April 2019

GS
SI

Theory of constrained optimization in infinite-dim. spaces.

Necessary conditions in finite-dimensions (KKT)

<u>Task.</u> (*) $\min_{x \in \sigma} f(x)$ $\qquad \sigma = \{ x \in \mathbb{R}^n : h(x) = 0, \ g(x) \geq 0 \}$

<u>Assumpt.</u> $f, g, h \in C^2$ with $h: \mathbb{R}^n \to \mathbb{R}^q$, $g: \mathbb{R}^n \to \mathbb{R}^m$, $f: \mathbb{R}^n \to \mathbb{R}$ with $q < n$.

<u>Constraint qualifications</u> (CQ)

Let $x^* \in \sigma$. x^* fulfills LICQ (linear independent CQ) if

$$\left(\nabla h_i(x^*), \nabla g_j(x^*) \right) \qquad i = 1, \dots, q, \ j \text{ s.t. } g_j(x^*) = 0.$$

are linear independent

Let $x^* \in \sigma$. x^* fulfills MFCQ (Mangasarin-Fromowitz)

if ① $\nabla h(x^*)$ has full rank $(=q)$

② $\exists d \in \mathbb{R}^n$: $\nabla g_j(x^*)^T d > 0 \quad \forall j: g_j(x^*) = 0$

and $\nabla h_j(x^*)^T d = 0 \quad \forall j$.

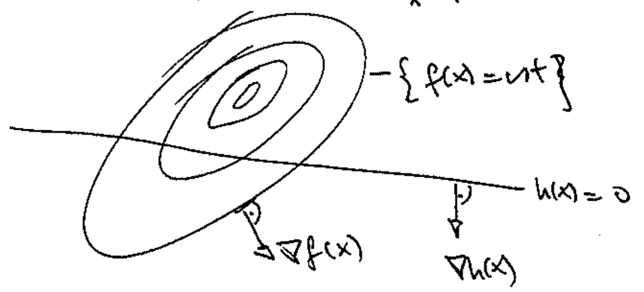
<u>Necessary optimality</u>:

Let x^* be a local Minimum of (*) and assume that either LICQ or MFCQ holds true. Then, there exists $\lambda^* \in \mathbb{R}^m_+$, $\mu^* \in \mathbb{R}^q$ s.t.

$$\left.\begin{aligned} \nabla f(x^*) - \nabla h(x^*)\mu^* - \nabla g(x^*)\lambda^* &= 0 \\ x^* &\in G \\ \lambda^* &\geq 0 \\ g_j(x^*)\lambda^*_j &= 0 \end{aligned}\right\} \underline{\text{KKT}}.$$

We prove in the case of equality constraints a corresponding infinite-dim. version and then discuss numerical methods.

Guiding example: $\min_x f(x)$ s.t. $G = \{x \in \mathbb{R}^n : h(x) = 0\}$



If x^* is local min. $\nabla f(x^*) \parallel \nabla h(x^*)$. This is only true if $\nabla h(x^*)$ does not degenerate at x^* (CQ).

Let $f : X \to \mathbb{R}$, X, Y Hilbert (or Banach space) and $H : X \to Y$.

<u>Def.</u> H is called Frechet differentiable at $x \in X$ if there exists a linear operator $DH \in L(X, Y)$ s.t.

$$H(x+h) = H(x) + DH(x)h + o(\|h\|) \qquad \forall h \in X$$

$DH(x)$ is unique and the Frechet derivative.

H is called Gateaux-diff. at $x \in X$ s.t. for all $h \in X$ we have

$$H(x+th) - H(x) - t\, D_G H(x)h + o(t) \qquad \forall t$$

Example: $B(y,z) = \int_\Omega \nabla y \nabla z \, dx \ : H^1_0 \times H^1_0 \mapsto \mathbb{R}$
is a bilinear operator on $H^1_0(\Omega)^2$. It's Frechet-derivative is

$$DB(y,z)[h_y,h_z] = \int_\Omega \nabla h_y \nabla h_z \, dx \quad \in L(H^1_0 \times H^1_0 ; \mathbb{R})$$

The op. is linear and bd due to Cauchy-Schwarz:

$$\| DB(y,z)[h_y,h_z] \|_{\mathbb{R}} \leq C \|h_y\|_{H^1_0} \|h_z\|_{H^1_0}$$

$\Rightarrow DB$ is the Frechet derivative.

Def (LICQ). Let $x \mapsto DH(x)$ be continuous and $DH(x)$ the Frechet derv. of H. If at $x_0 \in \mathcal{U}$ the map $DH(x_0)$ is surjective, then x_0 is called regular point.

Example. $DH(y) = -\Delta y$ as op. $DH: H^1_0 \mapsto H^{-1}$ i.e.

$$DH(y)\underset{\substack{\wedge \\ (H^{-1})^*}}{\phi} = \int_\Omega \nabla y \nabla \phi \, dx.$$

Then, $\forall f \in H^{-1}$ $\exists!\, y \in H^1_0$ s.t.

$$DH(y) = f \in H^{-1}$$

by standard PDE theory. Hence, $DH(y)$ is surjective.

Theorem. Let x_0 be a regular pt of $H: X \to Y$. Then, there exists a neighborhood $\mathcal{U}(y_0)$ of $y_0 = H(x_0)$ and a constant K s.t. $H(x) = y$ has a solution $\forall y \in \mathcal{U}(y_0)$ and $\|x - x_0\| \leq K \|y - y_0\|$.

<u>Lemma</u>. Let $f: X \to \mathbb{R}$, $H: X \to Y$ cont. Fréchet-differentiable in an open neighborhood $\mathcal{U}(x_0)$ and x_0 regular point. Let x_0 be a local min. of $\min f(x)$ sbj. $H(x) = 0$.

Then, $Df(x_0)h = 0 \quad \forall h$ st. $DH(x_0)h = 0$.

<u>Proof</u>. Define $T(x) = (f(x), H(x))$. Assume there ex. $h \in X$ st. $DH(x_0)h = 0$ but $Df(x_0)h \neq 0$. We have x_0 is regular point of $T(x)$ since $DT(x_0)[h_f, h] = Df(x_0)h_f + DH(x_0)h$ and $Df(x_0) \neq 0$ (w.l.o.g). Hence, the inverse fct thm. can be applied and we find for $\delta \in \mathbb{R}_+$ suff. small a sol. to $T(x) = \begin{pmatrix} f(x_0) - \delta \\ 0 \end{pmatrix}$. This contradicts x_0 local. min.

<u>Theorem</u>. (same assumpt. as in lemma).

There exists $y_0^* \in Y^*$ s.t.

$$\begin{cases} Df(x_0) - \langle y_0^*, DH(x_0) \rangle_{Y^*, Y} = 0 \\ H(x_0) = 0 \end{cases}$$

<u>Remark</u> $H(x_0) = 0$ is an equality in the space Y and the $Df(x_0) - \dots$ is an equality in the space $L(X;\mathbb{R})$. Hence:

$\forall h \in X$: $\quad Df(x_0)h - \underbrace{y_0^*}_{\in Y^*} \underbrace{DH(x_0)h}_{\in Y} = 0 \quad \forall h \in X$.

<u>Proof</u>. We have $Df(x_0) \in L(X;\mathbb{R}) = X^*$ is orthogonal to the kernel of $DH(x_0)$. Since $DH(x_0)$ is surjective the range is closed.

[Theorem of open mapping: $A \in L(X, Y)$, $R(A)$ closed $\Rightarrow$ $Ax = y$, $\|x\| \leq K\|y\|$ $\Rightarrow$ A has bd. inverse.]

GS
SI

The adjoint operator $DH^*(x_0) \in L(Y^*, X^*)$ is defined by $\langle y^*, DH(x_0)h \rangle =: \langle DH^*(x_0)y^*, h \rangle$.

Since $\mathcal{R}(DH(x_0))$ is closed we have $\mathcal{R}(DH^*(x_0)) = \ker(DH(x_0))^\perp$
[PROOF ON PAGE 5b]

Hence, $Df(x_0) \in \ker(DH(x_0))^\perp = \mathcal{R}(DH(x_0)^*) = DH(x_0)^* y_0^*$
for some y_0^*. Using the definition of the adjoint we obtain the multiplier result.

Example: Consider $\min \frac{1}{2}\|y - y_d\|_{L^2}^2 + \frac{\alpha}{2}\|u\|_{L^2}$ s.t. $-\Delta y = u$ in Ω,
$y = 0$ on $\partial\Omega$.

Define $x = (y,u) \in H_0^1 \times L^2$ with $X = H_0^1 \times L^2$ and $X^* \cong H_0^1 \times L^2$.
The operator $H: X \to Y$ with $Y = H^{-1} \cong H_0^1$ by
$$H(x)[v] = \int_\Omega \nabla y \nabla v - uv \, dx \quad \text{for all } v \in H_0^1.$$
and $f(x) := \int_\Omega \frac{1}{2}(y - y_d)^2 + \frac{\alpha}{2}u^2 \, dx$ as fct $f: X \mapsto \mathbb{R}$.

f, H are quadratic and Frechet differentiable. Also, as before DH is surjective for all $x_0 = (y_0, u_0)$ since
$DH(x_0) \in L(H_0^1 \times L^2, H_0^1)$ with $DH(x_0)[h_y, h_u][v] = \int_\Omega \nabla h_y \nabla v - h_u v \, dx$
and $DH(x_0)[h_y, h_u] = f$ in H_0^1

$\Leftrightarrow \quad \int_\Omega \nabla h_y \nabla v + h_u v - fv \, dx = 0 \quad \forall v \in H_0^1$

$\Rightarrow \quad \exists! \, (h_y, 0)$ for all $f \in H_0^1$.

$\Rightarrow \quad DH(x_0)$ is surjective.

GS
SI

Theorem

Let $T: X \mapsto Y$. Then, $\overline{R(T)} = (\ker T^*)^\perp$.

Proof. „$\subset$": Let $y = Tx \in R(T)$. If $y' \in \ker T^*$, then

$y'(y) = y'(Tx) = (T^* y')(x) = 0$ since $y' \in \ker T^*$. Hence, $R(T) \subset (\ker T^*)^\perp$. Also, $(\ker T^*)^\perp$ is closed.

„$\supset$": Let $U := \overline{R(T)}$. Then U is a closed subspace of Y. Let $\tilde{y} \notin U$. Due to Hahn-Banach separation theorem, there exists $y' \in Y'$ with $y' \not\equiv 0$ s.t. $y'(y) = 0$ $\forall y \in U$ and $y'(\tilde{y}) \neq 0$.

Hence, $y'(Tx) = 0$ for all $x \in X$, since $Tx \in U$, and therefore $(T^* y')x = 0$ and hence $y' \in \ker T^*$.

~~Hence,~~ Also, by construction $y'(\tilde{y}) \neq 0 \Rightarrow \tilde{y}^* \notin (\ker T^*)^\perp$

(because $(\ker T^*)^\perp = \{y \in Y : y'(y) = 0, y' \in \ker T^*\}$).

Hence, $\tilde{y} \notin U \Rightarrow \tilde{y} \notin (\ker T^*)^\perp$.

Remark. ① $Ty = y : C(0,1) \mapsto L^2(0,1)$. Then, $R(T)$ is not closed.

② Operators of the type $T = Id - S$, S compact have closed range.

Hence, there exists $y_0^* \in Y^* = H^1_0$ s.t.

$$Df(x_0)h_x - y_0^* DH(x_0)[h_x] = 0 \qquad \forall h_x \in X.$$

$$\overset{h_x=(h_y,h_u)}{\Leftrightarrow} \int_\Omega (y-y_d)h_y + \alpha u\, h_u - \int_\Omega \nabla h_y \nabla y_0^* + h_u y_0^* \, dx = 0$$

$$\Rightarrow \quad \begin{cases} \alpha u - y_0^* = 0 & \text{a.e. in } \Omega \\ +\Delta y_0^* + (y-y_d) = 0 & \text{a.e. in } \Omega \\ y_0^* = 0 & \text{on } \partial\Omega \end{cases}$$