LINMA2171 - Introduction: General framework for fitting a function to data

P.-A. Absil

December 3, 2024

1. Data:
   - Data domain: $\mathcal{X} \in \mathbb{R}^d$.
   - Data function: $f : \mathcal{X} \to \mathbb{R}$ in a set of admissible data functions $\mathcal{F} \subseteq \mathbb{R}^{\mathcal{X}}$, where $\mathbb{R}^{\mathcal{X}}$ denotes the set of all functions from $\mathcal{X}$ to $\mathbb{R}$.
2. Design:
   - Model domain: $\widehat{\mathcal{X}}$ such that $\mathcal{X} \subseteq \widehat{\mathcal{X}} \subseteq \mathbb{R}^d$.
   - Set of admissible model functions: $\widehat{\mathcal{F}} \subseteq \mathbb{R}^{\widehat{\mathcal{X}}}$.
   - Loss function: $\mathcal{L} : \widehat{\mathcal{F}} \times \mathcal{F} \to \mathbb{R}$. It is customary to require that

$$\mathcal{L}(\widehat{f}, f) \begin{cases} = 0 & \text{if } \widehat{f}\big|_{\mathcal{X}} = f, \\ > 0 & \text{otherwise.} \end{cases} \tag{1}$$

   - Regularizer: $\mathcal{R} : \widehat{\mathcal{F}} \to \mathbb{R}$, usually such that

$$\mathcal{R}(\widehat{f}) \begin{cases} = 0 & \text{if } \widehat{f} \text{ is “regular”}, \\ > 0 & \text{otherwise.} \end{cases} \tag{2}$$

3. Optimization problem:

$$\underset{\widehat{f} \in \widehat{\mathcal{F}}}{\arg\min}\ \mathcal{L}(\widehat{f}, f) + \lambda \mathcal{R}(\widehat{f}), \tag{3}$$

   where $\lambda \geq 0$ is a design parameter that tunes the balance between the usually conflicting goals of data attachment and regularity.
4. Optimization algorithm.

Exercise 1

List possible purposes for fitting a function to data. Illustrate those purposes by means of graphical examples. Think about how you would do the design part of the above framework for each purpose.