The Chinese University of Hong Kong, Shenzhen
School of Data Science

MAT3220 · Homework 2
Due: 23:59, October 7

Instructions:

- Assignment problems must be carefully and clearly answered to receive full credit. Complete sentences that establish a clear logical progression are highly recommended.
- You must submit your assignment in Blackboard. Please upload a pdf file with codes. The file name should be in the format **last name-first name-student ID-hw1**, e.g. **Zhang-San-123456789-hw1**.
- Please make your solutions legible and write your solutions in English. You are strongly encouraged to type your solutions in LaTeX/Markdown or others.
- Late submission will **not** be graded.
- Each student **must not copy** assignment solutions from another student or from any other source.
- For those questions that ask you to write MATLAB/Python/other codes to solve the problem. Please attach your code in the **pdf file**. You also need to clearly state (write or type) the optimal solution and the optimal value you obtained. However, you do not need to attach the outputs in the command window of MATLAB/Python/others.

Problem 1 Condition Number (10 pts).
Consider the minimization problem

$$\min\left\{x^TQx : x \in \mathbb{R}^2\right\},$$

where Q is a symmetric positive definite 2×2 matrix. Suppose we use the diagonal scaling matrix

$$D = \begin{pmatrix} Q_{11}^{-1} & 0 \\ 0 & Q_{22}^{-1} \end{pmatrix},$$

where $Q_{ij} \in \mathbb{R}$ denotes the (i, j)-element of Q. Show that the above scaling matrix improves the condition number of Q in the sense that

$$\kappa\left(D^{1/2}QD^{1/2}\right) \leq \kappa(Q).$$

Solution: To simplify, we denote the matrix Q as $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$ and $ac > b$, $a > 0$, $b > 0$ and W.L.O.G, $a \geq c$. Then, $D = \begin{pmatrix} \frac{1}{a} & 0 \\ 0 & \frac{1}{c} \end{pmatrix}$ and $D^{\frac{1}{2}}QD^{\frac{1}{2}} = \begin{pmatrix} 1 & \frac{b}{\sqrt{ac}} \\ \frac{b}{\sqrt{ac}} & 1 \end{pmatrix}$.
Firstly, if $b = 0$, $\kappa(Q) = \frac{a}{c} \geq 1 = \kappa(D^{\frac{1}{2}}QD^{\frac{1}{2}})$.

Secondly, if $b \neq 0$, we calculate the eigenvalue through the definition.
Solving the following linear system

$$Q\left(\begin{array}{c} x_1 \\ x_2 \end{array}\right)=\lambda\left(\begin{array}{c} x_1 \\ x_2 \end{array}\right)$$

we have the eigenvalue λ as the solution to $\lambda^2-(a+c)\lambda+ac-b^2=0$.
Hence, we get the condition numbers:

$$\begin{aligned} \kappa(Q) &= \frac{a+c+\sqrt{(a-c)^2+4b^2}}{a+c-\sqrt{(a-c)^2+4b^2}}=\frac{1+\sqrt{\frac{(a-c)^2}{(a+c)^2}+\frac{4b^2}{(a+c)^2}}}{1-\sqrt{\frac{(a-c)^2}{(a+c)^2}+\frac{4b^2}{(a+c)^2}}} \\ \kappa(D^{\frac{1}{2}}QD^{\frac{1}{2}}) &= \frac{1+\frac{b}{\sqrt{ac}}}{1-\frac{b}{\sqrt{ac}}} \end{aligned}$$

With the fact that $\frac{b}{\sqrt{ac}}<\sqrt{\frac{(a-c)^2}{(a+c)^2}+\frac{4b^2}{(a+c)^2}}$, and the function $f(t)=\frac{1+t}{1-t}$ is increasing on (0,1), we get our conclusion that

$$\kappa\left(D^{1/2}QD^{1/2}\right)\leq\kappa(Q)$$

Problem 2 Smooth Function (20 pts).

(a) Let $g(x)=x^TAx+2b^Tx+c$, where A is a symmetric $n\times n$ matrix, $b\in\mathbb{R}^n$, and $c\in\mathbb{R}$. Show that the smallest Lipschitz constant of ∇f is $2\|A\|$.

(b) Suppose $f:\mathbb{R}^d\to\mathbb{R}$ is an L-smooth function. Prove that

$$\|\nabla f(x)\|^2\leq 2L\left(f(x)-\bar{f}\right),$$

where $\bar{f}$ satisfies $f(x)\geq\bar{f}$ for any $x\in\mathbb{R}^d$.

Solution:

(a) The gradient of f is $\nabla f(x)=2Ax+2b$ and for arbitrary $x,y\in\mathbb{R}^n$

$$\|\nabla f(x)-\nabla f(y)\|=2\|A(x-y)\|\leq 2\|A\|\|x-y\|.$$

Since the above bound is tight (let $x-y$ be an eigenvector corresponding to $\lambda_{\max}(A)$), the proof is completed.

(b) (a) By descent lemma,

$$f(y)\leq f(x)+\langle\nabla f(x),y-x\rangle+\frac{L}{2}\|y-x\|^2,\ \forall x,y\in\mathbb{R}^d$$

Note that $f(x)\geq\bar{f}$ for any $x\in\mathbb{R}^d$. Setting $y=x-\nabla f(x)/L$ in the above inequality leads to

$$\bar{f}\leq f\left(x-\frac{1}{2L}\nabla f(x)\right)\leq f(x)-\frac{1}{2L}\|\nabla f(x)\|^2.$$

Rearranging terms yields the desired result.

Problem 3 Gradient Descent Method with Error (35 pts).

Consider the objective function $f \in C^L_{1,1}(\mathbb{R}^n)$ and the sequence $\{x_k\}_k$ generated by

$$x_{k+1} = x_k - \alpha_k g_k,$$

where $g_k = \nabla f(x_k) + e_k$, the term $e_k \in \mathbb{R}^n$ is the error from the computation of gradient, and the sequence of stepsizes $\{\alpha_k\}$ satisfies

$$0 < \alpha_k \le 1/L, \quad \sum_{k=0}^{\infty} \alpha_k = \infty, \quad \text{and} \quad \sum_{k=0}^{\infty} \alpha_k^2 < \infty.$$

Assume that $\inf_{x\in\mathbb{R}^n} f(x) > -\infty$ and the error terms satisfy $\langle e_k, \nabla f(x_k)\rangle = 0$ and

$$\|e_k\|^2 \le c_1[f(x_k) - f^*] + c_2, \ \forall k \in \mathbb{N},$$

where $c_1 \ge 0$ and $c_2 \ge 0$ are some constants and $f^* := \inf_x f(x)$. Prove:

(a) The approximate descent property of $f(x_k)$:

$$f(x_{k+1}) - f^* \le \left(1 + \frac{\alpha_k^2 L c_1}{2}\right)[f(x_k) - f^*] - \frac{\alpha_k}{2}\|\nabla f(x_k)\|^2 + \frac{\alpha_k^2 L c_2}{2}, \ \forall k \in \mathbb{N}. \tag{1}$$

(b) The sequence $\{f(x_k)\}_{k\ge 0}$ converges to some finite value and it holds $\sum_{k=0}^{\infty} \alpha_k \|\nabla f(x_k)\|^2 < \infty$.

(c) The term $\|x_{k+1} - x_k\|^2$ satisfies

$$\|x_{k+1} - x_k\|^2 \le \alpha_k^2 \|\nabla f(x_k)\|^2 + \alpha_k^2 c_1 [f(x_k) - f^*] + \alpha_k^2 c_2.$$

(d) Read Section 2 and Section 3.1 in [1] and show $\lim_{k\to\infty}\|\nabla f(x_k)\| = 0$. (Hint: You can refer to the proof of [1, Theorem 2.1] (Part I in Appendix A).)

Solution:

(a) By descent lemma,

$$\begin{aligned} f(x_{k+1}) &\le f(x_k) + \langle \nabla f(x_k), x_{k+1} - x_k\rangle + \frac{L}{2}\|x_{k+1} - x_k\|^2 \\ &= f(x_k) - \alpha_k \langle \nabla f(x_k), \nabla f(x_k) + e_k\rangle + \frac{L\alpha_k^2}{2}\|\nabla f(x_k) + e_k\|^2 \\ &\le f(x_k) - \alpha_k\left(1 - \frac{\alpha_k L}{2}\right)\|\nabla f(x_k)\|^2 + \frac{\alpha_k^2 L c_2}{2} + \frac{\alpha_k^2 L c_1}{2}[f(x_k) - f^*]. \end{aligned} \tag{2}$$

Adding $-f^*$ on both sides of (2) and noting $\alpha_k \le 1/L$ lead to the desired result.

(b) Note that

$$\Pi_{t=0}^{k}\left(1 + \frac{\alpha_t^2 L c_1}{2}\right) = \exp\left[\sum_{t=0}^{k} \ln\left(1 + \frac{\alpha_t^2 L c_1}{2}\right)\right] \le \exp\left(\sum_{t=0}^{k} \frac{\alpha_t^2 L c_1}{2}\right) < C < \infty, \tag{3}$$

where $C > 0$ is some finite constant.

Dividing $\Pi_{t=0}^{k}\left(1+\alpha_t^2 L c_1/2\right)$ on both sides of (1) leads to

$$\begin{aligned}\frac{f(x_{k+1})-f^*}{\Pi_{t=0}^{k}\left(1+\frac{\alpha_t^2 L c_1}{2}\right)} \leq \frac{f(x_k)-f^*}{\Pi_{t=0}^{k-1}\left(1+\frac{\alpha_t^2 L c_1}{2}\right)} - \frac{\alpha_k}{2\Pi_{t=0}^{k}\left(1+\frac{\alpha_t^2 L c_1}{2}\right)}\left\|\nabla f(x_k)\right\|^2 \\ + \frac{\alpha_k^2 L c_2}{2\Pi_{t=0}^{k}\left(1+\frac{\alpha_t^2 L c_1}{2}\right)},\ \forall k \in \mathbb{N}.\end{aligned} \tag{4}$$

We next consider the sequence

$$A_k := \frac{f(x_k)-f^*}{\Pi_{t=0}^{k-1}\left(1+\frac{\alpha_t^2 L c_1}{2}\right)} + u_k,\ u_k := \sum_{i=k}^{\infty}\frac{\alpha_i^2 L c_2}{2\Pi_{t=0}^{i}\left(1+\frac{\alpha_t^2 L c_1}{2}\right)}.$$

The sequence $\{u_k\}_k$ is finite and decreasing to zero.

Then, the recursion (4) can be rewritten as

$$A_{k+1} \leq A_k - \frac{\alpha_k}{2C}\left\|\nabla f(x_k)\right\|^2. \tag{5}$$

Hence, the sequence $\{A_k\}_k$ is non-increasing and bounded from below, we conclude that the sequence $\{A_k\}_k$ converges to some finite value. Therefore, $\{f(x_k)\}_{k\geq 0}$ converges to some finite value as $u_k \to 0$, $k \to \infty$ and (3).

Summing the above estimate over $k = 0, 1, \cdots, n$ and letting $n \to \infty$, we conclude that $\sum_{k=0}^{\infty}\alpha_k\left\|\nabla f(x_k)\right\|^2 < \infty$.

(c) From $x_{k+1} = x_k - \alpha_k g_k$, we have

$$\begin{aligned}\left\|x_{k+1}-x_k\right\|^2 &= \alpha_k^2\left\|\nabla f(x_k)+e_k\right\|^2 \\ &= \alpha_k^2\left\|\nabla f(x_k)\right\|^2 + \alpha_k^2\left\|e_k\right\|^2,\end{aligned}$$

where we invoke $\langle e_k, \nabla f(x_k)\rangle = 0$. Noting the condition on $\|e_k\|^2$ leads to the desired result.

(d) Let us assume on the contrary that $\lim_{k\to\infty}\|\nabla f(x_k)\| \neq 0$. Then there exists $\varepsilon > 0$ and infinite sequences $(t_i)_i$ and $(\ell_i)_i$ such that $\ell_i > t_i$ for all i and

$$\left\|\nabla f\left(x_{t_i}\right)\right\| \geq 2\varepsilon, \quad \left\|\nabla f\left(x_{\ell_i}\right)\right\| < \varepsilon, \quad \text{and} \quad \left\|\nabla f\left(x_k\right)\right\| \geq \varepsilon$$

for all $k = t_i + 1, \ldots, \ell_i - 1$. Using the result from (b), it follows

$$\infty > \sum_{k=0}^{\infty}\alpha_k\left\|\nabla f\left(x_k\right)\right\|^2 \geq \sum_{i=0}^{\infty}\sum_{k=t_i}^{\ell_i-1}\alpha_k\left\|\nabla f\left(x_k\right)\right\|^2 \geq \varepsilon^2\sum_{i=0}^{\infty}\sum_{k=t_i}^{\ell_i-1}\alpha_k$$

and consequently, setting $\beta_i := \sum_{k=t_i}^{\ell_i-1}\alpha_k$, we have $\beta_i \to 0$ as $i \to \infty$. Applying the Cauchy-Schwarz inequality, i.e.,

$$\sum_{k=1}^{K}|a_k b_k| \leq \left[\sum_{k=1}^{K}a_k^2\right]^{\frac{1}{2}}\left[\sum_{k=1}^{K}b_k^2\right]^{\frac{1}{2}}$$

we obtain

$$
\begin{aligned}
\|x_{\ell_i} - x_{t_i}\| &\leq \sum_{k=t_i}^{\ell_i - 1} \|x_{k+1} - x_k\| \leq \sum_{k=t_i}^{\ell_i - 1} \alpha_k (\|\nabla f(x_k)\| + \|e_k\|) \\
&= \sum_{k=t_i}^{\ell_i - 1} \sqrt{\alpha_k}\sqrt{\alpha_k} \|\nabla f(x_k)\| + \sum_{k=t_i}^{\ell_i - 1} \alpha_k \|e_k\| \\
&\leq \sqrt{\beta_i} \left[\sum_{k=t_i}^{\ell_i - 1} \alpha_k \|\nabla f(x_k)\|^2\right]^{\frac{1}{2}} + \sqrt{c}\beta_i \\
&\leq \sqrt{\beta_i} \left[\sum_{k=0}^{\infty} \alpha_k \|\nabla f(x_k)\|^2\right]^{\frac{1}{2}} + \sqrt{c}\beta_i,
\end{aligned}
$$

where the constant c is defined as $\|e_k\| \leq c_1[f(x_k) - f^*] + c_2 \leq c$ due to $\{f(x_k)\}_{k\geq 0}$ is convergent. Consequently, combining $\sum_{k=0}^{\infty} \alpha_k \|\nabla f(x_k)\|^2 < \infty$ and $\beta_i \to 0$, we can infer $\|x_{\ell_i} - x_{t_i}\| \to 0$. Finally, using the special definition of the sequences $(t_i)_i$ and $(\ell_i)_i$, the inverse triangle inequality, and the Lipschitz continuity of ∇f, we have

$$\varepsilon \leq \|\nabla f(x_{\ell_i})\| - \|\nabla f(x_{t_i})\| \leq \|\nabla f(x_{\ell_i}) - \nabla f(x_{t_i})\| \leq L \|x_{\ell_i} - x_{t_i}\| \to 0$$

as $i \to \infty$, which is a contradiction. Consequently, our assumption was wrong and it follows $\nabla f(x_k) \to 0$ as $k \to \infty$.

Problem 4 Second-Order Method (20pts).
Let f be a twice continuously differentiable function satisfying $LI \succeq \nabla^2 f(x) \succeq mI$ for some $L > m > 0$ and let x^* be the unique minimizer of f over $\mathbb{R}^n$.

(a) Show that

$$f(x) - f(x^*) \geq \frac{m}{2}\|x - x^*\|^2$$

for any $x \in \mathbb{R}^n$.

(b) Let $\{x_k\}_{k\geq 0}$ be the sequence generated by damped Newton's method with constant stepsize $t_k = \frac{m}{L}$. Show that

$$f(x_k) - f(x_{k+1}) \geq \frac{m}{2L} \nabla f(x_k)^T \left(\nabla^2 f(x_k)\right)^{-1} \nabla f(x_k),$$

and $x_k \to x^*$ as $k \to \infty$.

Solution:

(a) The condition $\nabla^2 f(x) \succeq mI$ implies strong convexity. Thus, the following inequality holds

$$f(x) - f(y) \geq \langle \nabla f(y), x - y\rangle + \frac{m}{2}\|x - y\|^2.$$

Taking $y = x^*$ and noting that $\nabla f(x^*) = 0$ complete the proof.

(b) Let us recall the update rule of damped Newton's method, i.e., $x_{k+1} = x_k - t_k d_k$, where $t_k = m/L,\ d_k = (\nabla^2 f(x_k))^{-1}\nabla f(x_k)$. Then, noting the Lipschitz continuity of ∇f and applying the descent lemma, i.e.,

$$f(x) - f(y) \le \langle \nabla f(y), x - y\rangle + \frac{L}{2}\|x-y\|^2,$$

with $x = x_{k+1}$ and $y = x_k$ yield

$$\begin{aligned} f(x_k) - f(x_{k+1}) &\ge \frac{m}{L}\nabla f\left(x_k\right)^T \left(\nabla^2 f\left(x_k\right)\right)^{-1}\nabla f\left(x_k\right) - \frac{m^2}{2L}\left\|\left(\nabla^2 f\left(x_k\right)\right)^{-1}\nabla f\left(x_k\right)\right\|^2 \\ &\ge \frac{m}{2L}\nabla f\left(x_k\right)^T \left(\nabla^2 f\left(x_k\right)\right)^{-1}\nabla f\left(x_k\right), \end{aligned}$$

where the second equality holds since

$$\begin{aligned} \left\|\left(\nabla^2 f\left(x_k\right)\right)^{-1}\nabla f\left(x_k\right)\right\|^2 &= \left\|\left(\nabla^2 f\left(x_k\right)\right)^{-1/2}\left(\nabla^2 f\left(x_k\right)\right)^{-1/2}\nabla f\left(x_k\right)\right\|^2 \\ &\le \left\|\left(\nabla^2 f\left(x_k\right)\right)^{-1/2}\right\|^2\left\|\left(\nabla^2 f\left(x_k\right)\right)^{-1/2}\nabla f\left(x_k\right)\right\|^2 \\ &\le \frac{1}{m}\nabla f\left(x_k\right)^T\left(\nabla^2 f\left(x_k\right)\right)^{-1}\nabla f\left(x_k\right). \end{aligned}$$

Unfolding the inequality by summing over $k = 0, 1, \ldots, n$, we get

$$f(x_0) - f(x_{n+1}) \ge \frac{m}{2L}\sum_{k=0}^{n}\nabla f\left(x_k\right)^T\left(\nabla^2 f\left(x_k\right)\right)^{-1}\nabla f\left(x_k\right) \ge \frac{m}{2L^2}\sum_{k=0}^{n}\|\nabla f(x_k)\|^2,$$

where the second inequality holds owing to $[\nabla^2 f(x)]^{-1} \succeq I/L$ for all $x \in \mathbb{R}^n$. By letting $n \to \infty$ and noting $f(x) \ge f(x^*)$ for all $x \in \mathbb{R}^n$, we obtain

$$\infty > f(x_0) - f(x^*) \ge \frac{m}{2L^2}\sum_{k=0}^{\infty}\|\nabla f(x_k)\|^2 \quad \Longrightarrow \quad \lim_{k\to\infty}\|\nabla f(x_k)\| = 0.$$

Moreover, recall $\nabla^2 f(x) \succeq mI$, which implies $\|\nabla f(x) - \nabla f(y)\| \ge m\|x-y\|$.[1] Hence, using the fact that $\nabla f(x^*) = 0$ and the estimate given above, we have

$$m\|x_k - x^*\| \le \|\nabla f(x_k) - \nabla f(x^*)\| = \|\nabla f(x_k)\| \to 0, \quad \text{as } k \text{ tends to infinity.}$$

Problem 5 Implementation of Gradient Method (15 pts).
Consider the quadratic minimization problem

$$\min\left\{x^T A x : x \in \mathbb{R}^5\right\},$$

where A is the 5×5 Hilbert matrix defined by

$$A_{i,j} = \frac{1}{i+j-1}, \quad i,j = 1,2,3,4,5.$$

The matrix can be constructed via the MATLAB command *A=hilb(5)*. Run the following methods and compare the number of iterations required by each of the methods when the initial vector is $x_0 = (1,2,3,4,5)^T$ to obtain a solution x with $\|\nabla f(x)\| \le 10^{-4}$:

[1] This inequality holds by one of the definitions of strongly convex function and Cauchy-Schwarz inequality. Check `https://en.wikipedia.org/wiki/Convex_function#Strongly_convex_functions`. For a rigorous proof, see Theorem 2.1.10 of the book *Lectures on Convex Optimization* by Yuri Nesterov.

(1.) gradient method with backtracking stepsize rule and parameters $\alpha = 0.5, \beta = 0.5$, and $s = 1$;

(2.) gradient method with backtracking stepsize rule and parameters $\alpha = 0.1, \beta = 0.5$, and $s = 1$;

(3.) gradient method with exact line search;

(4.) diagonally scaled gradient method with diagonal elements $D_{ii} = \frac{1}{A_{ii}}, i = 1, 2, 3, 4, 5$ and exact line search;

(5.) diagonally scaled gradient method with diagonal elements $D_{ii} = \frac{1}{A_{ii}}, i = 1, 2, 3, 4, 5$ and backtracking line search with parameters $\alpha = 0.1, \beta = 0.5,\ s = 1$.

References

[1] Xiao Li and Andre Milzarek. A unified convergence theorem for stochastic optimization methods. *Advances in Neural Information Processing Systems*, 35:33107–33119, 2022.