My Topic XYZ
Seminar: Foundations of Data Science

Jane Public, enrolment no.: 12345
Faculty of Statistics / Computer Science
TU Dortmund

October 22, 2020
winter term 2020/21

Abstract

A short summary of about two to three sentences that briefly and concisely outline the content ...

1 Introduction

This article deals with the fundamentals of data science. We are covering the topic XYZ in particular, which is very interesting and besides the theoretical depth has many practical applications.

We start with a definition, which plays a central role in this work.

Definition 1. *(Euclidean norm) Let $x \in \mathbb{R}^d$. We denote by*

$$\|x\| = \sqrt{\sum_{i=1}^{d} x_i^2}$$

the Euclidean norm of x.

2 Main Part

Here we work with the above definitions and notations and derive important results such as the following Theorem on the least-squares solution

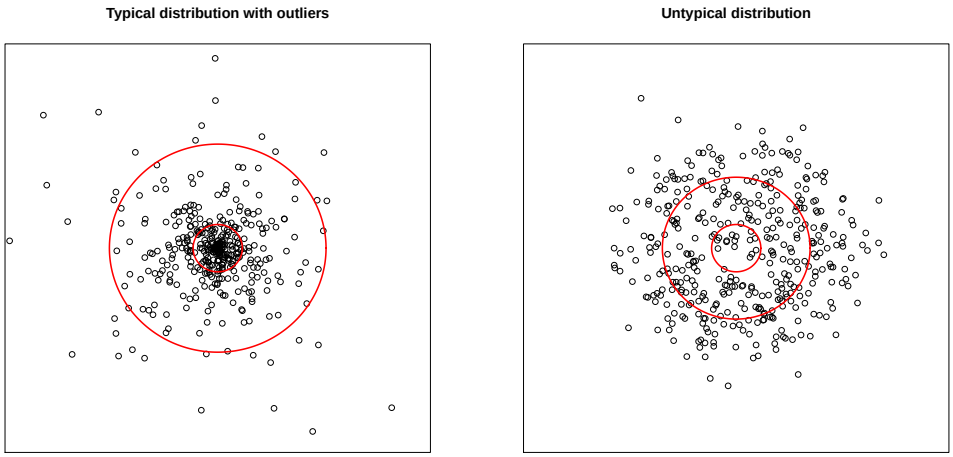

Figure 1: Distributions of points centered around their median. The caption should be placed below the figure.

Theorem 1. *(useful theorem) Let* $X \in \mathbb{R}^{n \times d}, Y \in \mathbb{R}^n$. *Further define* $\beta^* = \operatorname{argmin}_{\beta \in R^d} \|X\beta - Y\|^2$. *Then*

$$\|Y\|^2 = \|X\beta^*\|^2 + \|X\beta^* - Y\|^2.$$

Proof. The proof is left as an exercise. □

Sometimes figures help to illustrate a formalism. Figure 1 shows an example from the analysis of sampling algorithms for approximating the geometric median, which inappropriately is completely unrelated to Theorem 1.

3 Conclusion

Even after centuries of research in the field of data science, there is nothing more versatile than the useful theorem of chapter 2. It is used everywhere and has led to the greatest and most intriguing results, cf. [2]. By the way, the book for the seminar [1] is a great reference and should be cited. Further literature can be found in the respective *Bibliographic Remarks* sections and of course you are welcome to search and add your own references.

Note: BibTEX entries entries can often be found in the DBLP collection. Google Scholar also offers BibTEX entries, which can be copied into the .bib file and may need some minor adjustments.

References

[1] S. Shalev-Shwartz and S. Ben-David. *Understanding Machine Learning - From Theory to Algorithms*. Cambridge University Press, 2014.

[2] J. Someone and J. Someoneelse. Useful theorems, 2003.