Determining Governing Equations for Non-Linear Dynamical Systems from Measurement Data

Martin K. Steiger

University of Applied Sciences Upper Austria, Hagenberg

April 1, 2022

Problem Formulation

Let be given: Time Series IO Data
$y_t \in (y_1, ..., y_N)\,, u_t \in (u_1, ..., u_N)$

Goal: Approximate a Continuous Non-Linear Dynamical System that reproduces the measured IO behaviour

Hammerstein-Wiener Models have already been implemented for this purpose

$$y_t = f_2(G(z) * f_1(u_t)) \qquad \text{(discrete case)}$$
$$y(t) = f_2(H(s) * f_1(u(t))) \qquad \text{(continuous case)}$$

... where $f_1(\cdot)$ and $f_2(\cdot)$ are memory-less nonlinear functions and $G(z)$ resp. $H(s)$ represent the transfer function of a linear dynamical system

Problem Formulation

Question: can true physical behaviour be generated from the IO data instead of behavioral models?

Answer: yes, but not directly $\Rightarrow$ the formulated problem is under-determined, but we may use the behavioral model as a stepping stone

We assume that physical models may be formulated as a non-linear ordinary differential equation (ODE) system of 1st order

$$\dot{y}(t) = f(y(t), u(t), t; \Theta)$$

Possible approximation methods are AI approaches like Ordinary Differential Equation Networks (ODENets) or Sparse Identification of Non-Linear Dynamics (SINDy)...

Brief: Ordinary Differential Equation Networks

Goal: find an ODE system $\frac{d}{dt}x(t) = f(x(t), t, \Theta)$ that fits a given series of IO data

Method: Optimize the system parameters Θ using gradient-based methods. The required gradients are obtained through an arbitrary numerical solver and utilizing the adjoint method.

Advantages: The right-hand side may be chosen arbitrarily, but dense neural networks are common. Pure Black-Box method, which does not require a-priori knowledge of the approximated system. Results in compact systems in comparison to classical neural networks

Disadvantages: Not all numerical solvers provide gradients, Measures have to be taken to ensure the ODE stability

Brief: Sparse Identification of Nonlinear Dynamics

Goal: find an ODE system $\frac{d}{dt}x(t) = \Theta\ \Xi(x(t))$ that fits a given series of IO data, whereby $\Xi(x(t))$ is a library of elementary basis functions (polynomials, ...)

Method: Linear Regression Problem with Focus on Θ Sparsity using Lasso Regularization (L1-Norm) or Sequential Linear Least Squares Optimization. It is assumed that $x(t)$ can be measured and $\dot{x}(t)$ is determinable (through measurement or numerical differentiation)

Advantages: Fast Optimization Algorithms, Results in a Set of readable ODEs which may be interpreted easily, source terms can be easily included within $\Theta(x(t))$

Disadvantages: Requires prior knowledge or experience when selecting the basis functions, exponential computational effort for large systems or function libraries

Derivatives and Finite Differences

Starting from the Hammerstein-Wiener Difference Equation of the nth-order Linear Dynamic Section $G(z) = B(z)/A(z)$...

$$y_t = b_1 u_{t-1} + ... + b_n u_{t-n} - a_1 y_{t-1} - ... - a_n y_{t-n}$$

Goal: Difference Equation $\longleftrightarrow$ Differential Equation (Runge Kutta)

$$E : Ex_{t_n} = x_{t_n+1}, \Delta : \Delta = E - 1, \Delta x_t = x_{t_n+1} - x_{t_n}$$

$$D : Dx_{t_n} = \dot{x_{t_n}} = \frac{d}{dt} x(t)|_{t=t_n}$$

E and D are connected via a Taylor Series...

$$x_{t_n+1} = Ex_{t_n} = \sum_{k=0}^{\infty} \frac{h^k}{k!} D^k x_{t_n} = e^{hD} x_{t_n} \Rightarrow E = e^{hD}$$

$$D = \frac{1}{h} ln(E) = \frac{1}{h} ln(1 + \Delta) = \frac{1}{h} \sum_{k=0}^{\infty} (-1)^{k+1} \frac{\Delta^k}{k} \Rightarrow D^k \approx (\frac{\Delta}{h})^k$$

Derivatives and Finite Differences

Conclusion: If a difference equation contains the operator Δ^k, the corresponding ODE is also of order k

Problem: Higher Order ODEs can not be approximated by a single 1st order ODE of the form $\dot{y}(t) = f(y(t), u(t), t, \Theta)$

Solution: Any k-th order ODE may be formulated as system of k 1st order ODEs, but we lack the IO data to model the internal variables of the system...

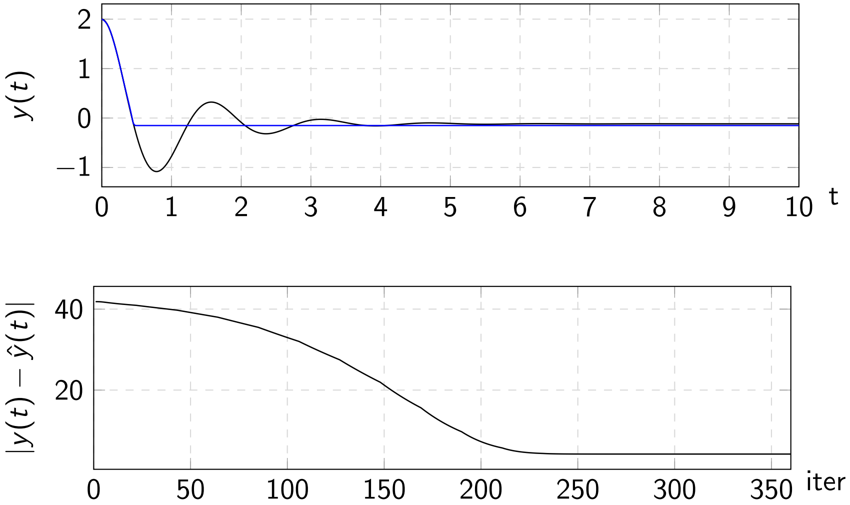
Alternatively we may utilize the behavioral models State-Space form with sampling rate T_s

$$\begin{array}{c} x_{t+1} = Ax_t + Bu_t \\ y_t = Cx_t + Du_t \end{array} \xrightarrow[s(z) \approx \frac{2}{T_s} \frac{z-1}{z+1}]{\text{bilinear}} \begin{array}{c} \dot{x}(t) = A_c x(t) + B_c u(t) \\ y(t) = C_c x(t) + D_c u(t) \end{array} \quad \cdots$$

which can be modeled without issues using classical approaches

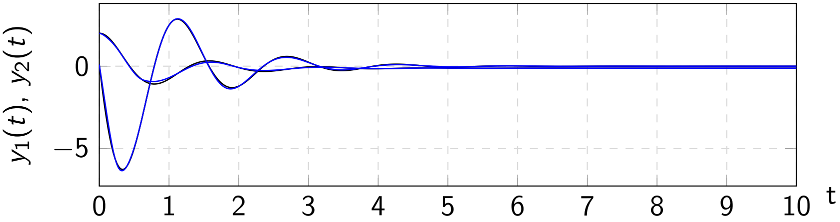
Visualizing the Approximation Problems

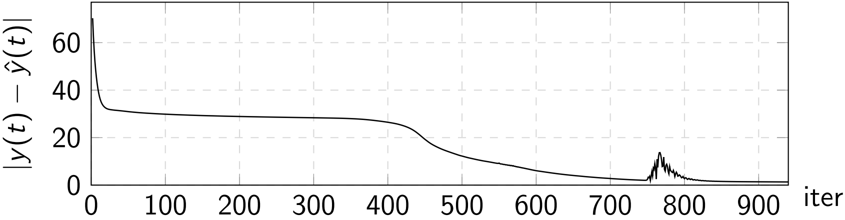
Example: $\ddot{y}(t) + 2\dot{y}(t) + 17y(t) = -2,\ y(0) = 2,\ \dot{y}(0) = 0$



Reformulating to 1st Order ODE System

Example: $\dot{y}_1(t) = y_2(t),\ \dot{y}_2(t) = -2y_2(t) - 17y_1(t) - 2$





Re-Defining the ODE Model Structure

Inspired by the discrete Hammerstein-Wiener Formulation, the following continuous model is proposed...

$$\sum_{k=0}^{K} \beta_k x^{(k)} = f(x(t), u(t), t, \Theta)$$
$$y(t) = x(t)$$

... which includes weighted higher-order derivatives on the left-hand side

Consequences

- ODENet approach may be neglected for behavioral models
- Parts of the SINDy-Methods are still relevant
- Non-Linearities can be encorporated explicitly within $f(\cdot)$

$\Rightarrow$ further investigation required