\textbf{L}_1 \textbf{Optimization under Linear Inequality Constraints}

Thomas J. Smith

University of Illinois at Urbana-Champaign

\textbf{Abstract:} An algorithmic approach is developed for the problem of \textit{L}_1 \text{ optimization}
under linear inequality constraints based upon iteratively reweighted iterative projection
(or IRIP). IRIP is compared to a linear programming (LP) strategy for \textit{L}_1 \text{ minimization}
(Späh 1987, Chapter 5.3) using the ultrametric condition as an exemplar class
of constraints to be fitted. Coded for general constraints, the LP approach proves to be faster.
Both methods, however, suffer from a serious limitation in being unable to
process reasonably-sized data sets because of storage requirements for the constraints.
When the simplicity of vector projections is used to allow IRIP to be coded for specific
(in this case, ultrametric) constraints, we obtain a fast and efficient algorithm capable of
handling large data sets. It is also possible to extend IRIP to operate as a heuristic search
strategy that simultaneously identifies both a reasonable set of constraints to impose
\textit{and} the optimally-estimated parameters satisfying these constraints. A few noteworthy
characteristics of \textit{L}_1 optimal ultrametrics are discussed, including other strategies for
reformulating the ultrametric optimization problem.

\textbf{Keywords:} \textit{L}_1-norm; Optimization; Ultrametric; Structural representation.

---

The author is most grateful to Lawrence Hubert for his assistance with this study, and to two anonymous
reviewers for their insightful comments. The suggestions of Phipps Arabie, Editor, were greatly appreciated
and led to a much improved paper.

Author's Address: Thomas J. Smith, Department of Educational Technology, Research and Assessment,
Northern Illinois University, DeKalb, IL 60115, U.S.A.; e-mail: tjsmith@niu.edu.
1. Introduction

The problem of minimizing an objective function subject to a collection of linear inequality constraints occurs with some regularity in the literature (e.g., see De Soete 1983; Hartigan 1967; Hubert, Arabie, and Meulman 1997; Irltl 1980; Lawson and Hanson 1974, Chapter 23). Optimization strategies may be used, for example, when a structural account of a data set is desired—such as the fitting of a circular structure, an additive tree, or an ultrametric tree—although they may also be used for what might be considered nonstructural aims as in the case of ordinary linear regression. The latter class of problems usually involve fairly simple and small sets of constraints, but with structural representations, relatively large and complex sets of constraints are typically involved (for example, the set of constraints specifying the fixed topology of a particular tree structure). The objective function chosen in such instances is traditionally least-squares (the $L_2$ norm) and defined by the sum of squared deviations of the fitted values from the input values, although an $L_1$ criterion consisting of the summed absolute values of such deviations is also natural to consider. In addition to providing an alternative to other clustering techniques (say, for validation purposes), the use of the $L_1$ criterion might be expected, as in the regression domain, to reduce the influence of outlying data values on the resultant structure.

This paper proposes a general approach to $L_1$ minimization under linear inequality constraints that combines an iterative projection least-squares minimization strategy (Dykstra 1983; Wollan and Dykstra 1987) with iteratively reweighted least-squares. The resulting method—iteratively reweighted iterative projection (IRIP)—is slower than the linear programming (LP) approach outlined by Späth (1987), but this issue becomes irrelevant when the resource limitations of either algorithm are considered; given the manner in which these problems are usually formulated and solved (by having the constraints stored in memory in matrix form), they are not capable of handling anything beyond small-to-moderate object set sizes. When the simplicity of vector projection is exploited, however, IRIP may be coded for specific (in this case, ultrametric) sets of constraints, obviating the necessity of explicit constraint storage. The resulting algorithm is very fast, efficient, and obtains optimal $L_1$ solutions for even large object set sizes. Moreover, phrasing IRIP according to specific constraints also allows a heuristic approach to identifying $L_1$ optimal structural representations, where a solution may be approximated by simultaneously identifying both a hopefully optimal set of constraints (rather than relying upon fixed constraints) and the optimally-estimated model parameters subject to these identified con-