Nonsmooth Nonconvex Optimization Problems

Kazufumi Ito, North Carolina State University

July 2, 2015, 9th Workshop on Control of Distributed Parameter Systems, Beijing Institute of Technology

Non-smooth Optimization Let $X \subset L^2(\Omega)$ be a Banach space.

$$\min \quad \int_\Omega (\ell(y) + h(u))\, d\omega$$

subject to the equality constraint for $y \in X$ and $u \in \mathcal{C}$

$$E(y) + f(y,u) = 0 \quad \text{in } X^*.$$

- $L^1(\Omega)$-minimization $h(u) = |u|$, Data Mining.
- $L^\infty(\Omega)$-minimization, $|u|_{L^\infty(\Omega)}$
- $L^0(\Omega)$-optimzation, $h(u) = |u|^0$ or $meas\{u \neq 0\}$=Ekeland metric, Volume control and Scheduling.
- Time optimal control with target constraint, State constraint.
- Binary, Mixed Integer optimization $h(u) = W(u)$.
- Mixed minimum effort and time optimal control problem.
- Stochastic control problem with $L^1(\Omega)$ control.

Inverse medium problem C: the observation map

$$\min \quad \int_\Gamma \frac{1}{2}|Cy-z|^2 + \beta \int_\Omega h(u)\,d\omega$$

with constraints

$$\Delta y + n^2(x)\,k^2 y = 0, \quad u = n^2 k^2 \ \text{(Defraction index in Helmhortz)}$$

$$-\Delta y + V(x) y = 0, \quad u = V \ \text{(Potential in Schrödinger)}$$

$$\mathrm{div}(\sigma\,\nabla u) = 0, \quad u = \sigma \ \text{(Conductivity in Tomography).}$$

Optimal Control Problem

$$\min \quad \int_0^T (\ell(y(t)) + h(u(t)))\,dt + G(y(T))$$

$$\frac{\partial}{\partial t} y - \mathrm{div}(\sigma\nabla y + b\,y) = 0$$

$$\nu\cdot(\sigma\nabla y + b\,y) = u$$

Multi-parameter regularization

$$|K(u)-y|_\Gamma^2+\alpha\int_\Omega |u|\,dx+\beta\int_\Omega |\nabla u|^2\,dx$$

$$|K(u)-y|^2+\alpha\int_\Omega |\nabla u|\,dx+\beta\int_\Omega |\Delta u|\,dx$$

Inverse medium problem and Direct sampling method

$$\Delta E+n^2\,k^2 E=0,\quad E^{inc}=e^{ik\,x\cdot d}$$

Unknown: Refractive index n^2, Observation: Near Field Scattering Data $E^s(x), x\in\Gamma$.

Distribution of $x_p\in\tilde{\Omega}\to\eta=n^2-1$

$$\Phi(x_p)=\frac{|\langle E^s(x),G(x,x_p)\rangle_{L^2(\Gamma)}|}{\|E^s(x)\|_{L^2(\Gamma)}\|G(x,x_p)\|_{L^2(\Gamma)}}.$$

Let $u=(n^2-1)k^2$ and define the induced current $I=u\,E$.

$$I(x)=u\,E^{inc}+u\int_{\Omega'} G(x,y)I(y)dy.$$

$$\min \quad \frac{1}{2}\int_\Gamma |Ku - E_s|^2\, ds + \int_{\Omega'} (\alpha\,|u|\,dx + \frac{\beta}{2}\,|\nabla u|^2)\,dx$$

where

$$Ku = \int_{\Omega'} G(x,y)\hat{E}^{tot}(y)\,u(y)\,dy,$$

$$K^*Ku - \beta\Delta u - K^*E_s + \alpha\,\lambda = 0, \qquad \lambda = \frac{\lambda + c\,u}{\max(1, |\lambda + c\,u|)}.$$

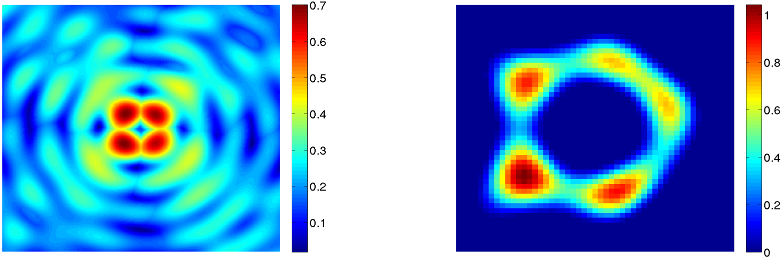

Figure: Classification Index and Reconstruction

Stochastic Control problem

$$\min \quad E^{0,x}\left[\int_0^{\tau} e^{-ct}(f(X_t)+|u_t|)\,dt\right]$$

subject to

$$dX_t = (b(X_t) + Gu_t)\,dt + \sigma\,dB_t \tag{1}$$

over $u_t \in \{\mathcal{F}_t-$adapted integrable control, $|u_t| \le \gamma$, a.e$\}$ and $\tau > 0$ is the exit time of the Ito diffusion X_t from a domain Ω. Let $\mathcal{L}$ is the generator of X_t

$$\mathcal{L}\phi = \frac{1}{2}a_{i,j}(x)\phi_{x_i x_j} + b_i(x)\phi_{x_i}.$$

Let V be the solution to HJB:

$$\mathcal{L}V - c\,V + f + \gamma\min(0, 1 - |G^t\nabla V|) = 0, \quad u = 0 \quad \text{at } x \in \partial\Omega.$$

Feedback synthesis

$$V_t = \alpha(x_t) = \begin{cases} -\gamma \frac{G^t \nabla V}{|G^t \nabla V|}, & x \in \{|G^t \nabla V| \geq 1\} \\ 0, & otherwise \end{cases}$$

If $\gamma \to \infty$ one can show that $V = \lim V_\gamma$ satisfies

Quasi-variational inequality

$$\max(-\mathcal{L}V + c\,V - f, |G^t \nabla V| - 1) = 0.$$

Let $F_0 : X \to R$ be C^1 and F_1 be convex and $\mathcal{C}$ is a closed convex set in X:

$$\min \quad F_0(x) + F_1(x) \quad \text{over} \;\; x \in \mathcal{C}.$$

Variational Inequality Suppose

$$F_0'(x^*)(x - x^*) + F_1(x) - F_1(x^*) \geq 0 \text{ for all } x \in \mathcal{C}.$$

If $\mathcal{C} = X$, then

$$-F_0'(x^*) \in \partial F_1(x^*)$$

The subdifferential of convex functional is defined by

$$\partial F_1(x^*) = \{p \in X^* : F_1(x) - F_1(x^*) \geq \langle p, x - x^* \rangle\}.$$

Pointwise minimum principle (let $p = -F_0'(x_0)$)

$$-pu + \alpha\,|u| \quad \text{over } u \in U.$$

is minimized at $u^*(x)$.

Quadratic Sparsity optimization

$$\min \quad J(x) = \frac{1}{2}|Ax - b|^2 + \alpha\,(Px, x) + \beta\, N_p(x)$$

$$N_p(x) = \sum_{k=1}^{\infty} |x_k|^p, \quad 0 \le p \le 1.$$

with

$N_p(x) \to N_0(x) =$ number of nonzero elements of x as $p \to 0^+$.

For $p = 1$

$$-A^*(Ax - b) - \alpha\, Px \in \beta \partial N_1(x)$$

Equivalently,

$$\begin{cases} A^*(Ax - b) + \alpha\, Px + \beta\, \lambda = 0 \\[2ex] \lambda = \dfrac{\lambda + c\, x}{\max(1, |\lambda + c\, x|)}. \end{cases}$$

Existence: R(A) closed, N(A) finite.

Necessary optimality ($p = 0,\ \ell^0-$optimization)

$$A^*(Ax - b) + \alpha\, Px + \lambda = 0,$$

with complementarity condition:

$$\begin{cases} \lambda_k = 0, & \text{if } k \in \{k : |\lambda_k + C_k x_k|^2 > 2\beta\, C_k\} \\ x_j = 0, & \text{if } j \in \{j : |\lambda_j + C_j x_j|^2 \le 2\beta\, C_j\} \end{cases}$$

Primal-Dual Active Set Method With $C_k = |A_k|_2^2 + \alpha\, P_{kk}$.

$$A^*(Ax^{n+1} - b) + \alpha\, Px^{n+1} + \lambda^{n+1} = 0, \qquad (2)$$

where

$$\begin{aligned} \lambda_k^{n+1} &= 0, \quad \text{if } k \in \{k : |\lambda_k^n + C_k x_k^n|^2 > 2\beta\, C_k\} \\ x_j^{n+1} &= 0, \quad \text{if } j \in \{j : |\lambda_j^n + C_j x_j^n|^2 \le 2\beta\, C_j\} \end{aligned} \qquad (3)$$

until a convergence criterion is satisfied.

Example Let A be the map from control $u \in L^2(0, T; R^2)$ to the final state $y(T, x) \in L^2(0,1)$ of the (normalized) one dimensional heat control system for $y = y(t, x)$:

$$y_t = y_{xx} + b_1(x)\, v_1(t) + b_2(x) v_2(t), \quad x \in (0,1)$$

with $y(t,0) = y(t,1) = 0$ and

$$b_1(x) = \chi_{(.3,.4)}, \quad b_2(x) = \chi_{(.6,.7)}.$$

The fourth order finite difference approximation with mesh-size $\Delta x = 1/50$ is used. Two piecewise constant controls $u = (v_1, v_2)$ is regularized by ℓ^0 norm in time. and We let the target function as a Gaussian distribution $y_d(x) = exp(-100(x-.7)^2)$ centered at $x = .7$.

β :	.001	.003	.005	.007	.01	.03	.05	.07	.1	.3	.5
iterates :	1	3	16	7	4	29	7	4	2	2	2
inactive :	98	95	79	72	68	22	13	9	7	2	1

$L^0(\Omega)$ Optimization Let $X \subset L^2(\Omega)$ be a Banach space.

$$\min \quad \int_\Omega (\ell(y) + h(u))\, d\omega$$

subject to the equality constraint for $y \in X$ and $u(\omega) \in U$

$$E(y) + f(y, u) = 0 \quad \text{in } X^*.$$

$$N_0(u) = \int_\Omega |u|^0\, d\omega \text{ with } |u|^0 = 1, \;\; u \neq 0, \;\; \text{and } |0|^0 = 0,$$

defines a complete (Ekeland) metric on measurable functions. Define $N_p(u)$

$$N_p(u) = \int_\Omega |u|^p\, d\omega$$

Then $N_0(u) = \lim_{p\to 0^+} N_p(u)$. $N_p,\ 0 \leq p < 1$ is ”not” weakly sequentially lower semi-continuous.

Define the Hamiltonian

$$\mathcal{H}(\omega, y, u, p) = \ell(\omega, y) + h(u) + p \cdot f(\omega, y, u).$$

Theorem (Pointwise Maximum Principle) Suppose $(y^*, u^*) \in X \times U_{ad}$ is optimal for problem and u^* is integrable, and that

$$E'(y^*) + f_y(y^*, u^*) : X \to X^* \text{ is bounded invertible.}$$

Let $p \in X$ satisfies the adjoint equation

$$(E'(y^*) + f_y(y^*, u^*))^* p + \ell'(y^*) = 0$$

Then, we have the necessary optimality

$$\mathcal{H}(y^*(\omega), u, p(\omega)) - \mathcal{H}(y^*(\omega), u^*(\omega), p(\omega)) \geq 0 \text{ for all } u \in U, \text{ a.e. in } \Omega.$$

That is, $u^*(\omega)$ minimizes $\mathcal{H}(y^*(\omega), u, p(\omega))$ over $u \in U$.

$$h(u)=\frac{\alpha}{2}|u|^2+\beta|u|^0, \quad f=f(y)+Bu$$

$$\Phi(q)=argmin_u\,(h(u)-(q,u))=\begin{cases}\frac{q}{\alpha} & \text{for } |q|>\sqrt{2\alpha\beta}\\ \frac{q}{\alpha},0 & \text{for } |q|=\sqrt{2\alpha\beta}\\ 0 & \text{for } |q|<\sqrt{2\alpha\beta}\end{cases}$$

Conjugate function $h^*(q)$

$$-h^*(q)=h(\Phi(q))-q\,\Phi(q)=\begin{cases}-\frac{1}{2\alpha}|q|^2+\beta & \text{for } |q|\geq\sqrt{2\alpha\beta}\\ \\ 0 & \text{for } |q|<\sqrt{2\alpha\beta}.\end{cases}$$

and h^* is convex.

Bi-conjugate function $h^{**}(u)$

$$h^{**}(u)=\begin{cases}\frac{\alpha}{2}|u|^2+\beta & |u|>\sqrt{\frac{2\beta}{\alpha}}\\ \\ \sqrt{2\alpha\beta}\,|u| & |u|\leq\sqrt{\frac{2\beta}{\alpha}}\end{cases}$$

h^{**} is a covexification of h. Clearly $\Phi : R \to R$ is monotone, but it is not maximal monotone.

$$\partial h^{**}(u) = \begin{cases} \alpha\, u & \text{for } |u| \geq \sqrt{\frac{2\beta}{\alpha}} \\ -\sqrt{2\alpha\beta} & \text{for } -\sqrt{\frac{2\beta}{\alpha}} \leq u < 0 \\ [-\sqrt{2\alpha\beta}, \sqrt{2\alpha\beta}] & \text{for } u = 0 \\ \sqrt{2\alpha\beta} & \text{for } 0 < u \leq \sqrt{\frac{2\beta}{\alpha}}. \end{cases}$$

The maximal monotone extension $\tilde{\Phi} = \partial h^{*}(q)$ of Φ and is given by

$$\tilde{\Phi}(q) = \begin{cases} \frac{q}{\alpha} & \text{for } |q| > \sqrt{2\alpha\beta} \\ 0 & \text{for } |q| < \sqrt{2\alpha\beta} \\ [\frac{q}{\alpha}, 0] & \text{for } q = -\sqrt{2\alpha\beta} \\ [0, \frac{q}{\alpha}] & \text{for } q = \sqrt{2\alpha\beta}. \end{cases}$$

Relaxed convex optimization

$$\min \quad J(y,u)=\int_\Omega (\ell(y)+h^{**}(u))\,d\omega \text{ over } u\in\mathcal{U}. \qquad (4)$$

subject to the equality constraint. Since h^{**} is convex, there exists a minimizer (y,u) to the relaxed problem and we have the necessary optimality

$$\begin{cases} Ey+f(y)-B\tilde{\Phi}(B^*p)=0 \\ \\ (E+f'(y))^*p+\ell'(\cdot,y)=0. \end{cases}$$

In a natural way $\tilde{\Phi}$ defines an operator from $L^2(\Omega)$ to itself, which is maximal monotone. It will be denoted by the same symbol.

Let $J(u)=\int_\Omega \ell(y(\omega))\,d\omega$ and

$$J(u)-J(u^*)-(B^*p,u-u^*)\geq \gamma\,|u-u^*|_{L^2(\Omega)}$$

for some $\gamma\geq 0$ in a neighborhood of u^*. For example, E is linear and ℓ is convex. Suppose there exists a solution to

$$\begin{cases} Ey-B\Phi(B^*p)=g \\ \\ E^*p+\ell_y=0. \end{cases}$$

For example, if $\ell_y=y$, then this reduces to

$$EE^*p+B\Phi(B^*p)+g=0.$$

Since $B\tilde{\Phi}(B^*p)$ is monotone (maximized), it has a unique solution. Then, the L^0 optimization has an optimal solution.

Example Let $\Omega = (0,1)$ and consider the control problem;

$$\min \quad \int_0^1 (\frac{1}{2}|x|^2 + \frac{\alpha}{2}|u|^2 + |u|^0)\, dt$$

subject to

$$\frac{dx}{dt} = u, \quad x(0) = x_0 > 0$$

Then, the necessary optimality is given by

$$\tfrac{dx}{dt} = -p \text{ on } [0, t^*) \quad \text{and} \quad 0 \text{ on } (t^*, 1]$$

$$\tfrac{dp}{dt} = -x, \quad p(t^*) = \sqrt{2\alpha} \text{ and } p(1) = 0.$$

where we used is the fact that $x,\ p \geq 0$.

If $\alpha = 1$ and $x_0 = 4$

$x(t) = Ae^t + Be^{-t}$ on $[0, t^*]$ and $x(t) = x^* = Ae^{t^*} + Be^{-t^*}$ on $[t^*, 1]$.

$p(t) = -Ae^t + Be^{-t}$ on $[0, t^*]$ and $p(t) = x^*(1 - t^*)$ on $[t^*, 1]$

where $A + B = x_0 = 4$. It reduces to

$$-Ae^{t^*} + (4 - A)e^{-t^*} = \sqrt{2}$$

$$(1 - t^*)(Ae^{t^*} + (4 - A)e^{-t^*}) = \sqrt{2}$$

and thus

$$A = \frac{4e^{-t^*} - \sqrt{2}}{e^{t^*} + e^{-t^*}}$$

$$(1 - t^*)(4e^{-t^*} + \tanh(t^*)(4e^{-t^*} - \sqrt{2})) = \sqrt{2}$$

Numerical calculation gives $t^* \sim 0.5077$.

$$\min \quad J(x) + \int_\Omega h^{**}(x(\omega))\, d\omega$$

Yoshida-Moreau approximation:

$$h_c(x,\lambda) = \inf_z \{h^{**}(z) + (\lambda, x - z)_H + \frac{c}{2}\,|x - z|^2\}.$$

$\lambda \in \partial h^{**}(x)$ iff $\lambda \in h'_c(x,\lambda)$ for all $c > 0$

$h'_c(x,\lambda) = p_\epsilon(\lambda) = \text{argmax}\{(p,x) - \frac{1}{2c}|p-\lambda|^2 - h^*(p)\}.$

The necessary optimality condition is written as

$$-J'(x) = \lambda$$

$$\lambda = h'_c(x,\lambda).$$

is the semi-smooth equation.

For the case of $h = \frac{\alpha}{2}|x|^2 + \beta|x|_0$

$$h'_c(x,\lambda) = \begin{cases} \lambda & |\lambda + cx| \le \sqrt{2\alpha\beta} \\ \sqrt{2\alpha\beta} & \sqrt{2\alpha\beta} \le \lambda + cx \le (1+\frac{c}{\alpha})\sqrt{2\alpha\beta} \\ -\sqrt{2\alpha\beta} & \sqrt{2\alpha\beta} \le -(\lambda + cx) \le (1+\frac{c}{\alpha})\sqrt{2\alpha\beta} \\ \frac{\lambda}{\alpha+c} & |\lambda + cx| \ge (1+\frac{c}{\alpha})\sqrt{2\alpha\beta}. \end{cases}$$

Thus, the semi-smooth Newton update: $-J'(x^{n+1}) = \lambda^{n+1}$ with

$$\begin{cases} x^{n+1} = 0 & |\lambda^n + cx^n| \le \sqrt{2\alpha\beta} \\ \lambda^{n+1} = \sqrt{2\alpha\beta} & \sqrt{2\alpha\beta} < \lambda^n + cx^n < (1+\frac{c}{\alpha})\sqrt{2\alpha\beta} \\ \lambda^{n+1} = -\sqrt{2\alpha\beta} & \sqrt{2\alpha\beta} < -(\lambda^n + cx^n) < (1+\frac{c}{\alpha})\sqrt{2\alpha\beta} \\ x^{n+1} = \frac{\lambda^{n+1}}{\alpha} & |\lambda^n + cx^n| \ge (1+\frac{c}{\alpha})\sqrt{2\alpha\beta}. \end{cases}$$

Primal-Dual Active Set Method

- Initialize $u^0 \in X$ and $\lambda^0 \in H$. Set $n = 0$.
- Solve for $(y^{n+1}, u^{n+1}, p^{n+1})$

  $$Ey^{n+1} + Bu^{n+1} = g, \quad E^* p^{n+1} + \ell'(y^{n+1}) = 0, \quad \lambda^{n+1} = -B^* p^{n+1}$$

  and

  $$u^{n+1} = -\frac{\lambda^{n+1}}{\alpha}, \text{ if } \omega \in \{|\lambda^n + c\, u^n| > (1 + \tfrac{c}{\alpha})\sqrt{2\alpha\beta}\}$$

  $$u^{n+1} = 0, \text{ if } \omega \in \{|\lambda^n + cu^n| \le \sqrt{2\alpha\beta}\}.$$

  $$\lambda^{n+1} = \pm\sqrt{2\alpha\beta}, \text{ if } \omega \in \{\sqrt{2\alpha\beta} \le |\lambda^n + cu^n| \le (1 + \tfrac{c}{\alpha})\sqrt{2\alpha\beta}\}.$$

- Stop, or set $n = n + 1$ and return to the second seep.

Extension:

$$\begin{cases} \cdot \;\; h(u) = \frac{\alpha}{2}\,|u|^2 + \beta\,|u|^p, \quad \alpha\, u + \beta p u^{1-p} = q. \\ \cdot \;\; \Phi(q) = argmin_{|u|\le 1}\,(\beta\,|u|^0 - (q, u)) = \begin{cases} \frac{u}{|u|} & \text{for } |q| \ge \beta \\ 0 & \text{for } |q| < \beta \end{cases} \end{cases}$$

Generalization Consider the minimization on a Banach space $X \subset L^2(\Omega)$;

$$\min \quad J(x) + \int_\Omega h(x(\omega))\,d\omega$$
$$\text{subject to } x \in \mathcal{C} = \{x(\omega) \in U \text{ a.e. }\}$$

Let h^{**} be the bi-conjugate function of h;

$$\begin{cases} h^*(\lambda) = \sup_{x \in U}(\lambda, x) - h(x) \\ h^{**}(x) = \sup_\lambda\{(x,\lambda) - h^*(\lambda)\}. \end{cases}$$

Let $\Phi(\lambda) = \text{argmax}_y\{(\lambda, y) - h(y)\}$, then

$$y \in \Phi(\lambda) = \partial h^*(\lambda) \text{ iff } \lambda \in \partial h(y) \quad \text{if } h \text{ is convex.}$$

In general $\partial h^*(\lambda)$ is the maximum monotone extension of Φ. Consider the relaxed problem

$$\min \quad J(y) + \int_\Omega h^{**}(y(\omega))\,d\omega. \tag{5}$$

$$-J'(y)(\omega) \in \partial h^{**}(y(\omega)).$$

Binary Optimization

$$\min \quad F(y) \quad \text{subject to } E(y,u)=0 \text{ and } u\in\{0,1\}^d$$

Transformation: for Q, a generator of Markov Chain and $W(u)=min(|u|,|u-1|)$

$$\min \quad F(y)+\beta\,(Qu,u)+c\,W(u) \quad \text{subject to } E(y,u)=0.$$

Algorithm

$$\begin{cases} E_y(y^n,u^n)p^n+F'(y^n)=0 \\ \dfrac{u^{n+1}-u^n}{\Delta t}+\beta\,Qu^{n+1}+cD(u^n)u^{n+1}+E_u(y^n,u^n)^*p^n=cf^n \end{cases}$$

Mixed Integer Optimization, Segmantatiom

$$W(u)=\min(|u|,|u-1|,\cdots,|u-N|) \text{ or } W(u)=|u||u-1|.$$

Ñ Coordinate-wise Descent method (Zhang Zhengru)

Data Mining

$$N_p(Au-b)+\beta\,(Qu,u) \quad \text{subject to } E(y,u)=0.$$

Non-smooth mathematical programming

$$\begin{cases} \min \ F(y) = F_0(y) + F_1(y) \\ \\ \text{subject to } G_1(y) = 0, \ \ G_2(y) \le 0, \ y \in \mathcal{C}, \end{cases} \tag{P}$$

where $F_0 : X \to R$ and $G_1 : X \to Y$ are C^1 mappings, $F_1 : X \to R$ and $G_2 : X \to Z$ are convex, and $\mathcal{C}$ is a closed convex set in X and denotes a set of additional constraints. Let $K \subset Z$ be a closed, convex cone with vertex at 0, which introduces an ordering on Z such that $u \le v$ if $u - v \in K$.

$$d_K(\hat{z}) = \inf_{z \in K} |z - \hat{z}|_Z.$$

It is convex and Lipschitz continuous. The convex sub-differential of d_K is defined by: for $\hat{z} \in Z$

$$\partial d_K(\hat{z}) = \{\xi \in Z^* : d_K(z) - d_K(\hat{z}) \ge \langle \xi, z - \hat{z} \rangle \text{ for all } z \in Z\}.$$

— Ekeland's variational principle.
For $\epsilon > 0$ define

$$J_\epsilon(y) = \Big(((F(y) - F(y^*) + \epsilon)^+)^2 + |G_1(y)|_Y^2 + d_K(G_2(y))^2 \Big)^{\frac{1}{2}},$$

There exists a $y^\epsilon \in \mathcal{C}$ such that

$$\begin{cases} J_\epsilon(y^\epsilon) \le J_\epsilon(y^*), \\ J_\epsilon(y) - J_\epsilon(y^\epsilon) \ge -\sqrt{\epsilon}\, d(y, y^\epsilon) \quad \text{for all } y \in \mathcal{C}, \\ d(y^\epsilon, y^*) \le \sqrt{\epsilon}. \end{cases} \tag{6}$$

— demi-continuity of the convex sub-differential $\partial d_K(z)$ and $\partial |y|_Y$. (sufficient to assume Y and Z are strictly convex). For example $L^p(\Omega)$ and $C(\Omega)$ (after re-normalization).

— Regular point condition (Extension of the Maurer-Zowe-Kurcyusz condition)

$$0 \in \text{int} \left\{ \begin{pmatrix} G_1'(y^*)(y - y^*) \\ G_2(y) - k \end{pmatrix} : y \in \mathcal{C} \cap domF_1,\ k \in K \right\}, \quad (7)$$

Theorem Let $y^* \in \mathcal{C}$ be a local minimum of (P). Then there exists a nontrivial $(\lambda_0, \mu_1, \mu_2) \in R^+ \times Y^* \times Z^*$ such that

$$\lambda_0 \left(F_0'(y^*)(y - y^*) + F_1(y) - F_1(y^*) \right) + \langle\, \mu_1, G_1'(y^*)(y - y^*)\, \rangle_{Y^*,Y}$$

$$+ \langle\, \mu_2, G_2(y) - G_2(y^*)\, \rangle_{Z^*,Z} \geq 0 \text{ for all } y \in \mathcal{C} \cap domF_1,$$

$$\langle \mu_2, z \rangle_{Z^*,Z} \leq 0 \text{ for all } z \in K,\ G_2(y^*) \leq 0,\ \langle\, \mu_2, G_2(y^*)\, \rangle_{Z^*,Z} = 0. \quad (8)$$

Moreover, $\lambda_0 \neq 0$.

Application: Semilinear Control problem

Let $X,\ U$ be Hilbert spaces. Consider the semilinear control system

$$\frac{d}{dt}x(t) = Ax(t) + F(x(t)) + B(x(t))u(t). \tag{9}$$

Assume that A is the infinitesimal generator of C_0 semigroup $S(t),\ t \geq 0$ on X and $F : X \to X$ is locally Lipschitz and $B(x) : X \to \mathcal{L}(U, X)$ is Lipschitz. There exists $\alpha \in R$ such that

$$(Ax + F(x), x) \leq \alpha\, |x|^2 \quad \text{for all } \ x \in dom(A), \tag{10}$$

$$|B(x)u|_X \leq (b_1 + b_2\, |x|_X)|u|_U \text{ for } b_1,\ b_2 \geq 0. \tag{11}$$

Given initial condition $x_0 \in X$ and control $u \in L^1(0, T; U)$, $x \in C(0, T; X)$ is a mild solution to (9) if x satisfies

$$x(t) = S(t)x_0 + \int_s^t S(t-s)(F(x(s)) + B(x(s))u(s))\, ds. \tag{12}$$

Theorem Given $x_0 \in X$ and $u \in L^1(0, T; U)$, (9) has a unique mild solution.

Minimum effort problem

$$\min \quad \int_0^T \ell(x(t)) + \psi(u(t))\,dt + \frac{\beta}{2}\,\gamma^2$$

subject to (12), and the control constraint

$$u \in U_{ad} = \{|u(t)|_\infty \leq \gamma\} \tag{13}$$

and the state constraint $g : X \to R^p$ is convex

$$\int_0^T w(t)g(x(t))\,dt \leq 0$$

Formulation $X = Y = L^2(0,1;X_0)$

$$G_1(x,u) = S(t)x_0 + \int_s^t S(t-s)(F(x(s)) + B(x(s))u(s))\,ds - x(t)$$

and $g : X \to R^p$ is convex

$$G_2(x) = \int_0^T w(t)g(x(t))\,dt \leq 0$$

By Theorem there exists a Lagrange multiplier $(\lambda, \vec{\mu}) \in L^2(0, T; X_0^*) \times (R^+)^p$ at (x^*, u^*). Define

$$p(t) = \int_t^T S^*(s-t)\lambda(s)\, ds.$$

Thus by Theorem

$$\int_0^T [\, \ell(x(t)) - \ell(x^*(t)) + \vec{\mu}^T w(t) \,(g(x(t)) - g(x^*(t))) \\ -\langle \lambda(t), x(t) - x^*(t)\, \rangle \,]\, dt$$

$$+ \int_0^T [\, \psi(v(t)) - \psi(u^*(t)) + (B^* p(t), v(t) - u^*(t))_U \,]\, dt \geq 0,$$

for all $x \in L^2(0, T; X_0)$ and $v \in U_{ad}$. Thus, for a.e. $t \in (0, 1)$

$$\ell(x) - \ell(x(t^*)) + \vec{\mu}^T w(t)(g(x) - g(x^*(t))) - \langle \lambda(t), x - x^*(t)\, \rangle \geq \; 0.$$

and thus

$$\lambda(t) \in \partial\big(\ell(x^*(t))\big) + \vec{\mu}^T\, w(t) \partial\big(g(x^*(t))\big),$$

That is, there exist

$$q(t) \in \partial \ell(x^*(t)), \quad z(t) \in \partial g(x^*(t)),$$

such that

$$\lambda(t) = q(t) + \vec{\mu}^T w(t)\, z(t)$$

and

$$-\frac{d}{dt}p = A^* p(t) + q(t) + \vec{\mu}^T w(t)\, z(t), \quad p(T) = 0.$$

$$\int_0^T [\, \psi(v(t)) - \psi(u^*(t)) + (B^* p(t), v(t) - u^*(t))_U \,]\, dt \geq 0$$

for all $v \in U_{ad}$, and therefore for a.e. $t \in (0, 1)$

$$u^*(t) = \mathrm{argmin}_{v \in U} \big(\, \psi(v) + (B^* p(t), v)_U \, \big).$$

Time optimal control

$$\min \quad \tau + \frac{\beta}{2}\gamma^2$$

subject to (12) and constraints (13) and the target constraint

$$C(x(T) - \bar{x} = 0, \quad |x(t) - \bar{x}|^2 \leq \delta^2. \tag{14}$$

Transformation Let $t = \tau\, s$ and $u = \gamma\, u$. We have the transformed (equivalent) formulation

$$\frac{dx}{ds} = \tau\,(Ax(s) + F(x(s)) + B(x(s))(\gamma\, v(t)))$$

subject to

$$v \in U_{ad} = \{|v(s)|_\infty \leq 1\}$$

and

$$C(x(1) - \bar{x})) = 0, \quad |x(1) - \bar{x}|_X^2 \leq \delta^2$$

In general one can consider the parameterized control

$$\frac{dx}{dt} = A(q)x(s) + F(x(s), q) + B(x(s), q)u(t)$$

for $q \in Q_{ad}$.

Technical Issues (1) The feasibility (constrained controllability) and the Regular point condition:
(2) Under the regularity of solution and/or the adjoint p.

$$(F_q + \int_0^1 (p(t), A_q x^*(t) + F_q(x^*(t)) + B_q(x^*(t))\,dt)(q - q^*) \geq 0$$

for all $q \in Q_{ad}$. Let $p(t)$ be the adjoint (for linear case) defined by

$$p(t) = S^*(1-t)C^*\mu$$

Necessary optimality condition:

$$1 + \int_0^1 (p(s), Ax(s) + F(x(s)) + Bv(s))\,dt = 0$$

reduces to the transversality condition $\mathcal{H}(t) = 0$ under assumption: either $x(s) \in C(0,1; dom(A))$ or $p(s) \in C(0,1; dom(A^*))$.

L^∞-norm Minimization

$$\min_{y\in X_0} \quad |\Lambda y|_{L^\infty} + F_0(y) \text{ subject to } G_1(y) = 0, \tag{15}$$

where $F_0 \in C^1(X_0, R),\ G_1 \in C^1(X_0, Y),\ \Lambda \in \mathcal{L}(X_0, L^2(\Omega))$.
Problem (15) can be equivalently be expressed as

$$\begin{aligned} &\min_{\gamma\in R^+, y\in X_0} \quad \gamma + F_0(y) \\ &\text{subject to } G_1(y) = 0,\ \text{and } |\Lambda y|_{L^\infty} \le \gamma. \end{aligned} \tag{16}$$

Transformation:

$$\begin{cases} \min_{\gamma\in R^+, z\in X_0} \quad \gamma + F_0(\gamma z) \\ \text{subject to } G_1(\gamma z) = 0, \quad z \in \mathcal{C}, \end{cases}$$

where $\mathcal{C} = \{z \in X_0 : |\Lambda z|_{L^\infty} \le 1\}$.

Formulation

$$x=(\gamma,y)\in X=R\times X_0,\quad F(\gamma,y)=\gamma+F_0(y).$$

Optimality condition: there exists $\mu\in Y^*$ such that for all $\gamma>0$

$$\langle F_0'(\gamma^*z^*)+G_1'(\gamma^*z^*)^*\mu,\phi-z^*\rangle_{X_0^*,X_0}\geq 0,\ \text{for all }\phi\in\mathcal{C}$$

$$\big(1+\langle F_0'(\gamma^*z^*),z^*\rangle_{X_0^*,X_0}+\langle\mu,G_1'(\gamma^*z^*)z^*\rangle_{Y^*,Y}\big)(\gamma-\gamma^*)\geq 0,$$

$$\langle F_0'(y^*)+G_1'(y^*)^*\mu,\phi-y^*\rangle_{X_0^*,X_0}\geq 0,\ \text{for all }|\Lambda\phi|_{L^\infty}\leq\gamma^*$$

$$\big(\gamma^*+\langle F_0'(y^*),y^*\rangle_{X_0^*,X_0}+\langle\mu,G_1'(y^*)y^*\rangle_{Y^*,Y}\big)(\gamma-\gamma^*)\geq 0.$$

If $\Lambda=Id$, $X_0=L^\infty$ and $F_0'(y^*)+G_1'(y^*)^*\mu\in L^1(\Omega)$ then the variational inequality in above can be expressed as

$$\begin{cases}(F_0'(y^*)+G_1'(y^*)^*\mu)(x)=0, & a.e.\text{ on }\{\,|y^*(x)|<\gamma^*\}\\ y^*(x)=-\gamma^*sgn(F_0'(y^*)+G_1'(y^*)^*\mu)(x)), & a.e.\text{ on }\{\,|y^*(x)|=\gamma^*\}.\end{cases}$$