OR review:

Consider a simple example

$$\min \ x$$
$$\text{s.t } x^2 \leq 0$$

Since it's impossible to find a solution for

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

Observation 1:

For a system $\min f(x)$

$$\text{s.t } h_i(x) \leq 0, \ i = 1 \sim m.$$
$$g_j(x) = 0, \ j = 1 \sim p.$$

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

To explore further, consider:

① Fritz John (eq).

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

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

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

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

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

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

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

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

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

Then, consider:

$$f(x) = f(x^*) - \varepsilon$$
$$h_i(x) = h_i(x^*), \ i=1\sim m.$$

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

$$f(x_\varepsilon) = f(x^*) - \varepsilon$$
$$h_i(x_\varepsilon) = h_i(x^*), \ i=1\sim m,$$

$\Rightarrow$ Contradiction with the local minima (let $\varepsilon \to 0^+$)

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

So this motivates a corrollary:

Corollary: let x^* be a local optima. of the problem, s.t.
min $f(x)$, s.t. $h_i(x) = 0$, $i=1\sim m$. (C^1 for all functions)
If $\nabla h_1(x^*) \sim \nabla h_m(x^*)$ are linearly independent, then there exists Lagrange multipliers λ_i, $i=1\sim m$, s.t

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

Now, it's natural to discuss inequality form:

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

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

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

proof: Firstly we show that:

lemma: $\left.\begin{cases} Bx < 0 & \text{(i)} \\ \exists y \geq 0, y \neq 0, y^T B = 0 & \text{(ii)} \end{cases}\right\}$ exactly one hold.

pf: $Bx < 0 \Leftrightarrow (B, I)\binom{x}{z} \leq 0$ for $z > 0$ $\quad (0^T y^T)\binom{x}{z} > 0$

$\Leftrightarrow (B, I)\binom{x}{z} \leq 0$ for $y^T z > 0$, $\forall y \neq 0, y \geq 0$.

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

Notice that $\begin{cases} \nabla f(x^*)^T d < 0 \\ \nabla g_i(x^*)^T d < 0, \ g_i \text{ is active} \end{cases}$ 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 NNAMCQ / NNAMOCQ)
We say NNAMCQ holds at x^* if there's no nonzero abnormal multipliers.

It's equivalent to:

(Positive Linear Independence Constraint Qualification)

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

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

(KKT under NNAMCQ / PLICQ)

Suppose that x^* is a local optima of:

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

If NNAMCQ / PLICQ holds at x^*, then : $\exists \lambda_i, i=1\sim p$ which aren't all zero, s.t.

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

proof: Consider the Fritz John thm with mixed constraints. Then use PLICQ / NNAMCQ

$\boxtimes$

Before we start, we propose a lemma.

lemma: For nonvacuous matrix $A_{p \times n}$ and matrix $D_{m \times n}$,

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

l.proof:

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

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

$\Leftrightarrow A^T u + D^T (v - w) = 0, \ u \geq 0$ 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 $I(x^*) = \{i = m \sim p \mid g_i(x^*) = 0\}$.
Then $\begin{cases} \nabla g_i(x^*)^T d = 0, & i = 1 \sim m-1 \\ \nabla g_i(x^*)^T d < 0, & i \in I(x^*) \end{cases}$ has a solution d.

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

$\nexists y, z$, 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 $\Longleftrightarrow$

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

And PLICQ can $\Rightarrow$ Ⓐ obviously

So can Ⓐ $\Rightarrow$ PLICQ $\therefore$ 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.

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

$$\left.\begin{aligned} &\min f(x) \\ &\text{s.t } g_i(x) \leq 0,\ i = 1 \sim m. \end{aligned}\right\} \text{differentiable}$$

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:

$\nabla g_1(x^*) \sim \nabla g_e(x^*)$ 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

$g_1(x_1, \dots, x_n) = -t$
$g_i(x_1, \dots, x_n) = 0$, $i = 2 \sim e$

Then, by Implicit function theorem. we get: $\exists\ \varepsilon > 0$, s.t a C^1 function $x_1(t) \sim x_n(t)$ is defined for $t \in [0, \varepsilon)$ with $x(0) = x^*$.

And for all $t \in [0, \varepsilon)$, $\begin{cases} g_1(x_1(t), \ldots, x_n(t)) = -t, \\ g_i(x_1(t), \ldots, x_n(t)) = 0, \ \forall i = 2 \sim e. \end{cases}$

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

Since x^* is local optima, we get:

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

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

Similarly, you can get the same result for all $i = 1 \sim e$. $\boxtimes$