Tutorial 11+Assignment 4

Question 1

Suppose f is strongly convex with $mI \preceq \nabla^2 f(x) \preceq MI$. Let Δx be a descent direction at x. Show that the backtracking stopping condition holds for

$$0 < t \le -\frac{\nabla f(x)^T \Delta x}{M \|\Delta x\|_2^2}$$

Use this to give an upper bound on the number of backtracking iterations.

Question 2

Let $f(x) = (1/2)(x_1^2 + \gamma x_2^2)$, with $\gamma > 0$. Suppose that the gradient descent method is used with exact line search, starting at $x^{(0)} = (\gamma, 1)$. Prove

$$x_1^{(k)} = \gamma \left(\frac{\gamma - 1}{\gamma + 1} \right)^k, \quad x_2^{(k)} = \left(-\frac{\gamma - 1}{\gamma + 1} \right)^k$$

How about Newton descent?

Question 3

Let δx_{nsd} and δx_{nsd} be the normalized and unnormalized steepest descent directions at x, for the norm $\| \cdot \|$ Prove the following identities.

1. $\nabla f(x)^T \Delta x_{nsd} = -\|\nabla f(x)\|_*$.

2. $\nabla f(x)^T \Delta x_{sd} = -\|\nabla f(x)\|_*^2$.

Question 4

Show that $f(x) = 1/x$ with domain $(0, 8/9)$ is self-concordant.

Question 5

Consider the unconstrained problem

$$\text{minimize} \quad f(x) = -\sum_{i=1}^{m} \log(1 - a_i^\top x) - \sum_{i=1}^{n} \log(1 - x_i^2),$$

with variable $x \in \mathbf{R}^n$, and dom $f = \{x | a_i^\top x < 1, i = 1, \ldots, m, |x_i| < 1, i = 1, \ldots, n\}$. This is the problem of computing the analytic center of the set of linear inequalities

$$a_i^\top x \le 1, i = 1, \ldots, m, \quad |x_i| \le 1, i = 1, \ldots, n.$$

Note that we can choose $x^{(0)} = 0$ as our initial point. You can generate instances of this problem by choosing a_i from some distribution on $\mathbf{R}^n$.

(a) Use the gradient method to solve the problem, using reasonable choices for the backtracking parameters, and a stopping criterion of the form $||\nabla f(x)||_2 \le \eta$. Plot the objective function and step length versus iteration number. (Once you have determined $p^\star$ to high accuracy, you can also plot $f - p^\star$ versus iteration.) Experiment with the backtracking parameters α and β to see their effect on the total number of iterations required. Carry these experiments out for several instances of the problem, of different sizes.

(b) Repeat using Newton's method, with stopping criterion based on the Newton decrement λ^2. Look for quadratic convergence. You do not have to use an efficient method to compute the Newton step; you can use a general-purpose dense solver, although it is better to use one that is based on a Cholesky factorization.

Hint. Use the chain rule to find expressions for $\nabla f(x)$ and $\nabla^2 f(x)$.

Question 6

Prove that the Newton Step and decrement for equality constrained problem are affine invariant.

Question 7

Implement the infeasible start Newton method for solving the centering problem arising in the standard form LP,

$$\text{minimize} \quad c^\top x - \sum_{i=1}^{n} \log x_i$$
$$\text{subject to} \quad Ax = b,$$

with variable x. The data are $A \in \mathbf{R}^{m \times n}$, with $m < n, c \in \mathbf{R}^n$, and $b \in \mathbf{R}^m$. You can assume that A is full rank. This problem cannot be solved when it is infeasible or unbounded below.

Your code should accept A, b, c, and x_0, and return $x^\star$, the primal optimal point, $v^\star$, a dual optimal point, and the number of Newton steps executed. The initial point $x^{(0)}$ must satisfy $x^{(0)} \succ 0$, but it need not satisfy the equality constraints.

Use the block elimination method to compute the Newton step. (You can also compute the Newton step via the KKT system, and compare the result to the Newton step computed via block elimination. The two steps should be close, but if any x_i is very small, you might get a warning about the condition number of the KKT matrix.)

Plot $||r(x, v)||_2$, the norm of the concatenated primal and dual residuals, versus iteration k for various problem data and initial points, to verify that your implementation achieves quadratic convergence. As stopping criterion, you can use $||r(x, v)||_2 \le 10^{-6}$ (which means the problem was solved) or some maximum number of iterations (say, 50) was reached, which means it was not solved (likely because the problem is either infeasible or unbounded below).

For a fixed problem instance, experiment with varying the algorithm parameters α and β, observing the effect on the total number of Newton steps required.

To generate problem data (*i.e.*, A, b, c, x_0) that are feasible, you can first generate A, then random positive vector p, and set $b = Ap$. You can be sure that the problem is not unbounded by making one row of A have positive entries. You may also want to check that A is full rank.

Test the behavior of your implementation on data instances that are not feasible, and also ones that are unbounded below.