MATH 4091/5091
SPRING 2011

LECTURE 17

NON LINEAR OPTIMIZATION WITH EQUALITY CONSTRAINTS (PART III)

EXAMPLES FOR FIRST ORDER OPTIMALITY CONDITIONS

(1)

$$\underset{x \in \mathbb{R}^2}{\text{minimize}} \quad \underbrace{x_1 + x_2}_{f(x)}$$

subject to

$$\underbrace{2x_1^2 - x_2 = 0}_{c(x)}$$

Note

$$\nabla f(x) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad J(x) = [4x_1 \ \ -1]$$

Recall that for this example constraint qualification holds.

First Order Optimality Condition

x_* is a local minimizer

$$\Longrightarrow$$

(i) $c(x_*) = 2(x_*)_1^2 - (x_*)_2 = 0$

and

(ii) $\nabla f(x_*) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \lambda J(x_*)^T = \lambda \begin{bmatrix} 4(x_*)_1 \\ -1 \end{bmatrix} \ \exists \lambda$

From (ii)

$* \ \lambda = -1$

$* \ -4(x_*)_1 = 1 \Longrightarrow (x_*)_1 = -1/4$

From (i)

$* \ 2(-1/4)^2 - (x_*)_2 = 0 \Longrightarrow (x_*)_2 = 1/8$

The point $x_* = \begin{bmatrix} -1/4 \\ 1/8 \end{bmatrix}$ is the unique local minimizer.

REMARK

* The point x_* satisfying

  (i) $\nabla f(x_*) = J(x_*)^T \lambda$ for some λ

  (ii) $c(x_*) = 0$

  does not have to be a local minimizer in general.

* For the example we know from the level-sets that problem has a local minimizer. Since there is only one x_* satisfying the necessary conditions, it must be the unique local minimizer.

② (LINEAR EQUALITY-CONSTRAINED PROGRAM)

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad \underbrace{c^T x}_{f(x)}$$

subject to

$$\underbrace{Ax = b}_{c(x) = Ax - b}$$

where

$$c \in \mathbb{R}^n, \quad b \in \mathbb{R}^m \text{ and } A \in \mathbb{R}^{m \times n}$$ (with [redacted] $m \leq n$)

Objective gradient and constraint Jacobian

$$\nabla f(x) = c \quad \text{and} \quad J(x) = A$$

For linear constraints constraint qualification always holds. (Soon as we will see.)

x_* is a local minimizer

$$\Longrightarrow$$

(i) $c(x_*) = 0 \Leftrightarrow Ax_* = b$

and

(ii) $\nabla f(x_*) = c = J(x_*)^T \lambda$ for some λ

$$= A^T \lambda$$

REMARK

(*) $A^T\lambda = c$ for some λ

is independent of x_*

Consequently

(i) if (*) holds for some λ all feasible points are local minimizers,

(ii) otherwise there are no local minimizers.

③

$$\underset{x\in\mathbb{R}^4}{\text{minimize}} \quad 3x_1 - x_2 + 2x_3 + x_4$$

subject to

$$x_1 + x_2 + x_3 + x_4 = 1$$

$$2x_1 - 2x_2 + x_3 = 3$$

A particular equality-constrained linear program with

$$c = \begin{bmatrix} 3 \\ -1 \\ 2 \\ 1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \text{ and } A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & -2 & 1 & 0 \end{bmatrix}$$

Since

$$J(x_*)^T \lambda = c \iff A^T \lambda = c$$

$$\iff \begin{bmatrix} 1 & 2 \\ 1 & -2 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} \lambda = \begin{bmatrix} 3 \\ -1 \\ 2 \\ 1 \end{bmatrix}$$

$$\iff \lambda = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

all feasible points x satisfying

$$Ax = b \iff \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & -2 & 1 & 0 \end{bmatrix} x = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

are local minimizers.

THE LAGRANGIAN

Suppose

* x_* is a local minimizer, and
* the constraint qualification holds at x_*.

Then

(i) $c_j(x_*) = 0 \quad j = 1, \ldots, m$

(ii) $\nabla f(x_*) = J(x_*)^T \lambda$ for some $\lambda \in \mathbb{R}^m$

- Lagrangian function

$$\mathcal{L}(x,\lambda) = f(x) - \sum_{j=1}^{m} \lambda_j c_j(x)$$

$$= f(x) - \begin{bmatrix} \lambda_1 & \dots & \lambda_m \end{bmatrix} \begin{bmatrix} c_1(x) \\ \vdots \\ c_m(x) \end{bmatrix}$$

$$= f(x) - \lambda^T c(x)$$

Derivative of the Lagrangian function

$$\nabla \mathcal{L}(x,\lambda) = \begin{bmatrix} \nabla_x \mathcal{L}(x,\lambda) \\ \nabla_\lambda \mathcal{L}(x,\lambda) \end{bmatrix}$$

$$= \begin{bmatrix} \nabla f(x) - \sum_{j=1}^{m} \lambda_j \nabla c_j(x) \\ -c(x) \end{bmatrix}$$

$$= \begin{bmatrix} \nabla f(x) - \begin{bmatrix} \nabla c_1(x) & \dots & \nabla c_m(x) \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{bmatrix} \\ -c(x) \end{bmatrix}$$

$$= \begin{bmatrix} \nabla f(x) - J(x)^T \lambda \\ -c(x) \end{bmatrix}$$

FIRST ORDER OPTIMALITY CONDITIONS IN TERMS OF LAGRANGIAN

x_* is a local minimizer

$\Longrightarrow$

There exists a λ_* such that (x_*, λ_*) is a stationary point of $\mathcal{L}(x,y)$, i.e.

EXAMPLES

① minimize $x_1 + x_2$

$2x_1^2 - x_2 = 0$

$$\mathcal{L}(x,\lambda) = (x_1 + x_2) - \lambda\,(2x_1^2 - x_2)$$

$$\nabla\mathcal{L}(x,\lambda) = \begin{bmatrix} 1 - \lambda 4x_1 \\ 1 + \lambda \\ x_2 - 2x_1^2 \end{bmatrix}$$

The only stationary point of the Lagrangian is $\left(-\frac{1}{4}, \frac{1}{8}, -1\right)$. Indeed $(-1/4, 1/8)$ is the unique local minimizer.

② minimize $c^T x$

$x \in \mathbb{R}^n$

subject to

$\underbrace{A}_{m \times n} x = \underbrace{b}_{\mathbb{R}^m}$

$$\mathcal{L}(x,\lambda) = c^T x - \lambda^T (Ax - b)$$

$$\nabla\mathcal{L}(x,\lambda) = \begin{bmatrix} c - A^T\lambda \\ b - Ax \end{bmatrix}$$

Any stationary point (x_*, λ_*) satisfies

(i) $A^T\lambda_* = c$ and (ii) $Ax_* = b$