arXiv:2301.02019v1 [math.NA] 5 Jan 2023

端口哈密顿系统的保结构辨识——一种基于灵敏度的方法

Michael G¨unther、Birgit Jacob 和 Claudia Totz eck

摘要 我们提出一种基于梯度的校准算法，用于根据给定的时域输入输出数据辨识端口哈密顿系统。该算法借助灵敏度计算梯度，并经过专门设计，确保端口哈密顿系统的系统矩阵结构（斜对称性和半正定性）在每次迭代中都得以保留。由于我们只需要输入输出数据，因此还需要校准端口哈密顿系统内部状态的初始条件。基于合成数据的数值结果表明，该方法具有可行性。

1 引言

在耦合动力系统的保结构建模中，端口哈密顿框架允许构建整体端口哈密顿系统（PHS），前提是：(a) 所有子系统都是端口哈密顿系统；(b) 子系统的输入与输出之间存在线性耦合 [4, 5, 7, 8]。在实际应用中，这种方法有其局限性：对于特定子系统，要么没有可用于定义基于物理的端口哈密顿系统的基于物理的知识，要么 (b) 必须使用用户指定的仿真软件包，而这些软件包不提供有关内在动力学的信息，因此只能获得输入输出特性。

在这两种情况下，解决此类子系统问题的方法如下：通过物理测量或运行仿真软件包生成输入输出数据，并据此推导出一个能够最佳拟合这些输入输出数据的端口哈密顿系统代理模型。随后可用该端口哈密顿系统代理模型对子系统进行建模，最终得到一个具有保结构性质的耦合端口哈密顿系统。

Michael G¨unther、Birgit Jacob 和 Claudia Totzeck，贝吉舍大学伍珀塔尔分校，IMACM，高斯街20号，D-42119伍珀塔尔，电子邮件：[guenther, bjacob,totzeck]@uni-wuppertal.de

我们的方法旨在一步构造出拟合效果最佳的端口哈密顿系统模型，无需先推导出拟合效果最佳的线性状态空间模型，再在后处理步骤中寻找最近的端口哈密顿实现，参见例如[2, 3]。与 [1] 等方法不同，我们采用时域方法。[9]也采用时域方法，通过对端口哈密顿系统类别进行参数化，使得辨识过程中可以使用无约束优化求解器。这里，一个具有n个状态以及k个输入和输出的端口哈密顿系统，由$n(\binom{3n+1}{2}+2k)+k^2$个参数表示。本文将开发一种基于梯度的校准算法，以根据给定的时域输入输出数据辨识端口哈密顿系统。梯度借助灵敏度计算。

因此，我们考虑如下给出的代理端口哈密顿系统

$$\frac{d}{dt}x = (J-R)Qx + Bu, \qquad x(0) = \hat{x}, \tag{1a}$$

$$y = B^\top Qx, \tag{1b}$$

其中 J、Q、$R \in \mathbb{R}^{n\times n}$满足 $J = -J^\top$，$Q > 0$，$R \geq 0$. 假设已知一些参考数据 y_{data} 和 $B \in \mathbb{R}^{n\times k}$，以及输入信号 u。

任务是将系统矩阵和初始条件 $v = (J、Q、R、\hat{x})$ 拟合到数据。为此，我们定义代价泛函

$$\mathscr{J}(x,v) = \frac{1}{2}\int_0^T |y(t) - y_{\text{data}}(t)|^2 dt = \frac{1}{2}\int_0^T |B^T Qx(t) - y_{\text{data}}(t)|^2 dt$$

从而得到校准问题

$$\min \mathscr{J}(x,v) \quad \text{subject to} \quad (1). \tag{P}$$

由于我们只关注系统的输入输出行为，因此可以从动力学中消去 Q。事实上，通过乔列斯基分解可得V，其中 $Q = VV^\top$。

$$w = V^\top x,\ \tilde{B} = V^\top B,\ \tilde{J} = V^\top JV,\ \tilde{R} = V^\top RV$$

得到该系统

$$\frac{d}{dt}w = (\tilde{J}-\tilde{R})w + \tilde{B}u, \qquad w(0) = \hat{w}(= V^\top \hat{x}), \tag{2}$$

$$y = \tilde{B}^\top w. \tag{3}$$

为便于后续使用，我们将与 (2) 对应的状态算子e 定义为

$$e(w,v) = \begin{pmatrix} \frac{d}{dt}w - (\tilde{J}-\tilde{R})w - \tilde{B}u \\ w(0) - w_0 \end{pmatrix}.$$

因此，(2) 等价于 $e(w,v) = 0$.

变换后的代价泛函为

$$\tilde{J}(w,v)=\frac{1}{2}\int_0^T |y(t)-y_{\text{data}}(t)|^2dt=\frac{1}{2}\int_0^T |\tilde{B}^T w(t)-y_{\text{data}}(t)|^2dt.$$

~ ~变换后，我们需要识别矩阵 J,R 和 w_0. 为方便记号，我们定义可容许控制空间

$$\mathscr{V}=\{(\tilde{J},\tilde{R},w_0)\in\mathbb{R}^{n\times n}\times\mathbb{R}^{n\times n}\times\mathbb{R}^n:\ \tilde{J}^\top=-\tilde{J},\ \tilde{R}\geq 0\}.$$

请注意，根据标准常微分方程理论，该微分方程组存在唯一解。因此，我们可以定义控制到状态映射

$$S\colon \mathscr{V}\mapsto C([0,T],\mathbb{R}^n),\quad S(v)=w.$$

此外，我们用S 定义约化代价泛函

$$\hat{J}(v):=\frac{1}{2}\int_0^T |\tilde{B}^T S(v)(t)-y_{\text{data}}(t)|^2dt.$$

接下来，我们旨在推导一种基于梯度的算法，以数值求解校准问题。特别是，我们需要计算J. 的梯度。详情将在下一节介绍。从现在起，为方便记号，我们只考虑变换后的系统，并省略 ~。

2 灵敏度方法

我们强调，系统矩阵J,R 以及初始条件x 都是有限维的。因此，可以采用灵敏度方法 [6]来解决校准问题。

计算灵敏度需要为 Gˆateaux 导数选取可容许方向。由于结构约束，J 只能沿满足$h_J\,h_J^\top=-h_J$ 的方向变化，而R 只能沿对称矩阵方向变化。

ˆJ沿$h=$方向（h_J,h_R,h_x）的方向导数为

$$d\hat{J}(v)[h]=\langle\hat{J}'(v),h\rangle=\langle d_wJ(w,v),S'(v)h\rangle+\langle d_vJ(w,v),h\rangle$$

为计算该值，我们需要$dw(v,h)=S'(v)h$，即所谓的灵敏度。这里我们使用状态方程$e(w,v)=0$. 事实上，有

$$e_w(w,v)dw(v,h)+e_v(w,v)=0\qquad\Leftrightarrow\qquad e_w(w,v)dw(v,h)=-e_v(w,v)h.\quad(4)$$

我们强调，为了识别梯度 ˆ$J'(v)$，我们需要计算相对于 $\mathscr{V}$.切空间中所有基元素的方向导数

3 梯度下降算法

在上一节中，我们为下文介绍的梯度下降算法奠定了理论基础。

从系统矩阵和初始条件 $v_0 = (J_0, R_0, \hat{x}_0)$ 的初始猜测出发，我们通过求解 (4)，计算 $\mathcal{Y}$ 的切空间所有基元素对应的灵敏度 $dw(v,h)$，并利用灵敏度信息计算梯度 $\hat{J}'(v_0)$.然后，我们使用 Armijo 规则 [6], 寻找一个可行步长 σ（参见算法 1 中的伪代码），并更新系统矩阵和初始条件$v_0 \leftarrow v_0 - \sigma \hat{J}'(v_0)$. 当代价泛函的值足够小时，校准过程即停止。校准算法的伪代码见算法 2。

算法 1 Armijo 步长搜索

输入： 梯度 g,初始步长 σ 和安全参数 γ
输出： 可行步长 σ,新参数集 v'
$v' \leftarrow v + \sigma g$
当$\hat{J}(v') - \hat{J}(v') > -\gamma\sigma\|g\|^2$ **时**
$\sigma \leftarrow 0.5\sigma$
$v' \leftarrow v - \sigma g$
结束循环

算法 2基于梯度的校准算法

输入： 初始猜测v_0 和附加参数
输出： 校准后的系统矩阵和初始条件 $v = (J, R, \hat{x})$
当 $\hat{J}(v_0) > \varepsilon_{\text{stop}}$ **do**
对 所有可容许方向h **执行**
通过求解 (4) 计算 $dw(v_0, h)$
结束循环
识别 $\hat{J}'(v_0)$
通过 Armijo 规则寻找可接受的步长 σ，参见算法 1
$v_0 \leftarrow v_0 - \sigma \hat{J}'(v_0)$
结束 while 循环

所提出的算法可用于数值研究。下面我们将讨论一个概念验证，其中状态为$x \in C([0,T], \mathbb{R}^2)$。

4 概念验证

下面我们将讨论一个概念验证，其状态为$x \in C([0,T], \mathbb{R}^2)$，输出为$y \in C([0,T], \mathbb{R})$。在二维情形下，切

空间 $\mathcal{Y}$的基元素是可处理的。事实上，我们有以下基元素

$$J_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, R_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, R_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, R_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, x_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, x_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

我们假设$B = \begin{pmatrix} 1 & 1 \end{pmatrix}$ 已知，并且时间步t_k 处的输入信号给定为$u(t_k) = 1 + 0.1$ $N(0,1)$，其中$N(0,1)$表示均值为0、标准差为1的正态分布随机变量的一次实现。

为简便起见，我们假设时间步$t_k, k = 1,\ldots,K$与用于求解状态常微分方程的欧拉离散化时间步重合。事实上，利用初始猜测，我们采用欧拉格式求解(2)。随后，根据 (3) 得到输出y，并用它计算初始猜测对应的代价泛函。如果代价高于容差 $\varepsilon_{\text{stop}}$，我们就启动校准过程。

为方便记号，我们将灵敏度$dw(v,h)$拆分为h_J, h_R和h_x.几部分。关于J的灵敏度通过求解$e_w(w,v)dw(v,h_J) = -e_v(w,v)h_J$得到，可将其显式写为

$$\frac{d}{dt}dw(v,h_J) - (J-R)dw(v,h_J) = h_J w, \qquad dw(v,h_J)(0) = 0.$$

在二维情况下，只有一个可行方向 $h_J = J_1$. 对于关于R 的灵敏度，我们求解

$$\frac{d}{dt}dw(v,h_R) - (J-R)dw(v,h_R) = -h_R w, \qquad dw(v,h_R)(0) = 0$$

对于 $h_R = \{R_1, R_2, R_3\}$. 对于初始条件，我们求解

$$\frac{d}{dt}dw(v,h_x) - (J-R)dw(v,h_x) = 0, \qquad dw(v,h_x)(0) = h_x$$

对于 时$_x = \{$次$_1$,次$_2\}$.

代价泛函的方向导数为

$$d\hat{J}(v)[h] = \langle B^\top S(v) - y_{\text{data}}, B^\top S'(v)h \rangle = \int_0^T B\big(B^\top S(v)(t) - y_{\text{data}}(t)\big), (S'(v)h)(t)dt,$$

我们可以利用上面计算出的灵敏度对其进行评估。注意，对于上面讨论的所有$h_\bullet$，都有$d\hat{}J(v)[h_\bullet] \in \mathbb{R}$。因此，梯度可按如下方式组装

$$\hat{J}'(v) = \begin{bmatrix} d\hat{J}(v)[J_1]J_1 & \sum_{\ell=1}^{3} d\hat{J}(v)[R_\ell]R_\ell & \sum_{\ell=1}^{2} d\hat{J}(v)[x_\ell]x_\ell \end{bmatrix}^\top$$

5 数值结果

作为概念验证，我们通过求解状态系统来生成合成数据，其中数据矩阵J_{data}，R_{data} 和初始条件$\hat{x}_{\text{data}}$. 固定。对于以下结果，我们选择

$$J_{\text{data}} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad R_{\text{data}} = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.3 \end{pmatrix}, \quad \hat{x}_{\text{data}} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}. \tag{5}$$

这些数据得出参考输出y_{data}，如图1（左）所示。

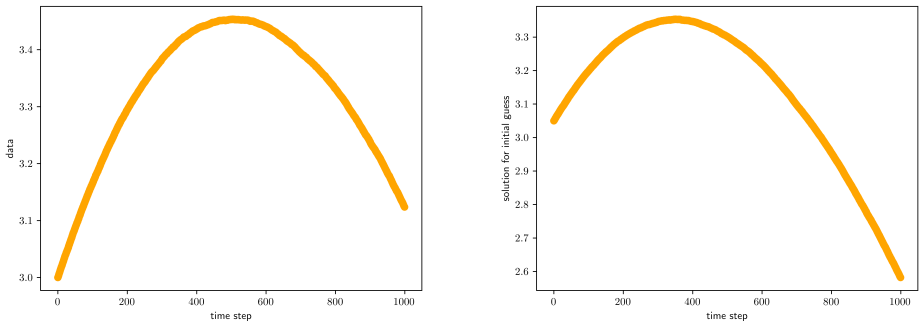

图1 左：与式（5）中给出的数据相对应的输出y_{data}。右：与初始猜测（6）相对应的输出y_0。

我们以如下初始猜测开始进行概念验证：

$$J_0 = \begin{pmatrix} 0 & 1.2 \\ -1.2 & 0 \end{pmatrix}, \quad R_0 = \begin{pmatrix} 0.4 & 0 \\ 0 & 0.4 \end{pmatrix}, \quad \hat{x}_0 = \begin{pmatrix} 1.1 \\ 1.95 \end{pmatrix} \tag{6}$$

由此得到图1（右）所示的输出。我们设定$T = 1$，并使用 1000 个时间步进行欧拉离散化。用于确定可接受步长的 Armijo 搜索以 $\sigma = 10$ 初始化；如果当前步长不可接受，则 $\sigma \leftarrow \sigma/2$。

算法 2 能够在 22 个梯度步内，以 $\varepsilon_{\text{stop}} = 1e^{-4}$ 重现输出 y_{data}。代价函数的变化如图2（左）所示，差值$y_{\text{data}} - y^{\text{opt}}$绘于图2（右）。校准矩阵和初始数据如下

$$J_{\text{opt}} = \begin{pmatrix} 0 & 1.073 \\ -1.073 & 0 \end{pmatrix}, \quad R_{\text{opt}} = \begin{pmatrix} 0.379 & -0.080 \\ -0.080 & 0.367 \end{pmatrix}, \quad \hat{x}_{\text{opt}} = \begin{pmatrix} 1.039 \\ 1.929 \end{pmatrix},$$

其中我们将数值舍入到精度 $1e^{-3}$.显而易见，R^{opt} 的非对角元素不为零。出于好奇，我们对R 仅限于对角矩阵的情况运行了同一个玩具问题。我们通过以下方式获得校准矩阵和初始数据：

$$J_{\text{opt},2} = \begin{pmatrix} 0 & 1.016 \\ -1.016 & 0 \end{pmatrix}, \quad R_{\text{opt},2} = \begin{pmatrix} 0.351 & 0 \\ 0 & 0.335 \end{pmatrix}, \quad \hat{x}_{\text{opt},2} = \begin{pmatrix} 1.023 \\ 1.973 \end{pmatrix}, \quad (7)$$

同样舍入到精度 $1e^{-3}$.R 中额外的结构信息总体上带来了更好的校准结果。参见图3，了解代价演化以及将R 限制为对角矩阵时校准所得输出的差异。

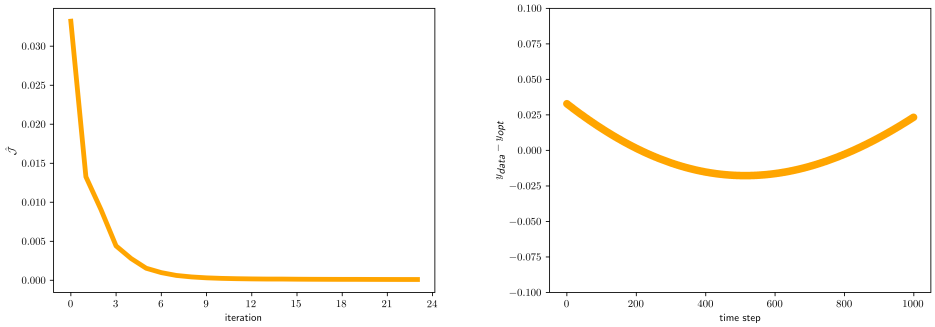

图2 左：与（5）中给出的数据相对应的输出y_{data}。右：参考输出与校准系统输出之差。

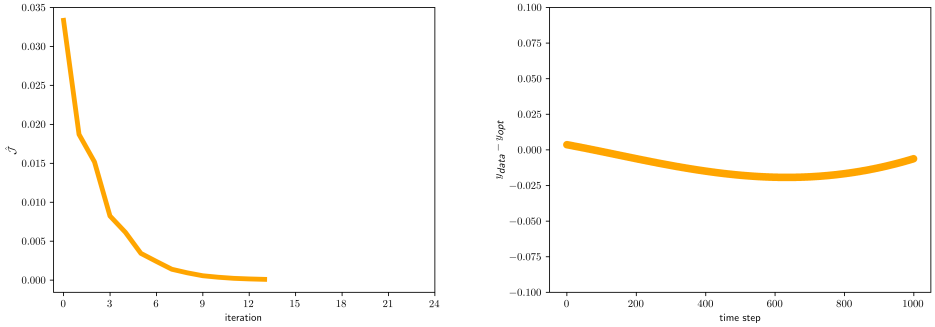

图3 左：与（5）中给出的数据相对应的输出y_{data}。右：参考输出与R 为对角矩阵的校准系统输出之差。

6 结论与展望

我们提出了一种基于梯度的算法，用于根据给定的输入—输出数据辨识由常微分方程构成的端口哈密顿系统。梯度借助灵敏度方法计算。概念验证表明，该方法具有可行性。

由于灵敏度方法的计算量随切空间基元素的数量增加，因此建议仅对小型系统采用所提出的校准算法。未来的工作中，我们将研究一种基于伴随的方法来计算梯度，以便为端口哈密顿输入—输出系统推导出一种保持结构的校准算法。

参考文献

1. Benner, P., Goyal, P., van Dooren, P.M.：基于频率响应数据的端口哈密顿系统辨识。《Systems & Control Letters》143(4):104741 (2020) 2. Cherifi, K., Goyal, P. K., Benner, P.：一种利用时域数据推断线性端口哈密顿实现的非侵入式方法。《Electronic Transactions on Numerical Analysis》：SciML 专刊，56，102-116 (2022)。3. Cherifi, K., Mehrmann, V., Hariche, K.：计算端口哈密顿系统最小实现的数值方法。arXiv:1903 1 4 07042. v. V. Duindam, A. Macchelli, S. Stramigioli, and H. Bruyninckx, 编，《复杂物理系统建模与控制》。德国：Springer，2009。5. D. Eberard, B. M. Maschke, and A. J. van der Schaft，“将哈密顿系统扩展至热力学相空间：迈向不可逆过程的几何学”，《Rep. Math. Phys.》，第60卷，第2期，第175–198页，2007。6. Hinze, M., Pinnau, R., Ulbrich, M., Ulbrich, S.：《带偏微分方程约束的优化》。Springer，柏林 (2009) 7. Mehrmann, V., Morandin, R.：端口哈密顿描述符系统的保结构离散化。2019 IEEE 第58届决策与控制会议 (CDC)，6863-6868 (2019)。8. A. Schaft，“端口哈密顿系统：入门综述”，《国际数学家大会论文集》第3卷，第1339-1366页，2006。9. Schwerdtner, P.：基于含噪频率响应数据的端口哈密顿系统辨识。arXiv:2106.11355。