

OR review:

Consider a simple example

$$\begin{aligned} & \min x \\ & \text{s.t. } x^2 \leq 0 \end{aligned}$$

Since it's impossible to find a solution for

$1 + \lambda \cdot 0 = 0$, we get:

Observation 1:

For a system $\min f(x)$

$$\text{s.t. } h_i(x) \leq 0, i=1 \dots m.$$

$$g_j(x) = 0, j=1 \dots p.$$

We get: sometimes we can't always write this system's optimal solution (even local solution) as "KKT" form.

To explore further, consider:

① First John (eq).

Let x^* be a local optimal solution of the problem

$$\begin{aligned} & \min f(x) \\ & \text{s.t. } h_i(x) = 0, i=1 \dots m. \end{aligned} \quad \left. \begin{array}{l} \text{differentiable and } C^1 \end{array} \right\}$$

Then there exists $\lambda_0 \geq 0, \lambda_i, i=1 \dots m$ not all zero, s.t.

$$0 = \lambda_0 \nabla f(x^*) + \sum_{i=1}^m \lambda_i \nabla h_i(x^*).$$

Proof: Consider $\nabla h_i(x^*)$, $i=1 \dots m$, we have:

① if $\nabla h_i(x^*)$, $i=1 \dots m$ are linearly dependent, then take $\lambda_0 = 0$.

② if $\nabla h_i(x^*)$, $i=1 \dots m$ are linearly independent, then

$\nabla h_i(x^*)$ and $\nabla f(x^*)$ are l. d. get it.

i.e. $\nabla h_i(x^*)$, $\nabla f(x^*)$ are l. id.

Then, consider:

$$f(x) = f(x^*) - \varepsilon$$

$$h_i(x) = h_i(x^*)$$

By implicit function theorem, for small enough ε , we have:

$$f(x_\varepsilon) = f(x^*) - \varepsilon \Rightarrow \text{Contradiction with the local minima}$$

$$h_i(x_\varepsilon) = h_i(x^*)$$

, $i=1 \dots m$. (let $\varepsilon \rightarrow 0^+$)

(Notice that at $(x, \varepsilon) = (x^*, 0)$, it admits a solution).

So this motivates a corollary:

Corollary: let x^* be a local optima. of the problem, s.t.
 $\min f(x)$, s.t. $h_i(x) = 0$, $i=1 \dots m$. (C^1 for all functions)

If $\nabla h_1(x^*) \wedge \nabla h_m(x^*)$ are linearly independent then there exists
Lagrange multipliers λ_i , $i=1 \dots m$, s.t

$$0 = \lambda_0 \triangleright f(x^*) + \sum_{i=1}^m \lambda_i \triangleright g_i(x^*) \text{ with } \lambda_0 \neq 0.$$

Now, it's natural to discuss. inequality form:

② (First John, ineq) Suppose that x^* is a local optimal solution of : $\min f(x)$, s.t. $g_i(x) \leq 0, i=1 \dots m$.
Here f, g are differentiable functions

Then there exists $\lambda_i, i=0 \dots m$ not all zero, s.t.

$$\begin{cases} 0 = \lambda_0 \triangleright f(x^*) + \sum_{i=1}^m \lambda_i \triangleright g_i(x^*) \\ 0 = \lambda_i g_i(x^*), \lambda_i \geq 0, i \geq 0 \text{ a.m.} \end{cases}$$

Proof: Firstly we show that:

Lemma: $\left\{ \begin{array}{l} Bx < 0 \\ \exists y \geq 0, y \neq 0, y^T B = 0 \end{array} \right. \begin{array}{l} \text{(i)} \\ \text{(ii)} \end{array} \right\}$ exactly one holds.

$$\begin{aligned} \text{pf: } Bx < 0 &\Leftrightarrow (B, I) \begin{pmatrix} x \\ z \end{pmatrix} \leq 0 \text{ for } z > 0 \quad (0^T y^T) \begin{pmatrix} x \\ z \end{pmatrix} \geq 0 \\ &\Leftrightarrow (B, I) \begin{pmatrix} x \\ z \end{pmatrix} \leq 0 \text{ for } y^T z > 0, \forall y \neq 0, y \geq 0. \end{aligned}$$

$$\begin{matrix} \text{Farkas} \\ \Leftrightarrow \end{matrix} \begin{pmatrix} B^T \\ I \end{pmatrix} x = \begin{pmatrix} 0 \\ y \end{pmatrix} \text{ has no solution for } \forall y \geq 0, y \neq 0.$$

Notice that $\begin{cases} \triangleright f(x^*)^T d < 0 \\ \triangleright g_i(x^*)^T d < 0, g_i \text{ is active} \end{cases}$ has no solution.

Hence, by lemma, we get our result.

(if g_i isn't active at x^* , assign $\lambda_i=0$)

However, notice that, we may have a bad situation here, i.e.

$\lambda_0=0$, which means we don't have any information about $f(x)$. We denote this situation by "abnormal"

This leads to the following definition.

(No nonzero abnormal constraint qualification NNAMC α / NNAM α_0)
We say NNAMC α holds at x^* if there's no nonzero abnormal multipliers.

It's equivalent to:

(Positive Linear Independence Constraint Qualification)

$$\left. \begin{array}{l} 0 = \sum_{i=1}^p \lambda_i g_i(x^*) \\ 0 = \lambda_i g_i(x^*), i=m-p \\ \lambda_i \geq 0, i \neq m-p \end{array} \right\} \Rightarrow \lambda_1 = \lambda_2 = \dots = \lambda_p = 0$$

You may wonder what's the p, m here, consider the following:

(KKT under NNAMC α / PLI α).

Suppose that x^* is a local optima of:

$$\min f(x), \text{s.t. } \begin{cases} g_i(x) = 0, i \in \{1, \dots, m-1\} \\ g_i(x) \leq 0, i = m-p. \end{cases} \quad \left. \begin{array}{l} \text{all differentiable} \end{array} \right\}$$

If NNA/MCO./PLICO. holds at x^* , then : $\exists \lambda_i, i \geq n$ which aren't all zero, s.t.

$$0 = \nabla f(x^*) + \sum_{i=1}^p \lambda_i \nabla g_i(x^*)$$

$$\lambda_i \geq 0, 0 = \lambda_i g_i(x^*), i \geq n.$$

Proof: Consider any eq can be written as 2 ineq.
Then use First John's combined form. Then Use our CO.

■

Before we start, we propose a lemma.

Lemma: For nonvacuous matrix A_{pxn} and matrix D_{mxn} ,

- (i) $Ax \leq 0, Dx = 0, x \in \mathbb{R}^n$
 - (ii) $A^T y + D^T z = 0, y \geq 0, y \neq 0$
- } exactly one holds.

L.proof:

$$\text{(i)} \Leftrightarrow \begin{pmatrix} A & I \\ D & 0 \\ -D & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \text{ has solution with } (0^T, y^T) \begin{pmatrix} x \\ y \\ z \end{pmatrix} \geq 0 \text{ for all } y \geq 0, y \neq 0$$

$$\Leftrightarrow \begin{pmatrix} A^T & D^T & -D^T \\ I & 0 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix} \text{ has no solution.}$$

$$\Leftrightarrow A^T U + D^T (V - W) = 0, U \geq 0 \text{ doesn't have solution, s.t. } U \geq 0, U \neq 0.$$

And we propose :

MFCQ: consider the problem above: we say it's Mangasarian Fromovitz constraint qualification iff

① $\nabla g_1(x^*) \sim \nabla g_{m-1}(x^*)$ are l.i.

② $I(x^*) = \{i = m-n-p \mid g_i(x^*) = 0\}$.

Then $\{\nabla g_i(x^*)\}^\top d = 0, i=1 \dots m-1$ has a solution d .
 $(\nabla g_i(x^*))^\top d \leq 0, i \in I(x^*)$.

By the lemma above, you can see that ② is equivalent to:

$\exists y, z, s.t. y \geq 0, y \neq 0$ with $\sum_{i \in I(x^*)} y_i \nabla g_i(x^*) + \sum_{i=1}^{m-1} z_i \nabla g_i(x^*) = 0$.

By ①, we can see that: ①, ② \Leftrightarrow

$\exists y, z \neq 0, s.t. y \geq 0$ with $\sum_{i \in I(x^*)} y_i \nabla g_i(x^*) + \sum_{i=1}^{m-1} z_i \nabla g_i(x^*) = 0$ ③

And PLI(2) can \Rightarrow ③ obviously.

So can ③ \Rightarrow PLI(0), the answer is yes

(Recall in complementary slackness, if not active, then $\lambda_i = 0$).

And now, we propose the final theorem to end the "necessary" part.

(CKT with inequation constraints): Let x^* be a local optimal solution of the problem.

$$\begin{aligned} & \min f(x) \\ & \text{s.t. } g_i(x) \leq 0, i \in [n], \end{aligned} \quad \left. \begin{array}{l} \text{differentiable} \end{array} \right\}$$

Suppose that x^* is a regular point, then the KKT condition holds at x^* .

proof: W.L.O.G. we assume the first e constraints

$g_1(x^*) \sim g_e(x^*)$ are active, with $g_{e+1}(x^*) \sim g_m(x^*)$ aren't.

Then due to regular point property, we get:

$\nabla g_i(x^*) \sim \nabla g_e(x^*)$ is linearly independent.

Since we are consider local solutions and g_i are cts, we can ignore $g_{e+1} \sim g_m$ in some open neighborhood of x^* .

So w.r.t the local solutions. we just need to consider the active constraint.

Similar to the first thm we prove, we have:

$$\nabla f(x^*) = \sum_{i=1}^e \lambda_i \nabla g_i(x^*).$$

Now we consider

$$g_1(x_1, \dots, x_n) = -t$$

$$g_i(x_1, \dots, x_n) \geq 0, i \in [2, e]$$

Then, by Implicit function theorem. we get: $\exists \varepsilon > 0$, s.t a C^1 function $x_i(t) \sim v_i(t)$ is defined for $t \in [0, \varepsilon]$ with $x(0) = x^*$.

And for all $t \in [0, \varepsilon]$, $\begin{cases} g_1(x_1(t), \dots, x_n(t)) = -t, \\ g_i(x_1(t), \dots, x_n(t)) = 0, \quad \forall i \geq 2 \text{ a.c.} \end{cases}$

Take derivative $\Rightarrow \begin{cases} \nabla g_1(x^*) \cdot x'(0) = -1 \\ \nabla g_i(x^*) \cdot x'(0) = 0, \quad \forall i \geq 2 \text{ a.e.} \end{cases}$

Since x^* is local optima, we get:

$$\frac{d}{dt} f(x(t)) = \nabla f(x^*) \cdot V \geq 0.$$

$$\text{with } \nabla f(x^*) \cdot V = \sum_{i=1}^n \lambda_i \nabla g_i(x^*) \cdot V = -\lambda_1 \geq 0 \Rightarrow \lambda_1 \leq 0$$

Similarly, you can get the same result for all $i \geq 1$ a.e. \square

