Inference and Synthesis of co-dimension one Bifurcations

Gregory Szep[1], Neil Dalchau[2] and Attila Csikasz-Nagy[1]

[1]*King's College London,* [2]*Microsoft Research Cambridge*

October 16, 2019

1 Introduction and Motivation

- qualitative equivalence more important
- fitting to time courses not appropriate
- gradient information makes optimisation tractable

2 Problem Statement

Suppose we would like to know the regions within paramter space $\theta \in \mathbb{R}^N$ where the steady state of a set of differential equations takes on a specified target form $t(p)$ along a known parameter $p \in \mathbb{R}$. The target is a label $t \in \{0,1\}$ that indicates whether the system is monostable or multistable at given p. We begin with a known form for the differential equations for $u \in \mathbb{R}^M$

$$\partial_t u = F(u|\,p,\theta) \tag{2.1}$$

and the set of steady states $U(\theta) := \left\{(u,p) \in \mathbb{R}^{M+1} : F(u|\,p,\theta) = 0\right\}$ can be found in a local region numerically using pseudo-arclength continuation along $u(s|\theta), p(s|\theta)$ [] where s parametrises the one co-dimensional curve – details in section 3. In order to perform an optimisation we need mapping from the states $U(\theta)$ to an output that can be compared to $t(p)$. Section 4 motivates this mapping and constructs the objective function.

3 Pseudo-arclength Continuation

- predictor-corrector algorithm

4 Objective Function

First we note that although $U(\theta)$ contains both u, p we are only interested in how many unique u there are for a given p. This can be done by taking a histogram of equally sized bins along p. Another way of getting a quantity that is proportional to count is by convolution $*$ of $p(s|\theta)$ along s with a symmetric kernel $\phi(p\,|\alpha)$ with a bandwidth α indicating bin width. This way we get a smooth and differentiable version of the histogram operation.

$$\phi(p\,|\alpha) * U(\theta) := \int_{\mathbb{R}} \phi(p - p(s|\theta)\,|\alpha)\,\mathrm{d}s \tag{4.1}$$

The result is now proportional to count, but needs to be mapped to the unit interval $[0,1]$ in such a way that high count regions that indicate multi-stability are mapped to one, and low counts which indicate monostability are mapped to zero. For this a sigmoidal activation function $\sigma(\frac{x-\mu}{\beta})$ with unknown smoothness β and threshold μ will do. We are now ready to write down the objective function

$$\mathcal{J}(\theta\,|\,\alpha,\beta,\mu) := \int_{\mathbb{R}} \left| \, t(p) - \sigma\left(\frac{\phi(p\,|\alpha) * U(\theta) - \mu}{\beta}\right) \right|^2 \mathrm{d}p \tag{4.2}$$

where θ are the parameters to be optimised and α, β, μ are hyper-parameters. Note that $U(\theta)$ is the only θ dependence in the expression. When we apply

$$\begin{aligned}
\partial_\theta \mathcal{J} &= -2 \int_{\mathbb{R}} \sigma'\, \partial_\theta \left(\frac{\phi(p\,|\alpha) * U(\theta) - \mu}{\beta}\right) \mathrm{d}p && (4.3)\\
&= -\frac{2}{\beta} \int_{\mathbb{R}} \sigma'\, \partial_\theta(\, \phi(p|\alpha) * U(\theta)\,)\,\mathrm{d}p && (4.4)\\
&= \frac{2}{\beta} \int_{\mathbb{R}} \sigma' \int_{\mathbb{R}} \phi'\, \partial_\theta p(s|\theta)\,\mathrm{d}s\,\mathrm{d}p && (4.5)
\end{aligned}$$

remarkably everything is differentiable including the algorithm in section 3. This gradient information can be used in a gradient descent optimization to minimize $\mathcal{J}(\theta)$.

5 Normal Forms

- saddle-node, pitchfork, transcritical
- optimisation landscape with changing hyperparams
- benchmarks against other algos

6 Chemical Reaction Networks

- toggle switch
- cell cycle
- application to structure $\rightarrow$ function (Luca)

7 Conclusions and Extensions

- hyperparam optimization
- hopf bifurcations
- pattern formation in pdes