \\ 0 \\ \vdots \\ 0 \\ 1\end{bmatrix}$$

备注 1. 由于 $\Lambda(s)$ 是用户定义的， $\mathcal{A}$ 和 b 是已知矩阵。

令 $\zeta = \text{col}(\zeta^1, \cdots, \zeta^{p+m})$ 和 $M = [M_1, \cdots, M_{p+m}]$, 其中 $M_i = [D_0 f_i, \cdots, D_{n-1}f_i]$，且 f_i为 B的第 i^{th}列，对应于 $i = 1,\cdots,m$和$(i-m)^{th}$ L的第$i = m+1\ \cdots\ m+p$ 10 11 $s = sI_n -$列。由于根据$()$和$()$, $\hat{x}() = ()^{-1}\left([B\ \ L]\begin{bmatrix}u \\ y\end{bmatrix}() \right.$ $\left. sI_n - \right] + \hat{}()\Big)$以及$()()$ $A + LC$ s s s ζ s u s y s $^{-1}[B\ \ L]\begin{bmatrix}u \\ y\end{bmatrix}() (s)$, 其中 $\hat{x}()()()()$ 分别是 $\hat{x}(t)$ $\zeta(t)$ $u(t)$ $y(t)$ 的拉普拉斯变换，我们有 $\lim_{t\to\infty}(M\zeta(t) - \hat{x}(t)) = 0$ 指数地趋于零。因此，$\lim_{t\to\infty}(M\zeta(t) - x(t)) = 0$ 指数地趋于零。

根据定义， ζ 的动力学可以写成以下形式:

$$\dot{\zeta} = (I_{m+p} \otimes \mathcal{A})\zeta + B_\zeta u + E_\zeta y \tag{12}$$

where $B_\zeta = \begin{bmatrix}I_m \otimes b \\ \mathbf{0}\end{bmatrix}$ and $E_\zeta = \begin{bmatrix}\mathbf{0} \\ I_p \otimes b\end{bmatrix}$.

Let $e_\zeta \triangleq M\zeta - x$。将 $y = x = CM\zeta - Ce_\zeta$代入式（12）可得

$$\dot{\zeta} = A_\zeta \zeta + B_\zeta u - E_\zeta C e_\zeta \tag{13}$$

where $A_\zeta = (I_{m+p} \otimes \mathcal{A}) + \begin{bmatrix} \mathbf{0} \\ I_p \otimes b \end{bmatrix} CM$.

Since e_ζ tends to exponentially, 经过有限时间 $t_0 > 0$后，e_ζ 的影响可以忽略。因此，式（13）可简化为以下形式：

$$\dot{\zeta} = A_\zeta \zeta + B_\zeta u, \quad t \geq t_0 \tag{14}$$

注意，由于 CM 未知， A_ζ 是一个未知矩阵，但 B_ζ 是已知的。

[12] 中的引理6和引理7表明，在假设1和假设2成立的条件下，矩阵对（A_ζ, B_ζ）是可稳定的；对于任意 $Q_y > 0$，矩阵对（$A_\zeta, \sqrt{Q_y CM}$）是可检测的。因此，以下代数黎卡提方程（15）存在唯一的半正定解 P_ζ^*：

$$A_\zeta^T P_\zeta^* + P_\zeta^* A_\zeta + Q_\zeta - P_\zeta^* B_\zeta R^{-1} B_\zeta^T P_\zeta^* = 0 \tag{15}$$

其中， $Q_\zeta = M^T Q M = M^T C^T Q_y C M \geq 0$、 Q_y和 R 与（2）中的相同。此外， [12] 中的定理2表明 $P_\zeta^* = M^T P^* M \geq 0$。因此，以下线性二次调节问题的解：

$$\begin{aligned} &\min_u \int_0^\infty (|\zeta|_{Q_\zeta} + |u|_R) dt, Q_\zeta \geq 0, R > 0 \\ &\text{subject to} \quad (14) \end{aligned} \tag{16}$$

由 $u_\zeta^* = K_\zeta^* \zeta$给出，其中 $K_\zeta^* = -R^{-1} B_\rho^T P_\zeta^*$。此外，可以证明 $K_\zeta^* = -R^{-1} B^T P^* M = K^* M$，其中$P^*$是（2）的唯一正定解。由于$M\zeta$指数地收敛到 x，$u_\zeta^* = K_\zeta^* \zeta = K^* M \zeta$将指数地收敛到（1）的最优控制器 $u^* = K^* x$。因此，（1）的输出反馈线性二次调节问题已转化为辅助系统（14）的状态反馈线性二次调节问题，且该辅助系统具有已知的输入矩阵。

令 $\bar{H}_k = A_\zeta^T \bar{P}_k + \bar{P}_k A_\zeta + Q_\zeta$, $\bar{K}_k = -R^{-1} B_\zeta^T \bar{P}_k$。那么，由 (14) 并利用关系式 $\zeta^T Q_\zeta \zeta = y^T Q_y y$，可得

$$\begin{aligned} |\zeta(t+\delta t)|_{\bar{P}_k} - |\zeta(t)|_{\bar{P}_k} + \int_t^{t+\delta t} |y|_{Q_y} d\tau \\ = \int_t^{t+\delta t} [|\zeta|_{\bar{H}_k} + 2u^T B_\zeta^T \bar{P}_k \zeta] d\tau \end{aligned} \tag{17}$$

注意到 $R\bar{K}_k = -B_\zeta^T \bar{P}_k$ 已知，由 (3) 和 (17) 可得

$$\tilde{\Psi} \text{vecs}(\bar{H}_k) = \tilde{\Phi}_k \tag{18}$$

其中 $\tilde{\Psi} = I_{\zeta\zeta}$ 和 $\Phi_k = \delta_\zeta \text{vecs}(P_k) + I_{yy} \text{vecs}(Q_y) + 2I_{\zeta u} \text{vec}(K_k)$。(18) 可解的条件由以下条件保证。

引理 2. 矩阵 Ψ 满列秩，当且仅当

$$\text{rank}(I_{\zeta\zeta}) = \frac{n_\zeta (n_\zeta + 1)}{2} \tag{19}$$

其中 $n_\zeta = (m+p)n$。

[12] 的基于输出的值迭代算法总结于算法2中，其中 ϵ_k 和 ε 的定义与第二节-A中的定义方式相同， $\{B_j\}_{j=0}^\infty$ 是 $\mathcal{P}^{n_\zeta}$ 中满足条件的有界子集的集合

$$\bar{B}_j \subseteq \bar{B}_{j+1},\ j \in \mathbb{N},\ \lim_{j\to\infty} \bar{B}_j = \mathcal{P}^{n_\zeta} \tag{20}$$

Algorithm 2 The Improved Output-based VI Algorithm [12]

Choose $\bar{P}_0 = (\bar{P}_0)^T \geq 0$. $k, j \leftarrow 0$.
loop
  Solve $\bar{H}_k$ from (18).
  $\tilde{\bar{P}}_{k+1} \leftarrow \bar{P}_k + \epsilon_k (\bar{H}_k - \bar{P}_k B_\zeta R^{-1} B_\zeta^T \bar{P}_k)$
  if $\tilde{\bar{P}}_{k+1} \notin \bar{B}_j$ then
    $\bar{P}_{k+1} \leftarrow \bar{P}_0$. $j \leftarrow j+1$.
  else if $||\tilde{\bar{P}}_{k+1} - \bar{P}_k|| / \epsilon_k < \varepsilon$ then
    return $\bar{P}_k$ and $-R^{-1} B_\zeta^T \bar{P}_k$ as approximations to P_ζ^* and K_ζ^*, respectively.
  else
    $\bar{P}_{k+1} \leftarrow \tilde{\bar{P}}_{k+1}$
  end if
  $k \leftarrow k+1$
end loop

备注2。算法2相较于 [16], [17] 中的方法有以下几点优势：首先， [16], [17]中的方法要求 M 满行秩，而算法2不依赖这一假设。其次， [16], [17] 中的方法需要求解一系列含有 $\frac{n_\zeta(n_\zeta+1)}{2} + mn_\zeta$个未知变量的方程。相比之下，利用矩阵 B_ζ 已知这一事实， [12] 将（18）中的未知变量数量减少至 $\frac{n_\zeta(n_\zeta+1)}{2}$，少于 [16],[17] 中的未知变量数量 mn_ζ。第三， [16], [17] 中的方法要求初始矩阵 P_0 正定，而 [12] 允许 P_0 半正定。第三个优势尤其值得关注，因为在许多情况下，黎卡提方程只能保证解为半正定。此外，由于现有的基于状态的 PI 和价值迭代方法仅适用于（A, B）可稳定且（A, $\sqrt{Q}$）可观测的情形，因此它们不能直接应用于（15），因为其中（A_ζ, B_ζ）可稳定，但（A_ζ, $\sqrt{Q_y CM}$）仅可检测。为此，他们进一步将基于状态的 PI 和价值迭代方法扩展到（A, B）可稳定且（A, $\sqrt{Q}$）可检测的情形。

III. 参数化输出反馈控制下的输出调节问题

本节首先总结基于 [7] 的内模方法解决输出调节问题的结果。然后，我们进一步提出一种设计输出反馈控制律的新方法，以处理输出调节问题。这种新方法为通过输出反馈控制律解决无模型输出调节问题奠定了基础。

A. 现有结果

考虑如下形式的连续时间线性系统:

$$\begin{aligned}\dot{x} =&Ax + Bu + Ev\\ y =&Cx\\ e =&Cx + Fv\end{aligned} \tag{21}$$

其中， $x \in \mathbb{R}^n$ 为系统状态， $u \in \mathbb{R}^m$ 为输入， $y \in \mathbb{R}^p$ 和 $e \in \mathbb{R}^p$ 分别为测量输出和受控输出。 $v \in \mathbb{R}^q$ 由以下自治系统生成:

$$\dot{v}(t) = Sv(t) \tag{22}$$

输出调节问题定义如下:

问题 1. 给定被控对象(21)和外系统(22)，设计状态反馈或输出反馈控制律，使闭环系统在 v 设为零时指数稳定，并且 $\lim_{t\to\infty} e(t) = 0$。

以下假设对于求解线性输出调节问题而言是相当标准的。

假设3. S 没有实部为负的特征值。

假设4。对于所有 $\lambda \in \lambda(S)$,

$$\operatorname{rank}\begin{bmatrix}A - \lambda I_n & B\\ C & \mathbf{0}\end{bmatrix} = n + p. \tag{23}$$

称 (G_1, G_2) 为 S 的最小 p 重内模，若

$$G_1 = \operatorname{blockdiag}\underbrace{(\beta, \cdots, \beta)}_{p-tuple}, G_2 = \operatorname{blockdiag}\underbrace{(\sigma, \cdots, \sigma)}_{p-tuple},$$

其中， β 是一个常数方阵，其特征多项式等于 S 的最小多项式； σ 是一个常数列向量，使得 (β, σ) 可控。如下形式的动态补偿器

$$\dot{z} =G_1 z + G_2 e \tag{25}$$

其中 $z \in \mathbb{R}^{n_z}$ 称为外系统的内模。内模的一个重要推论由 [7], 的引理1.26和1.27给出，现将其改述如下:

引理 3。在假设 1、3 和 4 下，设 (G_1, G_2) 为 S 的最小 -co 内模。则，
p py
(i) The pair $\begin{bmatrix}A & \mathbf{0}\\ G_2C & G_1\end{bmatrix}$ and $\begin{bmatrix}B\\ \mathbf{0}\end{bmatrix}$ is stabilizable.
(ii) F或任意,V e-cRow ng ma f xl等式 t i tion

$$ZS = G_1 Z + G_2 V \tag{26}$$

有一个解 $Z \in \mathbb{R}^{n_z \times q}$，那么 $V = \mathbf{0}$。

令 $\xi = \operatorname{col}(x, z)$。将 (21) 与 (25) 结合，得到以下所谓的增广系统:

$$\dot{\xi} = \begin{bmatrix}A & \mathbf{0}\\ G_2C & G_1\end{bmatrix}\xi + \begin{bmatrix}B\\ \mathbf{0}\end{bmatrix}u + \begin{bmatrix}E\\ G_2F\end{bmatrix}v \tag{27a}$$

$$e = \begin{bmatrix}C & \mathbf{0}\end{bmatrix}\xi + Fv \tag{27b}$$

设 $Y = \begin{bmatrix}A & \mathbf{0}\\ G_2C & G_1\end{bmatrix}$、 $J = \begin{bmatrix}B\\ \mathbf{0}\end{bmatrix}$ 和 $\bar{E} = \begin{bmatrix}E\\ G_2F\end{bmatrix}$。则 (27) 可写成以下紧凑形式:

$$\dot{\xi} =Y\xi + Ju + \bar{E}v \tag{28}$$

然后，根据 [7], 的引理 1.26 和注释 1.28，问题 1 的可解性可由以下定理概括。

定理 1. 在假设 1、3 和 4 下，设 (G_1, G_2) 为 S 的最小 p 重内部模型。设 $K = \begin{bmatrix}K_x & K_z\end{bmatrix}$ 满足 $Y + JK$ 为赫尔维茨矩阵。则问题 1 可由以下动态状态反馈控制律解决

$$u =K_x x + K_z z \tag{29a}$$

$$\dot{z} =G_1 z + G_2 e \tag{29b}$$

此外，设 L 满足 $A - LC$ 为赫尔维茨矩阵。则问题 1 可由以下动态输出反馈控制律解决

$$u = K_x\hat{x} + K_z z \tag{30a}$$

$$\dot{\hat{x}} = A\hat{x} + Bu + L(y - C\hat{x}) \tag{30b}$$

$$\dot{z} = G_1 z + G_2 e \tag{30c}$$

注释 3. 状态反馈控制律 (29) 仅涉及一个设计参数 K。对于任意 $Q_\xi = Q_\xi^T \geq 0$ $R = R^T > 0$，若 $(Y, \sqrt{Q_\xi})$ 可观测，则可以定义如下形式的黎卡提方程

$$Y^T P_\xi^* + P_\xi^* Y + Q_\xi - P_\xi^* J R^{-1} J^T P_\xi^* = 0 \tag{31}$$

由于 (Y, J) 可稳定，(31) 存在唯一的正定解 P_ξ^*。因此，稳定化反馈增益为 $K_\xi^* = -R^{-1}J^T P_\xi^*$。基于 (31) 和 (28)，还可以进一步提出一种数据驱动方法，在未知被控对象 [11] 参数的情况下求得 K。

注释 4. 由于在假设3下， S 和 $Y + JK$ 的谱不相交，根据 [7], 中的命题A.2，以下西尔维斯特方程存在唯一解 $X_c \in \mathbb{R}^{(n+n_z)\times q}$:

$$X_c S = (Y + JK)X_c + \bar{E} \tag{32}$$

将 X_c 分块为 $X_c = \begin{bmatrix}X\\ Z\end{bmatrix}$，其中 $X \in \mathbb{R}^{n\times q}$，且 $Z \in \mathbb{R}^{n_z \times q}$。则 (32) 表明

$$ZS =G_1 Z + G_2 V \tag{33}$$

其中

$$V = CX + F$$

由于方程 (33) 具有形式 (26)，根据引理 3，$V=\mathbf{0}$，即

$$\mathbf{0}=\begin{bmatrix}C & \mathbf{0}\end{bmatrix}X_c+F \tag{34}$$

定义 $\xi=\xi-X_cv$, $\bar{u}=u-KX_cv$。由 (34) 和 (32)，可得

$$\dot{\bar{\xi}}=Y\bar{\xi}+J\bar{u} \tag{35}$$

因此，$\bar{u}^*=K_\xi^*\xi$ 给出的控制律可解释为以下线性二次调节问题的解：

问题 2。

$$\begin{aligned}&\min_{\bar{u}}\int_0^\infty(|\bar{\xi}|_{Q_\xi}+|\bar{u}|_R)dt, Q_\xi\geq 0, R>0\\&\text{subject to}\quad (35)\end{aligned} \tag{36}$$

B. 通过参数化输出反馈控制实现输出调节问题

由于输出反馈控制律 (30) 中的观测器 (30b) 由未知矩阵 A, B, C 和 L 定义，因此无法实现控制律 (30)。在本小节中，我们将用 (12) 中给出的已知观测器替代 (30b)。结合基于变分不等式的学习方法，该控制律将提供一种无模型方法，通过输出反馈控制律解决输出调节问题。

令 $K=\begin{bmatrix}K_x & K_z\end{bmatrix}$ 满足 $Y+JK$ 为赫尔维茨矩阵，并令 L 满足 $A-LC$ 为赫尔维茨矩阵。我们考虑以下动态输出反馈控制律

$$u=K_xM\zeta+K_zz \tag{37a}$$
$$\dot{\zeta}=(I_{m+p}\otimes\mathcal{A})\zeta+B_\zeta u+E_\zeta y \tag{37b}$$
$$\dot{z}=G_1z+G_2e \tag{37c}$$

令 $x_c=\text{col}(x,\zeta,z)$。将式 (21) 与式 (37) 结合，得到以下闭环系统：

$$\dot{x}_c=A_cx_c+B_cv \tag{38a}$$
$$e=C_cx_c+Fv \tag{38b}$$

where $A_c=\begin{bmatrix}A & BK_xM & BK_z\\ E_\zeta C & I_{m+p}\otimes\mathcal{A}+B_\zeta K_xM & B_\zeta K_z\\ G_2C & \mathbf{0} & G_1\end{bmatrix}$, $W_c=\text{blockdiag}(I_n,M,I_{n_z})$, $B_c=\begin{bmatrix}E\\ \mathbf{0}\\ G_2F\end{bmatrix}$, and $C_c=\begin{bmatrix}C & \mathbf{0} & \mathbf{0}\end{bmatrix}$.

引理 4. 在假设 1–4 下，存在 K_x、K_z 和 L，使得 A_c 为赫尔维茨矩阵。

证明：根据引理 3，在假设 1、3 和 4 下，(Y J $K=$,) 对是可稳定的。因此，$\begin{bmatrix}K_x & K_z\end{bmatrix}$ 存在 a 使得 $Y+JK=\begin{bmatrix}A+BK_x & BK_z\\ G_2C & G_1\end{bmatrix}$ 是赫尔尔维茨的，并且在假设 2 下，存在 L 使得 $A-LC$ 是赫尔维茨的。令 $W_c=$ 分块对角矩阵(I_n,M,I_{n_z})，并且

$$A_\eta=\begin{bmatrix}A & BK_x & BK_z\\ LC & A-LC+BK_x & BK_z\\ G_2C & \mathbf{0} & G_1\end{bmatrix} \tag{39}$$

利用 [12] 中引理 5 的以下恒等式

$$M(I_{m+p}\otimes\mathcal{A})=(A-LC)M \tag{40}$$
$$M\begin{bmatrix}I_m\otimes b\\ \mathbf{0}\end{bmatrix}=B \tag{41}$$
$$M\begin{bmatrix}\mathbf{0}\\ I_p\otimes b\end{bmatrix}=L \tag{42}$$

得到

$$W_cA_c=\begin{bmatrix}A & BK_x & BK_z\\ LC & A-LC+BK_x & BK_z\\ G_2C & \mathbf{0} & G_1\end{bmatrix}W_c=A_\eta W_c \tag{43}$$

因此，令 $\eta=W_cx_c=\text{col}(x,M\zeta,z)$，得到

$$\begin{aligned}\dot{\eta}&=W_cA_cx_c+W_cB_cv\\&=A_\eta\eta+W_cB_cv\end{aligned} \tag{44}$$

Let

$$T=\begin{bmatrix}I_n & \mathbf{0} & \mathbf{0}\\ -I_n & I_n & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & I_{n_z}\end{bmatrix} \tag{45}$$

Then

$$TA_\eta T^{-1}=\begin{bmatrix}A+BK_x & BK_x & BK_z\\ \mathbf{0} & A-LC & \mathbf{0}\\ G_2C & \mathbf{0} & G_1\end{bmatrix} \tag{46}$$

由于 $\begin{bmatrix}A+BK_x & BK_z\\ G_2C & G_1\end{bmatrix}$ 和 $A-LC$ 均为赫尔维茨矩阵，A_η 也是赫尔维茨矩阵。因此，当 v 设为零时，$\eta=\text{col}(x,M\zeta,z)$ 将指数地收敛到原点。

根据 (37)，令 v 为零，ζ 满足

$$\begin{aligned}\dot{\zeta}&=(I_{m+p}\otimes\mathcal{A})\zeta+B_\zeta u+E_\zeta y\\&=(I_{m+p}\otimes\mathcal{A})\zeta+B_\zeta K_xM\zeta+B_\zeta K_zz+E_\zeta Cx\end{aligned} \tag{47}$$

由于 $(I_{m+p}\otimes\mathcal{A})$ 是赫尔维茨矩阵，且当 v 为零时，η 指数地衰减至零，因此 ζ 指数地衰减至零。因此，当 v 为零时，x_c 指数地衰减至零，这意味着 A_c 是赫尔维茨矩阵。□

现在我们可以证明以下结果。

定理2。在假设 1–4 下，设 $K=[K_x,K_z]$ 和L 满足条件，使得 $Y+JK$ 和 $A-LC$ 均为赫尔维茨矩阵。则动态输出反馈控制律 (37) 解决了问题 1。

证明：根据引理 4，存在 K_x、K_z 和 L，使得 A_c 是赫尔维茨矩阵。在假设3下，S 和 A_c 的谱不相交。因此，根据 [7]，中的命题A.2，~ 以下西尔维斯特方程有唯一解 X_c：

$$\tilde{X}_cS=A_c\tilde{X}_c+B_c \tag{48}$$

~将 $\tilde{X}_c$ 分块为 $\begin{bmatrix} X \\ \hat{Z} \\ \tilde{Z} \end{bmatrix}$，其中 $\tilde{X} \in \mathbb{R}^{n\times q}$、$\hat{Z} \in \mathbb{R}^{n(p+m)\times q}$、~和 $Z \in \mathbb{R}^{n_z\times q}$。那么，西尔维斯特方程 (48) 可展开如下：

$$\tilde{X}S = A\tilde{X} + B(K_x M\hat{Z} + K_z\tilde{Z}) + E \tag{49a}$$

$$\hat{Z}S = E_\zeta C\tilde{X} + (I_{m+p}\otimes \mathcal{A} + B_\zeta K_x M)\hat{Z} + B_\zeta K_z\tilde{Z} \tag{49b}$$

$$\tilde{Z}S = G_1\tilde{Z} + G_2\tilde{V} \tag{49c}$$

其中

$$\tilde{V} = C\tilde{X} + F$$

由于方程 (49c) 的形式为 (26)，根据引理 3，$V = \mathbf{0}$，这意味着

$$\mathbf{0} = C_c\tilde{X}_c + F. \tag{50}$$

~令 $\bar{x}_c = x_c - X_c v$。则利用 (48) 和 (50) 得到如下结果：

$$\dot{\bar{x}}_c = A_c\bar{x}_c \tag{51a}$$

$$e = C_c\bar{x}_c \tag{51b}$$

由于 A_c 是赫尔维茨的，$\bar{x}_c$，因此 e 将指数趋于零。证明完毕。□

IV. 主要结果

本节将进一步探讨如何设计一种无模型的基于输出的控制律，以解决如下所述的输出调节问题：

问题 3. 给定被控对象 (21) 和外系统 (22)，设计一个动态输出反馈控制律，使得将 v 置零时闭环系统指数稳定，并且 $\lim_{t\to\infty} e(t) = 0$，同时无需知道被控对象参数 A、B、C、E 和 F。

由于我们假设内模 (25) 已知，因此还需要如下另一个假设：

假设 5. S 的最小多项式已知。

A. 通过修正内模方法解决基于输出的输出调节问题

由于我们通过一种独立于 v 的控制律处理渐近跟踪和干扰抑制问题，因此 [3] 和 [23] 中提出的方法不适用于我们的问题。我们需要进一步赋予我们的方法两个新的特性。首先，注意到，由于持续外部信号 v 的存在，Luenberger观测器 (10) 的状态 $\hat{x}$ 不会收敛于被控对象的状态 x。事实上，有如下关系成立：

$$\dot{x} - \dot{\hat{x}} = (A - LC)(x - \hat{x}) + Ev \tag{52}$$

因此，$M\zeta$ 也不会收敛到被控对象的状态 x。所以，我们需要找出x 与 ζ之间的关系。第二点则更为微妙。对于输出调节问题，我们需要设计一种基于输出的控制律，以稳定增广系统 (27)，而不是原始被控对象。因此，我们需要为增广系统 (27) 推导出一个所谓的增广辅助系统，它对增广系统 (27) 的作用，与辅助系统 (14) 对问题 (1) 的线性二次调节问题所起的作用相同。

我们首先 выяс明 x 与 ζ之间的关系。在假设2下，存在一个 L 使得 $A - LC$为赫尔维茨矩阵。于是，根据 [7], 中的命题A.2，下面的西尔维斯特方程

$$X'S = (A - LC)X' + E \tag{53}$$

由于 S与 $A-LC$ 的谱不相交，因此存在唯一解 $X' \in \mathbb{R}^{n\times q}$。令 $e_x(t) = M\zeta(t)+X'v(t)-x(t)$。由(40)–(42)和(53)，可得

$$\begin{aligned}
\dot{e}_x =& M\dot{\zeta} + X'\dot{v} - \dot{x} \\
=& M((I_{m+p}\otimes\mathcal{A})\zeta + \begin{bmatrix} I_m\otimes b \\ \mathbf{0}\end{bmatrix} u + \begin{bmatrix} \mathbf{0} \\ I_p\otimes b\end{bmatrix} y) \\
&+ X'Sv - Ax - Bu - Ev \\
=& (A - LC)M\zeta + Bu + Ly + X'Sv - Ax - Bu - Ev \\
=& (A - LC)(M\zeta - x) + (X'S - E)v \\
=& (A - LC)(M\zeta - x) + (A - LC)X'v \\
=& (A - LC)e_x
\end{aligned} \tag{54}$$

因此，(54)是指数稳定的。

因此，我们得到以下关系：

$$x(t) = M\zeta(t) + X'v(t) - e_x(t) \tag{55}$$

其中，$e_x(t)$指数地收敛于零。

备注5。文献 [3] 仅考虑了 $E = 0$的跟踪问题。在这种情况下，$\hat{x}$ 将收敛于x。另一方面，文献 [23] 基于下面的Luenberger观测器构造了一个已知观测器：

$$\dot{\hat{x}} = A\hat{x} + Bu + Ev + L(y - C\hat{x}) \tag{56}$$

它取决于 v。因此，[23]中的控制律也取决于 v，且维数较高。更详细的比较将在注12和注13中给出。

然后，我们需要定义一个增广辅助系统，以便通过求解与该增广辅助系统相关联的黎卡提方程来获得所需的控制增益。为此，将 $y = Cx = CM\zeta + CX'v - Ce_x$ 代入（12），得到以下动力学：

$$\dot{\zeta} = A_\zeta\zeta + B_\zeta u + E_\zeta CX'v - E_\zeta Ce_x \tag{57}$$

此外，将（55）代入（29b）可得：

$$\dot{z} = G_2CM\zeta + G_1 z + G_2(CX' + F)v - G_2Ce_x \tag{58}$$

令 $\rho = \mathrm{col}(\zeta, z)$。那么，（57）和（58）的紧凑形式如下：

$$\dot{\rho} = A_\rho\rho + B_\rho u + E_\rho v + D_\rho e_x \tag{59}$$

其中 $\rho \in \mathbb{R}^{n_\rho}$ 与 $n_\rho = n(m + p) + n_z$，

$$A_\rho = \begin{bmatrix} A_\zeta & \mathbf{0} \\ G_2CM & G_1\end{bmatrix}, B_\rho = \begin{bmatrix} B_\zeta \\ \mathbf{0}\end{bmatrix}$$

$$E_\rho = \begin{bmatrix} E_\zeta CX' \\ G_2(CX' + F)\end{bmatrix}, D_\rho = \begin{bmatrix} -E_\zeta C \\ -G_2C\end{bmatrix}$$

值得注意的是，尽管 A_ρ 未知，但 B_ρ已知。接下来，系统 (59) 在求解 (21) 的输出调节问题时将发挥与 (13) 在求解 (1) 的线性二次调节问题时相同的作用。为此，我们先建立以下结果。

定理 3. 在假设 1-5 下，
(i) 该对 $(A \quad \rho, \quad B_\rho)$ 是可稳定的。
(ii) 令 K_ρ 为任意满足 $A_\rho+B_\rho K_\rho$ 为赫尔维茨且 L 满足 $A-LC$ 为赫尔维茨的矩阵。那么，问题 3 可由以下动态输出反馈控制律解决:

$$\begin{aligned} u =&K_\rho\rho \\ \dot{\zeta} =&(I_{m+p}\otimes\mathcal{A})\zeta + B_\zeta u + E_\zeta y \\ \dot{z} =&G_1 z + G_2 e \end{aligned} \tag{60}$$

证明: 第 (i) 部分。在假设 1 至 5 下，令 $\begin{bmatrix} K_x & K_z \end{bmatrix}$ 满足 $\begin{bmatrix} A & \mathbf{0} \\ G_2C & G_1 \end{bmatrix} + \begin{bmatrix} B \\ \mathbf{0} \end{bmatrix} \begin{bmatrix} K_x & K_z \end{bmatrix}$ 为赫尔维茨，并令 L 满足 $A-LC$ 为赫尔维茨。

令 $K_\rho = \begin{bmatrix} K_x & K_z \end{bmatrix} W$，其中 $W =$ 为分块对角矩阵(M, I_{n_z})。按此方式设计 K_ρ 后，控制律 (60) 与控制律 (37) 相同。因此，根据引理 4，当 v 设为零时， x、 ζ 和 z 均指数地收敛到原点。

现在将 $u = K_\rho\rho$ 代入（59）并令 $v = 0$，可得

$$\begin{bmatrix} \dot{\rho} \\ \dot{e}_x \end{bmatrix} = \begin{bmatrix} A_\rho + B_\rho K_\rho & D_\rho \\ \mathbf{0} & A - LC \end{bmatrix} \begin{bmatrix} \rho \\ e_x \end{bmatrix} \tag{61}$$

由于我们已经证明，当 v 设为零时， ζ、 z 和 e_x 均指数地收敛到原点，因此矩阵 $A_\rho + B_\rho K_\rho = \begin{bmatrix} A_\zeta + B_\zeta K_x M & B_\zeta K_z \\ G_2 CM & G_1 \end{bmatrix}$ 必为赫尔维茨矩阵。因此，(A_ρ, B_ρ) 必须是可稳定的。

（ii）由于 (A_ρ, B_ρ) 是可稳定的，因此存在 K_ρ，使得 $A_\rho+B_\rho K_\rho$ 为赫尔维茨矩阵。因此，通过令 $v = 0$, col(ρ, e_x) 满足（61）。由于 ζ、 z 和 e_x 均指数地收敛到原点，根据（55），当 v 设为 0 时， x 变为

$$x(t) = M\zeta(t) - e_x(t) \tag{62}$$

因此， x 也指数地衰减至 0。

令 $K_\rho = [K_{1\rho}, K_{2\rho}]$，其中 $K_{2\rho} \in \mathbb{R}^{n_z}$ 和 $x_c =$ col(x, ζ, z)，并将闭环系统表示如下:

$$\dot{x}_c = \bar{A}_c x_c + B_c v \tag{63a}$$

其中， $\bar{A}_c = \begin{bmatrix} A \quad e = C_c x_c B K_{1\rho} y & BK_{2\rho} \\ E_\zeta C & I_{m+p}\otimes\mathcal{A} + B_\zeta K_{1\rho} & B_\zeta K_{2\rho} \\ G_2 C & \mathbf{0} & G_1 \end{bmatrix}$、 (63b)

B_c 和 C_c 的定义见（38）。由于我们已证明，当 v 设为零时，x_c 会指数地趋于零，因此 A_c必为赫尔维茨矩阵。其余证明与定理2的证明相当类似，下面仅作简述。事实上，由于 A_c 是赫尔维茨矩阵，以下西尔维斯特方程ˆ存在唯一解 X_c:

$$\hat{X}_c S = \bar{A}_c \hat{X}_c + B_c \tag{64}$$

再次将ˆ$X_c = \begin{bmatrix} X \\ \hat{Z}' \\ \tilde{Z}' \end{bmatrix}$ ˆ按ˆ$X \in \mathbb{R}^{n\times q}$、ˆ$Z' \in \mathbb{R}^{n(p+m)\times q}$和 $\tilde{Z}' \in \mathbb{R}^{n_z\times q}$进行分块。则西尔维斯特方程（64）推出以下结果

$$\hat{X}S = A\hat{X} + B(K_{1\rho}\hat{Z}' + K_{2\rho}\tilde{Z}') + E \tag{65a}$$

$$\hat{Z}'S = E_\zeta C\hat{X} + (I_{m+p}\otimes\mathcal{A} + B_\zeta K_{1\rho})\hat{Z}' + B_\zeta K_{2\rho}\tilde{Z}' \tag{65b}$$

$$\tilde{Z}'S = G_1\tilde{Z}' + G_2 V' \tag{65c}$$

其中

$$V' = C_c\hat{X}_c + F$$

由于方程 (65c) 的形式为 (26)，根据引理 3，$V' = \mathbf{0}$，这意味着

$$\mathbf{0} = C_c\hat{X}_c + F. \tag{66}$$

令 $\tilde{x}_c = x_c -$ ˆ$X_c v$。则利用 (64) 和 (66) 可得:

$$\dot{\tilde{x}}_c = \bar{A}_c\tilde{x}_c \tag{67a}$$

$$e = C_c\tilde{x}_c \tag{67b}$$

¯由于 A_c 是赫尔维茨的， $\tilde{x}_c$，因此 e 指数地趋于零。证明完毕。□

注6. 由于 e_x 指数地衰减至 0，经过有限的 $t_0 > 0$ 后，(59) 中的项 $D_\rho e_x$ 可以忽略。因此，当 $t \geq t_0$ 时，(59) 可简化为

$$\dot{\rho} = A_\rho\rho + B_\rho u + E_\rho v \tag{68}$$

我们将 (68) 称为增广辅助系统。

定理 3 比定理2 更有意思之处在于，它不需要 $K_\rho = [K_x M, K_z]$，因为由于 M 未知， $K_\rho = [K_x M, K_z]$ 是无法实现的。因此，它将问题 3 转化为增广辅助系统 (68) 的稳定化问题，从而可通过求解与增广辅助系统相关的黎卡提方程，获得所需的控制增益。事实上，令 $Q_\rho \geq 0$ 满足 $(A_\rho, \sqrt{Q_\rho})$ 可观测，并定义如下代数黎卡提方程:

$$A_\rho^T P_\rho^* + P_\rho^* A_\rho + Q_\rho - P_\rho^* B_\rho R^{-1} B_\rho^T P_\rho^* = 0. \tag{69}$$

由于 $(A_\rho, \quad B_\rho)$ 可稳定，且 $(A_\rho, \quad \sqrt{Q_\rho})$ 可观测，因此 (69) 存在唯一的正定解 P_ρ^* [9]、 [21]。基于 (69) 的解 P_ρ^*，无需知道被控对象的参数即可获得稳定化增益 K_ρ^*，具体内容将在下一小节中详述。

注7. 与注释 4 类似，我们也可以将黎卡提方程 (69) 与如下定义的某个线性二次调节问题联系起来。事实上，由于在假设3下，S 和 $A_\rho + B_\rho K_\rho$ 的谱互不相交，根据 [7],中的命题A.2，以下西尔维斯特方程存在唯一解$X_\rho \in \mathbb{R}^{n_\rho\times q}$:

$$X_\rho S = (A_\rho + B_\rho K_\rho)X_\rho + E_\rho \tag{70}$$

将 X_ρ 分块为 $X_\rho = \begin{bmatrix} X_\zeta \\ X_z \end{bmatrix}$，其中 $X_\zeta \in \mathbb{R}^{n_\zeta \times q}$，且$X_z \in \mathbb{R}^{n_z \times q}$。则 (70) 意味着

$$X_z S = G_1 X_z + G_2(\begin{bmatrix} CM & \mathbf{0} \end{bmatrix} X_\rho + CX' + F) \tag{71}$$

由于方程 (71) 的形式与 (26) 相同，根据引理 3,

$$\mathbf{0} = \begin{bmatrix} CM & \mathbf{0} \end{bmatrix} X_\rho + CX' + F \tag{72}$$

令 $\bar{\rho} = \rho - X_\rho v$, $\bar{u} = u - K_\rho X_\rho v$。由 (72) 和 (70)，可得

$$\dot{\bar{\rho}} = A_\rho \bar{\rho} + B_\rho \bar{u} \tag{73}$$

因此，$\bar{u}^* = K_\rho^* \bar{\rho}$ 给出的控制律可解释为以下线性二次调节问题的解:

问题 4。

$$\begin{aligned} &\min_{\bar{u}} \int_0^\infty (|\bar{\rho}|_{Q_\rho} + |\bar{u}|_R) dt, Q_\rho \geq 0, R > 0 \\ &\text{subject to} \quad (73) \end{aligned} \tag{74}$$

尽管 ρ 的动力学未知，但 ρ 是可用的。因此，可以通过采用基于变分不等式的方法求解代数黎卡提方程 (69)，从而获得稳定化反馈增益K_ρ^* 。

B. 求解 (69) 的变分不等式方法

为得到求解 (69) 的基于变分不等式的算法，令 $H_k^\rho = A_\rho^T P_k^\rho + P_k^\rho A_\rho$。沿着 (68) 的解对 $d(\rho^T(t) P_k^\rho \rho(t))\, dt$ 积分，得到

$$\begin{aligned} &|\rho(t+\delta t)|_{P_k^\rho} - |\rho(t)|_{P_k^\rho} \\ &= \int_t^{t+\delta t} \left(|\rho|_{H_k^\rho} + 2u^T B_\rho^T P_k^\rho \rho + 2v^T E_\rho^T P_k^\rho \rho \right) d\tau \end{aligned} \tag{75}$$

对于任意向量 a、b 以及任意整数 $s \in \mathbb{N}_+$，令

$$\Gamma_{ab} = [\int_{t_0}^{t_1} a \otimes b d\tau, \int_{t_1}^{t_2} a \otimes b d\tau, \cdots, \int_{t_{s-1}}^{t_s} a \otimes b d\tau]^T \tag{76}$$

则 (75) 意味着

$$\Psi^\rho \begin{bmatrix} \mathrm{vecs}(H_k^\rho) \\ \mathrm{vec}(E_\rho^T P_k^\rho) \end{bmatrix} = \Phi_k^\rho \tag{77}$$

其中

$$\begin{aligned} \Psi^\rho &= [I_{\rho\rho}, 2\Gamma_{\rho v}] \\ \Phi_k^\rho &= \delta_\rho \mathrm{vecs}(P_k^\rho) - 2\Gamma_{\rho(B_\rho u)} \mathrm{vec}(P_k^\rho) \end{aligned} \tag{78}$$

(77) 的可解性由以下结果保证。

引理 5. 若满足以下条件，则矩阵 Ψ^ρ 满列秩

$$\mathrm{rank}([I_{\rho\rho}, \Gamma_{\rho v}]) = \frac{n_\rho(n_\rho+1)}{2} + q n_\rho \tag{79}$$

用于近似 $-R^{-1}B_\rho^T P_\rho^*$的VI算法总结于算法 3 中，其中 ϵ_k 和 ε 的定义方式与第二节-A中的相同，$\{B_j\}_{q=0}^\infty$ 是 $\mathcal{P}^{n_\rho}$ 中的一组有界子集，满足

$$\tilde{B}_j \subseteq \tilde{B}_{j+1},\ j \in \mathbb{N},\ \lim_{j\to\infty} \tilde{B}_j = \mathcal{P}^{n_\rho} \tag{80}$$

算法 3 求解 (69) 的无模型 VI算法

施加任意局部本质有界的初始输入 u^0。
选择适当的 $t_0 > 0$，使得 $\|e_x\|$ 足够小。
从 t_0 收集数据，直到满足秩条件 (79)
选择 $P_0^\rho = (P_0^\rho)^T > 0$。 $k, j \leftarrow 0$
循环
  将 P_k^ρ 代入（77）以求解 H_k^ρ 和 K_k^ρ $= -R^{-1}B_\rho^T P_k^\rho$.

$$\tilde{P}_{k+1}^\rho \leftarrow P_k^\rho + \epsilon_k (H_k^\rho + Q_\rho - (K_k^\rho)^T R K_k^\rho) \tag{81}$$

  如果 $\tilde{P}_{k+1}^\rho \notin B_j$ 则
    $P_{k+1}^\rho \leftarrow P_0^\rho$. $j \leftarrow j+1$。
  否则，如果 $||\tilde{P}_{k+1}^\rho - P_k^\rho|| / \epsilon_k < \varepsilon$ 则
    将 K_k^ρ 作为 $-R^{-1}B_\rho^T P_\rho^*$的近似
  else
    $P_{k+1}^\rho \leftarrow \tilde{P}_{k+1}^\rho$
  结束 if
  $k \leftarrow k+1$
结束循环

注释 8. 在推导（77）时，我们利用了 B_ρ 已知这一事实。因此，$K_k^\rho = -R^{-1}B_\rho^T P_k^\rho$ 也已知。由此，与 [3]等其他方法不同，我们无需在每次迭代时求解 K_k^ρ。

注释 9. 值 $\Gamma_{\rho v}$ 取决于外生信号 v。然而，根据假设 5，我们总可以假设 v 是已知的。事实上，正如 [6], 中假设 5 所指出的，存在一个已知向量 $\hat{v}(t)$、一个已知矩阵 S，其特征多项式等于 S的最小多项式，以及一个未知向量 C，它取决于 $v(0)$，使得，

$$\dot{\hat{v}} = \hat{S}\hat{v}, \quad v = \hat{C}\hat{v} \tag{82}$$

因此，方程 (21) 可改写如下:

$$\begin{aligned} \dot{x} &= Ax + Bu + E\hat{C}\hat{v} \\ y &= Cx \\ e &= Cx + F\hat{C}\hat{v} \end{aligned} \tag{83}$$

其中 $\hat{v}$ 是已知的。为简化符号，且不失一般性，本文假设 (21) 中的外生信号v 是已知的。

注释 10. 在学习阶段，初始输入的形式为 $u^0 = K_e \rho + \delta$，其中 K_e 是初始探索增益，δ 是探索噪声。如 [2], [8] 和 [11], 所述，秩条件的满足与探索噪声 δ的持续激励性质有关。探索信号 δ 的一种常见选择是多音正弦函数，其不同频率的数量取决于未知变量的数量。值得一提的是，通过减少未知变量的数量，更容易

满足改进算法所要求的秩条件。

C. 改进的 VI 算法

通过以下步骤，可以进一步降低计算复杂度，并放宽式（77）的可解条件：

首先，在式（77）中令 $k=0$，可得

$$\Psi^{\rho}\begin{bmatrix}\mathrm{vecs}(H_0^{\rho})\\ \mathrm{vec}(E_{\rho}^T P_0^{\rho})\end{bmatrix}=\Phi_0^{\rho} \tag{84}$$

在秩条件（79）下，求解式（84）可得 E_{ρ}。于是，式（77）中唯一的未知矩阵是 H_k^{ρ}。

因此，对于 $k\geq 1$，式（77）可简化为：

$$I_{\rho\rho}\mathrm{vecs}(H_k^{\rho})=\Phi_k^{\rho}-2\Gamma_{\rho v}\mathrm{vec}(E_{\rho}^T P_k^{\rho}) \tag{85}$$

矩阵 $I_{\rho\rho}$ 满列秩的条件是

$$\mathrm{rank}(I_{\rho\rho})=\frac{n_{\rho}(n_{\rho}+1)}{2} \tag{86}$$

显然，秩条件 (79) 保证矩阵 $I_{\rho\rho}$ 具有满列秩。

备注 11. 在我们的改进算法中，对于所有 $k\geq 1$，我们已将求解线性方程 (77) 的问题转化为求解 (85) 的问题。因此，未知变量的数量减少了 qn_{ρ}。根据 (84) 和 (85)，算法 3 可以修改为算法 4。

算法 4 求解 (69) 的改进的无模型 VI 算法 1：施加任意局部本质有界的初始输入 u^0。选择适当的 $t_0>0$，使 $\|e_x\|$ 足够小。从 t_0 收集数据，直至满足秩条件 (79)。2：选择任意正定矩阵 $P_0^{\rho}\in\mathbb{R}^{n_{\rho}\times n_{\rho}}$。由 (84) 求解 H_0^{ρ} 和 E_{ρ}。k，$j\leftarrow 0$。3：循环 4：将 P_k^{ρ} 代入 (85)，求解 H_k^{ρ} 和 $K_k^{\rho}=-R^{-1}B_{\rho}^T P_k^{\rho}$。5：$\tilde{P}_{k+1}^{\rho}\leftarrow P_k^{\rho}+\epsilon_k(H_k^{\rho}+Q_{\rho}-(K_k^{\rho})^T R K_k^{\rho})$。6：如果 $P_{k+1}^{\rho}\notin B_j$，则 7：$P_{k+1}^{\rho}\leftarrow P_0^{\rho}$。$j\leftarrow j+1$。8：否则，如果 $||P_{k+1}^{\rho}-P_k^{\rho}||/\epsilon_k<\varepsilon$，则 9：返回 K_k^{ρ}，作为 $-R^{-1}B_{\rho}^T P_{\rho}^*$ 的近似 10：否则 11：$P_{k+1}^{\rho}\leftarrow\tilde{P}_{k+1}^{\rho}$12：结束判断 13：$k\leftarrow k+1$14：结束循环 (87)

注12。将我们的控制律 (60) 与 [23]，中的控制律进行比较很有意思，后者具有以下形式：

$$\begin{aligned}u&=K_x M_x\mathrm{col}(\zeta,\zeta_v)+K_v z\\ \dot{\zeta}&=(I_{m+p}\otimes\mathcal{A})\zeta+B_{\zeta}u+E_{\zeta}y\\ \dot{\zeta}_v&=(I_q\otimes\mathcal{A})\zeta_v+(I_q\otimes b)v\\ \dot{z}&=G_1 z+G_2 e\end{aligned} \tag{88}$$

其中 $M_x=[M_1,\cdots,M_{p+q+m}]$ 其中 $M_i=[D_0 f_i,\cdots,D_{n-1}f_i]$，$f_i$ 是 i^{th} 中对应于$i=1$、$\cdots$、m的列，且 $(i-m)^{th}$ 是 L 中对应于 $i=m+1$、$\cdots$、$m+p$的列，而 $(i-m-p)^{th}$ 是 E 中对应于 $i=m+p+1$、$\cdots$、$m+p+q$的列。首先，(88) 依赖于 v，且不属于输出反馈形式。其次，(88) 包含额外的动力学 ζ_v。因此，(88) 的维数为 $n_{\gamma}=n(m+p+q)+n_z$，远大于我们控制律的维数 $n(p+m)+n_z$。第三，(88) 的综合基于 (A,B) 可控这一假设，而我们仅假设 (A,B) 可稳定。第四，[23] 的算法要求满足以下秩条件：

$$\mathrm{rank}([I_{\gamma\gamma},\Gamma_{\gamma u},\Gamma_{\gamma v}])=\frac{n_{\gamma}(n_{\gamma}+1)}{2}+(m+q)n_{\gamma}. \tag{89}$$

相比之下，我们在宽松得多的可解性条件（79）下，采用两种大幅简化的算法，即算法3和算法4，实现了我们的控制律。

注13. $E=0$ 的输出调节问题中的一个特殊情形已在 [3] 中研究。即使对于这一特殊情形，本文的方法也具有两项优势。首先，[3] 需要估计增广系统的整个状态$\xi=\mathrm{col}(x,z)$，而我们只需估计被控对象的状态 x，因为内模的状态 z 是已知的。因此，我们的控制律维数比 [3] 中的控制律小$n_z(m+p)$。其次，由于 [3] 没有利用 B_{ρ} 已知这一事实，因此 K_k^{ρ} 也是已知的，故 [3] 的方法需要求解一系列维数更大的方程。具体而言，[3] 中的方法要求满足以下秩条件：

$$\mathrm{rank}([I_{\chi\chi},\Gamma_{\chi u},\Gamma_{\chi\nu}])=\frac{n_{\chi}(n_{\chi}+1)}{2}+(m+p)n_{\chi} \tag{90}$$

其中，$\chi\in\mathbb{R}^{n_{\chi}}$ 是 [3] 中提出的增广辅助系统的状态，其维数为 $n_{\chi}=(n+n_z)m+2(n+n_z)p$，而 $\nu\in\mathbb{R}^p$ 是探索噪声。(90)中的秩条件比(79)更严格。为便于比较，考虑 $n=5$，$m=5$，$p=1$，$q=4$，$n_{\rho}=34$，$n_{\chi}=63$这一情形。表 I 比较了 [3], [23]和本文提出的各种算法的计算复杂度。

V. 增广辅助系统的线性二次调节问题

A. 问题 2 与问题 4 之间的关系

已知，对于任意满足 $(Y,\sqrt{Q_{\xi}})$ 可观测的 $Q_{\xi}\geq 0$，式 (31) 存在唯一的正定

表 I 秩条件比较

方法	方程标签	未知变量
[3] 中的算法	(90)	2394
[23] 中的算法	(89)	1971
我们的算法 3	(79)	731
我们的算法 4	(86)	595

解 P_ξ^*，用于求解问题 2 的控制律为

$$\bar{u}^* = K_\xi^* \bar{\xi} \tag{91}$$

其中 $K_\xi^* = -R^{-1}J^T P_\xi^*$。接下来，我们将建立式 (31) 的解 P_ξ^* 与式 (69) 的解 P_ρ^* 之间的关系，其中 $Q_\rho = W^T Q_\xi W$，且 $W = \text{blockdiag}(M, I_{n_z})$。为此，令 $\bar{C} = \begin{bmatrix} C & \mathbf{0} \\ \mathbf{0} & I_{n_z} \end{bmatrix}$ 和 $Q_\xi = \bar{C}^T \bar{Q} \bar{C}$，其中 $\bar{Q} \in \mathbb{R}^{p+n_z} > 0$。则 $\sqrt{Q_\xi} = \sqrt{\bar{Q}}\bar{C}$。由于假设2意味着对 $(Y, \bar{C})$ 是可观测的，因此对 $(Y, \sqrt{\bar{Q}}\bar{C})$ 也是可观测的。我们首先证明以下结果。

引理6。在假设2下，对 $(A_\rho, \sqrt{\bar{Q}}\bar{C}W)$ 是可检测的。

Proof: Let $H \in \mathbb{R}^{n_z \times n_z}$ be any Hurwitz matrix, and let $L_\rho = \begin{bmatrix} -E_\zeta & \mathbf{0} \\ -G_2 & H - G_1 \end{bmatrix} \sqrt{\bar{Q}}^{-1}$. Then

$$\begin{aligned} &A_\rho + L_\rho \sqrt{\bar{Q}}\bar{C}W \\ &= \begin{bmatrix} I_{m+p} \otimes \mathcal{A} + E_\zeta CM & \mathbf{0} \\ G_2 CM & G_1 \end{bmatrix} + \begin{bmatrix} -E_\zeta CM & \mathbf{0} \\ -G_2 CM & (H - G_1) \end{bmatrix} \\ &= \begin{bmatrix} I_{m+p} \otimes \mathcal{A} & \mathbf{0} \\ \mathbf{0} & H \end{bmatrix} \end{aligned}$$

由于 $\mathcal{A}$ 和 H 均为赫尔维茨矩阵，因此 $A_\rho + L_\rho \sqrt{\bar{Q}}\bar{C}W$ 为赫尔维茨矩阵。因此，$(A_\rho, \sqrt{\bar{Q}}\bar{C}W)$ 对是可检测的。□

我们现在可以给出本小节的主要结果。

定理4。在假设 1-5 下，对于任意 $Q > 0$ 和 $Q_\xi = \bar{C}^T \bar{Q} \bar{C}$，当 $Q_\rho = W^T Q_\xi W$ 时，代数黎卡提方程 (69) 存在唯一的半正定解 $P_\rho^* = W^T P_\xi^* W$，其中 P_ξ^* 是方程 (31) 的唯一正定解。因此，问题 4 的解为

$$\bar{u}^* = K_\rho^* \bar{\rho} \tag{92}$$

where $K_\rho^* = -R^{-1}B_\rho^T P_\rho^* = -R^{-1}B_\rho^T W^T P_\xi^* W = K_\xi^* W$.
证明：由于 (A_ρ, B_ρ) 可稳定，且 $(A_\rho, \sqrt{\bar{Q}}\bar{C}W)$ 可检测，带有 $Q_\rho = W^T Q_\xi W$ 的代数黎卡提方程 (69) 存在唯一的半正定解 P_ρ^*。我们只需证明 $P_\rho^* = W^T P_\xi^* W$，其中 P_ξ^* 是方程 (31) 的唯一正定解。由 (40)-(42) 可得

$$\begin{aligned} WA_\rho &= \begin{bmatrix} M & \mathbf{0} \\ \mathbf{0} & I_{n_z} \end{bmatrix} \begin{bmatrix} A_\zeta & \mathbf{0} \\ G_2 CM & G_1 \end{bmatrix} = \begin{bmatrix} MA_\zeta & \mathbf{0} \\ G_2 CM & G_1 \end{bmatrix} \\ &= \begin{bmatrix} A & \mathbf{0} \\ G_2 C & G_1 \end{bmatrix} \begin{bmatrix} M & \mathbf{0} \\ \mathbf{0} & I_{n_z} \end{bmatrix} = YW \end{aligned} \tag{93}$$

$$WB_\rho = \begin{bmatrix} M & \mathbf{0} \\ \mathbf{0} & I_{n_z} \end{bmatrix} \begin{bmatrix} B_\zeta \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} MB_\zeta \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} B \\ \mathbf{0} \end{bmatrix} = J \tag{94}$$

利用 (93) 和 (94)，可以验证

$$\begin{aligned} &A_\rho^T (W^T P_\xi^* W) + (W^T P_\xi^* W) A_\rho + Q_\rho \\ &- (W^T P_\xi^* W) B_\rho R^{-1} B_\rho^T (W^T P_\xi^* W) \\ =&W^T Y^T P_\xi^* W + W^T P_\xi^* Y W + W^T Q_\xi W \\ &- W^T P_\xi^* J R^{-1} J^T P_\xi^* W \\ =&W^T (Y^T P_\xi^* + P_\xi^* Y + Q_\xi - P_\xi^* J R^{-1} J^T P_\xi^*) W \\ =&0 \end{aligned} \tag{95}$$

因此，$W^T P_\xi^* W$ 是代数黎卡提方程 (69) 的一个半正定解。由于 P_ρ^* 是唯一的，$P_\rho^* = W^T P_\xi^* W$。

然后将 $P_\rho^* = W^T P_\xi^* W$ 代入 K_ρ^* 并使用式 (94)，可得

$$\begin{aligned} K_\rho^* =& -R^{-1}B_\rho^T P_\rho^* \\ =& -R^{-1}B_\rho^T W^T P_\xi^* W \\ =& -R^{-1}J^T P_\xi^* W \\ =& K_\xi^* W \end{aligned} \tag{96}$$

□

B. 问题 4 的变分不等式方法

定理4已建立求解问题 2 的控制律 (91) 与求解问题 4 的控制律 (92) 之间的关系，即若 $Q_\rho = W^T Q_\xi W$，则 $P_\rho^* = W^T P_\xi^* W$ 且 $K_\rho^* = K_\xi^* W$。然而，这种特定的 $Q_\rho = W^T Q_\xi W$ 不可行，因为 M，因此 $W =$ 分块对角矩阵(M, I_{n_z}) 是未知的。不过，对于 $E = \mathbf{0}$ 的特殊情形，可以绕过这一困难。为此，我们先建立以下结果。

引理 7. 在假设 1-5 下，若 $E = \mathbf{0}$，则对于 $\bar{Q} = \begin{bmatrix} Q_y & \mathbf{0} \\ \mathbf{0} & Q_z \end{bmatrix}$，其中 $Q_y > 0$ 且 $Q_z > 0$，$|\rho|_{Q_\rho}$ 指数地趋近于 $(|y|_{Q_y} + |z|_{Q_z})$。

证明：
首先注意，若 $E = \mathbf{0}$，则 (53) 中有 $X' = \mathbf{0}$。由 (55) 可知，$M\zeta(t)$ 指数地趋近于 $x(t)$。接下来，结合 $\bar{Q} = \begin{bmatrix} Q_y & \mathbf{0} \\ \mathbf{0} & Q_z \end{bmatrix}$，有 $Q_\xi = \begin{bmatrix} C^T Q_y C & \mathbf{0} \\ \mathbf{0} & Q_z \end{bmatrix}$。由于 $\xi = \text{col}(x, z)$，$|\xi|_{Q_\xi} = |Cx|_{Q_y} + |z|_{Q_z} = |y|_{Q_y} + |z|_{Q_z}$，其中 $y = Cx$。另一方面，由于 $Q_\rho = W^T Q_\xi W$，$|\rho|_{Q_\rho} = (M\zeta)^T C^T Q_y C M\zeta + |z|_{Q_z}$。由于 $M\zeta(t)$ 指数地趋近于 $x(t)$，$|\rho|_{Q_\rho}$ 指数地趋近于 $(|y|_{Q_y} + |z|_{Q_z})$。□

现在我们提出一种无模型 VI 方法来计算P_ρ^*和 K_ρ^*。为此，注意若 $E=\mathbf{0}$，则 (68) 可化简为以下形式

$$\dot{\rho}=A_\rho\rho+B_\rho u+\tilde{E}_\rho v. \tag{97}$$

where $\tilde{E}_\rho=\begin{bmatrix}\mathbf{0}\\ G_2F\end{bmatrix}$.

Integrating $\frac{d(\rho^T(t)P_k^\rho\rho(t))}{dt}$ along the solution of (97) gives

$$\begin{aligned}&|\rho(t+\delta t)|_{P_k^\rho}-|\rho(t)|_{P_k^\rho}\\&=\int_t^{t+\delta t}\left(|\rho|_{H_k^\rho}+2u^TB_\rho^TP_k^\rho\rho+2v^T\tilde{E}_\rho^TP_k^\rho\rho\right)d\tau\end{aligned} \tag{98}$$

(98) 可改写为如下形式:

$$\begin{aligned}&|\rho(t+\delta t)|_{P_k^\rho}-|\rho(t)|_{P_k^\rho}+\int_t^{t+\delta t}\left(|\rho|_{Q_\rho}\right)d\tau\\&=\int_t^{t+\delta t}\left(|\rho|_{\bar{H}_k^\rho}+2u^TB_\rho^TP_k^\rho\rho+2v^T\tilde{E}_\rho^TP_k^\rho\rho\right)d\tau\end{aligned} \tag{99}$$

其中 $\bar{H}_k^\rho=H_k^\rho+Q_\rho$。

根据引理 7，当 $Q_\xi=\begin{bmatrix}C^TQ_yC & \mathbf{0}\\ \mathbf{0} & Q_z\end{bmatrix}$ 时，$|\rho|_{Q_\rho}$ 指数地趋近于 $(|y|_{Q_y}+|z|_{Q_z})$。因此，(99) 可近似为以下方程:

$$\begin{aligned}&|\rho(t+\delta t)|_{P_k^\rho}-|\rho(t)|_{P_k^\rho}+\int_t^{t+\delta t}\left(|y|_{Q_y}+|z|_{Q_z}\right)d\tau\\&=\int_t^{t+\delta t}\left(|\rho|_{\bar{H}_k^\rho}+2u^TB_\rho^TP_k^\rho\rho+2v^T\tilde{E}_\rho^TP_k^\rho\rho\right)d\tau\end{aligned} \tag{100}$$

(100) 意味着

$$\tilde{\Psi}^\rho\begin{bmatrix}\mathrm{vecs}(\bar{H}_k^\rho)\\ \mathrm{vec}(\tilde{E}_\rho^TP_k^\rho)\end{bmatrix}=\tilde{\Phi}_k^\rho \tag{101}$$

其中

$$\begin{aligned}\tilde{\Psi}^\rho=&[I_{\rho\rho},2\Gamma_{\rho v}]\\ \tilde{\Phi}_k^\rho=&\delta_\rho\mathrm{vecs}(P_k^\rho)+I_{yy}\mathrm{vecs}(Q_y)+I_{zz}\mathrm{vecs}(Q_z)\\&-2\Gamma_{\rho(B_\rho u)}\mathrm{vec}(P_k^\rho)\end{aligned} \tag{102}$$

(101) 的可解性由以下引理保证。

~引理 8. 矩阵 Ψ^ρ 满列秩，当且仅当

$$\mathrm{rank}([I_{\rho\rho},\Gamma_{\rho v}])=\frac{n_\rho(n_\rho+1)}{2}+qn_\rho \tag{103}$$

与第 IV.C 节类似，我们可以进一步将 (101) 简化如下:

首先，在 (101) 中，令 $k=0$，可得

$$\tilde{\Psi}^\rho\begin{bmatrix}\mathrm{vecs}(\bar{H}_0^\rho)\\ \mathrm{vec}(\tilde{E}_\rho^TP_0^\rho)\end{bmatrix}=\tilde{\Phi}_0^\rho \tag{104}$$

~在秩条件 (103) 下，求解 (104) 可得 E_ρ。于是，(101) 中唯一的未知矩阵是 $\bar{H}_k^\rho$。

因此，对于 $k\geq 1$，(101) 可简化为:

$$I_{\rho\rho}\mathrm{vecs}(\bar{H}_k^\rho)=\tilde{\Phi}_k^\rho-2\Gamma_{\rho v}\mathrm{vec}(\tilde{E}_\rho^TP_k^\rho) \tag{105}$$

显然，秩条件 (103) 保证矩阵 $I_{\rho\rho}$ 满列秩。根据 (104) 和 (105)，算法 5 可修改为算法 6。

算法 5 使用 Q 求解 (69) 的 VI算法 $_\rho=W^TQ_\xi W$和 $E=0$

1: 应用任意局部本质有界的初始输入 u^0。
选择一个适当的 $t_0>0$，使得 $\|e_x\|$ 足够小。
从 t_0 收集数据，直到满足秩条件 (103) 为止。
2: 选择 $P_0^\rho=(P_0^\rho)^T\geq 0$。 $k,\ j\leftarrow 0$
3: 循环4: 将 P_k^ρ 代入 (101)，求解 $\bar{H}_k^\rho$ 和 $K_k^\rho=-R^{-1}B_\rho^TP_k^\rho$.
5:
$$\tilde{P}_{k+1}^\rho\leftarrow P_k^\rho+\epsilon_k(\bar{H}_k^\rho-(K_k^\rho)^TRK_k^\rho) \tag{106}$$
6: if $\tilde{P}_{k+1}^\rho\notin\bar{B}_j$ 则
7: $P_{k+1}^\rho\leftarrow P_0^\rho$. $j\leftarrow j+1$.
8: 否则，如果 $||P_{k+1}^\rho-P_k^\rho||/\epsilon_k<\varepsilon$ 则
9: 返回 K_k^ρ，作为 $-R^{-1}B_\rho^TP_\rho^*$的近似
10: else
11: $P_{k+1}^\rho\leftarrow\tilde{P}_{k+1}^\rho$
12: 结束条件判断
13: $k\leftarrow k+1$
14: 结束循环

算法 6 使用 $Q_\rho=W^TQ_\xi W$和 $E=0$求解 (69) 的改进的 VI 算法1: 施加任意局部本质有界的初始输入 u^0。选择适当的 $t_0>0$，使得 $\|e_x\|$ 足够小。从 t_0 收集数据，直到满足秩条件 (103)。2: 将 $P_0^\rho\in\mathbb{R}^{n_\rho\times n_\rho}$选为任意正定矩阵。根据 (104) 求解 $\bar{H}_0^\rho$和 $\tilde{E}_\rho$。 $k,\ j\leftarrow 0$。3: 循环4: 将 P_k^ρ 代入 (105)，求解 $\bar{H}_k^\rho$ 和 $K_k^\rho=-R^{-1}B_\rho^TP_k^\rho$。5:~ ~ 6: 如果 $P_{k+1}^\rho\notin B_j$，则7: $P_{k+1}^\rho\leftarrow P_0^\rho$。 $j\leftarrow j+1$。~8: 否则，如果 $||P_{k+1}^\rho-P_k^\rho||/\epsilon_k<\varepsilon$，则9: 返回 K_k^ρ作为 $-R^{-1}B_\rho^TP_\rho^*$的近似10: 否则11: $P_{k+1}^\rho\leftarrow\tilde{P}_{k+1}^\rho$12: 结束 if13: $k\leftarrow k+1$14: 结束循环

$$\tilde{P}_{k+1}^\rho\leftarrow P_k^\rho+\epsilon_k(\bar{H}_k^\rho-(K_k^\rho)^TRK_k^\rho) \tag{107}$$

VI. 数值示例

考虑如下线性系统 (21):

$$A=\begin{bmatrix}0&1&0\\0&0&0\\0&0&-1\end{bmatrix},B=\begin{bmatrix}0\\1\\0\end{bmatrix},S=\begin{bmatrix}0&1\\-1&0\end{bmatrix},$$
$$C=\begin{bmatrix}1&2&3\end{bmatrix},F=\begin{bmatrix}0.5&-0.8\end{bmatrix}.$$

由于（A, B）是可稳定但不可控的， [3] 和 [22] 中的现有方法无法处理此示例。该示例的内模如下:

$$G_1=\begin{bmatrix}0&1\\-1&0\end{bmatrix},G_2=\begin{bmatrix}0\\1\end{bmatrix}.$$

设 $L=\begin{bmatrix}-523&210&40\end{bmatrix}^T$，使得 $\lambda(A-LC)=\{-5,-6,-7\}$. 此外，我们有

$$M=\begin{bmatrix}M_u&M_y\end{bmatrix} \tag{108}$$

其中

$$\begin{aligned}M_u=&10^3\times\begin{bmatrix}1.1670&1.0470&0\\-0.5230&-0.4020&0.0010\\-0.0400&-0.0800&0\end{bmatrix}\\M_y=&10^3\times\begin{bmatrix}0.2100&-0.3130&-0.5230\\0.0000&0.2100&0.2100\\0.0000&-0.0000&0.0400\end{bmatrix}\end{aligned} \tag{109}$$

A. $E=0$的情形

此情形将由算法 6 处理。初始条件随机选取为 $v(0)=\text{col}(1,0.8)$, $x(0)=\text{col}(1,2,-0.8)$, $z(0)=\text{col}(0,0)$, $\zeta(0)=\text{col}(0,0,0,0,0,0)$。仿真采用探索初始输入 $u^0=K_0\zeta+\delta_1$进行，其中 $K_0=\begin{bmatrix}10&8&0&0&-4&-4\end{bmatrix}$ 且 $\delta_1=10(\sin(4t)+\sin(10t)+\sin(9t)-\sin(2t)-\sin(6t))$ 为探索噪声。观测器约需四秒才能收敛。令 $t_j=4+0.2j$，其中 $j=0$、1、$\cdots$、s 和 $s=120$。收集 4 秒至 28 秒的数据可保证满足式 (103)。为实现算法 6，设置各项参数为 $P_0^\rho=0.01I_8$、$B_\rho=\text{col}(0,0,1,0,0,0,0,0)$、$\tilde{B}_j=\{P\in\mathcal{P}^8:|P|\le 1000(j+20)\}$、$\epsilon_k=8/(k+10)$、$\varepsilon=0.05$、$Q_y=1$、$Q_z=I_2$ 和 $R=1$。随后，经过 8771 次迭代，收敛判据 $||\tilde{P}_{k+1}^\rho-P_k^\rho||/\epsilon_k<\varepsilon$得到满足。$P_k^\rho$和 K_k^ρ与期望值 P_ρ^*和 K_ρ^* 的比较结果如图1所示。

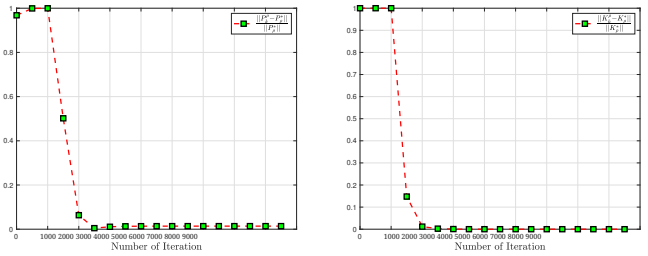


图1. P_k^ρ、 K_k^ρ与期望值 P_ρ^*、 K_ρ^*的比较

控制律 $u=K_{8771}^\rho\rho$ 于 $t=28$ 秒开始运行。图2显示了应用控制律后测量输出y 与参考信号 $-Fv$ 之间的误差。正如预期，控制器运行良好。

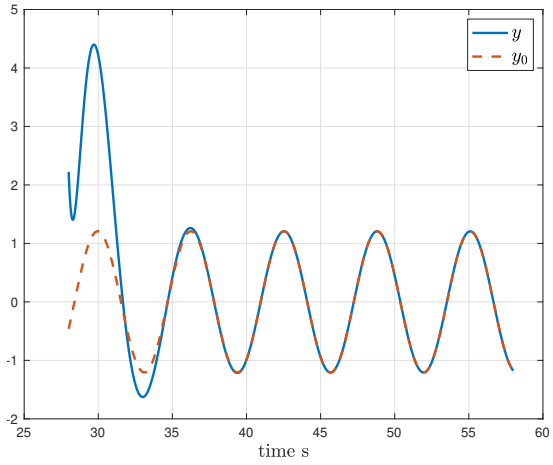


图2. 控制策略更新后系统输出 y 与参考信号 $-Fv$ 的曲线。

B. $E\neq 0$ 的情形

We apply Algorithm 4 to this scenario. Let $E=\begin{bmatrix}2&0\\0&1\\3&6\end{bmatrix}$. The initial conditions are randomly chosen as $v(0)=\text{col}(1,0.8)$, $x(0)=\text{col}(1,2,-0.8)$, $z(0)=\text{col}(0,0)$, $\zeta(0)=\text{col}(0,0,0,0,0,0)$。仿真采用探索初始输入 $u^0=K_0\zeta+\delta$，其中 $K_0=\begin{bmatrix}10&8&0&0&-4&-4\end{bmatrix}$ 和 $\delta=10(\sin(4t)+\sin(9t)+\sin(10t)-\sin(2t)-\sin(6t))$ 是探索噪声。观测器大约需要四秒才能收敛。令$t_j=4+0.2j$ 满足 $j=0$, 1, $\cdots$, s 和 $s=120$。收集 4 秒至 28 秒的数据可保证式 (79) 成立。

为实现算法 4，我们将各项参数设为 $P_0^\rho=0.1I_8$, $B_\rho=\text{col}(0,0,1,0,0,0,0,0)$, $\tilde{B}_j=\{P\in\mathcal{P}^8:|P|\le 1000(j+20)\}$, $\epsilon_k=20/(k+4000)$, $\varepsilon=0.01$,$Q_\rho=I_8$和 $R=1$。随后，经过 10602 次迭代，收敛判据 $||\tilde{P}_{k+1}^\rho-P_k^\rho||/\epsilon_k<\varepsilon$得到满足。图3展示了 P_k^ρ和 K_k^ρ与期望值 P_ρ^*和 K_ρ^* 的比较。

控制律 $u=K_{10602}^\rho\rho$ 在 $t=28$ 秒时开始运行。图4显示应用控制律后测量输出 y 与参考信号 $-Fv$ 之间的误差。正如预期，控制器运行良好。

注14。正如注12所指出的， [23] 的算法要求满足秩条件（89）。对于此示例，$n_\gamma=n(m+p+q)+n_z=14$。因此，秩条件（89）要求

$$\text{rank}([I_{\gamma\gamma},I_{\gamma u},I_{\gamma v}])=147 \tag{110}$$

然而，即使令 $s=200$，$\text{rank}([I_{\gamma\gamma},I_{\gamma u},I_{\gamma v}])=63<147$。

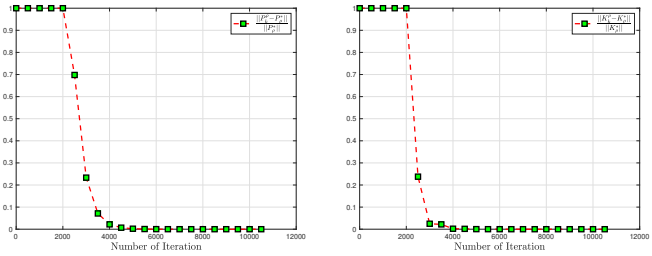


图3. P_k^ρ 与 K_k^ρ 和期望值 P_ρ^*、 K_ρ^*的比较

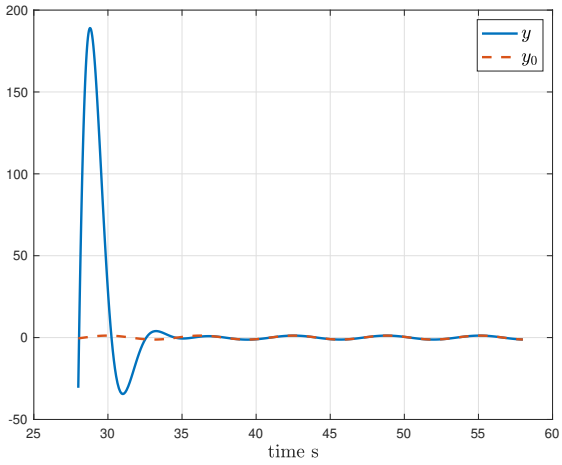


图4. 控制策略更新后系统输出 y 和参考信号 $-Fv$ 的曲线图。

VII. 结论

本文研究了通过值迭代实现的数据驱动输出方法解决未知线性系统的输出调节问题。首先，我们提出了一种新颖的输出反馈控制律，无需显式依赖观测器增益即可解决输出调节问题。随后，我们进一步证明，为给定对象设计输出反馈控制律的数据驱动方法，可以转化为为增广辅助系统设计状态反馈控制律的数据驱动方法，从而获得一种通过值迭代解决未知线性系统输出调节问题的系统化数据驱动方法。最后，我们在 LQR 意义下建立了数据驱动状态反馈控制律与数据驱动输出反馈控制律之间的关系。与现有结果相比，我们的控制律仅依赖于被控对象的输出，并可同时处理渐近跟踪和扰动抑制。此外，我们的算法显著降低了计算成本，并放宽了可解性条件。

参考文献

[1] D. Bertsekas，《强化学习与最优控制》，Athena Scientific，2019年。[2] T. Bian 和 Z. -P. Jiang，“面向数据驱动自适应最优控制设计的值迭代与自适应动态规划”，Automatica，第71卷，第348-360页，2016年。[3] C. Chen、L. Xie、K. Xie、F. L. Lewis 和 S. Xie，“基于输出反馈的强化学习实现连续时间系统的自适应最优输出跟踪”，Automatica，第146卷，110581，2022年。[4] E. J. Davison，“线性时不变多变量系统伺服机构问题的鲁棒控制”，IEEE自动控制汇刊，第21卷，第1期，第25-34页，1976年。[5] B. A. Francis 和 W. M. Wonham，“控制理论的内模原理”，Automatica，第12卷，第5期，第457-465页，1976年。[6] W. Gao 和 Z. P. Jiang，“线性系统的自适应动态规划与自适应最优输出调节”，IEEE自动控制汇刊，第61卷，第12期，第4164-4169页，2016年。[7] J. Huang，《非线性输出调节：理论与应用》，SIAM，2004年。[8] Y. Jiang 和 Z. P. Jiang，“完全未知动态特性的连续时间线性系统的计算自适应最优控制”，Automatica，第48卷，2704，2012 10, 2699 no. pp. - 。[9] V. Kucera，“对矩阵二次方程的贡献”，IEEE自动控制汇刊，第17卷，第3期，第344-347页，1972年。[10] F. L. Lewis、D. Vrabie 和 V. L. Syrmos，《最优控制》，美国新泽西州霍博肯：Wiley，2012年，第547页。[11] L. Lin 和 J. Huang，“基于内模原理的连续时间线性系统数据驱动最优输出调节”，IEEE自动控制汇刊，第70卷，第6期，第4202-4208页，2025年。[12] L. Lin、H. Lin 和 J. Huang，“连续时间线性系统数据驱动输出型线性二次调节问题的一种新方法”，IEEE自动控制汇刊，已获有条件接收。另见 arXiv预印本 arXiv:2509 18819，2025年。[13] Y. Liu 和 W. Gao，“基于内模原理的连续时间线性系统自适应最优输出调节”，2018年第九届IEEE年度普适计算、电子与移动通信会议（UEMCON），第38-43页，2018年。[14] S. Meyn，《控制系统与强化学习》，剑桥大学出版社，2022年。[15] H. Modares、F. L. Lewis 和 Z. P. Jiang，“利用离策略强化学习对未知连续时间线性系统进行最优输出反馈控制”，IEEE控制论汇刊，第46卷，第11期，第2401-2410页，2016年。[16] S. A. A. Rizvi 和 Z. Lin，《线性系统的输出反馈强化学习控制》，Birkhäuser，2023年。[17] S. A. A. Rizvi 和 Z. Lin，“利用动态输出反馈对连续时间系统进行基于强化学习的线性二次型调节”，IEEE控制论汇刊，第50卷，第11期，第4670-4679页，2019年。[18] R. S. Sutton 和 A. G. Barto，《强化学习：导论》，麻省理工学院出版社，2018年。[19] D. Vrabie、O. Pastravanu、M. Abu-Khalaf 和 F. Lewis，“基于策略迭代的连续时间线性系统自适应最优控制”，Automatica，第45卷，第2期，第477-484页，2009年。[20] P. Werbos，“用于控制与系统辨识的神经网络”，1989年IEEE决策与控制会议论文集，第260-265页，1969年。[21] W. M. Wonham，“关于随机控制的矩阵黎卡提方程”，SIAM控制学报，第6卷，697，1968 4, 681no. pp. - 。[22] K. Xie、X. Yu 和 W. Lan，“利用内模和自适应动态规划实现未知连续时间线性系统的最优输出调节”，Automatica，第146卷，110564，2022年。[23] K. Xie、Y. Zheng、W. Lan 和 X. Yu，“通过动态输出反馈和值迭代实现未知连续时间线性系统的自适应最优输出调节”，控制工程实践

，第141卷，105675，2023年。

[24] L. M. Zhu、H. Modares、G. O. Peen、F. L. Lewis 和 B. Yue，《利用积分强化学习的线性系统自适应次优输出反馈控制》，IEEE《控制系统技术汇刊》，第 23 卷第 1 期，264–273 页，2014年。



Haoyan Lin 于2022年获得中国广州华南理工大学自动化科学与工程专业工学学士学位。目前，她正在香港特别行政区香港中文大学机械与自动化工程系攻读博士学位。她目前的研究兴趣包括欧拉-拉格朗日系统、输出调节和数据驱动控制。



Jie Huang（IEEE终身会士）获得中国福州福州大学的文凭、中国南京南京理工大学的硕士学位，以及美国马里兰州巴尔的摩约翰斯·霍普金斯大学的博士学位。目前，他是香港中文大学（CUHK）机械与自动化工程专业的利荣森研究教授，同时担任香港中文大学工程学院副院长（科研）。他还

哈尔滨工业大学讲座教授。他的研究兴趣包括非线性控制理论及其应用、多智能体系统、博弈论以及飞行制导与控制。Huang博士是国际自动控制联合会、中国自动化学会和香港工程师学会会士，并且是中国工程院国际成员。