

Modellpunktsanalys

Maria Käki*

September 2013

Sammanfattning

I denna uppsats ska vi behandla ämnet Icke negativ Minstakvadrat Optimering, Nonnegative Least Squares (NNLS). Vi kommer att undersöka om det går att reducera antalet modellpunkter i två dataset. Syftet med att minska antalet modellpunkter är att de beräkningar som genomförs med dataseten går snabbare att genomföra med färre modellpunkter. Viktigt är också att inte förlora signifikant noggrannhet i resultaten då vi minskar antalet modellpunkter. I detta arbete kommer NNLS beskrivas kortfattat, resultat kommer att redovisas från NNLS optimeringen och diskuteras. NNLS fungerar genom att med hjälp av en algoritm hitta de värden på en vektor x som minimerar minstakvadratfunktionen och med hjälp av denna vektor skapa ett nytt viktat dataset med färre modellpunkter.

*Postadress: Matematisk statistik, Stockholms universitet, 106 91, Sverige.
E-post: maria.kakil@gmail.com. Handledare: Joanna Tyrcha.

Abstract

In this thesis we are going to use Nonnegative Least Squares Optimization to investigate if it is possible to reduce the number of model points in two datasets to be able to save time. We will start with a massive dataset from a insurance company, we are going to optimize this dataset and reduce the number of model point without losing significant accuracy. The aim of this thesis is to do this with the Nonnegative Least Squares (NNLS) method. We are going to describe NNLS briefly, results and conclusions from the NNLS optimization are going to be shown and discussed. NNLS works in the following way, given a dataset and a set of observed values we are going to find a nonnegative weight vector by an algorithm. Namely we are going to find the values of the vector \boldsymbol{x} which minimizes the least squares function and use these weights to create a new dataset with less model points.