1 Unconstrained Finite-Dimensional Optimization (May 19)

1.1 First Order Necessary Condition

Our main problem is

$$\min_{x\in\Omega} f(x) \qquad f : \mathbb{R}^n \supset \Omega \to \mathbb{R},$$

where Ω is one of the following three types:

- $\Omega = \mathbb{R}^n$.
- Ω open.
- Ω the closure of an open set.

We can consider minimization problems without any loss of generality, since any maximization problem can be converted to a minimization problem by taking the negative of the function in question: that is,

$$\max_{x\in\Omega} f(x) = \min_{x\in\Omega} -f(x).$$

Definition 1. *Given $\Omega \subset \mathbb{R}^n$ and a point $x_0 \in \Omega$, we say that the vector $v \in \mathbb{R}^n$ is a feasible direction at x_0 if there is an $\overline{s} > 0$ such that $x_0 + sv \in \Omega$ for all $s \in [0, \overline{s}]$.*

Theorem 1.1. *(First order necessary condition for a local minimum, or FONC) Let $f : \mathbb{R}^n \supset \Omega \to \mathbb{R}$ be C^1. If $x_0 \in \Omega$ is a local minimizer of f, then $\nabla f(x_0) \cdot v \geq 0$ for all feasible directions v at x_0.*

First we deduce a familiar case of the theorem - the one we know from second-year calculus.

Corollary 1.1.1. *If $f : \mathbb{R}^n \supset \Omega \to \mathbb{R}$ is C^1 and x_0 is a local minimizer of f in the interior of Ω, then $\nabla f(x_0) = 0$.*

Proof. If x_0 is an interior point of Ω, then all directions at x_0 are feasible. In particular, for any such v, we have $\nabla f(x_0) \cdot (v) \geq 0$ and $\nabla f(x_0) \cdot (-v) \geq 0$, which implies $\nabla f(x_0) = 0$ as all directions are feasible at x_0. $\square$

Now we prove the theorem.

Proof. Reduce to a single-variable problem by defining $g(s) = f(x_0 + sv)$, where $s \geq 0$. Then 0 is a local minimizer of g. Taylor's theorem gives us

$$g(s) - g(0) = sg'(0) + o(s) = s\nabla f(x_0) \cdot v + o(s).$$

If $\nabla f(x_0) \cdot v < 0$, then for sufficiently small s the right side is negative. This implies that $g(s) < g(0)$ for those s, a contradiction. Therefore $\nabla f(x_0) \cdot v \geq 0$. $\square$

1.2 Examples of using the FONC

1. Consider the problem
$$\min_{x\in\Omega} f(x,y) = x^2 - xy + y^2 - 3y \qquad \text{over } \Omega = \mathbb{R}^2.$$
By the corollary to the FONC, we want to find the points (x_0, y_0) where $\nabla f(x_0, y_0) = 0$. We have
$$\nabla f(x,y) = (2x - y, -x + 2y - 3),$$
so we want to solve
$$\begin{aligned} 2x - y &= 0 \\ -x + 2y &= 3, \end{aligned}$$
which has solution $(x_0, y_0) = (1, 2)$. Therefore $(1, 2)$ is the only *candidate* for a local minimizer. That is, if the function f has a local minimizer in $\mathbb{R}^2$, then it must be $(1, 2)$.
It turns out that $(1, 2)$ is a global minimizer for f on $\Omega = \mathbb{R}^2$. By some work, we have
$$f(x,y) = \left(x - \frac{y}{2}\right)^2 + \frac{3}{4}(y-2)^2 - 3.$$
In this form, it is obvious that a *global* minimizer occurs at the point where the squared terms are zero, if such a point exists. That point is $(1, 2)$.

2. Consider the problem
$$\min_{x\in\Omega} f(x,y) = x^2 - x + y + xy \qquad \text{over } \Omega = \{(x,y) \in \mathbb{R}^2 : x, y \geq 0\}.$$
We have
$$\nabla f(x,y) = (2x + y - 1, x + 1).$$
To apply the FONC, we'll divide the feasible set Ω into four different regions. Suppose that (x_0, y_0) is a local minimizer of f on Ω.
   (i) (x_0, y_0) is an interior point:
   By the corollary to the FONC, we must have $\nabla f(x_0, y_0) = 0$. Then $x_0 = -1$, which is not in the interior of Ω. This case fails.
   (ii) (x_0, y_0) on the positive x-axis:
   Then we are considering $(x_0, 0)$. The feasible directions at $(x_0, 0)$ are those vectors $v \in \mathbb{R}^2$ with $v_2 \geq 0$. The FONC tells us that $\nabla f(x_0, 0) \cdot v \geq 0$ for all feasible directions v. We then have
$$(2x_0 - 1)v_1 + (x_0 + 1)v_2 \geq 0$$
   for all v_1 and all $v_2 \geq 0$. In particular, this holds for $v_2 = 0$, so $(2x_0 - 1)v_1 \geq 0$ for all v_1, implying $x_0 = 1/2$. Therefore $(1/2, 0)$ is a candidate for a local minimizer of f on Ω - this is the only candidate for a local minimizer of f on the positive x-axis.

(iii) (x_0, y_0) on the positive y-axis:
Then we are considering $(0, y_0)$. The feasible directions here are $v \in \mathbb{R}^2$ with $v_1 \geq 0$. Then we have
$$(y_0 - 1)v_1 + v_2 \geq 0$$
for any v_2 and $v_1 \geq 0$. This is a contradiction if we take $v_1 = 0$, so f has no local minimizers along the positive y-axis.

(iv) (x_0, y_0) is the origin:
Then we are considering $(0, 0)$. The feasible directions here are $v \in \mathbb{R}^2$ with $v_1, v_2 \geq 0$. Then we have
$$-v_1 + v_2 \geq 0$$
for all $v_1, v_2 \geq 0$, a contradiction. Therefore the origin is not a local minimizer of f.

We conclude that the only candidate for a local minimizer of f is $(1/2, 0)$. It turns out that this is actually a global minimizer of f on Ω. (This is to be seen.)

1.3 Second Order Necessary Condition

Theorem 1.2. *(Second order necessary condition for a local minimum, or SONC) Let $f : \mathbb{R}^n \supset \Omega \to \mathbb{R}$ be C^2. If $x_0 \in \Omega$ is a local minimizer of f, then for any feasible direction v at x_0 the following conditions hold:*

(i) $\nabla f(x_0) \cdot v \geq 0$.

(ii) If $\nabla f(x_0) \cdot v = 0$, *then* $v^T \nabla^2 f(x_0) v \geq 0$.

Proof. Fix a feasible direction v at x_0. Then $f(x_0) \leq f(x_0 + sv)$ for sufficiently small s. By Taylor's theorem,
$$f(x_0 + sv) = f(x_0) + s\nabla f(x_0) + \frac{1}{2}s^2 v^T \nabla^2 f(x_0)v + o(s^2),$$
so by the FONC,
$$f(x_0 + sv) - f(x_0) = \frac{1}{2}s^2 v^T \nabla^2 f(x_0)v + o(s^2).$$
If $v^T \nabla^2 f(x_0)v < 0$, then for sufficiently small s the right side is negative, implying that $f(x_0 + sv) < f(x_0)$ for such s, which contradicts local minimality of $f(x_0)$. Therefore $v^T \nabla^2 f(x_0) \geq 0$. □

Corollary 1.2.1. *If $f : \mathbb{R}^n \supset \Omega \to \mathbb{R}$ is C^2 and x_0 is a local minimizer of f in the interior of Ω, then the following conditions hold:*

(i) $\nabla f(x_0) = 0$.

(ii) $\nabla^2 f(x_0)$ *is positive semidefinite.*

1.4 Sylvester's Criterion

Here's a useful criterion for determining when a matrix is positive definite or positive semidefinite.

Definition 2. *A principal minor of a square matrix A is the determinant of a submatrix of A obtained by removing any k rows and the corresponding k columns, $k \geq 0$. A leading principal minor of A is the determinant of a submatrix obtained by removing the last k rows and k columns of A, $k \geq 0$.*

Theorem 1.3. *(Sylvester's criterion for positive definite self-adjoint matrices) If A is a self-adjoint matrix, then $A \succ 0$ if and only if all of the leading principal minors of A are positive.*

Theorem 1.4. *(Sylvester's criterion for positive semidefinite self-adjoint matrices) If A is a self-adjoint matrix, then $A \succeq 0$ if and only if all of the principal minors of A are non-negative.*

1.5 Examples of using the SONC

1. Consider the problem

$$\min_{x \in \Omega} f(x,y) = x^2 - xy + y^2 - 3y \qquad \text{over } \Omega = \mathbb{R}^2.$$

Recall that $(1,2)$ was the only candidate for a local minimizer of f on Ω. We now check that the SONC holds. Since $(1,2)$ is an interior point of Ω, we must have $\nabla^2 f(1,2) \succeq 0$. We have

$$\nabla^2 f(1,2) = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

All of the leading principal minors of $\nabla^2 f(1,2)$ are positive, so $(1,2)$ satisfies the SONC by Sylvester's criterion.

2. Consider the problem

$$\min_{x \in \Omega} f(x,y) = x^2 - x + y + xy \qquad \text{over } \Omega = \{(x,y) \in \mathbb{R}^2 : x, y \geq 0\}.$$

Recall that $(1/2, 0)$ was the only candidate for a local minizer of f. We have

$$\nabla^2 f(1/2, 0) = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}.$$

To satisfy the SONC, we must have

$$v^T \nabla^2 f(1/2, 0) v \geq 0$$

for all feasible directions v at $(1/2, 0)$ such that $\nabla f(1/2, 0) \cdot v = 0$. We have

$$\nabla f(1/2, 0) = (0, 3/2),$$

so if $v = (v_1, 0)$, then v is a feasible direction at $(1/2, 0)$ with $\nabla f(1, 2, 0) \cdot v = 0$. Then

$$v^T \nabla^2 f(1/2, 0) v = \begin{pmatrix} v_1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ 0 \end{pmatrix} = \begin{pmatrix} v_1 & 0 \end{pmatrix} \begin{pmatrix} 2v_1 \\ v_1 \end{pmatrix} = 2v_1^2 \geq 0.$$

So the SONC is satisfied.

1.6 Completing the Square

Let A be a symmetric positive definite $n \times n$ matrix. Our problem is

$$\min_{x \in \Omega} f(x) = \frac{1}{2} x^T A x - b \cdot x \qquad \text{over } \Omega = \mathbb{R}^n.$$

The FONC tells us that if x_0 is a local minimizer of f, then since x_0 is an interior point, $\nabla f(x_0) = 0$. We thus have $Ax_0 = b$, so since A is invertible (positive eigenvalues), $x_0 = A^{-1}b$. Therefore $x_0 = A^{-1}b$ is the *unique* candidate for a local minimizer of f on Ω.

The SONC then tells us that $\nabla^2 f(x_0) = A$, so that $\nabla^2 f(x_0) \succ 0$, implying that $x_0 = A^{-1}b$ is a candidate for a local minimizer of f on Ω.

In fact, the candidate x_0 is a global minimizer. Why? We will "complete the square". We can write

$$f(x) = \frac{1}{2} x^T A x - b \cdot x = \frac{1}{2}(x - x_0)^T A (x - x_0) - \frac{1}{2} x_0^T A x_0;$$

this relies on symmetry. (Long rearranging of terms.) In this form it is obvious that x_0 is a global minimizer of f over Ω.