1 More on Inequality Constraints (June 9)

As before, we are working on an open $\Omega \subseteq \mathbb{R}^n$, and we want to optimize f subject to $h_1, \ldots, h_k = 0$ and $g_1, \ldots, g_l \leq 0$. The smoothness of our functions varies.

1.1 Second Order Conditions

One might guess that the second order conditions under inequality constraints will be the same thing as before. However, the tangent space on which we evaluate the positive-definiteness of the Lagrangian is slightly different (in a very obvious way).

Theorem 1.1. *Suppose $f, h_1, \ldots, h_k, g_1, \ldots, g_k \in C^2(\Omega)$, where $\Omega \subseteq \mathbb{R}^n$. Suppose x_0 is a regular point of the constraints. If x_0 is a local minimizer of f subject to the constraints, then*

(i) There are $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ and $\mu_1, \ldots, \mu_l \geq 0$ such that

$$\nabla f(x_0) + \sum_i \lambda_i \nabla h_i(x_0) + \sum_j \mu_j \nabla g_j(x_0) = 0,$$

and $\mu_j g_j(x_0) = 0$ for each j.

(ii) The matrix

$$L(x_0) = \nabla^2 f(x_0) + \sum_i \lambda_i \nabla^2 h_i(x_0) + \sum_j \mu_j \nabla^2 g_j(x_0)$$

is positive semi-definite on the tangent space $T_{x_0}\widetilde{M}$ to the active constraints at x_0. (Explicitly, $L(x_0)$ is positive semi-definite on the space

$$T_{x_0}\widetilde{M} = \{v \in \mathbb{R}^n : \nabla h_i(x_0) \cdot v = 0 \text{ for all } i, \text{ and } \nabla g_j(x_0) \cdot v = 0 \text{ for all } 1 \leq j \leq l'\},$$

where the active g constraints are indexed precisely by $1, \ldots, l'$.)

Proof. x_0 is a local minimizer of f subject to the constraints, so it is also a local minimizer of f subject to only the active constraints. Since the Lagrange multipliers of the inactive constraints are zero, our theory of equality-constrained minimization finishes the problem. □

1. Consider, for example, the problem

$$\begin{aligned} \text{minimize } & f(x,y) := -x \\ \text{subject to } & g_1(x,y) := x^2 + y^2 \leq 1 \\ & g_2(x,y) := y + x - 1 \leq 0 \end{aligned}$$

The feasible set is the closed unit ball $\overline{B_1(0)}$ with an open semicircle removed from the top right. Geometrically, it is clear that the minimizer should be the point $(1, 0)$. It is not hard to check that every feasible point is regular. Let's check that $(x_0, y_0) = (1, 0)$ satisfies the first order conditions. We look at

$$\nabla f(x_0, y_0) + \mu_1 \nabla g_1(x_0, y_0) + \mu_2 \nabla g_2(x_0, y_0) = (0, 0).$$

This becomes

$$\begin{pmatrix} -1 \\ 0 \end{pmatrix} + \mu_1 \begin{pmatrix} 2 \\ 0 \end{pmatrix} + \mu_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

or

$$\begin{aligned} 2\mu_1 + \mu_2 &= 1 \\ \mu_2 &= 0. \end{aligned}$$

So $\mu_1 = 1/2$. Also, $g_1(1, 0) = 1^2 + 0^2 - 1 = 0$ and $g_(1, 0) = 0$ as well, so the complementary slackness conditions are satisfied. Therefore $(1, 0)$ satisfies the Kuhn-Tucker conditions, and so it is a candidate local minimizer. What about the second order conditions?

$$L(1, 0) = \nabla^2 f(x_0, y_0) + \mu_1 \nabla^2 g_1(x_0, y_0) + \mu_2 \nabla^2 g_2(x_0, y_0),$$

or

$$L(1, 0) = I.$$

Clearly the second order necessary conditions are satisfied, but let's check the tangent space anyway. We have $\nabla g_1(1, 0) = (2, 0)$ and $\nabla g_2(1, 0) = (1, 1)$; they are linearly independent, so the tangent space is a point. Therefore the second order necessary conditions are satisfied.

2. Consider the problem

$$\begin{aligned} \text{minimize } f(x, y) &:= 2x^2 + 2xy + y^2 - 10x - 10y \\ \text{subject to } g_1(x, y) &= x^2 + y^2 - 5 \leq 0 \\ g_2(x, y) &:= 3x + y - 6 \leq 0 \end{aligned}$$

The Kuhn-Tucker conditions are

$$\begin{pmatrix} 4x + 2y - 10 \\ 2x + 2y - 10 \end{pmatrix} + \mu_1 \begin{pmatrix} 2x \\ 2y \end{pmatrix} + \mu_2 \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

as well as $\mu_1, \mu_2 \geq 0$ and $\mu_1(x^2 + y^2 - 5) = 0$ and $\mu_2(3x + y - 6) = 0$. We consider four cases:

(i) Suppose g_1 is inactive and g_2 is active. Then $\mu_1 = 0$. The equations become

$$\begin{aligned} 4x + 2y - 10 + 3\mu_2 &= 0 \\ 2x + 2y - 10 + \mu_2 &= 0 \\ \mu_2(3x + y - 6) &= 0 \end{aligned}$$

Subtracting the second equation from the first gives $x = -\mu_2$. If $\mu_2 = 0$, then $x = 0$, which implies that $y = 6$ since the second constraint is active. We get the point $(0, 6)$; but this does not satisfy the constraints. Therefore $\mu_2 \neq 0$. (After some work one can conclude that no such point here satisfies the constraints.)

(ii) Suppose g_1 is active and g_2 is inactive. Then

$$\begin{aligned} 4x + 2y - 10 + 2\mu_1 x &= 0 \\ 2x + 2y - 10 + 2\mu_1 y &= 0 \\ \mu_1(x^2 + y^2 - 5) &= 0 \\ \mu_1 &\geq 0 \end{aligned}$$

The solution is $(1, 2)$ and $\mu_1 = 1$. It is not hard to see that this point is regular. Therefore the point $(1, 2)$ is a candidate. The Lagrangian is, after some work,

$$L(1,2) = \begin{pmatrix} 6 & 2 \\ 2 & 4 \end{pmatrix}.$$

This matrix is clearly positive definite, so we conclude that the second order necessary (and, as we'll see later, sufficient) conditions are satisfied.

(iii) Suppose g_1 and g_2 are active.

(iv) And so on. (This problem was not completed during lecture.)

1.2 Second Order Sufficient Conditions

Theorem 1.2. *Suppose $f, h_1, \ldots, h_k, g_1, \ldots, g_k \in C^2(\Omega)$, where $\Omega \subseteq \mathbb{R}^n$ is open. Suppose that x_0 is feasible. If*

1. *There exist $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ and $\mu_1, \ldots, \mu_l \geq 0$ such that*
$$\nabla f(x_0) + \sum_i \lambda_i \nabla h_i(x_0) + \sum_j \mu_j \nabla g_j(x_0) = 0,$$
2. *$\mu_j g_j(x_0) = 0$ for each j.*
3. *The matrix*
$$L(x_0) = \nabla^2 f(x_0) + \sum_i \lambda_i \nabla^2 h_i(x_0) + \sum_j \mu_j \nabla^2 g_j(x_0)$$
is positive definite on the tangent space to the "strongly active constraints" at x_0. That is, it is positive definite on the space
$$\widetilde{\widetilde{T_{x_0}}} = \{v \in \mathbb{R}^n : \nabla h_i(x_0) = 0 \text{ for all } i, \text{ and } \nabla g_j(x_0) = 0 \text{ for all } 1 \leq k \leq l''\},$$
where $\{1, \ldots, l''\}$ is the set of all indices of active constraints whose Lagrange multipliers are positive.

Then x_0 is a strict local minimizer of f subject to the usual constraints.

Proof. Will be given on Thursday. (Copypaste it here?) $\square$

Let's consider some more examples.

1. Here's an example. Given (a, b) with $a, b > 0$ and $a^2 + b^2 > 1$. Consider the minimization problem:

$$\begin{aligned} \text{minimize } & f(x, y) := (x-a)^2 + (y-b)^2 \\ \text{subject to } & g_1(x, y) := x^2 + y^2 - 1 \leq 0 \end{aligned}$$

Our intuition says that the minimizer should be $\left(\frac{a}{\sqrt{a^2+b^2}}, \frac{b}{\sqrt{a^2+b^2}}\right)$. We have $\nabla g(x, y) = (2x, 2y)$, so clearly all feasible points are regular. The Kuhn-Tucker conditions are

$$\begin{pmatrix} 2(x-a) \\ 2(y-a) \end{pmatrix} + \mu \begin{pmatrix} 2x \\ 2y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

and $\mu g(x, y) = 0$. That is,

$$\begin{aligned} (1+\mu)x &= a \\ (1+\mu)y &= b \\ \mu(x^2 + y^2 - 1) &= 0, \mu \geq 0 \end{aligned}$$

Suppose $\mu = 0$. Then $x = a$ and $y = b$; since we assumed $a^2 + b^2 > 1$, we would have that (x, y) is not feasible. Therefore $\mu \neq 0$, and so $x^2 + y^2 = 1$ by the third equation. Squaring the first two equations and adding them gives

$$(1+\mu)^2(x^2 + y^2) = a^2 + b^2,$$

implying that $\mu = -1 + \sqrt{a^2 + b^2}$ - we took the positive root because $\mu > 0$. This is actually positive, since $a^2 + b^2 > 1$. Those first equations again give us

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \frac{1}{1+\mu} \begin{pmatrix} a \\ b \end{pmatrix} = \frac{1}{\sqrt{a^2 + b^2}} \begin{pmatrix} a \\ b \end{pmatrix},$$

as expected. What do the second order conditions tell us? The Lagrangian is

$$L(x_0, y_0) = 2I + 2\mu I = 2(1+\mu)I = 2\sqrt{a^2 + b^2} I,$$

which is everywhere positive definite. Therefore the second-order sufficient conditions are satisfied. For practice, however, let's compute the tangent space to the "strongly active constraints". The only constraint is g; since g is active and its Lagrange multiplier μ is positive, the constraint g is strongly active at (x_0, y_0). Therefore the tangent space we are interested in is the tangent space to S^1 at (x_0, y_0): that space is $\{v \in \mathbb{R}^2 : av_1 + bv_2 = 0\}$.

2. Consider the problem

$$\begin{aligned} \text{minimize } & f(x,y) := x^3 + y^2 \\ \text{subject to } & g(x,y) := (x+1)^2 + y^2 - 1 \leq 0. \end{aligned}$$

We have $\nabla g(x,y) = (2(x+1), 2y)$, which makes it clear that every feasible point is regular. The Kuhn-Tucker conditions are

$$\begin{pmatrix} 3x^2 \\ 2y \end{pmatrix} + \mu \begin{pmatrix} 2(x+1) \\ 2y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

with $\mu((x+1)^2 + y^2 - 1) = 0$ and $\mu \geq 0$.

Consider $(x_0, y_0) = (0,0)$. The Kuhn-Tucker conditions imply $\mu = 0$. In particular, g is active at $(0,0)$, but not strongly active there. The tangent space to the active constraint at $(0,0)$ is the y-axis. The Lagrangian at $(0,0)$ is

$$L(0,0) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix},$$

which is clearly positive definite on this tangent space. However, we cannot conclude anything, since the constraint g is not strongly active. In fact, it is clear that $(0,0)$ is not a local minimizer: for $x < 0$ sufficiently close to 0, $f(x,0)$ is negative, yet it is 0 at $(0,0)$.