ECE 490 (Introduction to Optimization) – Some Problems for Practice

Problem 1. Apply the optimality conditions to solve the following problems.

(a) Consider the unconstrained minimization problem

$$\text{minimize} \quad 2x^2 + 2xy + y^2 - 10x - 10y$$

Determine all the local mins for this problem.

(b) Consider the constrained minimization problem

$$\begin{aligned} \text{minimize} \quad & 2x^2 + 2xy + y^2 - 10x - 10y \\ \text{subject to} \quad & x^2 + y^2 = 5 \end{aligned}$$

Determine all the local mins for the above problem.

(c) Consider the constrained minimization problem

$$\begin{aligned} \text{minimize} \quad & x^2 + y^2 - 14x - 6y \\ \text{subject to} \quad & x + y \le 2 \\ & x + 2y \le 4 \end{aligned}$$

Determine all the local mins for the above problem.

Solutions

(a) The objective function can be rewritten as

$$f(x,y) = \frac{1}{2}\begin{bmatrix} x \\ y \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} 10 & 10 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Notice $\begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \succeq 0$. The objective function is convex. Setting the gradient to 0, we have

$$\begin{aligned} 4x + 2y &= 10 \\ 2x + 2y &= 10 \end{aligned}$$

Therefore, the only local (global) minimum is $(x,y) = (0,5)$.

(b) The Lagrangian function is given by $L(x,y,\lambda) = 2x^2 + 2xy + y^2 - 10x - 10y + \lambda(x^2 + y^2 - 5)$. By Lagrange multiplier theorem, we have

$$\begin{aligned} & 4x + 2y - 10 + 2\lambda x = 0 \\ & 2x + 2y - 10 + 2\lambda y = 0 \\ & x^2 + y^2 - 5 = 0 \end{aligned}$$

This is a set of nonlinear equations. We can use a numerical solver to obtain two solutions $(x,y,\lambda) = (1,2,1)$ and $(x,y,\lambda) = (-1.7246, -1.4233, -5.7246)$. Both points are regular since we have $\begin{bmatrix} 2x \\ 2y \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Now we check the second-order sufficient condition. For $(x,y,\lambda) = (1,2,1)$, the Hessian of L is given by

$$\begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix} + \lambda \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \succ 0$$

Therefore, the second-order sufficient condition is satisfied. Hence $(1, 2)$ is a local minimum.

For $(x, y, \lambda) = (-1.7246, -1.4233, -5.7246)$, the Hessian of L is given by

$$\begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix} + \lambda \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix} - 5.7246 \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \prec 0$$

Hence this point is not a local minimum.

(c) The Lagrangian function is given by $L(x, y, \mu) = x^2 + y^2 - 14x - 6y + \mu_1(x + y - 2) + \mu_2(x + 2y - 4)$. Now we can apply the KKT condition to get

$$\begin{aligned} &2x - 14 + \mu_1 + \mu_2 = 0 \\ &2y - 6 + \mu_1 + 2\mu_2 = 0 \\ &\mu_1 \geq 0, \quad \mu_2 \geq 0 \\ &\mu_1(x + y - 2) = 0 \\ &\mu_2(x + 2y - 4) = 0 \end{aligned}$$

Notice that $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ are linearly independent. Hence all the points are regular. There are four cases.

Case 1: Both inequality constraints are inactive. Then $\mu_1 = \mu_2 = 0$. We have $x = 7$ and $y = 3$. However, $x + y = 10 > 2$. This is not a feasible point.

Case 2: Only the first inequality constraint is inactive. We have $\mu_1 = 0$ but $\mu_2 > 0$. So $x + 2y = 4$. We can combine this with $2x - 14 + \mu_2 = 0$ and $2y - 6 + 2\mu_2 = 0$ to obtain $x = 5.2$, $y = -0.6$, and $\mu_2 = 3.6$. However, $x + y = 4.6 > 2$. This is not a feasible point.

Case 3: Only the second inequality constraint is inactive. We have $\mu_2 = 0$ but $\mu_1 > 0$. So $x + y = 2$. We can combine this with $2x - 14 + \mu_1 = 0$ and $2y - 6 + \mu_1 = 0$ to obtain $x = 3$, $y = -1$, and $\mu_1 = 8$. This is a feasible point since $x + 2y = 1 < 4$. The Hessian of the Lagrangian function is equal to $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \succ 0$. Hence this is a local minimum.

Case 4: Both inequality constraints are inactive. We have $\mu_1 > 0$ and $\mu_2 > 0$. We have $x + y - 2 = 0$ and $x + 2y - 4 = 0$. Hence we have $x = 0$ and $y = 2$. Then we have $\mu_1 + \mu_2 = 14$ and $\mu_1 + 2\mu_2 = 2$. This leads to $\mu_2 = -12 < 0$. This violates the KKT condition. Hence this case does not lead to any meaningful KKT point.

To summarize, the only local minimum is $(x, y) = (3, -1)$.

Problem 2. Consider the following constrained optimization problem

$$\begin{aligned} \text{minimize} \quad & \tfrac{1}{x_1} + \tfrac{1}{x_2} + \tfrac{1}{x_3} \\ \text{subject to} \quad & x_1 + x_3 \le 1 + \tfrac{1}{\sqrt{2}} \\ & x_2 + x_3 \le 1 + \tfrac{1}{\sqrt{2}} \\ & x_i > 0, \text{ for } i = 1, 2, 3 \end{aligned}$$

Use the KKT necessary conditions to find candidate points for the local minimum of this optimization problem. Then use the general sufficiency condition to find the global minimum for the optimization problem.

Solutions Define $g_1(x) = x_1 + x_3 - 1 - \frac{1}{\sqrt{2}}$ and $g_2(x) = x_2 + x_3 - 1 - \frac{1}{\sqrt{2}}$. We define the set S as $S = \{(x_1, x_2, x_3) : x_1 > 0, x_2 > 0, x_3 > 0\}$. Notice $\nabla g_1(x) = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}^\mathsf{T}$ and $\nabla g_2(x) = \begin{bmatrix} 0 & 1 & 1 \end{bmatrix}^\mathsf{T}$. Clearly, they are linearly independent, and all the points are regular.

Then we define the Lagrangian function as

$$L(x, \mu) = \frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \mu_1 g_1 + \mu_2 g_2.$$

From the KKT condition, we have

$$\begin{aligned} & -\frac{1}{x_1^2} + \mu_1 = 0 \\ & -\frac{1}{x_2^2} + \mu_2 = 0 \\ & -\frac{1}{x_3^2} + \mu_1 + \mu_2 = 0 \\ & \mu_1 \ge 0, \ \mu_2 \ge 0 \\ & \mu_1 g_1 = 0, \ \mu_2 g_2 = 0 \\ & x_1 > 0, x_2 > 0, x_3 > 0 \end{aligned}$$

From the first two equations, we know $\mu_1 > 0$ and $\mu_2 > 0$. Hence both inequality constraints are active. We have $g_1 = 0$ and $g_2 = 0$. Since $g_1 - g_2 = 0$, we have $x_1 = x_2$. Then we know $\mu_1 = \mu_2 = \frac{1}{x_1^2}$. Therefore, we have $x_3^2 = \frac{1}{2} x_1^2$. This leads to our final solution $x_1 = 1$, $x_2 = 1$, $x_3 = \frac{1}{\sqrt{2}}$, $\mu_1 = 1$, and $\mu_2 = 1$.

To verify the global optimality, notice that the Hessian of the cost function is given by

$$\begin{bmatrix} \frac{2}{x_1^3} & 0 & 0 \\ 0 & \frac{2}{x_2^3} & 0 \\ 0 & 0 & \frac{2}{x_3^3} \end{bmatrix} \succ 0, \text{ for } (x_1, x_2, x_3) \in S$$

Therefore, the cost function is strictly convex over S. This further implies that the Lagrangian $L(x, \mu)$ is strictly convex for $x \in S$ since the constraint functions g_i are linear. Therefore the point we found is the unique global minimum of $L(x, \mu)$ on S. By the general sufficiency condition we can conclude that the point $(1, 1, \frac{1}{\sqrt{2}})$ is the unique global minimum for the optimization problem.

Problem 3. Duality:

(a) Consider the standard linear programming problem

$$\begin{array}{rl} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array}$$

Show that the dual of the dual for the above problem is just the primal problem itself.

(b) Consider the following problem

$$\begin{array}{rl} \text{minimize} & x^2 + 2xy + y^2 \\ \text{subject to} & x^2 = 1 \\ & y^2 = 1 \end{array}$$

where x and y are scalar decision variables. What is the dual problem for the above problem?

(c) Consider the following problem

$$\begin{array}{rl} \text{minimize} & x^2 + y^2 \\ \text{subject to} & 1 - x - y - z \leq 0 \\ & 1 - x - 2y - z \leq 0 \\ & 1 - 2x - y + z \leq 0 \end{array}$$

where x, y, and z are scalar decision variables. What is the dual problem for the above problem?

Solution

(a) To formulate the dual problem, we first write out the Lagrangian:

$$L(x, \lambda, \mu) = c^\top x + \lambda^\top (Ax - b) + \mu^\top (-x) = \left(c^\top + \lambda^\top A - \mu^\top\right) x - \lambda^\top b$$

We have

$$D(\lambda, \mu) = \min_{x \in \mathbb{R}^p} L(x, \lambda, \mu) = \begin{cases} -\lambda^\top b & \text{if } c^\top + \lambda^\top A - \mu^\top = 0 \\ -\infty & \text{Otherwise} \end{cases}$$

Therefore, the dual problem is

$$\begin{array}{rl} \text{maximize} & -b^\top \lambda \\ \text{subject to} & c + A^\top \lambda - \mu = 0 \\ & \mu \geq 0 \end{array}$$

Notice we can eliminate μ by using the relation $\mu = c + A^\top \lambda$. The dual problem is then compactly rewritten as

$$\begin{array}{rl} \text{minimize} & b^\top \lambda \\ \text{subject to} & c + A^\top \lambda \geq 0 \end{array} \tag{1}$$

Now we try to derive the dual problem for (1). We denote the Lagrangian multiplier for this dual problem as x. Since the constraint in (1) is an inequality, we need $x \geq 0$. The dual function for (1) is equal to

$$\min_{\lambda} \left(b^\top \lambda - x^\top (c + A^\top \lambda)\right) = \begin{cases} -x^\top c & \text{if } b^\top - x^\top A^\top = 0 \\ -\infty & \text{Otherwise} \end{cases}$$

Therefore, the dual problem for (1) is

$$\begin{array}{rl} \text{maximize} & -x^\top c \\ \text{subject to} & b^\top - x^\top A^\top = 0 \\ & x \geq 0 \end{array}$$

which is equivalent to

$$\begin{array}{rl} \text{minimize} & c^\top x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array}$$

Therefore, the dual of the dual is the primal problem itself.

(b) We define the Lagrangian function as $L = x^2 + 2xy + y^2 + \lambda_1(x^2 - 1) + \lambda_2(y^2 - 1)$. We have

$$L = \begin{bmatrix} x \\ y \end{bmatrix}^\top \begin{bmatrix} 1+\lambda_1 & 1 \\ 1 & 1+\lambda_2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} - \lambda_1 - \lambda_2.$$

If $\begin{bmatrix} 1+\lambda_1 & 1 \\ 1 & 1+\lambda_2 \end{bmatrix}$ is positive semidefinite, then the minimum of L is achieved by choosing $x = y = 0$. Otherwise, we can choose $\begin{bmatrix} x \\ y \end{bmatrix}$ to be the eigenvector associated with the negative eigenvalue of $\begin{bmatrix} 1+\lambda_1 & 1 \\ 1 & 1+\lambda_2 \end{bmatrix}$ and drive the value of L to $-\infty$. Therefore, the dual function can be calculated as

$$D(\lambda_1, \lambda_2) = \min_{x,y} L(x, y, \lambda_1, \lambda_2) = \begin{cases} -\lambda_1 - \lambda_2 & \text{if } \begin{bmatrix} 1+\lambda_1 & 1 \\ 1 & 1+\lambda_2 \end{bmatrix} \succeq 0 \\ -\infty & \text{Otherwise} \end{cases}$$

The dual problem is

$$\begin{aligned} &\text{maximize} && -\lambda_1 - \lambda_2 \\ &\text{subject to} && \begin{bmatrix} 1+\lambda_1 & 1 \\ 1 & 1+\lambda_2 \end{bmatrix} \succeq 0 \end{aligned}$$

(c) We define the Lagrangian function as $L = x^2 + y^2 + \mu_1(1 - x - y - z) + \mu_2(1 - x - 2y - z) + \mu_3(1 - 2x - y + z)$. We require $\mu_i \geq 0$ for $i = 1, 2, 3$. If $\mu_1 + \mu_2 - \mu_3 \neq 0$, we can choose z to drive L to $-\infty$. If $\mu_1 + \mu_2 - \mu_3 = 0$, we have

$$L = x^2 + y^2 + 2\mu_3 - 3\mu_3 x - (\mu_1 + 2\mu_2 + \mu_3)y$$

which is strongly convex in (x, y). Setting gradient to 0, we get $2x - 3\mu_3 = 0$ and $2y = \mu_1 + 2\mu_2 + \mu_3 = \mu_2 + 2\mu_3$. Therefore, the dual function is

$$D(\mu_1, \mu_2, \mu_3) = \min_{x,y,z} L = \begin{cases} -\frac{9}{4}\mu_3^2 - \frac{1}{4}(\mu_2 + 2\mu_3)^2 + 2\mu_3 & \text{if } \mu_1 + \mu_2 - \mu_3 = 0 \\ -\infty & \text{Otherwise} \end{cases}$$

When $\mu_1 + \mu_2 - \mu_3 = 0$, the value of D actually does not depend on μ_1. Therefore, the dual problem is

$$\begin{aligned} &\text{maximize} && -\tfrac{9}{4}\mu_3^2 - \tfrac{1}{4}(\mu_2 + 2\mu_3)^2 + 2\mu_3 \\ &\text{subject to} && \mu_2 \geq 0,\ \mu_3 \geq 0. \end{aligned}$$

Problem 4. Consider the constrained minimization problem:

$$\begin{aligned} \text{minimize} \quad & x^2 + y^2 + 2z^2 \\ \text{subject to} \quad & x - 1 \geq 0 \\ & y + 1 \geq 0 \\ & z \geq 0 \end{aligned}$$

What is the optimal solution for this problem?

Now apply the barrier function method that iterates as

$$\begin{bmatrix} x_k \\ y_k \\ z_k \end{bmatrix} = \operatorname*{argmin}_{x,y,z} \left\{ x^2 + y^2 + 2z^2 - \epsilon_k \ln(x-1) - \epsilon_k \ln(y+1) - \epsilon_k \ln(z) \right\}$$

where ϵ_k decreases to 0 as k increases. Does the above barrier function converge to the optimal solution of the original problem? Prove your conclusion.

Solution Notice $x^2 \geq 1$ for any $x \geq 1$. We have $x^2 + y^2 + 2z^2 \geq 1 + 0 + 0 = 1$. The lower bound on the right side can be achieved using $x = 1$, $y = 0$, and $z = 0$. Notice this is a feasible point. Hence the global solution is provide by this point. If $x \neq 1$, then we have $x^2 > 1$. If $y \neq 0$ or $z \neq 0$, we will also have $x^2 + y^2 + 2z^2 > 1$. This means that $(1, 0, 0)$ is the unique global minimum for this problem. This problem is relatively simple so there is no need to use the KKT condition.

Now we apply the barrier method. For each k, setting the gradient to be 0 leads to

$$\begin{aligned} 2x - \frac{\epsilon_k}{x-1} &= 0 \\ 2y - \frac{\epsilon_k}{y+1} &= 0 \\ 4z - \frac{\epsilon_k}{z} &= 0 \end{aligned}$$

Since $x \geq 1$ and $2x^2 - 2x - \epsilon_k = 0$, we have $x = \frac{1+\sqrt{1+2\epsilon_k}}{2}$.

Since $y \geq -1$ and $2y^2 + 2y - \epsilon_k = 0$, we have $y = \frac{-1+\sqrt{1+2\epsilon_k}}{2}$.

Since $z \geq 0$ and $4z^2 - \epsilon_k = 0$, we have $z = \frac{\sqrt{\epsilon_k}}{2}$.

We have

$$\lim_{\epsilon_k \to 0} \begin{bmatrix} \frac{1+\sqrt{1+2\epsilon_k}}{2} \\ \frac{-1+\sqrt{1+2\epsilon_k}}{2} \\ \frac{\sqrt{\epsilon_k}}{2} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Therefore, the barrier method converges to the optimal solution of the original problem.

Problem 5. True or False. Provide reasons.

(a) Consider two closed intervals $C_1 \subset \mathbb{R}_+$ and $C_2 \subset \mathbb{R}_+$. Then $C_3 = \{xy : x \in C_1, y \in C_2\}$ is a convex set.

(b) Consider the point $(y, s) \in R^n \times \mathbb{R}_+$ with $\|y\| \geq s$. Then the projection of (y, s) on the set $\{(x, t) : \|x\| \leq t\}$ is $(s\frac{y}{\|y\|}, s)$.

(c) A nonconvex optimization problem can have zero duality gap.

(d) The subgradient of a convex function always gives a descent direction.

Solution

(a) False. Suppose $C_1 = [1, 1.01]$, and $C_2 = [10, 11] \cup [1000, 2000]$. Choose $x_1 = 1 \in C_1$ and $y_1 = 10 \in C_2$, we have $x_1 y_1 = 10 \in C_3$. Choose $x_2 = 1 \in C_1$ and $y_2 = 1000 \in C_2$, we have $x_2 y_2 = 2000 \in C_3$. The line segment between 10 and 2000 clearly contains some points which are not in C_3. For example, we know 20 is not in C_3.

(b) False. Suppose y is a scalar. Let's project $(2, 1)$ to the set $\{(x, t) : -t \leq x \leq t\}$. The resultant point is $(1.5, 1.5)$, not $(1, 1)$!

(c) True. Suppose we want to minimize $-x^2$ subject to the constraint $x^2 + x \leq 2$. Clearly the objective function is concave and hence not convex. The optimal value for the primal problem is achieved by $x = -2$. The optimal value is -4. The dual function is given by

$$D(\mu) = \min_x(-x^2 + \mu(x^2 + x - 2)) = \min_x(\mu - 1)x^2 + \mu x - 2\mu$$

If $0 \leq \mu \leq 1$, then $D(\mu) = -\infty$. If $\mu > 1$, we have $D(\mu) = -\frac{\mu^2}{4(\mu-1)} - 2\mu$. For $\mu > 1$, this is a concave function. The maximum value is achieved at $\mu = \frac{4}{3}$, and the maximum value of the dual function is -4. So there is zero duality gap.

(d) False. Consider $f(x) = |x|$. At $x = 0$, we know -1 is a subgradient. However, we have $|t| > 0$ for any non-zero t. This means that the subgradient -1 not a descent direction.