4M17 Coursework Assignment 1

Practical Optimisation

1. The simplest norm approximation problem is an unconstrained problem of the form

$$\text{minimise} \quad \|Ax - b\|, \tag{1}$$

where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ are given as problem data, $x \in \mathbb{R}^n$ is the variable, and $\|\cdot\|$ is a norm on $\mathbb{R}^m$.

(a) Define the l_1, l_2, and l_∞-norms on $\mathbb{R}^m$ and write down the norm approximation problems corresponding to each of these norms. In the case of the l_2-norm, show that the problem can be expressed as an optimisation problem with a convex quadratic function with an analytic solution, which amounts to solving a linear system of equations.

(b) Show that the norm approximation problems corresponding to the l_1 and l_∞-norms on $\mathbb{R}^m$ can be cast as linear programming (LP) problems of the form

$$\begin{aligned} \min_{\tilde{x}} \quad & \tilde{c}^T \tilde{x} \\ & \tilde{A}\tilde{x} \le \tilde{b}. \end{aligned} \tag{2}$$

Specify the dimensions and provide expressions for $\tilde{A}$, $\tilde{b}$, $\tilde{c}$ in terms of A and b for each of the two norms.

(c) In the coursework data folder, 5 pairs of problem data `(A1,b1)`, ..., `(A5,b5)` are provided for $m = 2n$ and $n = 16, 64, 256, 512, 1024$. Use a linear programming solver to solve the LP problems corresponding to l_1 and l_∞-norms, and a linear solver to minimise the l_2-norm. Produce a table showing the values of the minimised l_1, l_2, and l_∞ norms $\|Ax - b\|$ corresponding to each of the 5 pairs of data, as well as the running times of your algorithms in each case. You may use standard MATLAB functionalities or equivalent libraries in other programming languages to solve this problem.

(d) Plot the histogram of the residuals of the norm approximation problem (1) for the 5th pair of data `(A5,b5)` ($n = 1024$) for the l_1, l_2, and l_∞-norms. Comment on the shape of the distribution of the residuals in each case.

2. Consider the optimisation problem

$$\begin{aligned} \min_{x} \quad & f_0(x) \\ \text{subject to} \quad & f_i(x) \le 0, \quad i = 1, \dots, m \end{aligned} \tag{3}$$

where f_0, f_i are convex and twice continuously differentiable functions for $i = 1, \ldots, m$. The central path is the solution of the unconstrained minimisation

$$\min_x \; tf_0(x) + \phi(x) \tag{4}$$

for any $t \geq 0$, where ϕ is the logarithmic barrier function of the feasibility set $S = \{x : f_i(x) \leq 0\}$.

(a) Write down the explicit form of (4) corresponding to the LP formulation of the l_1-norm approximation problem of the first part of the coursework assignment. Calculate the expression for the gradient of the cost in (4) corresponding to a fixed $t \geq 0$.

(b) Apply a first order gradient method with backtracking linesearch to solve the problem in (a) for $t = 1$ in the case corresponding to the pair `(A3,b3)` in the data folder ($n = 256$).

(c) Perform a convergence analysis of the algorithm used in (b) by producing a semilogarithmic plot of the minimisation error against the number of iterations.

3. Consider the l_1-regularised least squares problem

$$\min_x \; \|Ax - b\|_2^2 + \lambda \|x\|_1 \tag{5}$$

where $\lambda > 0$ is a regularisation parameter.

(a) Show that (5) can be transformed to a convex quadratic problem with linear inequality constraints. Define a suitable logarithmic barrier function Φ that characterises the inequality constraints and show that the central path formulation of (5) takes the form

$$\phi_t(x, u) = t\|Ax - b\|_2^2 + t\lambda \sum_{i=1}^{n} u_i + \Phi(x, u) \tag{6}$$

where the parameter t varies from 0 to ∞.

(b) Derive expressions for the gradient and Hessian of ϕ_t.

(c) We consider a sparse signal recovery problem with a signal $x_0 \in \mathbb{R}^{256}$ which consists of 10 spikes of amplitude ± 1. The measurement matrix $A \in \mathbb{R}^{60 \times 256}$ and x_0 are given in the coursework data folder. The vector of observations is $b = Ax_0$. Perform a sparse signal reconstruction by applying a Newton interior-point method to (6) with regularisation parameter $\lambda = 0.01\lambda_{\max}$ with $\lambda_{\max} = \|2A^T b\|_\infty$, and plot the original and reconstructed signals.

(d) Compare the result to the minimum energy reconstruction, which is the point in the set $\{x \in \mathbb{R}^{256} : A^T Ax = A^T b\}$ that is closest to the origin in l_2 norm.

(e) (*Optional: not marked, but feedback may be provided*) The paper [1] discusses an efficient and scalable implementation of an interior point algorithm customised to (5). The Newton method is replaced by a truncated version (TNIPM) that does not require the calculation of the full Hessian (Section IV). The performance is illustrated in Section V.A. Compare the performance of your exact Newton algorithm on a larger size problem with that of TNIPM, which is available online.

Your Report

- The deadline for submitting your report on this assignment is 4pm on Friday the 11th of December 2020. You need to submit an electronic copy via Moodle, including a coursework cover sheet. **Undergraduates should only include their Coursework Candidate Numbers (CCN), and must ensure that their names do NOT appear in the report, or in any of the file names.** MPhil and postgraduate students should ensure that their names and CRSIDs are included on their reports.

- You are expected to include figures and associated explanations as part of your report. Plots without explanations will receive very few marks. Figures should be carefully designed and calibrated.

- Attach the source code of all programs you have used in an appendix to your report.

- Your report should not exceed 10 pages in length excluding source code.

- This assignment counts as 50% of your final grade and you are expected to spend at least 10-15 hours on it. Questions on this assignment should be directed to Dr Cyrus Mostajeran (csm54).

References

[1] Kim, Seung-Jean, Kwangmoo Koh, Michael Lustig, Stephen Boyd, and Dimitry Gorinevsky. "An interior-point method for large-scale ℓ_1-regularized least squares." IEEE journal of selected topics in signal processing 1, no. 4 (2007): 606-617.