Efficient Distributed SDCA

Jeff Hajewski and Mingrui Liu

April 30, 2017

University of Iowa

Overview

Problem Overview

- Implement SDCA using a GPU
- Two versions:
  - Sequential - uses GPU for vector operations
  - Distributed - perform updates on multiple dimensions in parallel

Issues We Encountered

Not as easy as we expected!

- Memory allocation
  - Initial implementation allocated memory on the GPU each time it was needed
  - Improved implementation used static pointers to locations on the GPU's memory
  - In a test dataset, this reduced the number of memory allocations from 200,000 to 3 allocations
- Algorithm verification
  - Code looks correct, results don't make sense
  - Unit tests can verify individual components (e.g., GPU dot product, data loading and parsing, etc.)

Results

- Code efficiency needs to improve