Streaming Symmetric Norms via Measure Concentration

Abstract:

A long line of research studies the space complexity of estimating a norm \( l(x) \) in the data-stream model, i.e., when \( x \) is the frequency vector of an input stream consisting of insertions and deletions of items of \( n \) types. I will focus on norms \( l \) in \( \mathbb{R}^n \) that are *symmetric*, meaning that \( l \) is invariant under sign-flips and coordinate-permutations, and show that the streaming space complexity is essentially determined by the measure-concentration characteristics of \( l \). These characteristics are known to govern other phenomena in high-dimensional spaces, such as the critical dimension in Dvoretzky's Theorem.

The family of symmetric norms contains several well-studied norms, such as all \( l_p \) norms, and indeed we provide a new explanation for the disparity in space complexity between \( p \leq 2 \) and \( p > 2 \). We also obtain bounds for other norms that are useful in applications.