Supplements 6 to Chapter III **E. Gekeler**

11/10/06

Local Lagrange Theory

Let $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$ be again real normed vector spaces, let $\mathcal{K} \subset \mathcal{Y}$ be an order cone with adjoined cone $\mathcal{K}_d$. We consider the general minimum problem (MP): Find $x^* \in \mathcal{X}$ such that

$$x^* = \arg\min\{f(x),\ x \in \mathcal{C},\ g(x) \leq 0,\ h(x) = 0\} \tag{1}$$

where $f : \mathcal{C} \to \mathbb{R},\ g : \mathcal{C} \to \mathcal{Y},\ h : \mathcal{C} \to \mathcal{Z}$.
The following *linearized* minimum problem is associated to the minimum problem (1):

$$\min\{\nabla f(x^*)(x - x^*),\ x \in \mathcal{C},\ g(x^*) + \nabla g(x^*)(x - x^*) \leq 0,\ \nabla h(x^*)(x - x^*) = 0\} \tag{2}$$

Definition 1 *Let $g : \mathcal{X} \supset \mathcal{C} \to \mathcal{Y}$ Fréchet-differentiable, $\mathcal{K} \subset \mathcal{Y}$ a positive cone, $g(x^*) \leq 0$, and let*

$$\mathrm{LC}(g, x^*) := \{v \in \mathcal{X},\ g(x^*) + \nabla g(x^*)v \leq 0\}$$

be the linearized *cone of the constraint $g(x) \leq 0$ in x^*. Then $g(x) \leq 0$ is locally solvable in x^* if*

$$\forall\, v \in \mathrm{LC}(g, x^*)),\ \exists\, \varepsilon > 0,\ \exists\, \varphi : \mathbb{R} \to \mathcal{X},\ \varphi(\alpha) = o(|\alpha|) :$$
$$g(x^*) + \nabla g(x^*)v \leq 0,\ 0 < \alpha \leq \varepsilon \Longrightarrow g(x^* + \alpha v + \varphi(\alpha)) \leq 0.$$

Theorem 1 *(Linearization Theorem) Let*

$$x^* = \arg\min\{f(x),\ x \in \mathcal{C},\ g(x) \leq 0,\ h(x) = 0\},$$

Let f, g, h F-differentiable, $\mathrm{int}(\mathcal{K}) \neq \emptyset$, *and let h in x^* w.r.t. $\mathcal{C}$ locally solvable. Further, let*

$$\begin{aligned} \mathcal{A} &= \{x \in \mathcal{C},\ \nabla h(x^*)(x - x^*) = 0\}, \\ \mathcal{B} &= \{x \in \mathcal{C},\ g(x^*) + \nabla g(x^*)(x - x^*) < 0\}, \\ \mathcal{D} &= \{x \in \mathcal{C},\ \nabla f(x^*)(x - x^*) < 0\}. \end{aligned}$$

Then

$$\mathcal{A} \cap \mathcal{B} \cap \mathcal{D} = \emptyset.$$

Proof. See [Craven78], p. 34. Suppose that there exists a $x \in \mathcal{A} \cap \mathcal{B} \cap \mathcal{D}$. Then

$$g(x^*) + \nabla g(x^*)(x - x^*) < 0,\ h(x^*) + \nabla h(x^*)(x - x^*) = 0.$$

(i) For sufficiently small $0 < \alpha$

$$\begin{aligned} &g(x^* + \alpha(x - x^*) + o(|\alpha|)) = g(x^*) + \alpha \nabla g(x^*)(x - x^*) + o(|\alpha|) \\ &= (1 - \alpha)g(x^*) + \alpha[g(x^*) + \nabla g(x^*)(x - x^*) + o(|\alpha|)/\alpha]. \end{aligned}$$

Both terms on the right side lie in $-\mathcal{K}$ by assumption for sufficiently small $\alpha > 0$ therefore

$$g(x^* + \alpha(x - x^*) + \varphi(\alpha)) \leq 0$$

for sufficiently small $\alpha > 0$ and arbitrary $\varphi(\alpha) = o(|\alpha|)$.
(ii) $h(x) = 0$ is in x^* w.r.t. $\mathcal{C}$ local solvable by assumption, therefore

$$\exists\, \varepsilon > 0,\ \exists\, \varphi \in o(|\alpha|) : 0 < \alpha < \varepsilon \Longrightarrow x^* + \alpha(x - x^*) + \varphi(\alpha) \in \mathcal{S}.$$

(iii) For sufficiently small $\alpha > 0$ by assumption

$$\begin{aligned} 0 &\leq \alpha^{-1}[f(x^* + \alpha(x - x^*) + \varphi(\alpha)) - f(x^*)] \\ &= \alpha^{-1}[\nabla f(x^*)\alpha(x - x^*) + \varphi(\alpha)] \\ &= \nabla f(x^*)(x - x^*) + \varphi(\alpha)/\alpha \to \nabla f(x^*)(x - x^*),\ \alpha \to 0, \end{aligned}$$

Therefore $\nabla f(x^*)(x - x^*) \geq 0$ hence $x \notin \mathcal{D}$ in contradiction to the assumption.

Definition 2 *Let $\mathcal{X}$ be a normed vector space and $\mathcal{C}, \mathcal{D} \subset \mathcal{X}$.*
(1°) $\operatorname{aff}(\mathcal{C})$ *is the smallest affine subspace of $\mathcal{X}$ which contains $\mathcal{C}$, $\mathcal{C} \subset \operatorname{aff}(\mathcal{C})$.*
(2°) *Let $x \in \mathcal{C} \cap \mathcal{D}$ then x is interior point of $\mathcal{C}$ relative to $\mathcal{D}$ if there exists a neighborhood of x in $\mathcal{D}$ which is entirely contained in $\mathcal{C}$:*

$$\exists\, \varepsilon > 0,\ \forall\, u \in \mathcal{D} : \|u - x\| \leq \varepsilon \Longrightarrow u \in \mathcal{C}.$$

(3°) $\operatorname{relint}(\mathcal{C})$ *is the set of interior points of $\mathcal{C}$ relative to* $\operatorname{aff}(\mathcal{C})$.

Let e.g. $h : \mathcal{C} \to \mathcal{Z}$ affine linear then $\operatorname{relint}(h(\mathcal{C})) \neq \emptyset$, if $\mathcal{Z}$ finite-dimensional or $\operatorname{relint}(\mathcal{C}) \neq \emptyset$. Cf. [Kirsch], p. 50.

Definition 3 *Let* $\operatorname{int}(\mathcal{K}) \neq \emptyset$.
(a) The pair (g, h) suffices the SLATER *condition (S) if*

$$\mathcal{A} := \{x \in \mathcal{C},\ g(x) < 0,\ h(x) = 0\} \neq \emptyset.$$

(b) (g, h) suffices the KARLIN *condition (K) if*

$$\mathcal{B} := \{(y, z) \in \mathcal{K}_d \times Z_d,\ \forall\, x \in \mathcal{C} : y \circ g(x) + z \circ h(x) \geq 0\} = \{(0, 0)\}.$$

$\neg(S)$ is therefore the condition $\mathcal{A} = \emptyset$ and $\neg(K)$ is the condition $\mathcal{B} \neq \{(0, 0)\}$.

Theorem 2 *Let $\mathcal{C} \subset X$ convex, g $\mathcal{K}$-convex, and h affine linear. Further, let*

$$\operatorname{int}(\mathcal{K}) \neq \emptyset,\ \operatorname{relint}(h(\mathcal{C})) \neq \emptyset.$$

Then
(1°) *(S) and $0 \in \operatorname{int}(h(\mathcal{C})) \Longrightarrow (K)$,*
(2°) *$\neg(S) \Longrightarrow \neg(K)$.*

Proof see [Kirsch], S. 50 ff. So (S) and (K) are nearly equivalent.

Theorem 3 *Let the minimum problem (1) be F-differenzierbar and suppose that:*
(1°)

$$x^* = \arg\min\{f(x),\ x \in \mathcal{C},\ g(x) \leq 0,\ h(x) = 0\},$$

(2°) $\operatorname{int}(\mathcal{K}) \neq \emptyset$,
(3°) $\operatorname{relint}(\nabla h(x^*)(\mathcal{C})) \neq \emptyset$.
(4°) *h in x^* w.r.t. $\mathcal{C}$ local solvable,*
Then there exists a triple $(0, 0, 0) \neq (\varrho^*, y^*, z^*) \in \mathbb{R}_{\geq 0} \times \mathcal{K}_d \times Z_d$ such that
(i)

$$\forall\, x \in \mathcal{C} :\ [\varrho^* \nabla f(x^*) + y^* \circ \nabla g(x^*) + z^* \circ \nabla h(x^*)](x - x^*) \geq 0,$$

(ii) $y^* \circ g(x^*) = 0$.
(iii) If there exists a $x \in \mathcal{C}$ so that

$$g(x^*) + \nabla g(x^*)(x - x^*) < 0,\ \nabla h(x^*)(x - x^*) = 0,$$

and if $x^* \in \operatorname{int}(\nabla h(x^*)(\mathcal{C}))$, then $\varrho^* = 1$ can be chosen and $y^* \neq 0$.

em Proof. See also [Kirsch]. Let $\mathcal{W} = \mathbb{R} \times Y$, $\mathcal{J} = \mathbb{R}_{\geq 0} \times \mathcal{K}_d$. Then $\mathcal{J} \subset \mathcal{W}$ is a convex cone with $\mathrm{int}(\mathcal{J}) = \mathbb{R}_{>0} \times \mathrm{int}(\mathcal{K}_d) \neq \emptyset$. Let $G : \mathcal{X} \to \mathcal{W}$ be defined by

$$G(x) = [\nabla f(x^*)(x - x^*),\ g(x^*) + \nabla g(x^*)(x - x^*)]$$

then G is $\mathcal{J}$-convex. Let $H : \mathcal{X} \to \mathcal{Z}$ definiert durch

$$H(x) = \nabla h(x^*)(x - x^*)$$

dann ist H affine linear. Then we obtain by the linearization theorem that

$$\mathcal{A} := \{x \in \mathcal{C},\ G(x) \overset{\mathcal{J}}{<} 0,\ H(x) = 0\} = \emptyset.$$

Therefore $\neg(S)$ for (G, H). Because $\mathrm{relint}(H(\mathcal{C})) = \mathrm{relint}(\nabla h(x^*)(\mathcal{C}) \neq \emptyset$, Theorem 2 supplies the exisrence of $(0, 0) \neq (\widetilde{y}^*, z^*) \in \mathcal{J}_d \times Z_d$ such that

$$\forall\ x \in \mathcal{C} :\ \widetilde{y}^* \circ G(x) + z^* \circ H(x) \geq 0.$$

Since $\widetilde{y}^* = (\varrho^*, y^*) \in \mathbb{R}_{\geq 0} \times \mathcal{K}_d'$, we obtain by this way directly that

$$\begin{aligned} &\forall\ x \in \mathcal{C} : \\ &\varrho^* \nabla f(x^*)(x - x^*) + y^* \circ [g(x^*) + \nabla g(x^*)(x - x^*)] + z^* \circ \nabla h(x^*)(x - x^*) \geq 0. \end{aligned} \tag{3}$$

Furthermore we have $g(x^*) \leq 0$ hence $y^* \circ g(x^*) \leq 0$. For $x = x^*$ (3) implies that $y^* \circ g(x^*) \geq 0$, hence together $y^* \circ g(x^*) = 0$. Now the assertion follows from (3).
Suppose $\varrho^* = 0$ then necessarily $(y^*, z^*) \neq (0, 0)$. It then follows from (3) that

$$\forall\ x \in \mathcal{C} : y^* \circ [g(x^*) + \nabla g(x^*)(x - x^*)] + z^* \circ \nabla h(x^*)(x - x^*) \geq 0.$$

By Ass. (iii) there exists a $x \in \mathcal{C}$ which fulfills this inequality. For this x we obtain by Lemma 1.26 of the section on convex sets that

$$y^* \circ [g(x^*) + \nabla g(x^*)(x - x^*)] + z^* \circ \nabla h(x^*)(x - x^*) < 0,$$

or $y^* = 0$. Suppose $y^* \neq 0$ then necessarily $\varrho^* \neq 0$. Suppose $y^* = 0$ and $\varrho^* = 0$ then necessarily $z^* \neq 0$, and it follows from (3) that $z^* \in (\nabla h(x^*)(\mathcal{C} - x^*))_d$. Because $x^* \in \mathrm{int}(\nabla h(x^*)(\mathcal{C}))$ then $0 \in \mathrm{int}(\nabla h(x^*)(\mathcal{C} - x^*))$. By Lemma 1.26 then $z^* = 0$. This is a contradiction to to $(\varrho^*, y^*, z^*) \neq (0, 0, 0)$ therefore $y^* \neq 0$ under the named additional assumptions. By division with ϱ^* then the last assertion of the Theorem is verfied.

Definition 4 *A feasible point x^* is a regular point if:*
(1°) *There exists a $x \in \mathcal{C}$ such that*

$$g(x^*) + \nabla g(x^*)(x - x^*) < 0,\ \nabla h(x^*)(x - x^*) = 0.$$

(2°) *h is in x^* w.r.t. $\mathcal{C}$ local solvable.*

Theorem 4 *(*Robinson*) Let $\mathcal{X}$ be a* Banach *space, $\mathcal{Y}$ a normed space, $\mathcal{K} \subset \mathcal{Y}$ closed and $g : \mathcal{X} \to \mathcal{Y}$ continuously F-differentiable. Then $g(x) \leq 0$ is local solvable $x^* \in \mathcal{X}$ if*

$$0 \in \mathrm{int}[g(x^*) + \nabla g(x^*)(\mathcal{X}) + \mathcal{K}].$$

Proof see [Robinson].

Theorem 5 (LJUSTERNIK) *Let $\mathcal{X}$, $\mathcal{Y}$ be* BANACH *spaces, $g : \mathcal{X} \to \mathcal{Y}$ continuously F-differentiable, $g(x^*) = 0$, and $\nabla g(x^*)(\mathcal{X}) = \mathcal{Y}$. Then $g(x) = 0$ is in x^* local solvable.*

Proof see [Ljusternik].

Apparently the theorem of LJUSTERNIK follows from the theorem of ROBINSON. Assume now that $\mathcal{Z} = \nabla h(x^*)(\mathcal{X})$ then relint$(\nabla h(x^*)(\mathcal{X}) = \text{int}(\mathcal{Z}) \neq \emptyset$, $0 \in \text{int}(\nabla h(x^*)(\mathcal{X})$ and h in x^* local solvable by ROBINSON's theorem if $h(x^*) = 0$. Theorem 3 then leads to the following result:

Corollary 1 *Let the minimum problem (1) be continuously F-differentiable and suppose that:*
(1°)
$$x^* = \arg\min\{f(x),\ x \in \mathcal{X},\ g(x) \leq 0,\ h(x) = 0\},$$
(2°) $\text{int}(\mathcal{K}) \neq \emptyset$,
(3°) $\nabla h(x^*) : \mathcal{X} \to \mathcal{Z}$ *surjective,*
(4°) $\exists\, x \in \mathcal{X} :\ g(x^*) + \nabla g(x^*)x < 0,\ \nabla h(x^*)x = 0.$
Then there exists a pair $(y^*, z^*) \in \mathcal{K}_d \times \mathcal{Z}_d$ with $y^* \neq 0$ such that
(i)
$$\nabla f(x^*) + y^* \circ \nabla g(x^*) + z^* \circ \nabla h(x^*) = 0.$$
(ii) $y^* \circ g(x^*) = 0$.

Theorem 6 *Let the minimum problem (MP) (1) for $\mathcal{C} = \mathcal{X}$ be convex (h affine linear), F-differentiable, $x^* \in \mathcal{S}$, and let the multiplier rule (MR) be fulfilled:*

$$\exists\, (y^*, z^*) \in \mathcal{K}_d \times \mathcal{Z}_d : \nabla_x L(x^*, y^*, z^*) = 0,\ \text{and}\ y^* \circ g(x^*) = 0.$$

Then x^* is solution of (MP).

Proof. Since f, g convex and h affine linear, we have

$$\begin{array}{rcl} f(x) & \geq & f(x^*) + \nabla f(x^*)(x - x^*), \\ g(x) & \geq & g(x^*) + \nabla g(x^*)(x - x^*), \\ h(x) & = & h(x^*) + \nabla h(x^*)(x - x^*). \end{array}$$

Then, by (MR) for $x \in \mathcal{X}$

$$\begin{array}{rcl} f(x) & \geq & f(x) + y^* \circ g(x) + z^* \circ h(x) \\ & \geq & f(x^*) + y^* \circ [g(x^*) + \nabla g(x^*)(x - x^*)] + z^* \circ [h(x^*) + \nabla h(x^*)(x - x^*)] \\ & = & f(x^*) + y^* \circ g(x^*) + z^* \circ h(x^*) + \nabla_x L(x^*, y^*, z^*)(x - x^*) \\ & = & f(x^*). \end{array}$$

Briefly: (MR) is necessary for (MP) and sufficient if (MP) convex.