Performance Analysis for Constrained Least Square Optimization

Linda Cai
Department of Computer Science, College of Engineering, University of Illinois at Urbana-Champaign

Introduction

Solving systems of linear equations is a common problem in science and engineering. The linear model

$$Ax = b + e$$

often arises from discretization of problems such as integral equations and differential equations, where e is the perturbation error.

A lower and upper bound for the solution x can often be obtained from the physical meaning of the solution.

When the underlying problem is ill posed, the discretization is ill conditioned. As a result, traditional least square solvers may produce result that has small Ax − b but not physically meaningful. To provide physically meaningful solutions, it is often useful to tighten the bound on x given an upper bound on the perturbation error e.

$$\|e^T S^{-2} e\| \leq \mu$$

The suboptimal method have been developed to provide a tighter bound containing x. In addition, general optimization methods such as interior point method and SQP can be used to obtain upper and lower bounds. For large-scale linear systems, it is important that the bounds is provided in a computationally efficient manner.

In this study I compare 3 methods: the suboptimal method, the interior point method and the SQP method for solving the following problem:

$$\min w^T x$$
$$(Ax-b)^T S^{-2}(Ax-b) \leq \mu^2$$
$$p \leq x \leq q$$
$$\text{where } w \in \{-e_i, e_i | i \in [1, n]\}$$

The suboptimal method tightens the bounds for each dimension holistically, while interior point method and SQP solves the sequence of optimization problems sequentially. It would be interesting to see if suboptimal method has significantly better running time on large scale problems.

Method Summaries

- **Interior Point Method**

The interior point method is a widely used method in convex optimization. It converges to the global optimum of a convex objective super-linearly. We can use interior point method to solve the sequence of nonlinear programming problems one by one.

The interior point method is derived from the Karush-Kuhn-Tucker(KKT) conditions for nonlinear programming problems.

For a optimization problem

$$\min f(x)$$
$$c(x) = 0$$
$$x \geq 0$$

The KKT conditions are

$$\nabla \mathcal{L}(x,\lambda,\mu) = \nabla f(x) + \lambda \nabla c(x) - \mu = 0$$
$$c(x) = 0$$
$$x_i \mu_i = 0 \ \forall i$$
$$x \geq 0$$
$$\mu \geq 0$$

$$\text{Denote } F(x,y,s) = \begin{bmatrix} \nabla f(x) + \lambda \nabla c(x) - \mu \\ c(x) \\ XUe \end{bmatrix} = 0$$

where $X = diag_{1 \leq i \leq n}(x_i)$ and $U = diag_{i \leq i \leq n}(\mu_i)$

The interior point method iteratively finds root to F(x,y,s) by computing the Jacobian matrix and taking a newton step at each iteration

- **Sequential Quadratic Programming**

The sequential quadratic programming method approximates a non-linear optimization problem with a quadratic programming problem (with quadratic objective and linear constraints) at each iteration. It is well known as an efficient method for nonlinear programming problems. Similar to interior point method, we can use SQP to solve our sequence of nonlinear programming problems one-by-one.

SQP method uses the fact that the gradient of Lagrangian function is 0 when the objective function is at a minimum or maximum.

$$\mathcal{L}(x,\lambda) = f(x) + \lambda^T c(x)$$

$$\nabla \mathcal{L}(x,\lambda) = \begin{bmatrix} \nabla f(x) + \mathcal{J}_c(x)\lambda \\ c(x) \end{bmatrix} = 0$$

$$H\mathcal{L}(x^k,\lambda^k) \cdot \begin{bmatrix} \delta x \\ \delta \lambda \end{bmatrix} = -\nabla \mathcal{L}(x,\lambda)$$

At each iteration, SQP finds the solution $\begin{bmatrix} \delta x \\ \delta \lambda \end{bmatrix}$ to a quadratic programming problem whose KKT condition correspond to the above equation.

- **Suboptimal Method**

he suboptimal method is a method that iteratively tightens the error bound by relaxing the constraints. The suboptimal method is different from the previous two methods in that in each iteration it tightens the bound on all dimensions of x simultaneously and thus more suited for large-scaled problems

The suboptimal method formulates the bound constraints on x as quadratic constraint

$$(x-d)^T Q^{-2}(x-d) \leq n$$
$$\text{where } Q = diag(\tfrac{p-q}{2})$$
$$d = diag(\tfrac{p+q}{2})$$

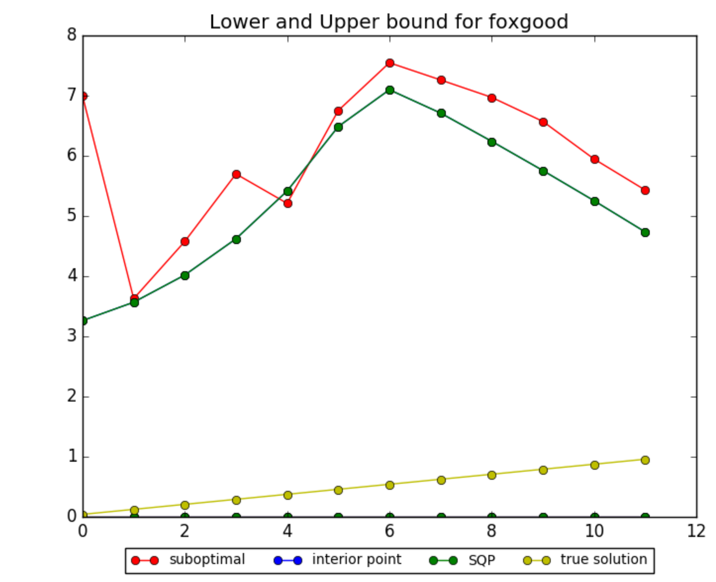
The suboptimal method then circumscribes the interior of the 2 quadratic constraints with a new ellipsoid

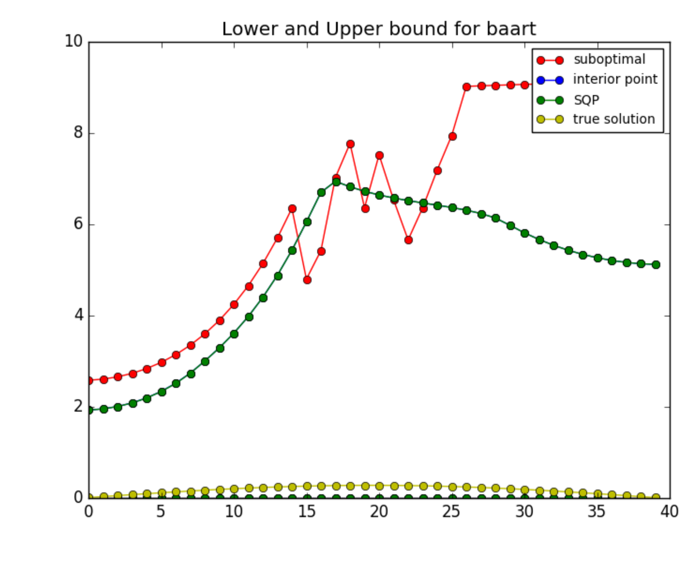
$$(Ax-b)^T S^{-2}(Ax-b) \leq \mu^2$$
$$(x-d)^T Q^{-2}(x-d) \leq n$$
$$\frac{\tau}{\mu^2}(Ax-b)^T S^{-2}(Ax-b) + (1-\tau)\frac{1}{n}(x-d)^T Q^{-2}(x-d) \leq 1$$

It is well known how to minimize or maximize a linear objective on a single ellipsoid, and the suboptimal method iteratively tightens the bound on x by computing the bounding box on the new ellipsoid and update bounds on x.

Experiments

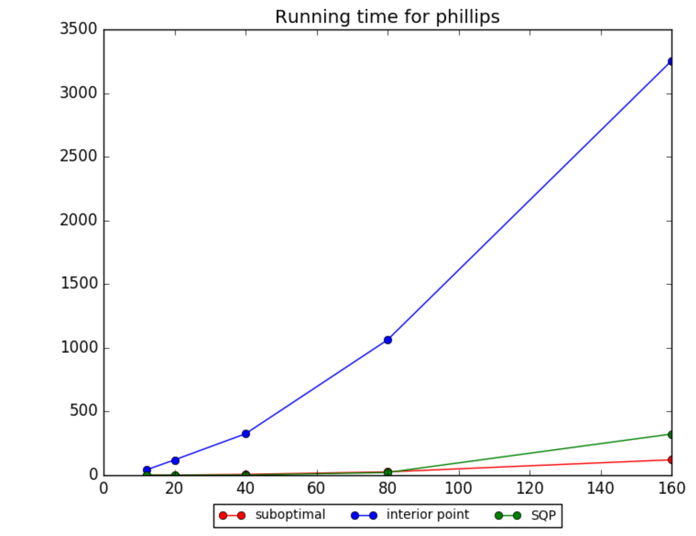
I compared the accuracy of the 3 methods on multiple ill conditioned test cases from the Regularization tools package by PC Hansen.

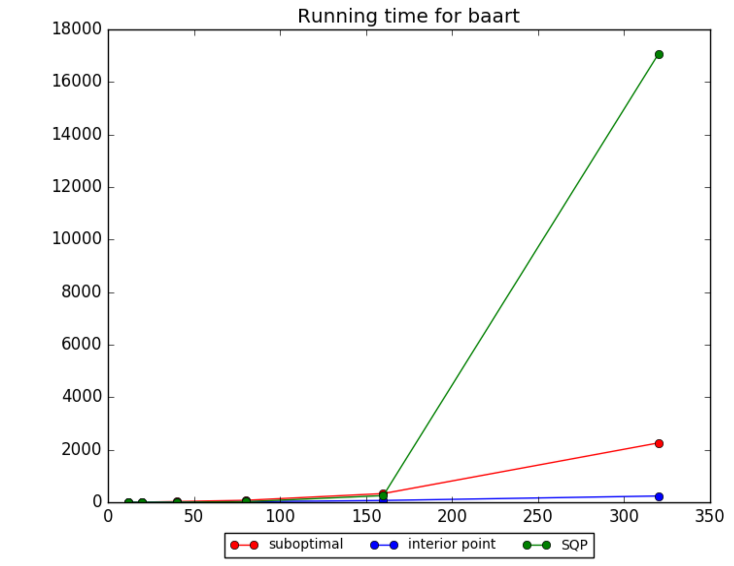




(note the blue line doesn't seem to exist in the plot because it coincides exactly with the green line)

As expected the suboptimal method gives a looser bound than interior point method and SQP. Both of the later methods make no relaxation on the problem and in a sense brute forces the solution on each dimension.





The running time of the methods highly depend on the problem. Both the interior point method and the SQP method have cases where they work extremely well and cases where they are the slowest method. The suboptimal method tend to have stable running time regardless of the problem.