MICHIGAN ENGINEERING
UNIVERSITY OF MICHIGAN

Memory Optimization in Quasi-Newton Methods

Hugh Van Deventer V, Itamar Pres

Introduction

- In class we learned Newton methods converge fast by using exact second-order information but are computationally expensive.
- Quasi-Newton methods approximate Hessians using curvature pairs, trading accuracy for speed.
- Memory (number of curvature pairs) can affect optimization performance.
- Goal: What is minimal memory size we can generally use while achieving strong L-BFGS performance.

Problems Considered

- Quadratic Problems (well and ill-conditioned)
- Quartic Functions (small and large quartic term)
- Rosenbrock Functions (2D and high-dimensional variants)
- Exponential Sums (10D and 100D)
- Data Fitting Problem (least squares structure)
- Genhumps Function (highly nonconvex)
- Full results in paper…

Algorithms Considered

- **Gradient Descent (GD)**
  - First-order method with backtracking and weak Wolfe line search variants.
- **Modified Newton's Method**
  - Second-order method using exact Hessian (regularized if necessary), with backtracking and weak Wolfe line search.
- **Trust Region Methods**
  - TRNewtonCG: Trust region with exact Hessian and Conjugate Gradient (CG) solver.
  - TRSR1CG: Trust region with SR1 Hessian approximation and CG solver.
- **Quasi-Newton Methods**
  - BFGS and DFP updates with both backtracking and Wolfe line searches.
  - L-BFGS (Limited-memory BFGS) with varying memory sizes.

Algorithms Performance

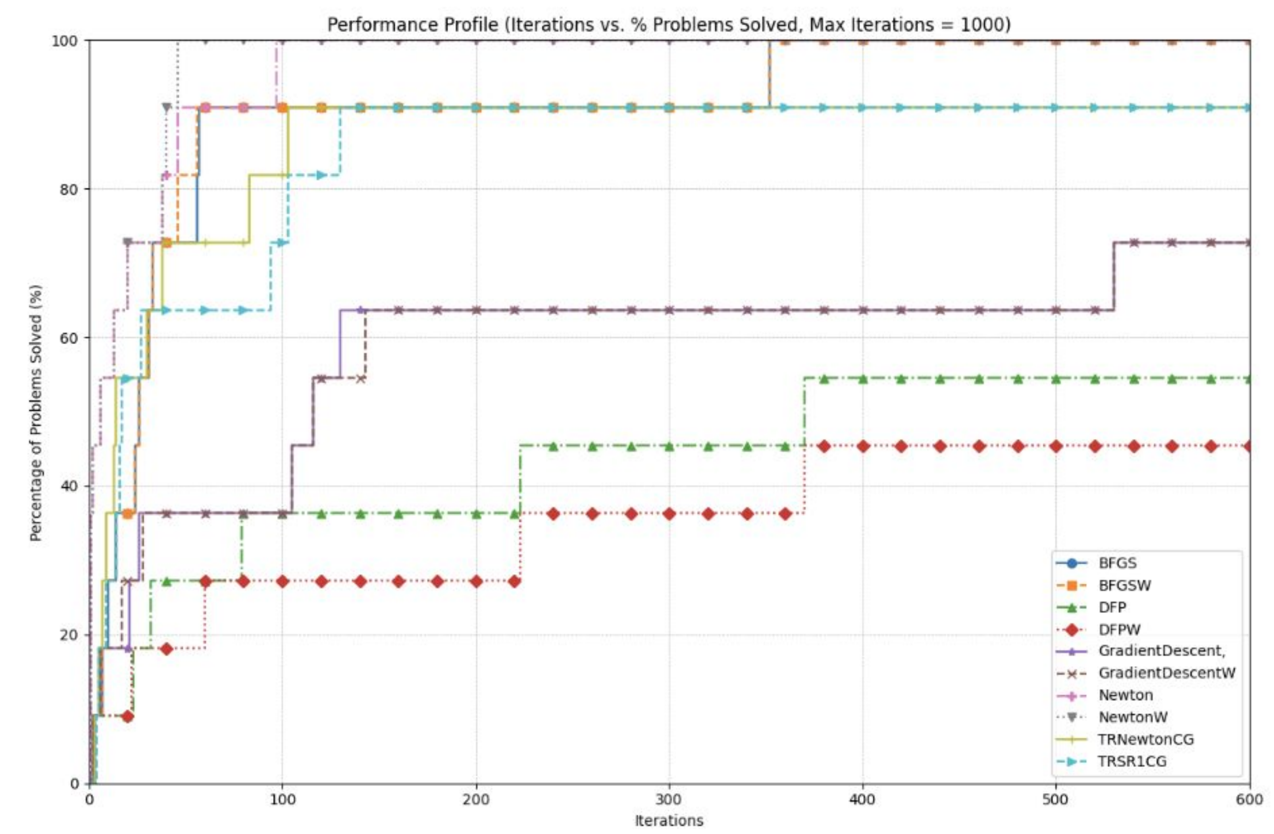

Figure 1: Performance Profile

L-BFGS Mem Exploration (Quadratic)

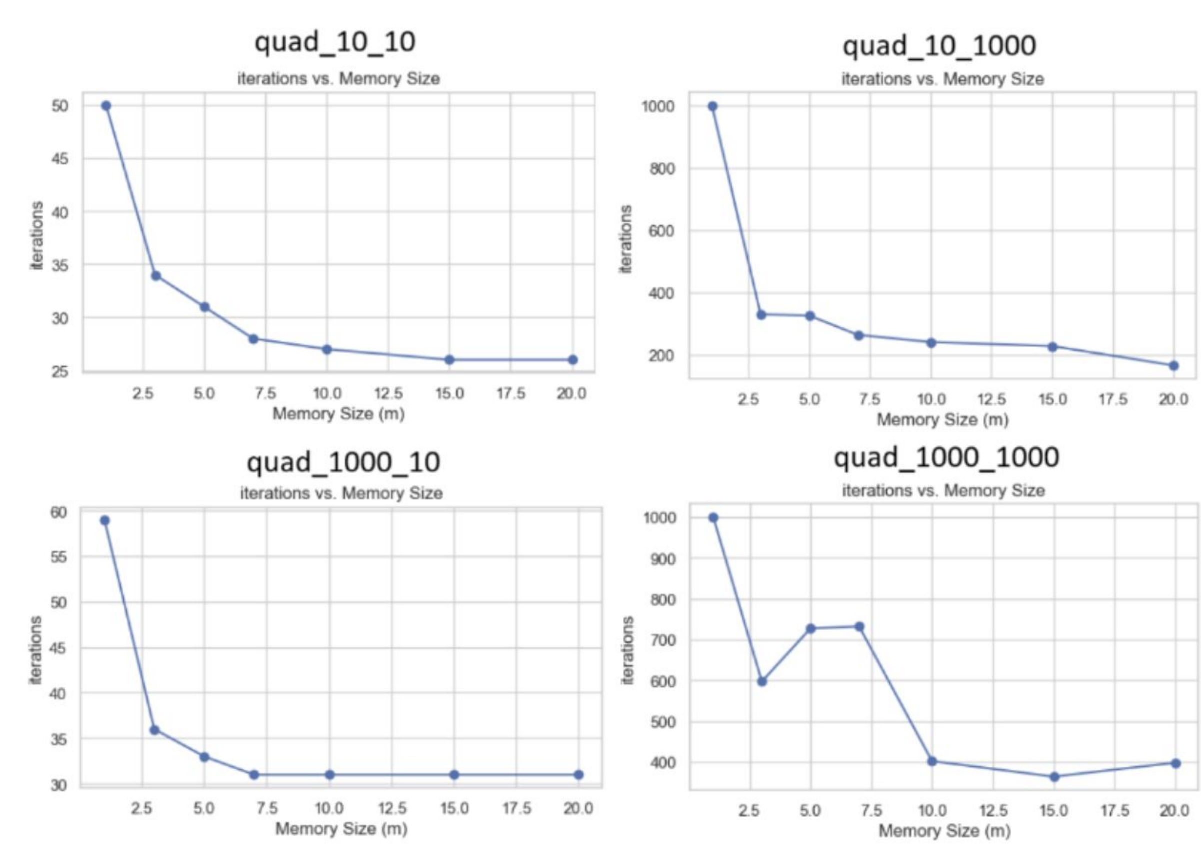

Figure 3: L-BFGS with varying memory sizes on the Quadratic problems

L-BFGS Mem Exploration (Rosenbrock)

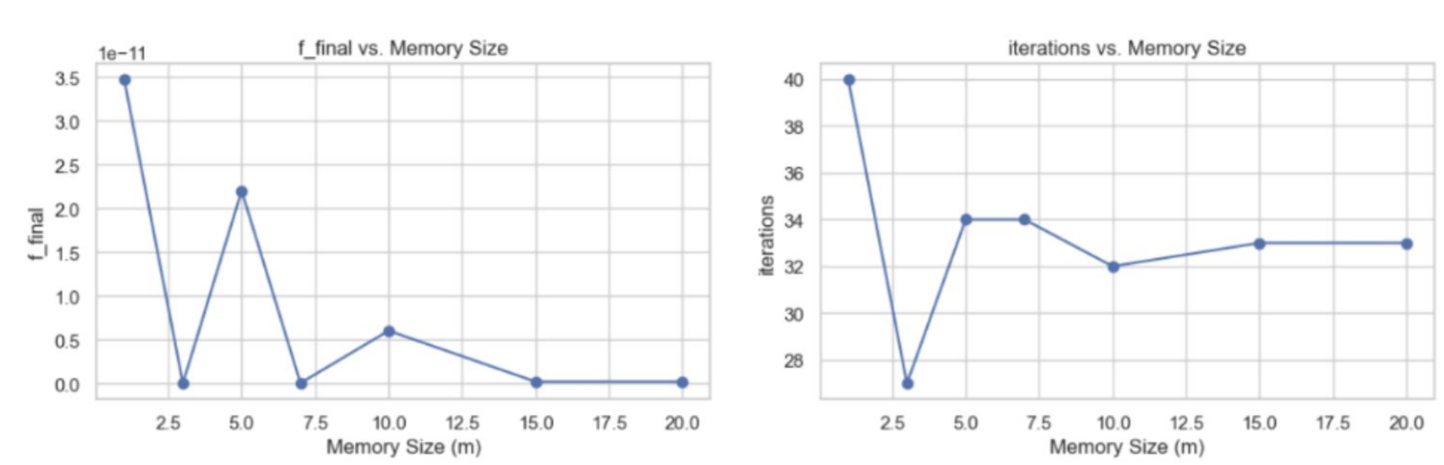

Figure 5: L-BFGS with varying memory sizes on the Rosenbrock problem (n=2)

L-BFGS Mem Exploration (Genhumps)

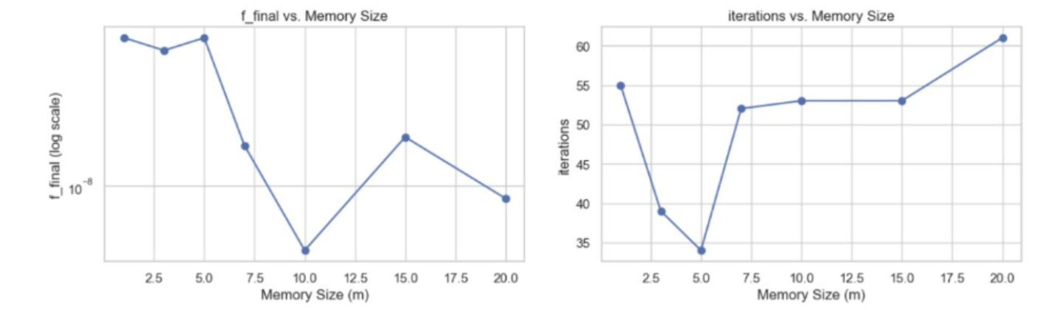

Figure 8: L-BFGS with varying memory sizes on the Genhumps_5 problem

L-BFGS Mem Exploration (Quartic)

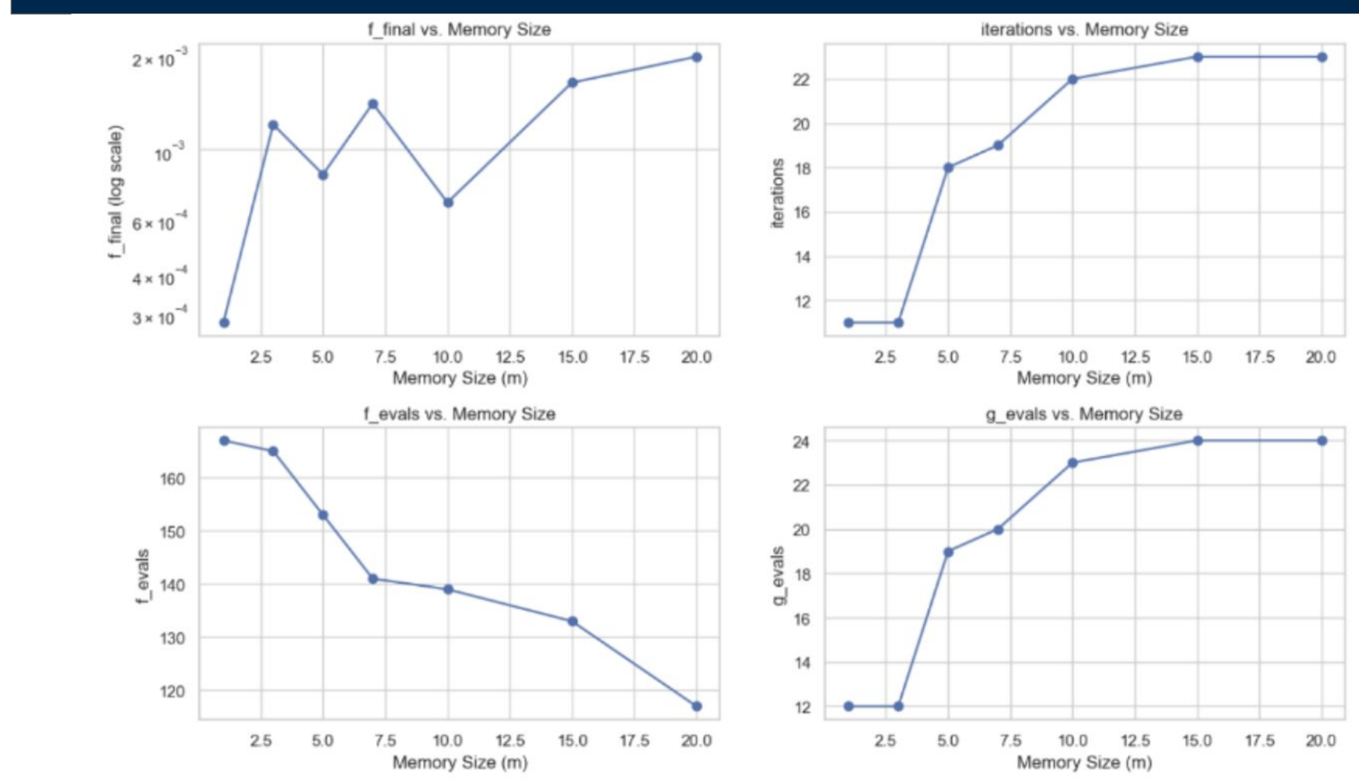

Figure 4: L-BFGS with varying memory sizes on the quartic_2 problem

Discussion

- **Quadratic Problems:**
  - Larger memory improves performance, especially on ill conditioned problems, but gains diminish.
- **Quartic_2 Problem:**
  - Higher memory sizes can hurt performance on strongly non-quadratic problems.
- **Rosenbrock 2D:**
  - Memory impacts convergence in a non-monotonic way; intermediate memory sizes can be better than extremes.
- **Genhumps_5 Problem:**
  - Solution quality and convergence speed are sensitive to memory size with moderate memory working well.
- **Overall Conclusion:**
  - No single memory size is universally best; moderate memory often balances speed and accuracy, but optimal settings depend heavily on problem structure.
- **See full-paper for rest of the problems!**