SI151A
Convex Optimization and its Applications in Information Science, Fall 2021
Homework 3

Due on Nov 1, 2021, 23:59 UTC+8

1. Find all of the stationary points of the following three functions. For each stationary point, determine if it is a local minimum, local maximum, or neither.

(1) $f_1(x, y) = x^2 - 4xy + 4y^2 + 2x + y$ on $\mathbb{R}^2$. (10 points)

(2) $f_2(x, y) = \frac{x^2}{y^4 - 4y^2 + 5}$ on $\mathbb{R}^2$. (10 points)

(3) $f_3(x, y) = 100(y - x^2)^2 - x^2$ on $\mathbb{R}^2$. (10 points)

Solution:

(1) We calculate the gradient:

$$\nabla f_1(x, y) = (2x - 4y + 2, 8y - 4x + 1).$$

The stationary points have to satisfy $\nabla f_1(x, y) = 0$, i.e. satisfy the linear system

$$\begin{pmatrix} 2 & -4 \\ 4 & -8 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \end{pmatrix}$$

However the matrix has rank 1 and the vector $(-2, -1)$ is not in its range, so that the linear system does not have a solution. This implies that f_1 does not have a stationary point.

(2) We calculate the gradient:

$$\nabla f_2(x, y) = \left(\frac{2x}{y^4 - 4y^2 + 5}, \frac{-4x^2 y(y^2 - 2)}{(y^4 - 4y^2 + 5)^2} \right).$$

Noting that $y^4 - 4y^2 + 5 = (y^2 - 2)^2 + 1 > 0$ for all y, the denominator is never zero, and so the gradient vanishes iff $x = 0$. Finally, note that since the denominator is always positive and the numerator is nonnegative, $f_2(x, y) \geq 0$, with equality iff $x = 0$. It follows that every stationary point is a local minimum.

(3) We calculate the gradient:

$$\nabla f_3(x, y) = (-400(y - x^2)x - 2x, 200(y - x^2)),$$

which vanishes only when $(x, y) = (0, 0)$. Therefore, $(0, 0)$ is the only stationary point. To characterize this point, we use the second derivative test and calculate the determinant of the Hessian

$$\begin{aligned} D(x, y) &= \frac{\partial^2 f_3}{\partial x^2} \frac{\partial^2 f_3}{\partial y^2} - \left(\frac{\partial^2 f_3}{\partial x \partial y} \right)^2 \\ &= 200(-400y + 1200x^2 - 2) - (-400x)^2. \end{aligned}$$

Since $D(0, 0) = -400 < 0$, the Hessian is indefinite, the point $(0, 0)$ is a saddle point, i.e. neither a local minimum nor a local maximum.

2. Consider the following optimization problem:

$$\begin{aligned} \underset{x_1,x_2}{\text{minimize}} \quad & x_1^2 + (x_2+1)^2 \\ \text{subject to} \quad & -1 \le x_1 \le 1, x_2 \ge 0 \end{aligned}$$

Use the optimality condition to show that the vector $(0,0)$ is a unique optimal solution. (15 points)

Solution:

By straightforward calculation we have

$$\nabla f(x_1,x_2) = (2x_1, 2x_2+2), \quad \nabla^2 f(x_1,x_2) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \succ 0$$

For any $\mathbf{y}$ satisfying $-1 \le y_1 \le 1, y_2 \ge 0$ and $\mathbf{x} = (0,0)$, we have

$$\nabla f(\mathbf{x})^\top (\mathbf{y} - \mathbf{x}) = 2y_2 \ge 0.$$

Therefore, $(0,0)$ is an optimal solution.
And by the strict convexity of f, $(0,0)$ is a unique optimal solution.

3. Consider the following constrained optimization problem:

$$\begin{aligned} \underset{x\in\mathbb{R}^n}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & x \in \mathcal{X}, \end{aligned}$$

where f is a convex and continuously differentiable function, and $\mathcal{X} \subseteq \mathbb{R}^n$ is a box constraint of the form

$$\mathcal{X} = \{x \in \mathbb{R}^n | a_i \le x_i \le b_i \text{ for all } i\},$$

for some scalars a_i and b_i. Using the optimality conditions verify that x^* is an optimal solution if and only if x^* satisfies the following relations for all $i = 1, \ldots, n$:

$$\begin{aligned} &\frac{\partial f(x^*)}{\partial x_i} \ge 0 \text{ if } x_i^* = a_i, \\ &\frac{\partial f(x^*)}{\partial x_i} = 0 \text{ if } a_i < x_i^* < b_i, \\ &\frac{\partial f(x^*)}{\partial x_i} \le 0 \text{ if } x_i^* = b_i. \end{aligned}$$

(15 points)

Solution:

$$x^* \text{ is an optimal solution} \iff \nabla f(x^*)(y - x^*) = \sum_{i=1}^n \frac{\partial f(x^*)}{\partial x_i}(y_i - x_i^*) \ge 0, \forall y \in \mathcal{X},$$

which is equivalent to the following conditions:
If $x_i^* = a_i$, then $y_i - x_i^* = y_i - a_i \ge 0$, which leads to $\frac{\partial f(x^*)}{\partial x_i} \ge 0$.
If $a_i < x_i^* < b_i$, then the sign of $y_i - x_i^*$ can not be determined, hence $\frac{\partial f(x^*)}{\partial x_i} = 0$ must hold.
If $x_i^* = b_i$, then $y_i - x_i^* = y_i - b_i \le 0$, which leads to $\frac{\partial f(x^*)}{\partial x_i} \le 0$.

4. Consider the problem

$$\begin{aligned} \underset{x\in\mathbb{R}^n}{\text{minimize}} \quad & \frac{\|Ax-b\|_1}{c^\top x + d} \\ \text{subject to} \quad & \|x\|_\infty \le 1, \end{aligned}$$

where $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m, c \in \mathbb{R}^n$, and $d \in \mathbb{R}$. We assume that $d > \|c\|_1$, which implies that $c^\top x + d > 0$ for all feasible x.

(1) Show that this is a quasi-convex optimization problem. (5 points)

(2) Show that it is equivalent to the convex optimization problem

$$\begin{array}{ll} \underset{y,t}{\text{minimize}} & \|Ay - bt\|_1 \\ \text{subject to} & \|y\|_\infty \leq t, \\ & c^\top y + dt = 1, \end{array}$$

with variables $y \in \mathbb{R}^n, t \in \mathbb{R}$. (10 points)

Solution:

(1) $f_0(x) \leq \alpha$ if and only if

$$\|Ax - b\|_1 - \alpha(c^\top x + d) \leq 0,$$

which is a convex constraint.

(2) Suppose $\|x\|_\infty \leq 1$. We have $c^\top x + d > 0$, because $d > \|c\|_1$. Define

$$y = \frac{x}{c^\top x + d}, \quad t = \frac{1}{c^\top x + d}.$$

Then y and t are feasible in the convex problem with objective value

$$\|Ay - bt\|_1 = \frac{\|Ax - b\|_1}{c^\top x + d}.$$

Conversely, suppose y, t are feasible for the convex problem. We must have $t > 0$, since $t = 0$ would imply $y = 0$, which contradicts $c^\top y + dt = 1$. Define

$$x = \frac{y}{t},$$

Then $\|x\|_\infty \leq 1$, and $c^\top x + d = \frac{1}{t}$, and hence

$$\frac{\|Ax - b\|_1}{c^\top x + d} = \|Ay - bt\|_1.$$

5. Consider the SDP

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & c^\top x \\ \text{subject to} & x_1 F_1 + x_2 F_2 + \cdots + x_n F_n + G \preceq 0, \end{array}$$

with $F_i, G \in \mathbb{S}^k, c \in \mathbb{R}^n$.

(1) Suppose $R \in \mathbb{R}^{k \times k}$ is nonsingular. Show that the SDP is equivalent to the SDP

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & c^\top x \\ \text{subject to} & x_1 \tilde{F}_1 + x_2 \tilde{F}_2 + \cdots + x_n \tilde{F}_n + \tilde{G} \preceq 0, \end{array}$$

where $\tilde{F}_i = R^\top F_i R, \tilde{G} = R^\top G R$. (10 points)

(2) Suppose there exists a nonsingular R such that $\tilde{F}_i$ and $\tilde{G}$ are diagonal. Show that the SDP is equivalent to an LP. (5 points)

(3) Suppose there exists a nonsingular R such that $\tilde{F}_i$ and $\tilde{G}$ have the form

$$\tilde{F}_i = \begin{bmatrix} \alpha_i I & a_i \\ a_i^\top & \alpha_i \end{bmatrix}, i = 1, \ldots, n, \quad \tilde{G} = \begin{bmatrix} \beta I & b \\ b & \beta \end{bmatrix},$$

where $\alpha_i, \beta \in \mathbb{R}$ and $a_i, b \in \mathbb{R}^{k-1}$. Show that the SDP is equivalent to an SOCP with a single second-order cone constraint. (10 points)

Solution:

(1) x is feasible iff we have

$$u^\top \left(\sum_{i=1}^{n} x_i F_i + G \right) u \leq 0, \forall u \in \mathbb{R}^k.$$

Since R is nonsingular, this is equivalent to the condition

$$(Ru)^\top \left(\sum_{i=1}^{n} x_i F_i + G \right) Ru \leq 0, \forall u \in \mathbb{R}^k,$$

which is the same as the condition

$$u^\top \left(\sum_{i=1}^{n} x_i \tilde{F}_i + \tilde{G} \right) u \leq 0, \forall u \in \mathbb{R}^k.$$

Since the feasible sets and objective functions of the two SDPs are the same, they are equivalent.

(2) Suppose $\tilde{G} = \mathrm{diag}(\tilde{g}_1, \cdots, \tilde{g}_k)$ and $\tilde{F} = \mathrm{diag}(\tilde{f}_{i1}, \cdots, \tilde{f}_{ik})$ for $1 \leq i \leq n$. Then x is feasible iff we have

$$\sum_{i=1}^{n} x_i \tilde{f}_{ij} + \tilde{g}_j \leq 0, \forall\, 1 \leq j \leq k.$$

This is because a diagonal matrix is positive semidefintie iff all its diagonal entries are nonnegative. Thus the SDP is equivalent to an LP.

(3) The LMI is equivalent to

$$\tilde{F}(x) = \begin{bmatrix} (\alpha^\top x + \beta) I & Ax + b \\ (Ax + b)^\top & (\alpha^\top x + \beta) I \end{bmatrix} \preceq 0,$$

where A has columns a_i.
Therefore, the constraint of the SDP is equivalent to $\|Ax + b\|_2 \leq \alpha^\top x + \beta$, which is a single second-order cone constraint.