MATH 173A: Homework #6

Due on Nov 26, 2024 at 23:59pm

Professor Cloninger

Ray Tsai

A16848188

Problem 1

Perform the conjugate gradient method by hand on the problem

$$\Phi(x) = \frac{1}{2}x^T \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} x \quad - \quad \sum_{i=1}^{2} x_i,$$

where $x \in \mathbb{R}^2$. Perform the algorithm either using version 0 or 1, where the conjugate directions are initialized and chosen algorithmically.

Proof. Let $A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and we have

$$\Phi(x) = \frac{1}{2}x^T Ax - b^T x,$$

Initialization:

$$x^{(0)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad r_0 = Ax^{(0)} - b = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \quad p_0 = -r_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Iteration 1:

$$\alpha_0 = \frac{r_0^T r_0}{p_0^T A p_0} = \frac{2}{3},$$

$$x^{(1)} = x^{(0)} + \alpha_0 p_0 = \begin{bmatrix} \frac{2}{3} \\ \frac{2}{3} \end{bmatrix},$$

$$r_1 = r_0 + \alpha_0 A p_0 = \begin{bmatrix} -1 \\ -1 \end{bmatrix} + \frac{2}{3}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ -\frac{1}{3} \end{bmatrix},$$

$$\beta_1 = \frac{r_1^T r_1}{r_0^T r_0} = \frac{1}{9},$$

$$p_1 = -r_1 + \beta_1 p_0 = \begin{bmatrix} -\frac{2}{9} \\ \frac{4}{9}. \end{bmatrix}$$

Iteration 2:

$$\alpha_1 = \frac{r_1^T r_1}{p_1^T A p_1} = \frac{3}{4},$$

$$x^{(2)} = x^{(1)} + \alpha_1 p_1 = \begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix},$$

$$r_2 = r_1 + \alpha_1 A p_1 = \begin{bmatrix} \frac{1}{3} \\ -\frac{1}{3} \end{bmatrix} + \frac{3}{4}\begin{bmatrix} -\frac{4}{9} \\ \frac{4}{9} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

$$\beta_2 = 0,$$

$$p_1 = 0$$

Thus, the conjugate gradient method converges to the solution $x^* = \begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}$ in 2 iterations. □

Problem 2

Here, we will prove the inequality used in class to prove fast convergence for strongly convex functions. Let $F(x)$ be a strongly convex function with constant c. Our goal is to show

$$F(x) - F(x^*) \leq \frac{1}{2c}\|\nabla F(x)\|^2 \qquad \text{for all } x \in \mathbb{R}^d. \tag{1}$$

(a) Fix $x \in \mathbb{R}^d$ and define the quadratic function

$$q(y) = F(x) + \nabla F(x)^T(y - x) + \frac{c}{2}\|x - y\|^2.$$

Find the y^* that minimizes $q(y)$.

Proof.

$$\nabla q(y) = \nabla F(x) - c(x - y) = 0 \implies y^* = x - \frac{1}{c}\nabla F(x).$$

□

(b) Show that $q(y^*) = F(x) - \frac{1}{2c}\|\nabla F(x)\|^2$

Proof.

$$q(y^*) = F(x) - \frac{1}{c}\|\nabla F(x)\|^2 + \frac{c}{2}\left\|\frac{1}{c}\nabla F(x)\right\|^2 = F(x) - \frac{1}{2c}\|\nabla F(x)\|^2.$$

□

(c) Use the above to deduce (1).

Proof. Since $F(x)$ is strongly convex, $F(y) \geq q(y)$ for all $y \in \mathbb{R}^d$, and thus

$$F(x^*) \geq q(x^*) \geq q(y^*) \geq F(x) - \frac{1}{2c}\|\nabla F(x)\|^2 \implies F(x) - F(x^*) \leq \frac{1}{2c}\|\nabla F(x)\|^2.$$

□

(d) Explain the proof technique in your own words to demonstrate understanding of what we did.

Proof. The strong convexity property of F yields $F \geq q$. Hence by minimizing q we can obtain a lower bound on F, and rearranging the equation yields the result. □

Problem 3

Indicate whether the following functions are strongly convex. Justify your answer.

(a) $f(x) = x$

Proof. Since $\nabla^2 f(x) = 0$, f is not strongly convex, as the Hessian is not positive definite. $\square$

(b) $f(x) = x^2$

Proof. Since $\nabla^2 f(x) = 2$, f is strongly convex with constant $c = 2$. $\square$

(c) $f(x) = \log(1 + e^x)$

Proof.

$$
\begin{aligned}
f'(x) &= \frac{e^x}{1+e^x} = \frac{1}{1+e^{-x}}, \\
f''(x) &= \frac{e^x}{(1+e^x)^2}.
\end{aligned}
$$

But then $\inf f''(x) = 0$, so f is not strongly convex. $\square$