Exercise 06

Saturday, November 18, 2023 9:38 PM

Exercise 20 (11 points)

Consider the optimization problem

$$\begin{array}{ll} \text{minimize} & (x-3)^2 + (y-1)^2 \\ \text{subject to} & y \geq x^2 + 1 \\ & y \leq -3x + 11 \end{array}$$

Write down the Lagrange function and the KKT conditions. Find all pairs $((x,y),\lambda)$ that satisfy the KKT conditions and find out which one(s) attain the minimum.

[Remark: You can go trough all the conditions and find out the points. Alternatively you can save yourself some work by checking whether the problem is convex, visually guessing the right subcase of the complementary slackness conditions and using what you learn on Tuesday: For a convex optimization problem the KKT conditions imply optimality.

At some point you might encounter the polynomial $4x^3 + 2x - 6 = (x-1)(4x^2+4x-6)$. You can use without further calculations that $x = 1$ is its only root.]

Feasibility: $g_1(x,y) = x^2 - y + 1 \leq 0$

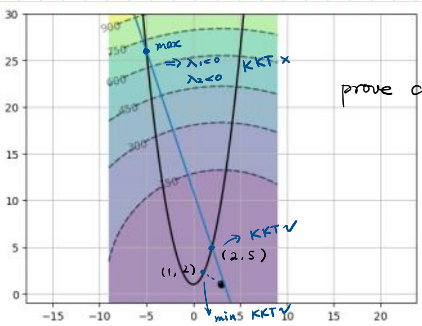
$g_2(x,y) = 3x + y - 11 \leq 0$

Dual feasibility: $\lambda_1 \geq 0$, $\lambda_2 \geq 0$

Lagrangian: $L(x, y, \lambda_1, \lambda_2)$

$= (x-3)^2 + (y-1)^2 + \lambda_1(x^2 - y + 1) + \lambda_2(3x + y - 11)$

$$\nabla_x L = \begin{pmatrix} 2(x-3) + 2\lambda_1 x + 3\lambda_2 \\ 2(y-1) - \lambda_1 + \lambda_2 \end{pmatrix} = \begin{pmatrix} 2x + 2\lambda_1 x + 3\lambda_2 - 6 \\ 2y - \lambda_1 + \lambda_2 - 2 \end{pmatrix}$$



prove convexity: $Hf(x) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \Rightarrow Hf$ is positive definite $\Rightarrow f$ is convex

$Hg_1(x) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \Rightarrow Hg_1$ is positive semi-definite $\Rightarrow g_1$ is convex

$g_2(x) = \underbrace{(3 \;\; 1)}_{A} \cdot \underbrace{\begin{pmatrix} x \\ y \end{pmatrix}}_{x} + \underbrace{(-11)}_{b}$ is affine $\Rightarrow g_2$ is convex

Since $\lambda_1 \geq 0, \lambda_2 \geq 0 \Rightarrow L$ is a non-negative combination of convex functions, therefore L is also convex.

We may find by the graph that, the minimum is achieved at $\lambda_1 > 0$, $\lambda_2 = 0$ (Case 1)

$$\Rightarrow \left. \begin{array}{l} g_1 = x^2 - y + 1 = 0 \\ \nabla_x L = \begin{pmatrix} 2x + 2\lambda_1 x - 6 \\ 2y - \lambda_1 - 2 \end{pmatrix} = 0 \end{array} \right\}$$

$\Rightarrow 2x + 2(2y-2)x - 6 = 0$

$2x + 2(2x^2)x - 6 = 0$

$4x^3 + 2x - 6 = 0$

$\Rightarrow x = 1$

$\Rightarrow y = 2$, $\lambda_1 = 2 > 0$

$\boxed{((1,2),2,0)}$ suffices the KKT condition

$\Rightarrow x = 1$

$\Rightarrow y = 2, \lambda_1 = 2 > 0$

$\boxed{((1,2),2,0)}$ surfices the KKT condition

Case 2: $\lambda_1 = \lambda_2 = 0$, according to the graph: $x > 0, y > 0$ achieves minimum

$$\left.\begin{array}{l} g_1 = x^2 - y + 1 = 0 \\ g_2 = 3x + y - 11 = 0 \end{array}\right\} \Rightarrow \begin{cases} x = 2 \\ y = 5 \end{cases} \Rightarrow \boxed{((2,5),0,0)} \text{ surfices the KKT condition}$$

Case 3: $\lambda_1 = 0, \lambda_2 > 0 \Rightarrow g_2 = 0, g_1 < 0$

on the line of $g_2 = 0$, ∇f descents along the line until $g_1 = 0$, to case 2

Case 4: $\lambda_1 > 0, \lambda_2 > 0$, $\Rightarrow$ minimum within the convex hull of g_1 and g_2

since the unconstrained minimum of f is outside the region,

there is no such a case.

Exercise 21 (12 points)

Consider the optimization problem

$$\begin{array}{rl} \text{minimize} & f(x,y) := -4xy + 3x^2 + 2x + 4y \\ \text{subject to} & (x-1)^2 \le 4 \\ & y \ge 0 \end{array}$$

$x = 1, y = 2$

$-1 \le x \le 3$

$-8 + 3 + 2 + 8 = 5$

(a) (2 points) Show that the set of feasible points is convex.

Let $P = \{(x,y) \in \mathbb{R}^2 \mid (x-1)^2 \le 4, y \ge 0\}$

P is the set of feasible points

Suppose $(x_1, y_1), (x_2, y_2) \in P$ are arbitrary.

for $\forall \theta \in [0,1]$,

① Certainly $(\theta x_1 + (1-\theta)x_2, \theta y_1 + (1-\theta)y_2) \in \mathbb{R}^2$ is in the domain of f

② $(\theta x_1 + (1-\theta)x_2 - 1)^2 = (\theta(x_1 - 1) + (1-\theta)(x_2 - 1))^2$

$= \theta^2 (x_1 - 1)^2 + (1-\theta)^2 (x_2 - 1)^2 + 2\theta(1-\theta)\underbrace{(x_1 - 1)}_{a}\underbrace{(x_2 - 1)}_{b}$ $\quad 2ab \le a^2 + b^2$

$\le \theta^2 (x_1 - 1)^2 + (1-\theta)^2 (x_2 - 1)^2 + \theta(1-\theta)[(x_1 - 1)^2 + (x_2 - 1)^2]$

$\le 4\theta^2 + 4(1-\theta)^2 + 8\theta(1-\theta)$

$= 4 \cdot (\theta + 1 - \theta)^2 = 4$

③ $$\left.\begin{array}{l} \theta > 0, 1-\theta \ge 0 \\ y_1 \ge 0, y_2 \ge 0 \end{array}\right\} \Rightarrow \theta y_1 + (1-\theta) y_2 \ge 0$$

So $\forall (x_1, y_1), (x_2, y_2) \in P$, $\theta(x_1, y_1) + (1-\theta)(x_2, y_2) \in P$

P is a convex set $\square$

(b) (1 point) Write down the Lagrangian and the KKT conditions.

Feasibility $\quad g_1(x,y) = (x-1)^2 - 4 = x^2 - 2x - 3 \le 0$

$(-4y + 6x + 2)$

Feasibility $\quad g_1(x,y) = (x-1)^2-4 = x^2-2x-3 \le 0$

$\quad g_2(x,y) = -y \le 0$

$$\nabla_x f(x) = \begin{pmatrix} -4y+6x+2 \\ -4x+4 \end{pmatrix} \quad Hf(x) = \begin{pmatrix} 6 & -4 \\ -4 & 0 \end{pmatrix}$$

f is not convex !

Dual Feasibility: $\lambda_1 \ge 0,\ \lambda_2 \ge 0$

Complementary Slackness: $\lambda_i g_i(x,y) = 0 \quad i=1,2$

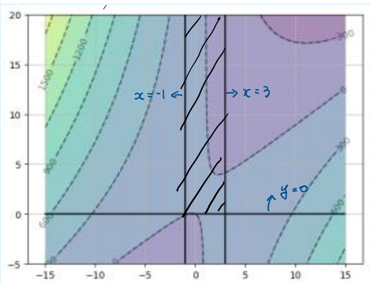
Lagrangian: $\mathcal{L}(x,y,\lambda_1,\lambda_2)$

$$= -4xy + 3x^2 + 2x + 4y + \lambda_1(x^2-2x-3) - \lambda_2 y$$

Gradient Condition:

$$\nabla_x \mathcal{L} = \begin{pmatrix} -4y+6x+2+2\lambda_1 x - 2\lambda_1 \\ -4x+4-\lambda_2 \end{pmatrix} = 0$$

(c) (4 points) Find all pairs $((x,y),\lambda)$ that satisfy the KKT conditions.



Case 1 $\lambda_1 = \lambda_2 = 0$

$$\nabla_x \mathcal{L} = \begin{pmatrix} -4y+6x+2 \\ -4x+4 \end{pmatrix} = 0$$

$$\Rightarrow \begin{cases} x=1 \\ y=2 \end{cases}$$

$((1,2),0,0)$ satisfies the KKT condition

Case 2: $\lambda_1 > 0,\ \lambda_2 = 0$

$g_1(x) = x^2-2x-3 = 0 \Rightarrow \begin{cases} x_1 = 3 \\ x_2 = -1 \end{cases}$

$(x-3)(x+1) = 0$

$$\nabla_x f = \begin{pmatrix} -4y+6x+2+2\lambda_1 x - 2\lambda_1 \\ -4x+4 \end{pmatrix} = 0$$

$-4x_1+4 = -8 \ne 0$! $\Rightarrow$ No points satisfies KKT condition

$-4x_2+4 = 8 \ne 0$ when $\lambda_1>0,\ \lambda_2=0$

Case 3: $\lambda_1 = 0,\ \lambda_2 > 0$

$g_2(x) = y = 0$

$$\nabla_x f = \begin{pmatrix} 6x+2 \\ -4x+4-\lambda_2 \end{pmatrix} = 0 \Rightarrow \begin{cases} x = -\frac{1}{3} \\ \lambda_2 = \frac{16}{3} > 0 \end{cases}$$

So $((-\frac{1}{3},0),0,\frac{16}{3})$ satisfies the KKT condition

Case 3: $\lambda_1>0,\ \lambda_2>0$

$$\begin{cases} x^2-2x-3=0 \\ y=0 \end{cases} \Rightarrow \begin{cases} x_1=3,\ y_1=0 \\ x_2=-1,\ y_2=0 \end{cases}$$

for $x_1=3,\ y_1=0$

$$\nabla_x f = \begin{pmatrix} 18+2+6\lambda_1-2\lambda_1 \\ -12+4-\lambda_2 \end{pmatrix} = 0 \quad \lambda_1 = -5 < 0 \quad \times$$

for $x_2=-1,\ y_2=0$

$$\nabla_x f = \begin{pmatrix} -6+2-2\lambda_1-2\lambda_1 \\ \end{pmatrix} = 0 \Rightarrow \lambda_1 = -1 > 0 \quad \times$$

for $x_2 = -1,\ y_2 = 0$

$$\nabla_x f = \begin{pmatrix} -6+2-2\lambda_1-2\lambda_1 \\ 4+4-\lambda_2 \end{pmatrix} = 0 \Rightarrow \lambda_1 = -1 > 0 \quad \times$$

(d) (1 points) Is one of the points (x, y) belonging to a KKT pair $((x, y), \lambda)$ an optimal point? Justify your answer.

For $\forall\, 1 \le x \le 3$, of course for $\forall\, y \ge 0$, $(x, y) \in \tilde{f}$

$f(x,y) = 4y(1-x) + 3x^2 + 2x$, with $1 - x \le 0$

$\Rightarrow \lim_{y\to\infty} f(x,y) = -\infty$, $f(x,y)$ has no lower bound even with feasibility constraints

$\Rightarrow p^* = -\infty$, infimum of $f(x,y)$ does not exists in the domain

No points in the KKT pair are optimal points.

e) (1 points) Is the objective function convex? Justify your answer.

[Hint: On Tuesday we will see that for a convex optimization problem the KKT conditions imply optimality.]

$$\nabla_x f(x) = \begin{pmatrix} -4y+6x+2 \\ -4x+4 \end{pmatrix} \quad Hf(x) = \begin{pmatrix} 6 & -4 \\ -4 & 0 \end{pmatrix}$$

$$\chi_\lambda(H_f) = \begin{pmatrix} \lambda-6 & 4 \\ 4 & \lambda \end{pmatrix} = \lambda^2 - 6\lambda - 16 = (\lambda-8)(\lambda+2)$$

$\lambda_{1,2} = -2, 8 \Rightarrow Hf(x)$ is not positive definite!

therefore $f(x)$ is not convex.

* There should be a saddle point

(f) (3 points) State the dual problem and show that it has no feasible points.

[Remark: The dual problem will be introduced on Tuesday. Dual feasibility also requires that the dual function $g(\lambda, \nu)$ is defined, i.e. that the infimum in question exists.]

The dual problem is:

maximize $g(\lambda, \nu) = \inf_{x\in\mathbb{R}^2} L(x, y, \lambda_1, \lambda_2)$

$$= \inf_{x\in\mathbb{R}} \left(-4xy + 3x^2 + 2x + 4y + \lambda_1(x^2-2x-3) - \lambda_2 y\right)$$

subject to $\lambda_1 \ge 0,\ \lambda_2 \ge 0$

From (b) we know for $\forall \lambda_1 \ge 0, \lambda_2 \ge 0$, $\exists\, x \in [1,3]$, $\lim_{y\to+\infty} L = -\infty$, $\Rightarrow g(\lambda, \nu) = \inf_{x\in\mathbb{R}} (f(x,y)) = -\infty$

$\Rightarrow g(\lambda, \nu)$ has no feasible points

Exercise 22 (9 points)

Let $A \in \mathbb{R}^{n\times m}$ and $b \in \mathbb{R}^n$. For the Euclidean norm $\| - \|_2$ on $\mathbb{R}^n$ consider the unconstrained optimization problem

$$\text{minimize} \quad \|Ax - b\|_2$$

Since this problem is unconstrained, the dual function is constant with value p^* (think about why this is true!), so the dual problem won't help us here.

Now consider the related problem

$$\begin{array}{ll} \text{minimize} & \|z\|_2^2 \\ \text{subject to} & Ax - b = z \end{array} \qquad \|z\|_2^2 = \sum z_i^2$$

[Remark: For this to make sense, you have to consider the objective function as a function $\mathbb{R}^{n+m} \to \mathbb{R}$, $(x_1, \ldots, x_m, z_1, \ldots, z_n) \mapsto \|z\|_2$ If $(x_1, \ldots, x_m, z_1, \ldots, z_n)$ is an optimal point for this new problem, the vector $(x_1, \ldots, x_m)$ will be an optimal point for the previous problem.]

(a) (6 points) Show that the Lagrange dual function for this problem is given by

$$g\colon \operatorname{dom} g \to \mathbb{R}, \qquad \nu \mapsto -\frac{1}{4}\|\nu\|_2^2 - b^T\nu$$

with $\operatorname{dom} g = \{\nu \mid \nu^T A = 0\} \subseteq \mathbb{R}^n$.

[Remark: The domain of the dual function consists of the points where the infimum in question exists. It is helpful here to write the summand of the Lagrangian coming from the constraints as $\nu^T(Ax - b - z)$ for a vector ν of the right size.]

(a) Lagrangian: $L(x, z, \nu) = \|z\|_2^2 + \nu^T(Ax - b - z) \qquad \nu = (\nu_1, \cdots, \nu_n)^T$

$= z^T I z + \nu^T(Ax - b - z)$

Feasibility: $h(x, z) = Ax - b - z = 0$

Gradient: $\frac{\partial L}{\partial x} = \nu^T A = 0$

$\frac{\partial L}{\partial z} = 2z - \nu = 0 \quad \Rightarrow \quad z = \frac{\nu}{2}$

since I is positive definite, $Ax - b - z$ is affine $\Rightarrow L$ is convex $\Rightarrow$ strong duality!

$(\|z\|_2^2 = z^T I z$ is convex)

For $\forall z \in \mathbb{R}^n$

$$\inf_{z \in \mathbb{R}^n} L(z, \nu) = \begin{cases} z^T I z + \nu^T(-b - z), & \nu^T A = 0, \ z = \frac{\nu}{2} \\ -\infty & \nu^T A \neq 0 \ (x \to -\infty) \end{cases}$$

this is also validated in the gradient condition, therefore $\operatorname{dom} g = \{\nu \in \mathbb{R}^n \mid \nu^T A = 0\}$

so $g(\nu)$

$$= \inf_{\substack{\nu \in \mathbb{R}^n \\ \nu^T A = 0}} L(z = \tfrac{\nu}{2}, \nu) = \frac{\nu^T}{2} \cdot I \cdot \frac{\nu}{2} + \nu^T\left(-b - \frac{\nu}{2}\right)$$

$$= \frac{1}{4}\|\nu\|_2^2 - \frac{1}{2}\|\nu\|_2^2 - \nu^T b$$

$$= -\frac{1}{4}\|\nu\|_2^2 - b^T\nu$$

(b) (3 points – slightly tricky, do at your own risk) Now consider the problem

$$\begin{array}{ll} \text{minimize} & \|z\|_2 \\ \text{subject to} & Ax - b = z \end{array}$$

Show that the Lagrange dual function for this problem is given by

$$g\colon \operatorname{dom} g \to \mathbb{R}, \qquad \nu \mapsto -b^T\nu$$

with $\operatorname{dom} g = \{\nu \mid \nu^T A = 0 \text{ and } \|\nu\| = 1\} \subseteq \mathbb{R}^n$

Lagrangian: $L(x, z, \nu) = \|z\|_2 + \nu^T(Ax - b - z) \qquad \nu = (\nu_1, \cdots, \nu_n)^T$

$= \sqrt{z^T z} + \nu^T(Ax - b - z)$

Feasibility: $h(x, z) = Ax - b - z = 0$

Gradient: $\frac{\partial L}{\partial x} = \nu^T A = 0, \quad \Rightarrow \nu^T A = 0$, proved in (a)

$\frac{\partial L}{\partial z} = \frac{z}{\sqrt{z^T z}} - \nu = 0 \quad \Rightarrow \quad \frac{z}{\sqrt{z^T z}} = \nu \qquad z = \nu\sqrt{z^T z}$

So $g(\nu) = \inf L(x, z, \nu) = \inf\left((1 - \nu^T\nu)\sqrt{z^T z} - \nu^T b\right)$, since $z \in \mathbb{R}^n$,

$$\Rightarrow \quad g(\nu) = \begin{cases} -\nu^T b & \text{for } \nu^T\nu = 1 \iff \|\nu\| = 1 \\ -\infty & \text{for } \nu^T\nu \neq 1 \text{ or} \end{cases}$$

The constraint can also be derived from gradient condition:

$$\nu^T\nu = \frac{z^T z}{z^T z} = 1 \iff \|\nu\| = 1$$

So the dual problem is $g(\nu) = -\nu^T b$, where $\operatorname{dom} g = \{\nu \in \mathbb{R}^n \mid \|\nu\| = 1,\ \nu^T A = 0\}$