Insert here your thesis' task.

FACULTY
OF INFORMATION
TECHNOLOGY
CTU IN PRAGUE

Bachelor's thesis

Probabilistic algorithms for computing the LTS estimate

Martin Jenč

Department of Applied Mathematics
Supervisor: Ing. Karel Klouda, Ph.D.

March 4, 2019

Acknowledgements

THANKS to everybody

Declaration

I hereby declare that the presented thesis is my own work and that I have cited all sources of information in accordance with the Guideline for adhering to ethical principles when elaborating an academic final thesis.

I acknowledge that my thesis is subject to the rights and obligations stipulated by the Act No. 121/2000 Coll., the Copyright Act, as amended, in particular that the Czech Technical University in Prague has the right to conclude a license agreement on the utilization of this thesis as school work under the provisions of Article 60(1) of the Act.

In Prague on March 4, 2019 .....

Czech Technical University in Prague
Faculty of Information Technology

Citation of this thesis

Jenč, Martin. *Probabilistic algorithms for computing the LTS estimate*. Bachelor's thesis. Czech Technical University in Prague, Faculty of Information Technology, 2019.

Abstrakt

V několika větách shrňte obsah a přínos této práce v českém jazyce.

Klíčová slova LTS odhad, lineÃąrnÃŋ regrese, optimalizace, nejmenÅąÃŋ usekanÃĺ ÄĎtvrece, metoda nejmenÅąÃŋch ÄĎtvercÅŕ, outliers

Abstract

The least trimmed squares (LTS) method is a robust version of the classical method of least squares used to find an estimate of coefficients in the linear regression model. Computing the LTS estimate is known to be NP-hard, and hence suboptimal probabilistic algorithms are used in practice.

Keywords LTS, linear regressin, robust estimator, least trimmed squares, ordinary least squares, outliers, outliers detection

Contents

List of Figures

Introduction

Chapter 1

The Least trimmed squares

1.0.1 Objective function

1.0.1.1 something

1.1 Robust statistics

1.2 Description

1.3 Computation

Chapter 2

Algorithms

2.1 FAST-LTS

In this section we'll describe FAST-LTS algorithm and it's main properties. The main idea of this algorthm is based on the fact that from one approximation of the algorithm we can compute another which can have lower objective function.

Thoerem 1: [1] Let w0 ... wp be the LTS estimate. for each data sample we can compute —y-wx—

Hlavni myslenka tohoto algoritmu spociva ve faktu,

Vyberte si Åąablonu podle druhu prÃące (bakalÃąÅŹskÃą, diplomovÃą), jazyka (ÄŀeÅątina, angliÄŀtina) a kÃşdovÃąnÃŋ (ASCII, UTF-8, ISO-8859-2 neboli latin2 a nebo Windows-1250).

V ÄŀeskÃľ variantÄŻ naleznete Åąablony v souborech pojmenovanÃ · ch ve formÃątu prÃące_kÃşdovÃąnÃŋ.tex. Typ prÃące mÅŕÅ"e bÃ · t:

BP bakalÃąÅŹskÃą prÃące,

DP diplomovÃą (magisterskÃą) prÃące.

KÃşdovÃąnÃŋ zdrojovÃľho souboru (LaTeX), ve kterÃľm chcete psÃąt, mÅŕÅ"e bÃ · t:

UTF-8 kÃşdovÃąnÃŋ Unicode,

ISO-8859-2 latin2,

Windows-1250 znakovÃą sada 1250 Windows.

V pÅŹÃŋpadÄŻ nejistoty ohlednÄŻ kÃşdovÃąnÃŋ doporuÄŀujeme nÃąsledujÃŋcÃŋ postup:

1. OtevÅŹete Åąablony pro kÃşdovÃąnÃŋ UTF-8 v editoru prostÃľho textu, kterÃ · chcete pro psanÃŋ prÃące pouÅ"Ãŋt – pokud mÅŕÅ"ete texty s diakritikou normÃąlnÄŻ pÅŹeÄŀÃŋst, pouÅ"ijte tuto Åąablonu.

2. V opaÄĎnÃľm pÅŹÃŋpadÄŻ postupujte dÃąle podle toho, jakÃ· operaÄĎnÃŋ systÃľm pouÅ"ÃŋvÃąte:
   - v pÅŹÃŋpadÄŻ Windows pouÅ"ijte Åąablonu pro kÃşdovÃąnÃŋ Windows-1250,
   - jinak zkuste pouÅ"Ãŋt Åąablonu pro kÃşdovÃąnÃŋ ISO-8859-2.

V anglickÃľ variantÄŻ jsou Åąablony pojmenovanÃľ podle typu prÃące, moÅ"nosti jsou:

bachelors bakalÃąÅŹskÃą prÃące,

masters diplomovÃą (magisterskÃą) prÃące.

2.2 Exact algorithm

2.3 Feasible solution

2.4 MMEA

2.5 Branch and bound

2.6 Adding row

Chapter 3

Experiments

3.1 Data

3.2 Performance

3.3 Outlier detection

Conclusion

Bibliography

[1] Rybicka, J. *LaTeX pro začátečníky*. Brno: Konvoj, third edition, ISBN 80-7302-049-1.

Appendix A

Datasets

GUI Graphical user interface

XML Extensible markup language

Appendix B

Contents of enclosed CD

- `readme.txt` ..... the file with CD contents description
- `exe` ..... the directory with executables
- `src` ..... the directory of source codes
  - `wbdcm` ..... implementation sources
  - `thesis` ..... the directory of LaTeX source codes of the thesis
- `text` ..... the thesis text directory
  - `thesis.pdf` ..... the thesis text in PDF format
  - `thesis.ps` ..... the thesis text in PS format