MSBD 5007 HW2

RONG Shuo

March 27, 2025

Question1

Determine the convexity of the following functions, where $\boldsymbol{x} \in \mathbb{R}^n$ and $\boldsymbol{X} \in \mathbb{S}^n_{++}$(the set of symmetric positive definite matrices). Justify your answer.

(a) $f(\boldsymbol{x}) = \log(e^{x_1} + e^{x_2} + \cdots + e^{x_n})$.

(b) $f(\boldsymbol{X}) = \text{logdet}(\boldsymbol{X})$.

Question2

Consider the linear system $\boldsymbol{Ax} = \boldsymbol{b}$, with $\boldsymbol{A} = \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}$ and $\boldsymbol{b} \in \begin{bmatrix} 2 \\ -4 \end{bmatrix}$, and the initial guess $\boldsymbol{x}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

(a) Present the first two update iterations using the **steepest descent algorithm**.

(b) Present the first two updated iterations using the **conjugate gradient algorithm**.

Question3

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a twice continuously differentiable functions. Suppose that for every $\boldsymbol{x} \in \mathbb{R}^n$, the eigenvalues of the Hessian matrix $\nabla^2 f(\boldsymbol{X})$ lie uniformly in the interval $[m,\ M]$ with $0 < m \leq M < \infty$.

Prove that:

(a) The function f has a unique global minimizer $\boldsymbol{x}^*$.

(b) For all $\boldsymbol{x} \in \mathbb{R}^n$, the following inequality holds:

$$\frac{1}{2M}\|\nabla f(x)\|^2 \leq f(\boldsymbol{x}) - f(\boldsymbol{x}^*) \leq \frac{1}{2m}\|\nabla f(\boldsymbol{x})\|^2$$

Question4

Consider the optimization problem $\min_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x})$, where $f : \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function. To develop a weighted gradient descent method, let $\boldsymbol{W} \in \mathbb{R}^{n\times n}$ be a symmetric positive definite (SPD) matrix. Denote by $\boldsymbol{W}^{\frac{1}{2}}$ the unique SPD square root of $\boldsymbol{W}$ (i.e., $(\boldsymbol{W}^{\frac{1}{2}})^2 = \boldsymbol{W}$) and by $\boldsymbol{W}^{-\frac{1}{2}}$ its inverse. Given the current iterate $\boldsymbol{x}^{(k)}$, define the next iterate $\boldsymbol{x}^{(k+1)}$ as the solution of the following constrained optimization problem:

$$\begin{gathered} \min_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x}^{(k)}) + \langle \nabla f(\boldsymbol{x}^{(k)}), \boldsymbol{x} - \boldsymbol{x}^{(k)} \rangle \\ \text{subject to } \|\boldsymbol{W}^{\frac{1}{2}}(\boldsymbol{x} - \boldsymbol{x}^(k))\|_2 \leq \alpha_k \|\boldsymbol{W}^{-\frac{1}{2}}\nabla f(\boldsymbol{x}^(k))\|_2 \end{gathered}$$

where $\alpha_k > 0$ is a step-size parameter.

Answer the following questions:

(a) Derive an explicit formula for $\boldsymbol{x}^{(k+1)}$

(b) Prove that $\boldsymbol{x}^{(k+1)}$ is equivalently the unique minimizer of the unconstrained quadratic problem:

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \left\{ \nabla f(\boldsymbol{x}^{(k)}) + \langle f(\boldsymbol{x}^{(k)}), \boldsymbol{x} - \boldsymbol{x}^{(k)} \rangle + \frac{1}{2\alpha_k}\|\boldsymbol{W}^{\frac{1}{2}}(\boldsymbol{x} - \boldsymbol{x}^{(k)})\|_2^2 \right\}$$