SSY230 Learning dynamical systems using system identification

Hand In 2

Alfred Aronsson

May 14, 2024

1 Introduction

Second hand in task in SSY230 focused on maximum likelihood estimations.

2 Problem definition

Given N observations of $y(t)$ and $\phi(t)$, the model

$$y(t) = \theta^T \phi(t) + e(t).$$

and the assumption that the noise $e(t)$ is iid according to

$$f(x) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}, \quad x > 0.$$

Formulate the Maximum-Likelihood estimate of σ^2 and θ.

Note: An explicit solution is not likely since the PDF of the noise is not Gaussian. Hence, your answer can be stated as the equations which need to be solved numerically, given the data.

3 Solution

In the problem definition we are given the model,

$$y(t) = \theta^T \phi(t) + e(t) \tag{1}$$

with noise that is iid according to

$$f(x) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}, \quad x > 0 \tag{2}$$

where we want to find the Maximum-Likelihood estimate of σ^2 and θ.

1. **Likelihood Function**
   We can begin with defining the likelihood function $L(\theta, \sigma^2)$. The likelihood function for all N observations is the product of the likelihood of each observation. Given eq. (1) and (2) the likelihood for a single observation can be expressed as follows:

$$f(e(t)) = \frac{e(t)}{\sigma^2} e^{-\frac{e^2(t)}{2\sigma^2}}, \quad e(t) > 0 \tag{3}$$

The total likelihood is thereby:

$$L(\theta, \sigma^2) = \prod_{t=1}^{N} \left(\frac{y(t) - \theta^T \phi(t)}{\sigma^2} \right) e^{-\frac{(y(t)-\theta^T\phi(t))^2}{2\sigma^2}} \tag{4}$$

2. **Logarithmization**
   The next step is to transform the likelihood function $L(\theta, \sigma^2)$ to the log-likelihood function $\log L(\theta, \sigma^2)$. The log-likelihood simplifies future calculations greatly. Logarithmization of (4) looks like this:

$$\log L(\theta, \sigma^2) = \sum_{t=1}^{N} \left(\log \left(\frac{y(t) - \theta^T \phi(t)}{\sigma^2} \right) - \frac{(y(t) - \theta^T \phi(t))^2}{2\sigma^2} \right) \tag{5}$$

3. **Partial Derivatives**
   In order to find the MLE of θ and σ^2 we need to take partial derivatives of (5) and set them to zero.

   Beginning with θ we get:

$$\begin{aligned} \frac{\partial}{\partial \theta} \log L(\theta, \sigma^2) = \sum_{t=1}^{N} \Bigg[& \frac{-1}{y(t) - \theta^T \phi(t)} \cdot \frac{\partial}{\partial \theta}(y(t) - \theta^T \phi(t)) - \\ & \frac{1}{\sigma^2} \cdot 2(y(t) - \theta^T \phi(t)) \cdot \frac{\partial}{\partial \theta}(y(t) - \theta^T \phi(t)) \Bigg] \end{aligned} \tag{6}$$

   We also know that,

$$\frac{\partial}{\partial \theta}(y(t) - \theta^T \phi(t)) = -\phi(t) \tag{7}$$

   which simplifies (6) to:

$$\frac{\partial}{\partial \theta} \log L(\theta, \sigma^2) = \sum_{t=1}^{N} \left[\frac{\phi(t)}{(y(t) - \theta^T \phi(t))} - \frac{2(y(t) - \theta^T \phi(t))}{\sigma^2} \phi(t) \right] \tag{8}$$

   Continuing with σ^2 we get:

$$\frac{\partial}{\partial \sigma^2} \log L(\theta, \sigma^2) = \sum_{t=1}^{N} \left[\frac{-1}{\sigma^2} - \frac{-1}{2\sigma^4}(y(t) - \theta^T \phi(t))^2 \right] \tag{9}$$

4. **Numerical Solution**
   When setting (8) and (9) to zero in order to get the MLE for θ and σ^2 we quickly realise that the equations can not be solved analytically. In order to solve this task we would have to adapt a numerical approach such as the Newton method or gradient descent.