Homework 1

Zhen Haoyu

September 25, 2021

Problem 1

$f(x) = 2x_1^2 + x_1x_2 + x_2^2 - 3x_1 - 5x_2$

(a) Because

$$2x_1^2 + x_1x_2 + x_2^2 - 3x_1 - 5x_2 \geq \frac{3}{2}x_1^2 + \frac{1}{2}x_2^2 - 3x_1 - 5x_2$$

And if

$$|x_1|, |x_2| > 100,$$

then

$$f(x) \geq \frac{1}{2}x_1^2 + \frac{1}{4}x_2^2.$$

Thus we get: when $||x|| \to +\infty$, $f(x) \to +\infty$, which means that $f(x)$ is coercive.

(b) The maximun of $f(x)$ does **not** exist.

The local minimun (x_1, x_2) should satisfy:

$$\frac{\partial f}{\partial x_1} = 4x_1 + x_2 - 6 = 0$$

and

$$\frac{\partial f}{\partial x_2} = x_1 + 2x_2 - 5 = 0.$$

Then we get $(x_1, x_2) = (\frac{1}{7}, \frac{17}{7})$ and $\nabla^2 f$ is positive definite.

Thus the minimum is $-\frac{44}{7}$.

Problem 2

$$f(\boldsymbol{\omega}) = \sum_{i=1}^{m} \log(1 + e^{-y_i \boldsymbol{x}_i^T \boldsymbol{\omega}})$$

(a) By $\log(1 + e^x) > 0$, we have:

$$\forall \boldsymbol{\omega},\ f(\boldsymbol{\omega}) > 0.$$

And

$$\lim_{k \to +\infty} f(k\boldsymbol{\omega}_0) = \lim_{k \to +\infty} \sum_{i=1}^{m} \log(1 + e^{-k y_i \boldsymbol{x}_i^T \boldsymbol{\omega}_0}) = 0$$

Thus f does not have a global minimum.

(b) **Proof by step:**

(i)

$$\begin{aligned} f(\boldsymbol{\omega}) &= \sum_{i=1}^{m} \log(1 + e^{-y_i \boldsymbol{x}_i^T \boldsymbol{\omega}}) \\ &\geq \sum_{y_i \boldsymbol{x}_i^T \boldsymbol{\omega} < 0} \log(1 + e^{-y_i \boldsymbol{x}_i^T \boldsymbol{\omega}}) \\ &\geq \sum_{y_i \boldsymbol{x}_i^T \boldsymbol{\omega} < 0} -y_i \boldsymbol{x}_i^T \boldsymbol{\omega} \\ &\geq \max_{y_i \boldsymbol{x}_i^T \boldsymbol{\omega} < 0} -y_i \boldsymbol{x}_i^T \boldsymbol{\omega} \\ &= \max_{1 \leq i \leq m} -y_i \boldsymbol{x}_i^T \boldsymbol{\omega} \end{aligned}$$

(ii) Set S is bounded and closed while the function $h(\boldsymbol{\omega})$ is continuos.
By Extreme Value Theorem, $h(\boldsymbol{\omega})$ has a global minimum.

(iii) $\forall \boldsymbol{\omega} \in S$:

$$\frac{h(\boldsymbol{\omega})}{||\boldsymbol{\omega}||} \geq \max_{1 \leq i \leq m} -y_i \boldsymbol{x}_i^T \frac{\boldsymbol{\omega}}{||\boldsymbol{\omega}||} \geq C$$

Thus

$$h(\boldsymbol{\omega}) \geq C||\boldsymbol{\omega}||$$

(iv) By (i), (ii) and (iii), we have:

$$\exists C \in \mathbb{R},\ f(\boldsymbol{\omega}) \geq C||\boldsymbol{\omega}||$$

So when $||\boldsymbol{\omega}|| + \infty$, $f(\boldsymbol{\omega}) \to +\infty$. Then function f is coercive and continuous.
Thus, f has a global minimum.

(c)

$$\nabla f = \sum_{i=1}^{m} \frac{-y_i \boldsymbol{x}_i e^{-y_i \boldsymbol{x}_i^T \boldsymbol{\omega}}}{1 + e^{-y_i \boldsymbol{x}_i^T \boldsymbol{\omega}}}$$

Problem 3

(a) Given $\boldsymbol{x}$, let

$$R(\boldsymbol{d}) = f(\boldsymbol{x}+\boldsymbol{d}) - f(\boldsymbol{x}) - \nabla f(\boldsymbol{x})^T \boldsymbol{d}$$

Then

$$\nabla R(\boldsymbol{d}) = \nabla f(\boldsymbol{x}+\boldsymbol{d}) - \nabla f(\boldsymbol{x})$$

By Lagrange mean value theorem, $\exists \boldsymbol{\xi} \in (\boldsymbol{x},\ \boldsymbol{x}+\boldsymbol{d})$:

$$\nabla f(\boldsymbol{x}+\boldsymbol{d}) - \nabla f(\boldsymbol{x}) = \nabla^2 f(\boldsymbol{\xi})\boldsymbol{d}$$

Lemma 1 *If $\nabla f(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$ and $\boldsymbol{A}$ is symmetric, then $f(x) = \frac{1}{2}\boldsymbol{x}^T A\boldsymbol{x} + C$, where $\boldsymbol{A} = (a_{ij})_{n\times n}$, $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T$ and C is a const.*

Proof 1

$$f(x) = \int \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \mathrm{d}x_i = \int \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_j \mathrm{d}x_i = \sum_{i\neq j} a_{ij}x_i x_j + \sum_{i=1}^{n} \frac{1}{2} a_{ii}x_i^2 + C = \frac{1}{2}\boldsymbol{x}^T A\boldsymbol{x} + C$$

Because $\nabla^2 f(\boldsymbol{\xi})$ is symmetric, we have:

$$\begin{aligned} R(\boldsymbol{d}) &= \frac{1}{2}\boldsymbol{d}^T \nabla^2 f(\boldsymbol{\xi})\boldsymbol{d} + C \\ &= \frac{1}{2}\boldsymbol{d}^T \nabla^2 f(\boldsymbol{x}+t\boldsymbol{d})\boldsymbol{d} + C, \end{aligned}$$

where $t = \dfrac{\boldsymbol{\xi}-\boldsymbol{x}}{d} \in (0,\ 1)$.

By $R(\mathbf{0}) = 0$, we have: $C = 0$ and finnaly get:

$$R(\boldsymbol{d}) = \frac{1}{2}\boldsymbol{d}^T \nabla^2 f(\boldsymbol{x}+t\boldsymbol{d})\boldsymbol{d}$$

(b) **Proof:** By the Newton-Leibniz formula, $\forall i \in 1, 2, \ldots, n$:

$$\int_0^1 \left[\nabla f_i(\boldsymbol{x}+t\boldsymbol{d})\right]\boldsymbol{d}\ \mathrm{d}t = \int_0^1 \sum_{j=1}^{n} f_{ij}(\boldsymbol{x}+t\boldsymbol{d})\mathrm{d}(td_j) = f_i(\boldsymbol{x}+\boldsymbol{d}) - f_i(\boldsymbol{x})$$

where $f_i = \dfrac{\partial f}{\partial x_i}$, $f_{ij} = \dfrac{\partial^2 f}{\partial x_i \partial x_j}$ and $\boldsymbol{d} = (d_1, d_2, \ldots, d_n)^T$.

Also, we have

$$\int_0^1 \nabla^2 f(\boldsymbol{x}+t\boldsymbol{d})\boldsymbol{d}\ \mathrm{d}t = \begin{pmatrix} \displaystyle\int_0^1 \left[(\nabla f_1)(\boldsymbol{x}+t\boldsymbol{d})\right]\boldsymbol{d}\ \mathrm{d}t \\ \displaystyle\int_0^1 \left[(\nabla f_2)(\boldsymbol{x}+t\boldsymbol{d})\right]\boldsymbol{d}\ \mathrm{d}t \\ \vdots \\ \displaystyle\int_0^1 \left[(\nabla f_n)(\boldsymbol{x}+t\boldsymbol{d})\right]\boldsymbol{d}\ \mathrm{d}t \end{pmatrix}.$$

Therefore,

$$\int_0^1 \nabla^2 f(\boldsymbol{x}+t\boldsymbol{d})\boldsymbol{d}\ \mathrm{d}t = \nabla f(\boldsymbol{x}+\boldsymbol{d}) - \nabla f(\boldsymbol{x})$$

Problem 4

Solve:

$D_1(\boldsymbol{A}) = 6$, $D_2(\boldsymbol{A}) = 26$ and $D_3(\boldsymbol{A}) = 30$. So $\boldsymbol{A}$ is positive definite.

$D_1(\boldsymbol{B}) = 1$, $D_2(\boldsymbol{B}) = -2$ and $D_3(\boldsymbol{B}) = 5$. So $\boldsymbol{B}$ is indefinite.

$\boldsymbol{C}$'s eigenvalues are $3,\ 3,\ 0$. So $\boldsymbol{C}$ is positive semidefinite.