Optimization II

Dietmar Hömberg

NTNU 2023

Optimization II

Lecture 2, Jan. 12

2. Finite-dim. Case

admissible set $\mathcal{U}_{ad} \subset \mathbb{R}^m$
closed, bounded, convex

$A \in \mathbb{R}^{n,n}$ regular, rank n

$B \in \mathbb{R}^{n,m}$ $\quad m \leq n$

- state equation $\quad Ay = Bu \quad$ (SE)
- $y \in \mathbb{R}^n$ state
- $u \in \mathbb{R}^m$ control
- cost functional

$$J(y,u) = \underbrace{\frac{1}{2}|y - y_d|^2}_{\text{tracking}} + \underbrace{\frac{\gamma}{2}|u|^2}_{\text{regularization}}$$

y_d desired state

(CP) $\quad \min J(y,u)$

subject to $Ay = Bu$ (SE)

and $u \in \mathcal{U}_{ad}$ control constr.

A regular $\Rightarrow$ $y = A^{-1}Bu$

define $S: \mathbb{R}^m \to \mathbb{R}^n \quad u \mapsto y(u)$

solution operator

here: $S = A^{-1}Bu$

now, consider reduced cost functional

$$f(u) = J(y(u),u) = \frac{1}{2}|Su - y_d|^2 + \frac{\delta}{2}|u|^2$$

(CP) $\Longleftrightarrow$ $\boxed{\min\limits_{u \in \mathcal{U}_{ad}} f(u)}$ $(\tilde{CP})$

Q1: Existence of sol. to (CP)

f continuous, $\mathcal{U}_{ad}$ compact

Weierstrass theorem $\Rightarrow$ $\exists$ sol. to (CP)

Q 2: Optimality conditions

let $\bar{u} \in \mathcal{U}_{ad}$ is a sol to (CP)

take $u \in \mathcal{U}_{ad}$ then

$$\bar{u} + t(u-\bar{u}) = tu + (1-t)\bar{u} \in \mathcal{U}_{ad}$$

for $t \in [0,1]$

$$\lim_{t \searrow 0} \frac{f(\bar{u}+t(u-\bar{u})) - f(\bar{u})}{t} \geq 0 \quad \text{for } t \in [0,1]$$

$$\boxed{\nabla f(\bar{u})^T (u-\bar{u}) \geq 0 \quad \forall u \in \mathcal{U}_{ad}}$$

recall $f(u) = \frac{1}{2} | Su - y_d |^2 + \frac{\gamma}{2} |u|^2$

$$\frac{\partial f}{\partial u_k} = \frac{\partial}{\partial u_k} \left(\frac{1}{2} \sum_{i=1}^{n} [Su - y_d]_i^2 \right) + \frac{\gamma}{2} \frac{\partial}{\partial u_k} \sum_{i=1}^{m} u_i^2$$

$$= \sum_{i=1}^{n} [Su - y_d]_i \underbrace{\frac{\partial}{\partial u_k} \left(\sum_{j=1}^{n} S_{ij} u_j \right)}_{S_{ik} = S^T_{ki}} + \gamma u_k$$

$$= \left[S^T (Su - y_d) \right]_k + \gamma u_k$$

$$\Rightarrow \quad \nabla f(u) = S^T (Su - y_d) + \gamma u$$

since $S = A^{-1} B$, we get

$$\nabla f(u) = B^T \underbrace{(A^T)^{-1} (y - y_d)}_{=: p} + \gamma u$$

to simplify the expression, we

introduce $p = (A^T)^{-1}(y - y_d)$

the adjoint state, and

$$A^T p = y - y_d \quad \text{(adjoint equation)}$$

<u>Theorem 1:</u> Let $\bar{u}$ be a sol. to (CP)

then, there exists an adj. state p s.th.

(1) $A\bar{y} = B\bar{u}$ (SE)

(2) $A^T p = \bar{y} - y_d$ (AE)

(3) $\langle \underbrace{B^T p + \gamma \bar{u}}_{\nabla f(\bar{u})}, u - \bar{u} \rangle \geq 0$

$\forall u \in \mathcal{U}_{ad}$

- Introduce Lagrangian

$$\mathcal{L}(y, p, u) = J(y, u) - \langle Ay - Bu, p \rangle$$

$$= \frac{1}{2} | y - y_d |^2 + \frac{\gamma}{2} |u|^2 - \langle Ay - Bu, p \rangle$$

$$\mathcal{L}_y = y - y_d - A^T p \overset{!}{=} 0 \rightsquigarrow (AE)$$

$$\mathcal{L}_p = -(Ay - Bu) \overset{!}{=} 0 \rightsquigarrow (SE)$$

$$\mathcal{L}_u = \gamma u + B^T p$$

In other words $\mathcal{L}_p(\bar{y}, \bar{p}, \bar{u}) = 0$ (SE)

$\mathcal{L}_y(\bar{y}, \bar{p}, \bar{u}) = 0$ (AE)

$\langle \mathcal{L}_u(\bar{y}, \bar{p}, \bar{u}), u - \bar{u} \rangle \geq 0$ (VI)

Box constraints and KKT conditions

$$\mathcal{U}_{ad} = \{ u \in \mathbb{R}^m \mid u^a \leq u \leq u^b \}$$

with $u^a \leq u^b$ and $u^a, u^b \in \mathbb{R}^m$

to be understood componentwise, i.e.

$$u_i^a \leq u_i \leq u_i^b, \quad 1 \leq i \leq m$$

(CP) $\quad \min f(u)$

$$\text{s.t.}\quad \left.\begin{aligned} c^1(u) &= u - u^a \geq 0 \\ c^2(u) &= u^b - u \geq 0 \end{aligned}\right\} \begin{matrix} 2m \\ \text{ineqs.} \end{matrix}$$

LICQ (Linear indep. constraint qualif.)

$\{ \nabla c_i, i \in A(\bar{u}) \}$ l.i.

$$\frac{\partial c_i^1}{\partial u_k} = \delta_{ik} \qquad \frac{\partial c_i^2}{\partial u_k} = -\delta_{ik}$$

we can have at most $\begin{pmatrix} \pm 1 & & 0 \\ & \ddots & \\ 0 & & \pm 1 \end{pmatrix}$

$\Longrightarrow$ (LICQ) holds

introduce Lagr.

$$\mathcal{L}(u, \mu^a, \mu^a) = f(u) - \langle \mu^a, u - u^a \rangle$$
$$- \langle \mu^b, u^b - u \rangle$$

i) compute Lagrange multipliers

$$\nabla_u \mathcal{L}(\bar{u}, \mu^a, \mu^b) = 0 = \nabla f(\bar{u}) - \mu^a + \mu^b$$

$$\Rightarrow \quad \frac{\partial f}{\partial u_k} = \mu_k^a - \mu_k^b \quad \text{and}$$

$$\mu^a \text{ and } \mu^b \geq 0$$

$$\frac{\partial f}{\partial u_k} \geq 0 \Rightarrow \mu_k^a = \frac{\partial f}{\partial u_k}$$

$$\mu_k^b = 0$$

$$\frac{\partial f}{\partial u_k} \leq 0 \Rightarrow \mu_k^a = 0$$

$$\mu_k^b = -\frac{\partial f}{\partial u_k}$$

$$\mu^a = [\nabla f(\bar{u})]_+ \quad \mu^b = [\nabla f(\bar{u})]_-$$

where $[x]_+ = \max\{0, x\}$

$[x]_- = -\min\{0, x\}$

and $\quad x = [x]_+ - [x]_-$

(ii) Complementarity conditions

exploit var. ineqn.

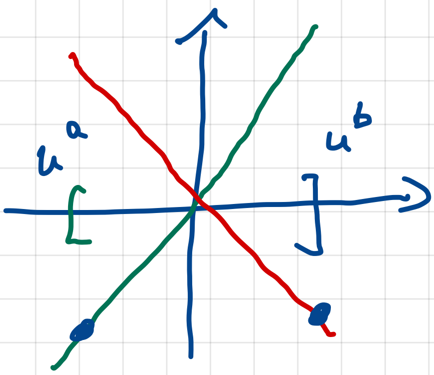
$$\langle \nabla f(\bar u), \tilde u - \bar u \rangle \geq 0 \quad \forall\, \tilde u \in \mathcal{U}_{ad}$$

$$\tilde u_i = \begin{cases} \bar u_i, & \text{for } i \neq k \\ u, & \text{for } i = k \end{cases} \quad \in \mathcal{U}_{ad}$$

$$\frac{\partial f(\bar u)}{\partial u_k}(u - \bar u_k) \geq 0 \quad \forall\, u \in [u_k^a, u_k^b]$$

$$\Longleftrightarrow \quad \frac{\partial f(\bar u)}{\partial u_k}\,\bar u_k \leq \frac{\partial f(\bar u)}{\partial u_k}\,u \quad \forall\, u \in [u_k^a, u_k^b]$$

$$\Longleftrightarrow \quad \frac{\partial f(\bar u)}{\partial u^k}\,\bar u_k = \min_{u \in [u_k^a, u_k^b]} \frac{\partial f(\bar u)}{\partial u_k}\,u$$



$$\Longleftrightarrow \quad \bar u_k = \begin{cases} u_k^a & \text{if } \frac{\partial \bar f}{\partial u_k} > 0 \\ u_k^b, & \text{if } \frac{\partial \bar f}{\partial u_k} < 0 \end{cases} \qquad (*)$$

now, consider compl. cond. :

$$\mu_k^a (\bar{u}_k - u_k^a) = 0$$

if $\bar{u}_k > u_k^a \overset{(*)}{\Longrightarrow} \frac{\partial f(\bar{u})}{\partial x_k} \leq 0$

$$\Longrightarrow \mu_k^a = 0$$

if $\mu_k^a > 0 \Longrightarrow \frac{\partial f(\bar{u})}{\partial u_k} > 0$

$$\overset{(*)}{\Longrightarrow} \bar{u}_k = u_k^a$$

$\Longrightarrow$ complement. cond. holds,

all in all, we have shown

<u>Theorem 2:</u> If $\bar{u}$ is sol. to (P) with box constr., with p adj. state, A regular, then there exist Lagr. multipliers μ^a, μ^b satisfying

$$\nabla_y \mathcal{L}(\bar{y}, \bar{u}, \bar{p}, \mu^a, \mu^b) = 0$$

$$\nabla_u \mathcal{L}(\bar{y}, \bar{u}, \bar{p}, \mu^a, \mu^b) = 0$$

$$\mu^a \geq 0, \quad \mu^b \geq 0$$

$$\langle u^a - \bar{u}, \mu^a \rangle = \langle \bar{u} - u^b, \mu^b \rangle = 0$$