CS5321 Numerical Optimization Homework 3

Annabella Putri Dirgo - 109006238

1. (20%) In the trust region method (unit 3), we need to solve the model problem m_k

$$\min_{\vec{p}} m_k(\vec{p}) = f_k + \vec{g}_k^T \vec{p} + \frac{1}{2}\vec{p}^{\,T} B_k \vec{p}.$$

$$\text{s.t.} \|\vec{p}\| \leq \Delta$$

Show that $\vec{p}*$ is the optimal solution if and only if it satisfies

$$(B_k + \lambda I)\vec{p}* = -\vec{g}$$

$$\lambda(\Delta - \|\vec{p}*\|) = 0$$

where $B_k + \lambda I$ is positive definite. (Hint: using KKT conditions.)

Answer:

2. (15%) Prove that for the matrix $\begin{bmatrix} G & A^T \\ A & 0 \end{bmatrix}$, if A has full row-rank and the reduced Hessian $Z^T G Z$ is positive definite, where span$\{Z\}$ is the null space of span$\{A^T\}$ then the matrix is nonsingular. (You may reference Lemma 16.1 in the textbook.)

Answer: Proof. Suppose there are vectors w and v such that

$$\begin{bmatrix} G & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} w \\ v \end{bmatrix} = 0$$

Since $Aw = 0$, we have from (16.8) that

$$0 = \begin{bmatrix} w \\ v \end{bmatrix}^T \begin{bmatrix} G & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} w \\ v \end{bmatrix} = w^T G w$$

Since w lies in the null space of A, it can be written as $w = Zu$ for some vector $u \in R^{n-m}$. Therefore, we have

$$0 = w^T G w = u^T Z^T G Z u$$

which by positive definiteness of $Z^T G Z$ implies that $u = 0$. Therefore, $w = 0$, and by vector above, $A^T v = 0$. Full row rank of A then implies that $v = 0$. We conclude that equation is satisfied only if $w = 0$ and $v = 0$, so the matrix is nonsingular, as claimed.

3. (30%) Consider the problem

$$\begin{aligned} \min_{x_1, x_2} \quad & (x_1 - 3)^2 + 10x_2^2 \\ \text{s.t.} \quad & x_1^2 + x_2^2 - 1 \leq 0 \end{aligned} \tag{1}$$

(a) Write down the KKT conditions for (1).

(b) Solve the KKT conditions and find the optimal solutions, including the Lagrangian parameters.

(c) Compute the reduced Hessian and check the second order conditions for the solution.

Answer:

(a) We will write the constraint as $g_1(x_1, x_2) = x_1^2 + x_2^2 - 1 \leq 0$. Thus the KKT conditions could be written as:

i. $\Rightarrow 2(x_1 - 3) + 2x_1\lambda_1 = 0$
$\Rightarrow 20x_2 + 2x_2\lambda_1 = 0$

ii. $\lambda_1(x_1^2 + x_2^2 - 1) = 0$

iii. $x_1^2 + x_2^2 - 1 \leq 0$

iv. $\lambda_1, \lambda_2 \geq 0$

(b) The form of Lagrangian: $L(x, \lambda) = (10x_2^2 + (x_1 - 3)^2) + \lambda(x_1^2 + x_2^2 - 1)$
The first-order partial derivatives:

$$\frac{\partial}{\partial x_1} = ((10x_2^2 + (x_1 - 3)^2) + \lambda(x_1^2 + x_2^2 - 1)) = 2\lambda x_1 + 2x_1 - 6$$

$$\frac{\partial}{\partial x_2} = ((10x_2^2 + (x_1 - 3)^2) + \lambda(x_1^2 + x_2^2 - 1)) = 2x_2(\lambda + 10)$$

$$\frac{\partial}{\partial \lambda} = ((10x_2^2 + (x_1 - 3)^2) + \lambda(x_1^2 + x_2^2 - 1)) = x_1^2 + x_2^2 - 1$$

Next we want to solve the system:

$$\begin{cases} \frac{\partial L}{\partial x_1} = 0 \\ \frac{\partial L}{\partial x_2} = 0, \text{ or} \\ \frac{\partial L}{\partial \lambda} = 0 \end{cases} \quad \begin{cases} 2\lambda x_1 + 2x_1 - 6 = 0 \\ 2x_2(\lambda + 10) = 0 \\ x_1^2 + x_2^2 - 1 = 0 \end{cases}$$

The system has the following solution: $(x_1, x_2) = (-1, 0), (x_1, x_2) = (1, 0), (x_1, x_2) = (\frac{-1}{3}, \frac{-2\sqrt{2}}{3}), (x_1, x_2) = (\frac{-1}{3}, \frac{2\sqrt{2}}{3})$
$f(-1, 0) = 16$
$f(1, 0) = 4$
$f(\frac{-1}{3}, \frac{-2\sqrt{2}}{3}) = 20$
$f(\frac{-1}{3}, \frac{2\sqrt{2}}{3}) = 20$
Thus, $\lambda = 2$ and $(x_1, x_2) = (1, 0)$

(c) The bordered Hessian formula looked as

$$|H| = \begin{bmatrix} 0 & g_x & g_y \\ g_x & L_{xx} & L_{xy} \\ g_y & L_{yx} & L_{yy} \end{bmatrix}$$

We annotate x_1 as x and x_2 as y and we have,
$g_x = 2$
$g_y = 2$
$L_{xx} = 2$
$L_{xy} = 0$

$L_{yx} = 0$
$L_{yy} = 20$
Hence,

$$|H| = \begin{bmatrix} 0 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 20 \end{bmatrix}$$

and the second order condition is

$$|H_2| = \begin{bmatrix} 2 & 0 \\ 0 & 20 \end{bmatrix}$$

If we compute it's determinant we get $|H_2| = 0(40-0) - 2(40-0) - 2(0-4) = 0 - 80 + 8 = -72 < 0$ satisfies the sufficient condition for relative minimum.

4. (20%) Consider the following constrained optimization problem

$$\begin{aligned} \min_{x_1, x_2} \quad & x_1^3 + 2x_2^2 \\ \text{s.t.} \quad & x_1^2 + x_2^2 - 1 = 0 \end{aligned}$$

(a) What is the optimal solution and the optimal Lagrangian multiplier?
(b) Formulate this problem to the equation of augmented Lagrangian method, and derive the gradient of Lagrangian.
(c) Let $\rho_0 = -1, \mu_0 = 1$. What is x_1 if it is solved by the augmented Lagrangian problem.
(d) To make the solution of augmented Lagrangian method exact, what is the minimum ρ should be?

Answer:

(a) Form of the Lagrangian $L(x, \lambda) = (x_1^3 + 2x_2^2) + \lambda(x_1^2 + x_2^2 - 1)$ The first order partial derivative:

$$\frac{\partial}{\partial x_1} = ((x_1^3 + 2x_2^2) + \lambda(x_1^2 + x_2^2 - 1)) = x_1(2\lambda + 3x_1)$$
$$\frac{\partial}{\partial x_2} = ((x_1^3 + 2x_2^2) + \lambda(x_1^2 + x_2^2 - 1)) = 2x_2(\lambda + 2)$$
$$\frac{\partial}{\partial x_2} = ((x_1^3 + 2x_2^2) + \lambda(x_1^2 + x_2^2 - 1)) = x_1^2 + x_2^2 - 1$$

Next, we will solve the system

$$\begin{cases} \frac{\partial L}{\partial x_1} = 0 \\ \frac{\partial L}{\partial x_2} = 0, \\ \frac{\partial L}{\partial \lambda} = 0 \end{cases} \text{or} \quad \begin{cases} x_1(2\lambda + 3) = 0 \\ 2x_2(\lambda + 2) = 0 \\ x_1^2 + x_2^2 - 1 = 0 \end{cases}$$

The system has the following real solutions: $(x, y) = (-1, 0), (x, y) = (0, -1), (x - y) = (0, 1), (x, y) = (1, 0)$ $f(-1, 0) = -1$
$f(0, -1) = 2$
$f(0, 1) = 2$
$f(1, 0) = 1$
Thus, it will reach its minimum value when $x^* = (-1, 0)$ and $\lambda^* = \frac{3}{-2}$

(b) Recall that we have Lagrangian formula: $L(x, \lambda) = (x_1^3 + 2x_2^2) + \lambda(x_1^2 + x_2^2 - 1)$

We have the Lagrangian $L(x, \lambda) = (x_1^3 + 2x_2^2) + \lambda(x_1^2 + x_2^2 - 1)$, $x^* = (-1, 0)$, and $\lambda^* = \frac{3}{-2}$

Thus, the augmented Lagrangian method is

$$L_a(x, \lambda) = (x_1^3 + 2x_2^2) + \lambda(x_1^2 + x_2^2 - 1) + \frac{\mu}{2}(x_1^2 + x_2^2 - 1)^2$$

Gradient of augmented Lagrangian derivation

$$\nabla_x L_a(x, \lambda) = \nabla f(x) + \nabla h(x)(\lambda + \mu h(x))$$

Thus, we have

$$\nabla_x L_{aug}(x, \lambda) = (3x_1 + 4x_2) + (2x_1 + 2x_2)(\lambda + \mu(x_1^2 + x_2^2 - 1))$$

(c) With respect to Augmented Lagrangian, we need to have a strict local minimum correspond x^* and λ^*. If we have $\rho_0 = -1$ and $\mu_0 = 1$. Then we have x_1 as:

$$L_a(x, \lambda) = (x_1 + 0) + \frac{3}{-2}(x - 1^2 + 0 - 1) + \frac{1}{2}(-1^2 + 0 - 1)^2$$

5. (15%) Consider the following constrained optimization problem

$$\begin{aligned} \min_{x_1, x_2} \quad & -3x_1 + x_2 \\ \text{s.t.} \quad & 2x_1 + x_2 \leq 20 \\ & x_1 + 2x_2 \leq 16 \\ & x_1, x_2 \geq 0 \end{aligned}$$

Formulate this problem to the equation of the interior point method, and derive the gradient and Jacobian.

Answer:

(a) This inequality-constrained optimization problem is then solved by converting it into an unconstrained objective function whose minimum we hope to find efficient. Thus,we use the Frisch barrier function to yield the unconstrained problem:

$$\begin{aligned} \min_{x_1, x_2} \quad & -3x_1 + x_2 \\ \text{s.t.} \quad & -20 - 2x_1 + x_2 - x_3 = 0 \\ & -16 - x_1 + 2x_2 - x_4 = 0 \\ & x_1, x_2 \geq 0 \\ & x_3, x_4 \geq 0 \end{aligned}$$

(b) derive the gradient

$$\nabla_x C \begin{bmatrix} -x_2 & -2 \\ -2 & -x_1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix}$$

derive the Jacobian

$$f(x,y) = \left[\begin{array}{c} -2x - y - 20 \\ -x - 2y - 16 \end{array} \right]$$

$$f_1(x,y) = -2x - y - 20$$

$$f_2(x,y) = -x - 2y - 16$$

$$J_f(x,y) = \left[\begin{array}{cc} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{array} \right]$$

$$J_f(x,y) = \left[\begin{array}{cc} -2 & -1 \\ -1 & -2 \end{array} \right]$$