Proof of KKT

Monday, 11 April 2022 2:03 PM

Let $f, h, g \in C^1$

Let $\underline{x}^*$ be a regular point and a local minimizer for the minimization of $f : \mathbb{R}^n \to \mathbb{R}$ s.t. $\underbrace{\underline{h}(x) = 0, \ g(x) \leq 0}_{\text{Primal feasibility}}$.

Then $\exists \ \underline{\lambda}^* \in \mathbb{R}^m$ and $\underline{\mu}^* \in \mathbb{R}^p$ such that
$(m \leq n)$.

$\boxed{1}$ $\underline{\mu}^* \geq 0$ (dual feasibility) $\to$ KKT ×

$\boxed{2}$ $Df(\underline{x}^*) + \underset{\text{Lagrange} \times}{\underline{\lambda}^{*T}} D\underline{h}(\underline{x}^*) + \underline{\mu}^{*T} D\underline{g}(\underline{x}^*) = \underline{0}^T$ [Stationarity]

$\boxed{3}$ $\underline{\mu}^{*T} g(\underline{x}^*) = 0$ [Complementarity Slack].

Proof: Step 1

$$S = \{ \underline{x} : \underline{h}(x) = 0, \ \underline{g}(x) \leq 0 \}$$

$$S_1 = \{ \underline{x} : \underline{h}(x) = 0, \ g_j(\underline{x}) = 0, \ j \in \underbrace{J(\underline{x}^*)}_{\text{active set}} \}$$

If $\underline{x}^*$ is a local minimizer of f on S, $\underline{x}^*$ is also a local minimizer of f on S_1. $\boxed{\text{Ex. 21.16}}$

[$\underline{x}^*$ is a regl. local minimum on a constrained set $\Rightarrow$ it is also a regl. local minimum on a subset defined by setting the active constraints to 0].

Step 2. Lagrange Thm $\Rightarrow \exists \ \underline{\lambda}^* \in \mathbb{R}^m, \ \underline{\mu}^* \in \mathbb{R}^p$ s.t.

$$Df(\underline{x}^*) + \underline{\lambda}^{*T} D\underline{h}(\underline{x}^*) + \underline{\mu}^{*T} D\underline{g}(\underline{x}^*) = 0^T$$

[Choose $\mu_j^* = 0 \quad \forall j \notin J(x^*)$]

	Active	Inactive
Constraints	$g_j(x^*) = 0$	$g_j(x^*) < 0$
Multipliers	μ_j^*	$\mu_j^* = 0$

Step gives us the condition 3

$\boxed{\text{Step gives us the condition (3)}}$

Multipliers | μ_j^* | $\mu_j^* = 0$

(Lagrange Thm) (Choose).

Step 3:

$$\underline{\mu_1^* g_1(x^*) + \mu_2^* g_2(x^*) + \cdots + \mu_p g_p(x^*) = 0}$$

$j \notin J(x^*)$, $\mu_j^* = 0$; $\mu_j^* g_j(x^*) = 0$

$j \in J(x^*)$, $g_j(x^*) = 0$

$\mu_j g_j(x^*) = 0$

Condtn (3) also holds.

Step 4: $\mu_j^* \geq 0 \quad j = 1, \ldots p$.

To S.T. $\mu_j^* \geq 0$ for $j \in J(x^*)$. [active sets; $g_j(x^*) = 0$].

Proof by contradiction: If possible, let $\exists\, j \in J(\underline{x}^*)$ for which $\mu_j^* < 0$.

Define $\hat{S} = \{ \underline{x} : h(\underline{x}) = 0;\ g_i(\underline{x}) = 0,\ i \in J(x^*),\ i \neq j \}$.

$\hat{T}(\underline{x}^*) = \{ \underline{y} : Dh(x^*)\underline{y} = 0,\ Dg_i(\underline{x}^*)y = 0,\ i \in J(x^*),\ i \neq j \}$.

Step 4a

Claim: $\exists\ \underline{y} \in \hat{T}(\underline{x}^*)$ s.t. $Dg_j(x^*)\underline{y} \neq 0$. [$\underline{x}^*$ is a regular point].

Justify: If possible, $\nabla g_j(\underline{x}^*)^T \underline{y} = Dg_j(\underline{x}^*)\underline{y} = 0$

$$\nabla g_j(x^*) \in \hat{T}(\underline{x}^*)^\perp = \hat{N}(\underline{x}^*)$$

$$= \text{Span}\{\nabla h_1(x^*), \ldots \nabla h_m(x^*), \nabla g_i(x^*)\ \ i \neq j\}.$$

Contradiction to the fact that $\underline{x}^*$ is a regular point. [Defn of regular point].

[Defn of regular point].

$\boxed{\text{Step 4b}}$ W.l.o.g. assume that $\boxed{Dg_j(x^*)\,y < 0}$ $(y \in \hat{T}(x^*))$

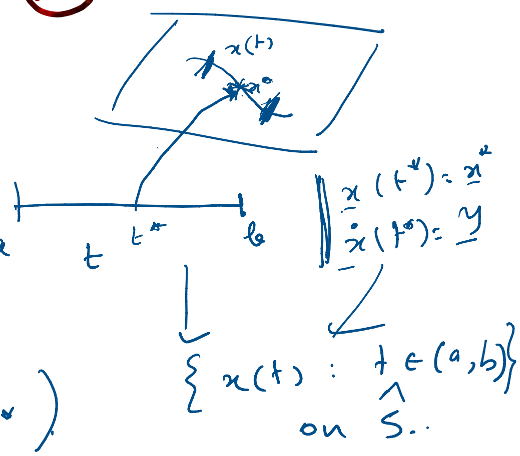
From Lagrange condn,

$$Df(x^*)\underline{y} + \underbrace{\lambda^{*T} Dh(x^*)\underline{y}}_{=0} + \mu_j^* Dg_j(x^*)\underline{y} + \overbrace{\sum_{i \neq j} \mu_i^* Dg_i(x^*)\underline{y}}^{0} = 0$$

as $y \in \hat{T}(x^*)$.

$$\boxed{Df(x^*)\,\underline{y}} = -\overset{<0}{\mu_j^*}\;\overbrace{Dg_j(x^*)\underline{y}}^{<0} < 0 \quad \text{for } y \in \hat{T}(x^*) \qquad (A)$$

Thm 20.1 Let $\underline{x}^* \in \hat{S}$ be a regular pt and $\hat{T}(\underline{x}^*)$ is the tangent space at $\underline{x}^*$. Then $\underline{y} \in \hat{T}(\underline{x}^*) \iff \exists$ a diff. curve in $\hat{S}$ s.t. $\dot{x}(t^*) = \underline{y}$, $x(t^*) = x^*$)

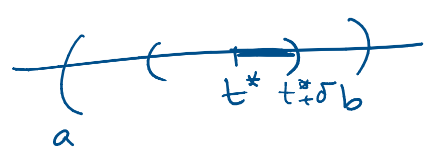


$$\frac{d}{dt}\Big(f(\underline{x(t^*)}) \Big) = \underbrace{Df(x(t^*))^T \overbrace{\dot{x}(t^*)}^{\underline{y}}} < 0$$

$x(t)$

$$= Df(x^*)\,\underline{y} < 0 \quad [\text{from } (A)]$$

$\underline{x'(t^*) < 0}$



$$\Rightarrow \exists\, \delta > 0 \text{ s.t. } \forall\, t \in (t^*, t^*+\delta),$$

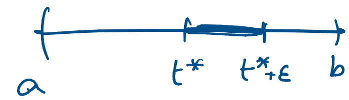
$$f(x(t)) < f(x(t^*)) = f(x^*).$$

$$\frac{d}{dt}\Big(g_j(\underline{x(t^*)}) \Big) = Dg_j(x^*)\,\underline{y} < 0$$

$\frac{d}{dt}\left(g_j(x(t)) \right) = -g_j$ —

$\Rightarrow \exists\, \varepsilon > 0$ and $t \in [t^*, t^* + \varepsilon]$ s.t.

$$g_j(x(t)) < 0$$



$\forall\, t \in [t^*, t^* + \min\{\varepsilon, \delta\}]$,

$$\left.\begin{array}{l} g_j(x(t)) < 0 \\ f(x(t)) < f(x^*) \end{array}\right]$$

The point $x(t) \in \hat{S}$, $t \in [t^*, t^* + \min\{\varepsilon, \delta\}]$ are feasible points with lower values of the objective function f (than $f(x^*)$); this contradicts the assumption that x^* is a local minimizer.

$\Rightarrow$ This assumption ($\mu_j < 0$) is wrong.

SYMQBD