DSA3102 Convex optimization

HW 2 Solution

1. (a)The gradient is

$$\nabla l(w) = -\frac{1}{n}\sum_{i=1}^{n} \frac{y_i}{1+\exp(y_i w^T x_i)} x_i.$$

(b)The Hessian is given by

$$H_l(w) = \frac{1}{n}\sum_{i=1}^{n} \frac{x_i x_i^T \exp(y_i w^T x_i)}{(1+\exp(y_i w^T x_i))^2} y_i^2.$$

Given any vector v

$$v^T H_l(w) v = \sum_{i=1}^{n} \frac{(v^T x_i)^2 y_i^2 \exp(y_i w^T x_i)}{(1+\exp(y_i w^T x_i))^2} \geq 0.$$

So $H_l(w)$ is PSD and l is convex.
(c) (d): Please see the uploaded MATLAB codes, HW2Q1.

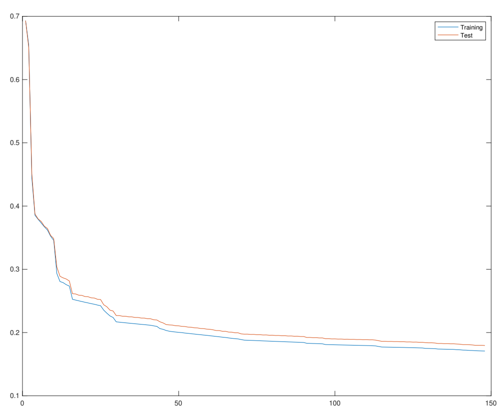

Figure 1: Q1

2. Consider the following minimization problem in $\mathbf{R}^3$:

$$\begin{aligned} \min \quad & f(\mathbf{x}) := x_1 - x_2 \\ \text{s.t.} \quad & g_1(\mathbf{x}) := x_1^2 + x_2^2 - 2 = 0 \\ & g_2(\mathbf{x}) := x_2 - x_3^3 = 0 \end{aligned}$$

(a) Let S be the feasible set. Find a positive number $\alpha > 0$ such that

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in S \implies -\alpha \le x_i \le \alpha, \ \forall\, i = 1, 2, 3.$$

(b) Explain why the minimization problem has an optimal solution.

(c) Does the regularity condition hold at every feasible point? Justify your answer.

(d) Find all the optimal solutions of the minimization problem.

Solution. (a) $g_1(\mathbf{x}) = 0 \ \Rightarrow \ |x_1|, |x_2| \le \sqrt{2}$. Now $x \in S \ \Rightarrow \ |x_3| = |x_2|^{1/3} \le 2^{1/6} \le \sqrt{2}$. We take the required $\alpha = \sqrt{2}$.

(b) S is closed and bounded, and f is continuous on S. By Weierstrass Theorem, the NLP has a global minimizer.

(c) Now

$$\left[\nabla g_1(\mathbf{x}) \ \nabla g_2(\mathbf{x})\right] = \begin{bmatrix} 2x_1 & 0 \\ 2x_2 & 1 \\ 0 & -3x_3^2 \end{bmatrix} \quad \text{which has rank 2 unless the columns are parallel.}$$

We will show that the columns cannot be parallel by contradiction, and hence conclude that all the feasible points are regular. Suppose the columns are parallel. That is, there is a scalar ρ such that

$$\begin{bmatrix} 2x_1 \\ 2x_2 \\ 0 \end{bmatrix} = \rho \begin{bmatrix} 0 \\ 1 \\ -3x_3^2 \end{bmatrix} \quad \Rightarrow \ x_1 = 0, \ x_2 = \rho/2.$$

Now $g_1(\mathbf{x}) = 0 \Rightarrow \rho^2 = 8$. Since $\rho \ne 0$, from the third component, we get $x_3 = 0$. But $x_3 = 0$ and $g_2(\mathbf{x}) = 0 \Rightarrow x_2 = 0$. This contradicts $x_2 = \rho/2 \ne 0$.

(d) For $x \in S$, the KKT conditions are

$$\nabla f(\mathbf{x}) + \lambda_1 \nabla g_1(\mathbf{x}) + \lambda_2 \nabla g_2(\mathbf{x}) = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + \lambda_1 \begin{bmatrix} 2x_1 \\ 2x_2 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 \\ 1 \\ -3x_3^2 \end{bmatrix} = \mathbf{0}$$

$$\Rightarrow \begin{aligned} 1 + 2\lambda_1 x_1 &= 0 \\ -1 + 2\lambda_1 x_2 + \lambda_2 &= 0 \\ -3\lambda_2 x_3^2 &= 0 \end{aligned} \tag{1}$$

From the third equation in (1), either $x_3 = 0$ or $\lambda_2 = 0$. We consider 2 cases.

(i) Suppose $x_3 = 0$. Then $g_2(\mathbf{x}) = 0 \Rightarrow x_2 = 0$. Hence second equation in (1) $\Rightarrow \lambda_2 = 1$. Now $g_1(\mathbf{x}) = 0 \Rightarrow x_1 = \pm\sqrt{2}$, and the first equation in (1) $\Rightarrow \lambda_1 = \mp 1/(2\sqrt{2})$. Thus $\mathbf{x}^{(1)} = [-\sqrt{2}; 0; 0]$ is a KKT solution with $f(\mathbf{x}^{(1)}) = -\sqrt{2}$; and $\mathbf{x}^{(2)} = [\sqrt{2}; 0; 0]$ is a KKT solution with $f(\mathbf{x}^{(2)}) = \sqrt{2}$.

(ii) Suppose $\lambda_2 = 0$. Then the first 2 equations in (1) give

$$x_1 = \frac{-1}{2\lambda_1}, \quad x_2 = \frac{1}{2\lambda_1}.$$

Now $g_1(\mathbf{x}) = 0 \Rightarrow \ \lambda_1 = \pm 1/2$.

If $\lambda_1 = 1/2$, then $x_1 = -1$, $x_2 = 1$, and $g_2(\mathbf{x}) = 0 \Rightarrow x_3 = 1$. If $\lambda_1 = -1/2$, then $x_1 = 1$, $x_2 = -1$, and $g_2(\mathbf{x}) = 0 \Rightarrow x_3 = -1$.

Thus $\mathbf{x}^{(3)} = [-1; 1; 1]$ is a KKT solution with $f(\mathbf{x}^{(3)}) = -2$; Thus $\mathbf{x}^{(4)} = [1; -1; -1]$ is a KKT solution with $f(\mathbf{x}^{(4)}) = 2$.

The global minimizer of the NLP is $\mathbf{x}^{(3)}$.

3. Consider the the scalar minimization problem

$$\begin{aligned} \min \quad & x \\ \text{s.t.} \quad & x^2 \geq 0 \\ & x + 1 \geq 0, \end{aligned}$$

for which the solution is $x^* = -1$.

Let

$$F^< = \{x \in R : x^2 > 0, \ x + 1 > 0\}.$$

Write down an explicit expression of the set $F^<$.

Define $P(\cdot, \mu) : R \to R$ by

$$P(x, \mu) = f(x) + \mu B(x) = x - \mu(\log(x^2) + \log(x + 1))$$

For each $\mu > 0$, find all local minimizers x_μ of $P(\cdot, \mu)$ in $F^<$. Find the limits of all convergent subsequences of $\{x_\mu : \mu \to 0\}$.

Solution. $F^< = (-1, 0) \cup (0, +\infty)$.

For any $x \in F^<$,

$$P'(x, \mu) = 1 - \mu\left(\frac{2x}{x^2} + \frac{1}{x + 1}\right) = 1 - \mu\frac{3x + 2}{x^2 + x} = \frac{x^2 + (1 - 3\mu)x - 2\mu}{x^2 + x}.$$

$P' = 0$ gives rise to two stationary points:

$$\begin{aligned} x_+(\mu) &= \frac{-(1 - 3\mu) + \sqrt{(1 - 3\mu)^2 + 8\mu}}{2} \in (0, +\infty) \\ x_-(\mu) &= \frac{-(1 - 3\mu) - \sqrt{(1 - 3\mu)^2 + 8\mu}}{2} \in (-1, 0) \end{aligned}$$

$$P''(x,\mu) = \mu \frac{3x^2+4x+2}{(x^2+x)^2} > 0, \quad \forall\, x \in F^<.$$

Thus, both $x_+(\mu)$ and $x_-(\mu)$ are local minimizers.

As $\mu \to 0$, $x_+(\mu) \to 0$ and $x_-(\mu) \to -1 = x^*$.

4. You are given 20 cm of wire and 16 cm^2 of special paper to make a rectangular box with a bottom and a lid. Find the dimension of the box to maximize its volume. All the wire and paper must be used.

   (i) Formulate the problem as a nonlinear programming problem (NLP).

   (ii) By using mathematical arguments, show that your NLP has an optimal solution.

   (iii) Find the optimal solution.

   Solution. (i) Let the dimension of the box be x_1, x_2, x_3. Then we want to solve the following NLP:

$$\begin{aligned} \max \quad & f(\mathbf{x}) := x_1x_2x_3 \\ \text{s.t.} \quad & g_1(\mathbf{x}) := x_1 + x_2 + x_3 - 5 = 0 \\ & g_2(\mathbf{x}) := x_1x_2 + x_1x_3 + x_2x_3 - 8 = 0, \quad x_1, x_2, x_3 \geq 0. \end{aligned}$$

   (ii) Let $S = \{\mathbf{x} \in \mathbf{R}^3 : g_1(\mathbf{x}) = 0,\ g_2(\mathbf{x}) = 0,\ \mathbf{x} \geq \mathbf{0}\}$. The set S is closed and bounded since $\mathbf{0} \leq \mathbf{x} \leq [5;5;5]$. Since f is continuous on S, Weierstrass Theorem implies that the NLP has a global maximizer.

   (iii) (a) Claim: all feasible points are regular.

$$[\nabla g_1(\mathbf{x}) \ \ \nabla g_2(\mathbf{x})] = \begin{bmatrix} 1 & x_2 + x_3 \\ 1 & x_1 + x_3 \\ 1 & x_1 + x_2 \end{bmatrix}.$$

   The matrix has rank $= 2$, unless $x_2 + x_3 = x_1 + x_3 = x_1 + x_2 \Leftrightarrow 5 - x_1 = 5 - x_2 = 5 - x_3 \Leftrightarrow x_1 = x_2 = x_3$. Thus $\mathbf{x}$ is regular, unless $x_1 = x_2 = x_3$. We show that the latter situation cannot happen by contradiction. Suppose $x_1 = x_2 = x_3$. Then $g_1(\mathbf{x}) = 0 = g_2(\mathbf{x}) \Rightarrow x_1 = 5/3$ and $x_1 = \sqrt{8/3}$ (impossible!).

   (b) Find the optimal solution from the KKT points.
   The KKT conditions are:

$$\nabla f(\mathbf{x}) + \lambda_1 \nabla g_1(\mathbf{x}) + \lambda_2 \nabla g_2(\mathbf{x}) = \mathbf{0}$$

$$\Longleftrightarrow \quad \begin{aligned} x_2x_3 + \lambda_1 + \lambda_2(x_2 + x_3) &= 0 \\ x_1x_3 + \lambda_1 + \lambda_2(x_1 + x_3) &= 0 \\ x_1x_2 + \lambda_1 + \lambda_2(x_1 + x_2) &= 0 \end{aligned} \qquad (2)$$

Note that summing all the equations in (2) implies that $8 + 3\lambda_1 + 10\lambda_2 = 0$. The difference between the equations in (2) gives

$$\begin{aligned} (x_1 - x_2)(x_3 + \lambda_2) &= 0 \\ (x_2 - x_3)(x_1 + \lambda_2) &= 0 \\ (x_1 - x_3)(x_2 + \lambda_2) &= 0 \end{aligned} \tag{3}$$

If $x_1 - x_2 \neq 0$, $x_2 - x_3 \neq 0$ and $x_1 - x_3 \neq 0$, then $x_1 = x_2 = x_3 = -\lambda_2$, contradicting that $x_1 \neq x_2$, $x_2 \neq x_3$, $x_1 \neq x_3$. Thus $x_1 - x_2 = 0$, or $x_2 - x_3 = 0$, or $x_1 - x_3 = 0$. It is enough to consider the case $x_1 - x_2 = 0$ since (2) remains unchanged with respect to permutation in x_1, x_2, x_3.

Assume that $x_1 = x_2$. Then $g_1(\mathbf{x}) = 0 = g_2(\mathbf{x}) \Rightarrow$

$$\left.\begin{aligned} 2x_1 + x_3 &= 5 \\ x_1^2 + 2x_1x_3 &= 8 \end{aligned}\right\} \Rightarrow (3x_1 - 4)(x_1 - 2) = 0 \Rightarrow x_1 = 4/3 \text{ or } x_1 = 2.$$

Thus one of the KKT solution is: $\mathbf{x}^{(1)} = [4/3; 4/3; 7/3]$, and (3) $\Rightarrow \lambda_2 = -x_1 = -4/3$, and hence $\lambda^{(1)} = [16/9; -4/3]$. The other KKT solution is: $\mathbf{x}^{(2)} = [2; 2; 1]$, and (3) $\Rightarrow \lambda_2 = -x_1 = -2$, and hence $\lambda^{(2)} = [4; -2]$. Note that $f(\mathbf{x}^{(1)}) = 112/27$, $f(\mathbf{x}^{(2)}) = 4$. Thus, the optimal solution is $\mathbf{x}^{(1)} = [4/3; 4/3; 7/3]$.