Optimal control of state constrained PDEs system with Sparse controls.

Kazufumi Ito
Dept. of Math, North Carolina State University

Abstract: In this talk we discuss point-wise state constraint problems with sparse controls for a general class of optimal control problems. We use the penalty formulation and derive the necessary optimality condition based on the Lagrange multiplier theory. The existence of Lagrange multiplier associated with the point-wise state constraint as a measure is established. Also we develop a semi-smooth Newton method for the penalty formulation. Numerical tests are presented for parabolic and elliptic control problems. The results show that the state constraint optimal control method enables us to develop a much powerful and useful control law. For example one can minimize the extreme events and prevent blow up solutions for semi-linear equations. We also extend the results for point-wise gradient constraint

State constraint optimal control problem

We discuss a point-wise state constraint problem with sparse controls for a general class of optimal control problems, including semi linear elliptic, parabolic equations and wave equation. State constrained optimal control problems have presented a challenge for some time. Earlier works focused on the derivation of first order optimality conditions. Lavrentiev-type regularization approach is used for a entire domain control case in to the lack of the regularity and existence of the multiplier associated with the point-wise state constrain. It deduces to the mixed control-state type point wise constraints.

In this talk the control inputs enters in a partial domain or are of a finite-rank and thus it is is much more general and practical. In order to remedy of such problems.

- We use the penalty formulation [Ito-Kunisch] and derive the corresponding necessary optimality condition for a wide class of point-wise state constraint.
- Then, we establish a necessary optimality condition with measure-valued Lagrange multiplier associated with the point-wise state constraint by letting penalty parameter to the infinity. The penalty formulation is very general and powerful.
- The uniform L^1 estimate of Lagrange multiplier for the penalty formulation is established for a general class of point-wise state constraint optimal control problems. It is the one of our major contribution.
- The other contribution is to use the semi-smooth Newton method for the saddle point problem for the penalized problem. It requires to solve a sequence of symmetric saddle point linear systems.. An Algorithm based on our formulation is very general to solve a wide class of point-wise sate constraint problems and it is observed that it is rapidly convergent and globally convergent. Numerical tests demonstrate the feasibility of the penalty formulation and the algorithm. It enables us to develop a unified approach for a state constraint optimal control problem.

Problem formulation Consider the state constraint optimal control problem

$$\min \quad \int_0^T \int_\Omega (H(u)+J(y))\,dxdt$$

subject to $E(y,u)=0$ and $y\le\psi$ (point-wise state constraint) E is an equality constraint as below and $\psi\ge 0$ is an obstacle function. J and H are cost functionals for state y and control u, respectively. Our approach is based on the penalty formulation: for $\epsilon>0$

$$\min \quad \int_0^T \int_\Omega (H(u)+J(y))+\frac{|(y-\psi)^+|^2}{2\epsilon})\,dxdt. \tag{0.1}$$

where $(y-\psi)^+=\max(0,y-\psi)$. One uses $H(u)=|u|_1$ (L^1) and $H(u)=|u|_0$ (L^0) for the sparsity. $Bu=I_{\tilde{\Omega}}u$ with a subdomain $\tilde{\Omega}$ for example.

- The sparsity method can determine an active sparse domain.
- One can minimize an extreme event by minimizing

  $$\min \quad \int_0^T (H(u)+J(y))\,dxdt+\frac{1}{2}\gamma^2 \text{ subject to } |y|\le\gamma$$

  and the penalty formulation is given by

  $$\min \quad \int_0^T \int_\Omega (H(u)+J(y))\,dxdt+\int_0^T \int_\Omega \frac{|(|y|-\gamma)^+|^2}{2\epsilon})\,dxdt+\frac{1}{2}\gamma^2$$

  The necessary condition gives $\gamma=\displaystyle\int_0^T\int_\Omega \frac{(y-\gamma)^+}{\epsilon}\,dxdt.$

- Norm constraint case (e.g., Navier Stokes)

  $$\min \quad \int_0^T (H(u)+J(y))\,dxdt+\int_0^T \frac{|(|\nabla y|_{L^2}-\gamma)^+|^2}{2\epsilon})\,dt+\frac{1}{2}\gamma^2$$

  $$\gamma=\frac{1}{|\nabla y|_{L^2}}\frac{(|\nabla y|_{L^2}-\gamma)^+}{\epsilon}$$

Examples Let A be the second order elliptic operator:

$$Ay = \nabla \cdot (A_0(x)\nabla y) + b(x) \cdot \nabla y + c(x)y$$

where matrix $A_0(x)$ is symmetric and uniformly positive definite, $b(x)$ and $c(x)$ are uniformly bounded. We use the fact that

$$(Ay, y^+) \leq 0 \text{ for } y \in H^1(\Omega). \tag{0.2}$$

For a bounded domain Ω we consider the Dirichlet, Neumann boundary conditions for state variable y

Parabolic semilinear equations

$$E(y,u) = \frac{d}{dt}y - Ay - F(y) - Bu - f, \quad y(0) = y_0 \text{ in } \Omega,$$

Elliptic semilinear equations

$$E(y,u) = Ay - F(y) - Bu - f \text{ in } \Omega, ,$$

Wave equation

$$E(y,u) = \frac{d^2}{dt^2}y - \Delta y - Bu - f, \quad y(0) = y_0, \ \ \frac{d}{dt}y(0) = v_0 \text{ in } \Omega.$$

Here, F is a semilinear operator which is a sum of monotone and Lipschitz operators and B is a bounded input operator from $U \to L^2(\Omega)$, where U is a Hilbert space. f is a souse function in $L^\infty(0,T;L^2(\Omega)$.For example a partial domain control $Bu = \chi_{\tilde{\Omega}}u$. The well-posedness of the control problems, i.e., a unique mild solution exists given initial conditions and control functions have been well studied in the literatures, e.g., via the semigroup approach, e.g., [Ito-Kappel]. For example, one can show that $y \in L^\infty((0,T) \times \Omega)$ given $u \in L^2(0,T,U)$ and $y_0 \in L^\infty(\Omega)$ for parabolic and elliptic case.

Parabolic case For the simplicity we consider $H(u) = \frac{1}{2}|u|^2$ (can be L^1 or L^0 control) Let $X_0 = H_0^1(\Omega)$ for the Dirichlet boundary and $X_0 = H^1(\Omega)$ for the Neumann boundary. Let $\langle \cdot, \cdot \rangle_{X_0^*, X_0}$ be the dual product. Let $y \in W = H^1(0, T, X_0^*) \cap L^2(0, T, X_0)$ is a weak solution to

$$\langle \frac{d}{dt} y - Ay - F(y) - Bu - f, \psi \rangle = 0 \text{ for all } \psi \in X_0 \text{ and a.e. } t \in (0, T). \tag{0.3}$$

Here,

$$\langle Ay, \psi \rangle = \int_\Omega (A_0(x) \nabla y, \nabla \psi) + b(x) \cdot \nabla y \psi + c(x) y \psi) \, dx$$

and it defines a bounded bilinear elliptic form on $X_0 \times X_0$. We assume a scalar value function F satisfies

$$(F(y_1) - F(y_2), y_1 - y_2) \leq \delta \, |y_1 - y_2|^2 \text{ all } y_1, \; y_2 \in X_0$$

It follows from e.g, [Ito-Kapple] that (0.3) has a unique solution $y \in W$, given $y_0 \in L^2(\Omega)$ and $u \in L^2(0, T; X)$.

Define the Lagrangian functional

$$L(y,u,p) = J(y)+\frac{1}{2}|u|^2+\frac{1}{2\epsilon}|(y-\psi)^+|^2_{L^2}+\int_0^T \langle Ay+Bu+f-\frac{d}{dt}y,p\rangle\,dt.$$

It follows from the Lagrange multiplier theory, e.g., [Ito-Kunisch] that the necessary optimality condition is given by

$$u(t) = -B^*p(t)$$

where $y,\ p\in W$ satisfies the state and adjoint equation:

$$\begin{aligned} &\tfrac{d}{dt}y = Ay+F(y)+Bu+f, \quad y(0)=y_0 \\ &-\tfrac{d}{dt}p = A^*p+F'(y)^*p+\lambda_\epsilon+J'(y), \quad p(T)=0 \end{aligned} \tag{0.4}$$

with

$$\lambda_\epsilon(t) = \frac{(y(t)-\psi)^+}{\epsilon} \geq 0..$$

We use the notation $s^+ = \max(0,s)$. Here, $(u_\epsilon, y_\epsilon, p_\epsilon)$ depends on $\epsilon > 0$ but we drop the dependency for the simplicity. Note that if there is a feasible control $u \in L^2(0,T;U)$ so that $y(t) \leq \psi$, then

$$\sqrt{\epsilon}\,p(t) \in W$$

uniformly in $\epsilon > 0$. In fact,

$$J(y_\epsilon)+\frac{|(y_\epsilon(t)-\psi)^+|^2}{\epsilon}+\frac{1}{2}|u_\epsilon|^2$$

is uniformly bounded in $\epsilon > 0$ and thus

$$\int_0^T |\sqrt{\epsilon}\lambda_\epsilon|^2\,dt \text{ is uniformly bounded in } \epsilon > 0.$$

Since the adjoint equation is linear, we consider the decomposition $p = p_1 + p_2$ where p_2 satisfies

$$-\frac{d}{dt}p_2 = A^*p_2+F'(y)^*p_2+J'(y).$$

$$-\frac{d}{dt}p_1 = A^*p_1 + F'(y)^*p_1 + \lambda_\epsilon$$

Multiplying $e^{2\omega t}p_1^-$ we have

$$-\frac{1}{2}\frac{d}{dt}(e^{2\omega t}|p_1^-|^2)+e^{2\omega t}\omega\,|p_1(t)^-|^2 = e^{2\omega t}\,((Ap_1^-,p_1^-)+(F'(y)^*p_1^-,p^-)+(p^-,$$

If $F'(y)-\omega \leq 0$,

$$e^{2\omega t}|p_1^-|^2 \leq 0.$$

Thus, $p_1(t) \geq 0$ and $p_2 \in H^1(0,T;dom(A^*))$ uniformly in $\epsilon > 0$. First, consider the Neumann boundary condition case $n \cdot (A_0\nabla y) = g$ and thus $n \cdot (A_0\nabla p) = 0$ at the boundary $\partial\Omega$. Without loss of generality we assume $A_0 = I$ for the clarity of our presentation through the remaining of the talk. We assume $b(x) = 0$ at $\partial\Omega$.

Next, we show

$$\int_0^T\int_\Omega \lambda_\epsilon\,dxdt \text{ is uniformly bounded in } \epsilon > 0.$$

First, integrating the adjoint equation, we have

$$\int_\Omega p_1(0)\,dx = \int_0^T\int_\Omega (F'(y),p_1(t)+\lambda_\epsilon(t))\,dxdt \tag{0.5}$$

Next, multiplying $-p$ and $y-\psi$ to the first and second equation of (0.4), respectively

$$\int_\Omega (y_0-\psi)p(0)\,dx + \int_0^T (\frac{\partial y-\psi}{\partial n},p)\,dt$$

$$= \int_0^T (-(A\psi+f,p) + \frac{|(y-\psi)^+|^2}{\epsilon} + |B^*p|^2$$

$$+(J'(y),y-\psi)+(-F(y)+F'(y)(y-\psi),p))\,dxdt. \tag{0.6}$$

Assume that for $\delta > 0$

$-F(y)+F'(y)(y-\psi) \geq -2\omega(y-\psi)-(A\psi+f) \geq \delta$ for $y \leq \psi$,

$y_0 - \psi < -\delta < 0$ and $\frac{\partial y-\psi}{\partial n} \leq 0$ at $\partial\Omega$. Since

$$((y-\psi)^+, p) \leq \frac{|(y-\psi)^+|^2}{2\epsilon} + \frac{1}{2}|\sqrt{\epsilon}p|^2,$$

we obtain

$$(p_1(0), \psi - y_0) + \int_0^T (-F(y) + F'(y)(y-\psi) - (A\psi + f))p_1(t)\, dt$$

$$+ \int_0^T (|u|^2 + \frac{|(y-\psi)^+|^2}{2\epsilon})\, dt \tag{0.7}$$

is uniformly bounded in $\epsilon > 0$.

Remark Concerning with our assumption on $\psi$and f we have the followings.

(1) Let $\bar{y}$ satisfy

$$A\bar{y} + f = 0.$$

If we define $\tilde{y} = y - \bar{y}$. then $\tilde{y}$ satisfies

$$\frac{d}{dt}\tilde{y} = A\tilde{y} + Bu, \quad \tilde{y}(0) = y - 0 - \bar{y}.$$

(2) If we discuss $e^{\omega t} y(t) \leq \psi$, then $\hat{y} = e^{\omega t} y$ satisfies

$$\frac{d}{dt}\hat{y} = (A + \omega)\hat{y} + B\hat{u}, \quad \hat{y}(0) = y_0.$$

where we assume $f = 0$.

(3) For $F(y) = -y^3$

$$-F(y) + F'(y)(y - \psi) = y^3 - 3 - 3y^2(y - \psi) = y^2(3\psi - 2y).$$

(4) For $\psi = c_1 + c_2\,(1 - x)x$ on $x \in \Omega = (0, 1)$

$$\Delta\psi = -c_2, \quad \frac{\partial\psi}{\partial n} = -c_2.$$

Assume that $u_\epsilon \to u$ in $L^2(0, T; U)$ and $y_\epsilon - \psi \to y - \psi$ in $L^\infty(0, T, \Omega)$. In fact, since $\lambda_\epsilon \in L^1((0, T \times \Omega)$ uniformly, it follows from the second equation in (0.4) that $\{p_\epsilon\}$ is a compact in $L^\infty(0, T; L^2(\Omega))$ and $u_\epsilon = -B^* p_\epsilon$ has a strong convergence subsequence $u_n \to u$ in $L^\infty(0, T; U)$. Thus, the corresponding sequence $y_n - \psi$ converges strongly to $L^\infty((0, T) \times \Omega)$. Then, there exits $\lambda \in L^\infty(\Omega)^*$ such that

$$\begin{cases} \frac{d}{dt}y = Ay + Bu + f, \quad u = -B^*p \\ \\ -\frac{d}{dt}p = A^*p + \lambda + J'(y), \quad \langle \lambda, y - \psi \rangle = 0 \\ \\ y \leq \psi \text{ and } \lambda \geq 0 \end{cases} \tag{0.8}$$

That is, since $\frac{|(y-\psi)^+|^2}{\epsilon}$ is bounded and $y \leq \psi$ and since $\lambda_\epsilon \geq 0$ and $L^1((0,T) \times \Omega)$ bounded $\lambda \geq 0$ in $L^\infty((0,T) \times \Omega)^*$. In fact, let $X = L^\infty((0,T) \times \Omega$. It follows from the fact that closed balls in X^* are w^*-compact (by the Banach-Alaoglu theorem there exists a w^*-accumulation point λ of λ_ϵ. The sequence $y_\epsilon - \psi$ and $y - \psi$ generate a separable subspace of X. Therefore we can assume without restriction that X itself is separable. Using the fact that for separable spaces the w^*-topology on w^*-compact subsets of X^* is metrizable, we conclude that there exists a subsequence λ_n such that $w^* - \lim \lambda_n = \lambda$. This implies

$$\lim_{n \to \infty} \langle \lambda_n, y_n - \psi \rangle = \langle \lambda, y - \psi \rangle$$

Since

$$0 \leq (\lambda_n, y_n - \psi) = (\lambda_n, y_\epsilon - y) + (\lambda_n - \lambda, y - \psi) \to 0$$

the complementarity $\langle \lambda, y - \psi \rangle = 0$ holds.

Dirichlet boundary case We consider the Dirichlet boundary condition $y = g(x)$ at $\partial\Omega$. Since $p(0) = 0$ at $\partial\Omega$, $\frac{\partial}{\partial n} p_1(t) \le 0$ at $\partial\Omega$. (0.5) becomes

$$\int_\Omega p_1(0)\, dx = \int_0^T \Big(\int_{\partial\Omega} \frac{\partial}{\partial n} p_i(t) + \int_\Omega \lambda_\epsilon\Big)\, dt$$

and (0.7) becomes

$$(p(0), y_0 - \psi)_\Omega - \int_0^T \Big(g - \psi, \frac{\partial p}{\partial n}\Big)_{\partial\Omega} + \int_0^T \Big(|u|^2 + \frac{|(y-\psi)^+|^2}{\epsilon}$$

$$+(J'(y), y - \psi) - (A\psi + f, p) + (-F(y) + F'(y)(y - \psi), p)\, dt.$$

Assume that $g - \psi \le -\delta < 0$. Using the same arguments as above we have

$$\int_\Omega p_1(0)(\psi - y_0)\, dx - \int_0^T \Big(\int_{\partial\Omega} (\psi - g)\frac{\partial}{\partial n} p_1\, ds dt\Big)$$

$$+ \int_0^T (-F(y) + F'(y)(y - \psi) - (A\psi + f))p_1(t)\, dt$$

$$+ \int_0^T \Big(|u|^2 + \frac{|(y-\psi)^+|^2}{2\epsilon} \text{ is uniformly bounded in } \epsilon > 0,$$

Hence, we have

$$\int_0^T \int_\Omega \lambda_\epsilon\, dx dt \text{ is uniformly bounded in } \epsilon > 0.$$

Thus, the optimality condition (0.8) holds for (y, u, p, λ).

Boundary optimal control case Consider a boundary control of the form

$$\frac{\partial}{\partial n}y + G(y) = Bu$$

at boundary $\partial\Omega$ with $B \in \mathcal{L}(U, L^2(\partial\Omega))$. Then, we have

$$\frac{\partial}{\partial n}p + G'(y)p = 0.$$

for the adjoint p. We assume that $-G(y) + G'(y)(y - \psi) \geq \delta$ for $y \leq \psi$. Note that

$$(y - \psi, A^*p)_\Omega = (-G'(y)(y - \psi), p)_{\partial\Omega} - (-G(y) + Bu, p)_{\partial\Omega} - (A(y - \psi$$

Using the similar arguments above $p_1(t) \geq 0$ and we have

$$(p(0), y_0 - \psi)_\Omega - \int_0^T (-G(y) + G'(y)(y - \psi), p)_{\partial\Omega} + (-F(y) + F'(y)(y -$$

$$= \int_0^T (|u|^2 + \frac{|(y - \psi)^+|^2}{\epsilon} + (J'(y), y - \psi) - (A\psi, p).$$

and

$$(\psi - y(0), p(0))_\Omega + \int_0^T (-G(y) + G'(y)(y - \psi)), p_1(t))_{\partial\Omega}\, dt$$

$$+ \int_0^T (-F(y) + F'(y)(y - \psi) - (A\psi + f), p_1) + |u|^2 + \frac{|(y - \psi)^+|^2}{2\epsilon})\, dt$$

is uniformly bounded in $\epsilon > 0$. Under the same assumption as above, we have

$$\int_0^T \int_\Omega \lambda_\epsilon \, dxdt \text{ is uniformly bounded in } \epsilon > 0.$$

Algorithm (semi-smooth Newton method The optimality system (0.8) is a general saddle point problem. Note that a Newton derivative N [Ito-Kunisch] of the max function $s \to \max(0, s)$ is given by

$$N(s) = \begin{cases} 0 & s \leq 0 \\ \\ 1 & s > 0 \end{cases}$$

Thus, for the linear case the semi-smooth Newton method [Ito-Kunisch] is of the form for the new update (p^+, y^+) given the current integrate (p, u)

$$\begin{pmatrix} BB^* & D \\ \\ D^* & -Q - \dfrac{\chi_{y>\psi}}{\epsilon} \end{pmatrix} \begin{pmatrix} p^+ \\ \\ y^+ \end{pmatrix} = \begin{pmatrix} f \\ \\ -Qy_d - \dfrac{\chi_{y>\psi}}{\epsilon}\psi \end{pmatrix} \tag{0.9}$$

where

$$J(y) = (Q(y - y_d), y - y_d).$$

Here, χ_S is the indicator function of set S and y_d is a desired state. and the solution operator D is given by

$$D = -A, \quad D = \frac{d}{dt} - A, \quad D = \frac{d^2}{dt^2} - \Delta,$$

for elliptic, parabolic and wave equation cases, respectively. One can prove the local super linear convergence and the global monotone convergence [Ito-Kunisch] . The details will be presented in a forthcoming paper.

For our tests we discretize A by the standard central difference method for A and the implicit Euler method for $\frac{d}{dt}$ and the explicit central difference method for $\frac{d^2}{dt^2}$ in time under CFL condition. One can solve (0.9) for the discretized system by MINRES (symmetric indefinite system) or a reduced order CG for a large scale system.

Numerical tests We consider two dimensional heat equation

$$y_t = \Delta y + \chi_S(x)u(t) + f(x), \quad y(0) = y_0$$

in square domain $\Omega = (0,1)^2$ with homogeneous boundary condition $y = 0$, where $S = (.4,.6)^2$ is the control subdomain.

$$\frac{y^n_{j,i} - y^n_{j,i}}{\Delta t} = \frac{4y^n_{j,i} - y^n_{i,i+1} - y^n_{j,i-1} - y^n_{j+1,i} - y^n_{j-1,i}}{h^2} + B_{j,i}u_{j.i} + f_{j,i}, \quad 1 \le i \le$$

with

$$y_{j,0} = y_{j,N} = 0, \;\; 1 \le j \le N-1 \quad y_{0,i} = y_{N,i} = 0, \;\; 1 \le i \le N-1$$

and $h = \frac{1}{N}$. Thus,

$$Dy = \frac{y^n - y^{n-1}}{\Delta t} - Hy^n, \;\; B = \chi_S.$$

with $H \in R^{N-1,N-1}$ is the central difference matrix.

We set $N = 10$ and $\Delta = \frac{1}{50}$ and $y_0 = exp(-100((x_1 - .5)^2 + (x_2 - .5)^2))$ and $\psi = .02$, $f = 1$ and $y_d = 0$. We let $y(t) \le \psi$ on $.1 \le t \le T = 1$.

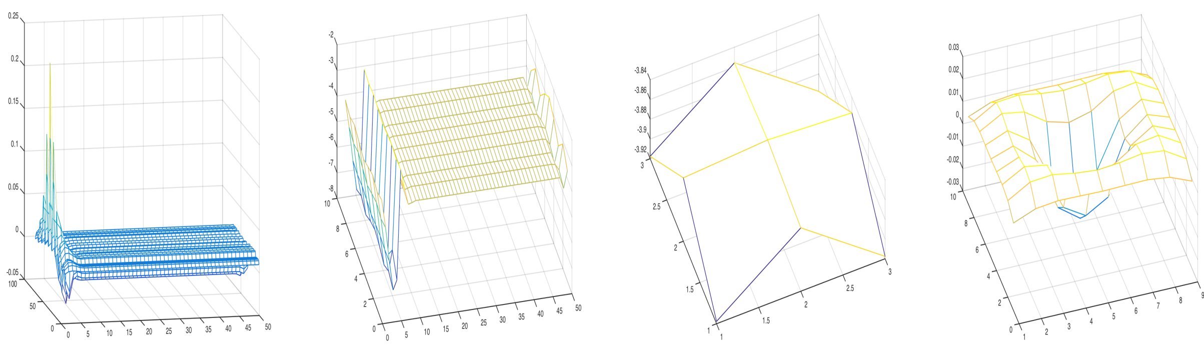

Figure 1: 2D heat equation

The first two figures are for state and control trajectory on $(0, T)$. (state vector $y(t) \in R^{81,1}$ and control vector $u(t) \in R^{9,1}$. The last two figures show control function on S and state function $y(t)$ on Ω at $t = .6$. The control trajectory achieves the state constraint on $t \geq .1$ and the control becomes stationary as t increases as we discussed.