Using Coresets And Sketches To Maintain Accuracy And Improve Run-Time of Least-Mean-Squares Solvers

Joseph Gonzalez

6/4/2021

Abstract

In the early 1900s, Constantin Carathéodory and Ernst Steinitz developed and proved Carathéodory's theorem. This theorem states that we can write each point contained in a convex hull of n points in R^d as a convex combination of at most d+1 points. These details suggest that we can maintain a $d^2 + 1$ scaled set of points that we can use to calculate the covariance matrix for a design matrix X and, with some other calculations, solve the LMS equation. While this method avoids updating the linear combinations of all n points and reduces numerical errors, it is not popular for its run time of $O(n^2 d^2)$ or $O(nd^3)$.

In this paper, we implement a novel Least-Mean-Squares Solvers from scratch that combines sketches and coresets. This method modifies Carathéodory's theorem to improve the runtime and maintain accuracy. For the application, we show that this solver can boost the performance of existing solvers such as those in the scikit-learn library.

Introduction

Methodologies

Results

Conclusion

Appendix

Appendix :

Appendix :