Sufficient Optimality in a Parabolic Control Problem

King Fahd University
of Petroleum and Minerals
May 2005

H. D. Mittelmann
Department of Mathematics and Statistics
Arizona State University

Our work in PDE-constrained Optimization

(I) H. Maurer and H. D. Mittelmann,
Optimization Techniques for Solving Elliptic Control Problems with Control and State Constraints.
Part 1: Boundary Control
Comp. Opt. Applic. 16, 29-55 (2000)
Part 2: Distributed Control
Comp. Opt. Applic. 18, 141-160 (2001)

(II) H. D. Mittelmann and H. Maurer,
Interior Point Methods for Solving Elliptic Control Problems with Control and State Constraints. Boundary and Distributed Control,
J. Comp. Math. Applic. 120, 175-190 (2001)

(III) H. D. Mittelmann,
Verification of Second-Order Sufficient Optimality Conditions for Semilinear Elliptic and Parabolic Control Problems.
Comp. Opt. Applic. 20, 93-110 (2001)

(IV) H. D. Mittelmann,
Sufficient Optimality for Discretized Parabolic and Elliptic Control Problems, in Fast solution of discretized optimization problems, K.-H. Hoffmann, R.H.W. Hoppe, and V. Schulz (eds.), ISNM 138, Birkhäuser, Basel, 2001

(V) H. D. Mittelmann and F. Tröltzsch,
Sufficient Optimality in a Parabolic Control Problem. in Trends in Industrial Mathematics, A.H. Siddiqi and M. Kocvara (eds), Kluwer, Dordrecht, The Netherlands, 2002

all papers accessible (PS/PDF/HTML) through:

http://plato.asu.edu/papers.html

Main content of these papers:

Papers I,II:

- elliptic cases only
- formal derivation of necessary optimality conditions
- establish relationship between continuous and discrete conditions
- solve discretized problems (full discretization)
- verify necessary conditions
- includes active state constraints, bang-bang

Papers III, IV:

- consider also parabolic (1d) problems
- including one with known exact solution
- derive and evaluate second order conditions (SSC)
- parabolic problem has globally spd Hessian of the Lagrangian

Paper V:

- consider specific parabolic problem
- problem has known solution and state constraints
- show that continuous SSC are satisfied
- show that Hessian is not globally spd
- verify both numerically

The Parabolic Problem Class

(P) **Minimize**

$$\begin{aligned} J(y,u) &= \tfrac{1}{2}\int_Q \alpha(x,t)(y(x,t)-y_d(x,t))^2\,dxdt \\ &+ \tfrac{\nu}{2}\int_0^T u^2(t)\,dt \\ &+ \int_0^T [a_y(t)y(l,t)+a_u(t)u(t)]\,dt \end{aligned} \qquad (1)$$

subject to

$$\begin{aligned} y_t - y_{xx} &= e_Q && \text{in } Q \\ y(x,0) &= 0 && \text{in } (0,l) \\ y_x(0,t) &= 0 && \text{in } (0,T) \\ y_x(l,t)+y^2(l,t) &= e_\Sigma(t)+u(t) && \text{in } (0,T) \end{aligned} \qquad (2)$$

and to

$$u_a \le u(t) \le u_b, \quad \text{a.e. in } (0,T), \qquad (3)$$

$$\int_Q y(x,t)\,dxdt \le 0. \qquad (4)$$

Space of admissible controls

$U_{ad} = \{u \in L^\infty(0,T) \,|\, u_a \le u \le u_b, \quad \text{a.e. in } (0,T)\}$

- Problem (P) **not convex** since state equation (2) semilinear

- quadratic nonlinearity not **monotone**, no standard existence&uniqueness results

- will construct explicit solution $(\bar{y}, \bar{u})$ of (2)

- linearization of (2) at $\bar{y}$ **uniquely** solvable for all $u \in U_{ad}$

- implicit function theorem guarantees existence&uniqueness of the solution $y = y(u)$ of (2) in $W(0,T) \cap C(\bar{Q})$ for all u in a sufficiently small L^p-neighborhood of $\bar{u}$, $W(0,T)$ $= \{y \in L^2(0,T; H^1(0,l)) \,|\, y_t \in L^2(0,T; H^1(0,l)^*)\}$

Plan: To construct an instance of (P) for which

- the SSC are fulfilled
- second order derivative of Lagrange function not positive definite on the whole space
- consider strongly active control constraints
- example more involved than the one from Arada, Raymond, Tröltzsch (to app.)
- that one is coercive on whole space (III: H.D.M. 2001, IV: H.D.M. 2001)

First order necessary conditions

Let $\bar{u}$ be locally optimal for (P) with associated state $\bar{y}$

$$J(y,u) \geq J(\bar{y},\bar{u})$$

for all (y,u) satisfying (2-4) for u near $\bar{u}$. Then there exist Lagrange multipliers $\bar{p} \in W(0,T) \cap C(\bar{Q})$ (the adjoint state) and $\bar{\lambda} \geq 0$ such that the **adjoint equation**

$$\begin{array}{rcllll} -\bar{p}_t - \bar{p}_{xx} & = & \alpha(\bar{y} - y_d) + \bar{\lambda} & \text{in} & Q \\ \bar{p}(x,T) & = & 0 & \text{in} & (0,l) \\ \bar{p}_x(0,t) & = & 0 & \text{in} & (0,T) \\ \bar{p}_x(l,t) + 2\bar{y}(l,t)\,\bar{p}(l,t) & = & a_y(t) & \text{in} & (0,T) \end{array} \tag{5}$$

the **variational inequality**

$$\int_0^T (\nu\bar{u}(t) + \bar{p}(l,t) + a_u(t))(u(t) - \bar{u}(t))\,dt \geq 0 \quad \forall u \in U_{ad} \tag{6}$$

and the **complementary slackness condition**

$$\bar{\lambda} \int_Q \bar{y}(x,t)\,dxdt = 0 \tag{7}$$

are fulfilled.

(6) is equivalent to the projection property

$$\bar{u}(t) = \Pi_{[u_a,u_b]}\{-\frac{1}{\nu}(\bar{p}(l,t) + a_u(t))\}$$

Proof:
Derive from variational principles applied to **Lagrange function** $\mathcal{L}$

$$\begin{aligned}\mathcal{L}(y,u,p,\lambda) &= J(y,u) - \int_Q (y_t - y_{xx} - e_Q)\, dxdt + \\ &\quad \int_Q \bar{\lambda} y \, dxdt - \int_0^T (y_x + y^2 - u - e_\Sigma) p \, dt\end{aligned}$$

(homogeneous initial/boundary conditions of y formally included in the state space)

(5-6) follow from

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

for all admissible increments y and

$$\mathcal{L}_u(\bar{y},\bar{u},\bar{p},\bar{\lambda})(u - \bar{u}) \geq 0 \quad \forall u \in U_{ad}$$

Let $\tau > 0$ be given. We define

$$\begin{aligned} A^+(\tau) &= \{t \in (0,T) \,|\, \nu\bar{u}(t) + \bar{p}(l,t) + a_u(t) \leq -\tau\} \\ A^-(\tau) &= \{t \in (0,T) \,|\, \nu\bar{u}(t) + \bar{p}(l,t) + a_u(t) \geq \tau\} \end{aligned}$$

It holds $\bar{u} = u_b$ on A^+ and $\bar{u} = u_a$ on A^-. These sets indicate strongly active control constraints.

Second order sufficient optimality conditions

$$\mathcal{L}''(\bar{y},\bar{u},\bar{p},\bar{\lambda})[y,u]^2 = \int_Q \alpha\, y^2\, dxdt + \nu \int_0^T u^2\, dt + 2\int_0^T \bar{p}y^2\, dt$$

Assume state-constraint (4) active at $\bar{y}$ and $\bar{\lambda} = 1$

(SSC) There exist positive δ and τ such that

$$\mathcal{L}''(\bar{y}, \bar{u}, \bar{p}, \bar{\lambda})[y, u]^2 \geq \delta \int_0^T u^2 \, dt \qquad (8)$$

holds for all $y \in W(0, T)$, $u \in L^2(0, T)$ such that

$$\begin{array}{rcl} y_t - y_{xx} & = & 0 \\ y(x, 0) & = & 0 \\ y_x(0, t) & = & 0 \\ y_x(l, t) + 2\bar{y}(l, t) y(l, t) & = & u(t) \end{array} \qquad (9)$$

and

$$u(t) = 0 \qquad \text{on } A^+(\tau) \cup A^-(\tau) \qquad (10)$$

$$u(t) \geq 0 \quad \bar{u}(t) = u_a, \quad t \notin A^-(\tau) \qquad (11)$$

$$u(t) \leq 0 \quad \bar{u}(t) = u_b, \quad t \notin A^+(\tau) \qquad (12)$$

$$\int_Q y(x, t) \, dx dt = 0 \qquad (13)$$

We shall require (8) for all (y, u), which satisfy only (9-10)

The Test Example

$$T = 1,\ l = \pi,\ u_a = 0,\ u_b = 1,\ \nu = 0.004$$

$$\alpha(x,t) = \begin{cases} \alpha_o \in \mathbb{R}, & t \in [0,1/4] \\ 1, & t \in (1/4,1], \end{cases}$$

$$z(x,t) = \tfrac{1}{\alpha(x,t)}(1-(2-t)cosx),$$

$$y_d(x,t) = \begin{cases} z(x,t), & t \in [0,1/2] \\ z(x,t) + (t-1/2)^2 cosx, & t \in (1/2,1], \end{cases}$$

$$a_y(t) = \begin{cases} 0, & t \in [0,1/2] \\ 2(t-1/2)^2(1-t), & t \in (1/2,1], \end{cases}$$

$$a_u(t) = \nu + 1 - (1+2\nu)t,$$

$$e_Q(t) = \begin{cases} 0, & t \in [0,1/2] \\ (t^2+t-3/4)cosx, & t \in (1/2,1], \end{cases}$$

$$e_\Sigma(t) = \begin{cases} 0, & t \in [0,1/2] \\ (t-1/2)^4 - (2t-1), & t \in (1/2,1]. \end{cases}$$

The quantities

$$
\begin{aligned}
\bar{u} &= max\{0, 2t-1\} \\
\bar{y} &= \begin{cases} 0, & t \in [0,1/2] \\ (t-1/2)^2 cosx, & t \in (1/2,1] \end{cases} \\
\bar{p} &= (1-t)cosx \\
\bar{\lambda} &= 1
\end{aligned}
$$

satisfy the first order necessary conditions

Checking the SSC:
$\bar{u} = 0$ is strongly active on $[0, 1/2)$: If $b < 1/2$ is given, then

$$\nu\bar{u}+\bar{p}(\pi,.)+a_u = \bar{p}(\pi,.)+a_u = -\nu(2t-1) > \nu(2b-1)$$

holds for $t \in [0, b]$. Therefore, $t \in A^{-}(\tau)$ for $\tau = |\nu(2b-1)|$.

It suffices to confirm the coercivity condition (8) for all pairs (y, u) coupled through the linearized equation (9) and satisfying

$$u = 0 \text{ on } [0, b]$$

($0 < b < 1/2$ being arbitrary but fixed)

Let α have the form (14), where $b \in [0, 1/2)$.

$$\alpha(x,t) = \begin{cases} \alpha_0, & 0 \le t \le b \\ 1, & b < t \le 1. \end{cases} \tag{14}$$

Then the SSC are satisfied by $(\bar{y}, \bar{u}, \bar{p}, \bar{\lambda})$ for arbitrary α_0.

Proof:

Let u vanish on $[0, b]$ and let y solve (9). Then $y(x,t) = 0$ on $[0, b]$. For $\mathcal{L}''$ we get

$$\begin{aligned} \mathcal{L}''(\bar{y}, \bar{u}, \bar{p}, \bar{\lambda})[y,u]^2 &= \int_0^\pi \int_0^b \alpha_0 \cdot 0 \, dx dt + \int_0^\pi \int_b^1 y^2 dx dt + \\ & \quad \nu \int_0^1 u^2 dt - 2 \int_0^1 \bar{p}(\pi,t) y^2(\pi,t)\, dt \\ &\ge \nu \int_0^1 u^2 dt + 2 \int_0^1 (1-t) y^2(\pi,t)\, dt \\ &\ge \nu \int_0^1 u^2 dt. \end{aligned} \tag{15}$$

Hence the coercivity condition (8) is satisfied. If $\alpha_0 \ge 0$, then $\mathcal{L}''$ is obviously coercive on the whole space. However, we might find negative values for α_0 such that $\mathcal{L}''$ is partially indefinite.

If $\alpha_0 < 0$ is sufficiently small, then a pair (y, u) exists, such that $u \geq 0$, y solves the linearized equation (9), and

$$\mathcal{L}''(\bar{y}, \bar{u}, \bar{p}, \bar{\lambda})[y, u]^2 < 0. \tag{16}$$

We take an arbitrary but fixed $b < 1/2$ and set

$$u(t) = \begin{cases} 1 & \text{on } [0, b] \\ 0 & \text{on } (b, 1]. \end{cases}$$

Then $\int_0^\pi \int_0^b y^2 \, dxdt$ is positive. Hence

$$\alpha_0 \int_0^\pi \int_0^b y^2 \, dxdt \to -\infty$$

as $\alpha_0 \to -\infty$.

Therefore, the expression (14) becomes negative for sufficiently small α_0, if y^2 is substituted there for 0 in the first integral.

To get a negative value of $\mathcal{L}''(\bar{y}, \bar{u}, \bar{p}, \bar{\lambda})[y, u]^2$, we must have

$$\alpha_0 \int_0^\pi \int_0^b y^2\,dxdt + \int_0^\pi \int_b^1 y^2\,dxdt + \int_0^1 2(1-t)y^2(\pi,t)\,dt + \nu \int_0^1 u^2 dt < 0,$$

hence

$$|\alpha_0| > \frac{\int_0^\pi \int_b^1 y^2\,dxdt + \int_0^1 2(1-t)y^2(\pi,t)\,dt + \nu \int_0^1 u^2 dt}{\int_0^\pi \int_0^b y^2\,dxdt} = \frac{I_1 + I_2 + I_3}{I_0} \qquad (17)$$

must hold.
Here, $b \in [0, 1/2)$ can be chosen arbitrarily. We take the value $b = 1/4$. Thus, we evaluate the integrals I_j for

$$u(t) = \begin{cases} 1 & \text{on } [0, 1/4] \\ 0 & \text{on } (1/4, 1] \end{cases}$$

and the associated state y.

The state y solves the homogeneous heat equation subject to homogeneous initial condition, homogeneous boundary condition at $x = 0$ and

$$y_x(\pi, t) = \begin{cases} 1 & \text{on } [0, 1/4] \\ -2\bar{y}(\pi, t) & \text{on } (1/4, 1]. \end{cases}$$

We have evaluated the integrals I_j, $j = 0, .., 3$ numerically. The result is

$$\begin{array}{ll} I_0 = .0103271, & I_1 = .0401844 \\ I_2 = .0708107, & I_3 = .001 \end{array}$$

$$\frac{I_1 + I_2 + I_3}{I_0} = 10.845. \qquad (18)$$

Recall problem (P)

(P) **Minimize**

$$
\begin{aligned}
J(y,u) = \; & \frac{1}{2}\int_Q \alpha(x,t)(y(x,t) - y_d(x,t))^2 \, dxdt \\
& + \frac{\nu}{2}\int_0^T u^2(t)\, dt \\
& + \int_0^T [a_y(t) y(l,t) + a_u(t) u(t)]\, dt
\end{aligned}
$$

subject to

$$
\begin{array}{rcll}
y_t - y_{xx} & = & e_Q & \text{in} \quad Q \\
y(x,0) & = & 0 & \text{in} \quad (0,l) \\
y_x(0,t) & = & 0 & \text{in} \quad (0,T) \\
y_x(l,t) + y^2(l,t) & = & e_\Sigma(t) + u(t) & \text{in} \quad (0,T)
\end{array}
$$

and to

$$u_a \le u(t) \le u_b, \quad \text{a.e. in } (0,T),$$

$$\int_Q y(x,t)\, dxdt \le 0.$$

We **finite-difference discretize** this as:

minimize $\quad f_h(y_h, u_h) =$

$$\frac{dxdt}{2} \sum_{i=0}^{m} \sum_{j=0}^{n} \alpha_{j,i}\beta_j\gamma_i(y_{j,i} - y_d(x_j, t_i))^2)$$

$$+\frac{\nu dt}{2}(\sum_{i=0}^{m} \gamma_i u_i^2) + \frac{dt}{2}(\sum_{i=0}^{m} \gamma_i(a_y(t_i)y_{n,i} + a_u(t_i)u_i))$$

subject to $\qquad\qquad\qquad\qquad$ **(P$_h$)**

$$\frac{y_{j,i+1} - y_{j,i}}{dt} = (y_{j-1,i} - 2y_{j,i} + y_{j+1,i} + y_{j-1,i+1}$$

$$-2y_{j,i+1} + y_{j+1,i+1})/(2dx^2) + e_Q(x_j, t_{i+\frac{1}{2}})$$

$$i = 0, \ldots, m-1, \quad j = 1, \ldots, n-1$$

$$y_{j,0} = 0, \quad j = 0, \ldots, n$$

$$y_{2,i} - 4y_{1,i} + 3y_{0,i} = 0, \quad i = 1, \ldots, m$$

$$(y_{n-2,i} - 4y_{n-1,i} + 3y_{n,i})/(2dx) + y_{n,i}^2$$

$$= u_i + e_\Sigma(t_i), \quad i = 1, \ldots, m$$

$$u_a \leq u_i \leq u_b, \quad i = 0, \ldots, m$$

$$dxdt \sum_{i=0}^{m} \sum_{j=0}^{n} \beta_j\gamma_i y_{j,i} \leq 0.$$

Here $x_j = jdx$, $dx = \pi/n$, $t_i = idt$, $dt = T/m, \beta_0 = \beta_n = \frac{1}{2}, \beta_j = 1$ otherwise; analogously for γ.

(P_h) is a **nonlinear optimization problem**

$$\min F^h(z) \quad \text{subject to} \quad G^h(z) = 0, \quad H^h(z) \le 0$$

- z the discretized control and state variables
- $G^h(z)$ the state equation and boundary conditions
- $H^h(z)$ the control bounds and the integral state constraint

Assume $z \in \mathbf{R}^{N_h}$, $G^h : \mathbf{R}^{N_h} \to \mathbf{R}^{M_h}$, $M_h < N_h$. Let z^* be an admissible point satisfying the first-order necessary optimality conditions with associated Lagrange multipliers μ^* and λ^*. Let further

$$N(z^*) = (\nabla G^h(z^*), \nabla H_a(z^*))$$

be a column-regular $N_h \times (M_h + P_h)$ matrix where $M_h + P_h < N_h$ and $\nabla H_a(z^*)$ denotes the gradients of the P_h **active** inequality constraints with **positive** Lagrange multipliers.

We have

$$N_h = (m+1)(n+2),\ M_h = (m+1)(n+1)$$

The **maximal** $m+1$ degrees of freedom are further reduced by one through the active integral nonnegativity constraint on y and by any active bounds on u.

Let finally $N = QR$ be a **QR decomposition** and $Q = (Q_1, Q_2)$ a splitting with Q_1 of size $N_h \times (M_h + P_h)$.

The point z^* is a strict local minimizer if a $\gamma_h > 0$ exists such that

$$\lambda_{\min}(L_2(z^*)) = \gamma_h > 0.$$

Here $L_2(z^*)$ is the projected Hessian of the Lagrangian

$$L_2(z^*) = Q_2^T(\nabla^2 F^h(z^*) - \mu^{*T}\nabla^2 G^h(z^*))Q_2.$$

To verify that the problem is not **globally coercive** we compute the smallest eigenvalue

$$\delta_h$$

of another projected Hessian obtained by splitting

$$Q = (Q_1, Q_2)$$

where now

$$Q_1 \text{ is } N_h \times (M_h)$$

corresponding to the **equality constraints** only.

Numerical Implementation

- NLP solver called through AMPL

- output solution z^*, Lagrange multipliers

- utilize AMPL routines for gradient evaluation

- QR decompose $N(z^*)$ with **sparse** QR method

- check first order necessary conditions

- utilize AMPL for Lagrangian times vector

- apply this with columns of Q_2

- call LAPACK routine to compute γ

Errors, Eigenvalues, CPU times

n	m	$\|u-\bar{u}\|_\infty$	$\|u-\bar{u}\|_2$	$\|y-\bar{y}\|_\infty$	$\|y-\bar{y}\|_2$
127	41	3.74e-2	4.59e-5	6.10e-3	2.73e-7
192	61	1.33e-2	1.15e-5	1.77e-3	1.85e-7

m	α_0	γ_h	δ_h
41	-11	5.80e-5	-3.85e-5
	-10.5	5.80e-5	4.32e-5
	-10	5.80e-5	4.88e-5
61	-30	3.59e-5	-1.78e-3
	-15	3.59e-5	-3.87e-4
	-11	3.59e-5	-2.12e-5
	-10.5	3.59e-5	2.34e-5
	-10	3.59e-5	3.28e-5
	1	3.59e-5	3.28e-5

var	lc	nlc	KNI	LAN	MIN	SNO
5290	5208	41	40	1587	9	94
11835	11775	61	185	14762	72	1291
32549	31917	101	635	>60000	unb	5001

Optimal control and state