GÖTEBORGS UNIVERSITET

Efficient training of interpretable, non-linear regression models

OSKAR ALLERBO

Institutionen för matematiska vetenskaper
Naturvetenskapliga fakulteten

Akademisk avhandling för filosofie doktorsexamen i matematisk statistik, som med tillstånd från Naturvetenskapliga fakulteten kommer att offentligt försvaras **fredagen den 22 september 2023, kl. 13.15 i Pascal**, Institutionen för matematiska vetenskaper, Chalmers tvärgata 3, Göteborg.
Fakultetsopponent är professor Henrik Hult, KTH.

ISBN 978-91-8069-337-0 (TRYCKT)
ISBN 978-91-8069-338-7 (PDF)

Efficient training of interpretable, non-linear regression models

Oskar Allerbo

Division of Applied Mathematics and Statistics
Department of Mathematical Sciences
University of Gothenburg and Chalmers University of Technology

Abstract

Regression, the process of estimating functions from data, comes in many flavors. One of the most commonly used regression models is linear regression, which is computationally efficient and easy to interpret, but lacks in flexibility. Non-linear regression methods, such as kernel regression and artificial neural networks, tend to be much more flexible, but also harder to interpret and more difficult, and computationally heavy, to train.

In the five papers of this thesis, different techniques for constructing regression models that combine flexibility with interpretability and computational efficiency, are investigated. In Papers I and II, sparsely regularized neural networks are used to obtain flexible, yet interpretable, models for additive modeling (Paper I) and dimensionality reduction (Paper II). Sparse regression, in the form of the elastic net, is also covered in Paper III, where the focus is on increased computational efficiency by replacing explicit regularization with iterative optimization and early stopping. In Paper IV, inspired by Jacobian regularization, we propose a computationally efficient method for bandwidth selection for kernel regression with the Gaussian kernel. Kernel regression is also the topic of Paper V, where we revisit efficient regularization through early stopping, by solving kernel regression iteratively. Using an iterative algorithm for kernel regression also enables changing the kernel during training, which we use to obtain a more flexible method, resembling the behavior of neural networks.

In all five papers, the results are obtained by carefully selecting either the regularization strength or the bandwidth. Thus, in summary, this work contributes with new statistical methods for combining flexibility with interpretability and computational efficiency based on intelligent hyperparameter selection.

Keywords: sparse regression, kernel regression, neural network regression, early stopping, bandwidth selection