\textbf{Abstract}

We consider the problem of finding an approximate solution to \( \ell_1 \) regression while only observing a small number of labels. Given an \( n \times d \) unlabeled data matrix \( X \), we must choose a small set of \( m \ll n \) rows to observe the labels of, then output an estimate \( \hat{\beta} \) whose error on the original problem is within a \( 1 + \epsilon \) factor of optimal. We show that sampling from \( X \) according to its Lewis weights and outputting the empirical minimizer succeeds with probability \( 1 + \epsilon \).

There is a full training set \( \text{Active LAD Regression} \).

Subspace Embeddings and Regression

- Can use \textit{Lewis weight subsampling} to get \textit{subspace embedding}:
  \[ (1 - \epsilon) \| X \beta \|_1 \leq \| S X \beta \|_1 \leq (1 + \epsilon) \| X \beta \|_1 \]
  for all \( \beta \).
  
  - Need \( m = O \left( \frac{d \log d}{\epsilon \delta} \right) \) sampled rows [1]
  
  - Show that we can use this to sample well for regression as well! (without even seeing \( y \)).

\textbf{Proof Approach}

- Cannot show \( \| \beta^* - \hat{\beta} \|_1 \) is small
  - Because \( \ell_1 \) minimizer is not unique
  
  - Instead:
  \[ (\| S X \beta - S y \|_1 - \| S X \beta - S y^\star \|_1) - (\| X \beta' - y \|_1 - \| X \beta - y \|_1) \leq \epsilon \| X \beta' - X \beta \|_1 \]
  
  - So in figure 3, we would be showing that the two blue distances are close compared to \( \| X \beta' - X \beta \|_1 \).
  
  - The effects of the difficult \( y \) cancel in each term, and now we can show this using the subspace embedding property for Lewis weight sampling.

\textbf{Results}

- \textbf{Upper bound}: Need \( O \left( \frac{d \log d}{\epsilon \delta} \right) \) rows!
- \textbf{Lower bound}: No algorithm can do fewer than \( \Omega \left( \frac{d}{\epsilon^2} + \frac{\log d}{\epsilon} \right) \)

\textbf{References}