

Dynamical Lasso

Facundo Sapienza^{*1}

¹Department of Statistics, University of California, Berkeley (USA)

August 26, 2023

^{*}Corresponding author: fsapienza@berkeley.edu

1 Introduction

We are interested in the problem of fitting a time series such that it remains *constant* over some unknown periods of time. Here the notion of constant is dictated by the dynamics of the system, which we will assume is governed by a series of continuous or discrete differential equation.

Consider observations

$$y_i \sim N(x_i, \sigma_i^2) \quad i = 1, 2, \dots, N \quad (1)$$

where $x_i = x(t_i; \theta) \in \mathbb{R}^n$ are some latent variables with observation times t_i following some **continuous** dynamics given by the differential equation

$$\frac{dx}{dt} = f(x; \theta(t), t), \quad x(t_0) = x_0 \quad (2)$$

where $\theta = \theta(t) \in \mathbb{R}^p$ is a continuous parameter of the system. Notice that we had emphasized the potential time dependency of the parameter θ as a function of time. Our goal is to find $\theta(t)$ such that the model approximates well enough the observations y_i at the same time we impose the constraint that $\theta(t)$ is a piecewise constant function. This dynamics results adequate for physical systems that evolve following some predicted dynamics until an abrupt episode changes the default configuration of the system, making the system to evolve following a new value of the system parameter $\theta(t)$.

A different description of the dynamics can be given with a set of **discrete** equations. This can arise by directly discretizing the continuous differential equation using a numerical solver, or just by describing the system as a discrete time series. In both cases, latent variables are generated by the equation

$$x_{i+1} = f_i(x_i, \theta_i), \quad (3)$$

where f_i is an evolution function. For many cases, this function will be linear and can be directly be described as

$$x_{i+1} = A_i(\theta_i) x_i + b_i(\theta_i) \quad (4)$$

with $A_i \in \mathbb{R}^{n \times n}$ and $b \in \mathbb{R}^n$.

A first approach to solve this problem is by to directly solve the continuous minimization problem

$$\min_{\theta(\cdot)} \sum_{i=1}^N \|y_i - \text{ODESolve}(t_i; \theta(\cdot))\|_2^2 + \lambda \int \left\| \frac{d\theta}{dt}(\tau) \right\|_2 d\tau \quad (5)$$

where $\text{ODESolve}(t; \theta(\cdot))$ is the numerical solution of the differential equation (2) evaluated at time t with time dependent parameter $\theta(\cdot)$. The second term penalizes changes in the function $\theta(\cdot)$, encouraging functions that stay constant over long period of times. Solving such optimization problem will require to be able to compute the gradient of the numerical solution with respect to the parameter θ . Furthermore, we need to parametrize the space of possible functions $\theta(\cdot)$.

A second approach consists in working directly with the discrete dynamical system and solve the problem¹

$$\min_{\theta_1, \theta_2, \dots, \theta_N} \sum_{i=1}^N \|y_i - x_i\|_2^2 + \lambda \sum_{i=1}^{N-1} \|\theta_{i+1} - \theta_i\|_2 \quad (6)$$

subject to the constraint

2 Examples

Let us consider some relevant simple examples.

2.1 Piecewise linear function

Consider the simple case where $f(x; \theta(t), t) = \theta(t)$. If $\theta(t) = \theta_0$ constant, then the solution $x(t)$ are simply linear functions. If our notion of smoothness is dictated by linear functions, that is, linear functions are consider smooth, we can then aim to find piecewise constant functions $\theta(t)$. Assuming I

2.2 Apparent polar wander paths

¹Notice that in this case we have the problem of what to do with the location of the changepoints, since changes in the value of the parameter θ don't necessary have to coincide with the t_i . We can instead consider a similar loss function but where the empirical error just evaluates in a subset of indices $I \subset \{1, 2, \dots, N\}$ where we actually have observations, leaving the $1, 2, \dots, N$ for the latent variables where the solution is evaluated.