Notes

1 Introduction

2 Constraints

2.1 Bounded Constraints

2.2 Simplex Constraints

2.3 Definite Matrix constraints

3 Softmax Transform

The softmax function can be understood from Multinomial Logistic Regression employed for predicting probabilities of a Categorically distributed variable. Geometrically it maps R^K to the boundary of a unit K-simplex(it is simply the convex hull of $k + 1$ affinely independent points in R^K. Essentially it transforms a vector of size K to another vector of size K where the outputs sum to 1. It is worth noting that the mapping is actually from R^K to R^{K-1}, so when a vector of size K is transformed the $K_t h$ vector is simply 1- sumofk-1 vectors R