Numerical Optimization - HW3

Due is May 6 (Tu) but it is collected at 13:00 May 8 (Thu). No late HW will be accepted. For computer coding problems, you have to submit a **hard copy** of your codes and the computational result. Also you have to send them to **csenahw@gmail.com** using zip compression format.

---

1. Consider the quadratic norm $\|z\|_P = (z^T P z)^{1/2} = \|P^{1/2}\|_2$ with $P \in S^n_{++}$.

   (a) Show that $\|z\|_* = \|z\|_{P^{-1}} = \|P^{-1/2}\|_2$.

   (b) Show that the normalized steepest descent direction is

   $$\Delta x_{\text{nsd}} = -(\nabla f(x)^T P^{-1} \nabla f(x))^{-1/2} P^{-1} \nabla f(x).$$

2. Let f be a smooth function and consider the Newton method:

   $$x^+ = x + h \quad \text{where} \quad f(x) + f'(x)h = 0,$$

   to solve $f(x) = 0$. Let x^* be a root of f and assume $f'(x^*) \neq 0$. We want to show that it has a quadratic convergence near x^*.

   (a) Define

   $$\phi_{Newt}(x) = x - \frac{f(x)}{f'(x)}$$

   and show that $\phi_{Newt}(x^*) = x^*$ and $\phi'_{Newt}(x^*) = 0$.

   (b) Apply the following Taylor theorem:

   $$g(x) = g(x_0) + g'(x_0)(x - x_0) + \frac{1}{2} g''(\tilde{x})(x - x_0)^2$$

   for some $\tilde{x}$ between x_0 and s to ϕ_{Newt} at x^* to deduce

   $$|x_{k+1} - x^*| \leq C|x_k - x^*|^2$$

   near x^*.

3. Let $f : \mathbb{R}^n \to \mathbb{R}$ be given by

   $$f(x) = \frac{1}{2} x^T A x - x^T b + c,$$

   where A is an $n \times n$ symmetric positive definite matrix, b is an n-vector and c is a scalar.

   (a) If the gradient descent method with the exact line search is applied, how many iteration is required if the starting value x_0 is such that $x_0 - x^*$ is an eigenvector of A, where x^* is the solution?

   (b) If $f(x) = \frac{1}{2} x^T x - x^T b + c$, show that the number of iteration is 1 under the method in (a), using the result in (a).

4. (**Computer problem**) Newtons method with fixed step size $t = 1$ can diverge if the initial point is not close to x^*, a minimizer. In this problem we consider two examples.

(a) $f(x) = \log(e^x + e^{-x})$ has a unique minimizer $x^* = 0$. Run Newton's method with fixed step size $t = 1$, starting at $x(0) = 1$ and at $x^{(0)} = 1.1$.

(b) $f(x) = -\log x + x$ has a unique minimizer $x^* = 1$. Run Newton's method with fixed step size $t = 1$, starting at $x^{(0)} = 3$.

Plot f and f', and show the first few iterates and discuss.

5. (**Computer problem**) Consider the unconstrained problem

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

with variables $x \in \mathbb{R}^{100}$ and $\text{dom } f = \{x \mid a_i^T x < 1, \ i = 1, \ldots, m, \ |x_i| < 1, \ i = 1, \ldots n\}$. This is the problem of computing the analytic center of the set of linear inequalities

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

Choose $x^{(0)} = 0$ as an initial point and a^i from Normal distribution on $\mathbb{R}^n$.

(a) Use the gradient descent method to solve the problem, using reasonable choices for the backtracking parameters, and a stopping criterion of the form $\|\nabla f(x)\| \leq \eta$. Plot the objective function and step length versus iteration number. Once you have determined p^* to high accuracy, you also plot $f - p^*$ versus iteration and make a table. See Figure 9.4. 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. The form of table might be

k	$f(x_k) - p^*$	$(f(x_k) - p^*)/(f(x_{k-1}) - p^*)$
0	*	
1	*	*
$\vdots$	$\vdots$	$\vdots$

Do not make a big table. The purpose of the table and plot is to check the properties of the gradient descent method, linear convergence, for example.

(b) Repeat using Newton's method, with stopping criterion based on the Newton decrement λ^2. Look for quadratic convergence. Plot the objective function and step length versus iteration number likewise Figure 9.20 and 9.22. Also make a table and discuss damped Newton phase and quadratically convergent phase.

The form of table might be

k	$f(x_k) - p^*$	$(f(x_k) - p^*)/(f(x_{k-1}) - p^*)$	$(f(x_k) - p^*)/(f(x_{k-1}) - p^*)^2$
0	*		
1	*	*	*
$\vdots$	$\vdots$	$\vdots$	$\vdots$