

NON LINEAR OPTIMIZATION WITH
EQUALITY CONSTRAINTS

(PART III)

EXAMPLES FOR FIRST ORDER OPTIMALITY
CONDITIONS

①

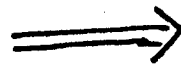
minimize $x_1 + x_2$
 $x \in \mathbb{R}^2$
subject to $2x_1^2 - x_2 = 0$

Note

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

Recall that for this example
constraint qualification holds.

First Order Optimality Condition
 x_* is a local minimizer



(i) $c(x_*) = 2(x_{*1})^2 - (x_{*2}) = 0$
and

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

From (ii)

$$* \lambda = -1$$

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

From (i)

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

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)

$$\begin{array}{l} \text{minimize} \\ x \in \mathbb{R}^n \\ \text{subject to} \end{array} \quad \underbrace{c^T x}_{f(x)}$$

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

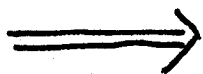
$$c \in \mathbb{R}^n, \quad b \in \mathbb{R}^m \quad \text{and} \quad A \in \mathbb{R}^{m \times n} \quad \left(\begin{array}{l} \text{with} \\ \text{---} \\ m \leq n \end{array} \right)$$

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



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

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

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

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

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

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

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

Lagrangian function

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

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

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

Derivative of the Lagrangian function

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

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

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

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

FIRST ORDER OPTIMALITY CONDITIONS
IN TERMS OF LAGRANGIAN

x_* is a local minimizer

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

EXAMPLES

$$\textcircled{1} \quad \begin{array}{l} \text{minimize } x_1 + x_2 \\ 2x_1^2 - x_2 = 0 \end{array}$$

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

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

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

$$\textcircled{2} \quad \begin{array}{l} \text{minimize } c^T x \\ x \in \mathbb{R}^n \\ \text{subject to} \\ \underbrace{Ax}_{m \times n} = \underbrace{b}_{\mathbb{R}^m} \end{array}$$

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

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

Any stationary point (x_*, λ_*) satisfies
(i) $A^T \lambda_* = c$ and (ii) $Ax_* = b$