INTRODUCTION TO THE OPTIMAL CONTROL OF SYSTEMS WITH DISTRIBUTED PARAMETERS V. GRADIENT PROJECTION METHOD IN OPTIMAL CONTROL OF PARABOLIC PDES

UGUR ABDULLA

1. Idea of the Gradient Method

This is the fifth lecture on the optimal control of systems with distributed parameters. In this talk, we will discuss gradient methods for the solution of minimization problems in Hilbert spaces.

The idea of the gradient method is based on the following observation. Let H be an Hilbert space with inner product $\langle \cdot, \cdot \rangle_H$. If a functional $J : H \to \mathbb{R}$ is Frechet differentiable, then

$$J(v + h) - J(v) = \langle J'(v), h \rangle_H = o(h, v)$$

where

$$\frac{o(h, v)}{\|h\|} \to 0 \text{ as } \|h\| \to 0.$$

If $J'(v) \neq 0$, the main part of the increment is the differential $dJ(v) = \langle J'(v), h \rangle$. By CBS inequality,

$$-\|J'(v)\| \, \|h\| \leq \langle J'(v), h \rangle_H \leq \|J'(v)\| \, \|h\|$$

Moreover, the right-hand inequality becomes equality if and only if $h = \alpha J'(v)$, $\alpha \geq 0$, and the left-hand inequality becomes equality if and only if $h = -\alpha J'(v)$, $\alpha \geq 0$; this implies that if $J'(v) \neq 0$ then the direction of fastest decrease is the anti-gradient direction, and the direction of fastest increase is in the gradient direction.

Assuming that some initial control v_0 is chosen, the gradient method consists of an iterative process

$$v_{k+1} = v_k - \alpha_k J'(v_k), \;\; \alpha_k > 0, \;\; k = 1, 2, \dots$$

where α_k is the step length; if $J'(v_k) \neq 0$, then α_k can always be chosen such that $J(v_{k+1}) < J(v_k)$. Indeed,

$$\begin{aligned} J(v_{k+1}) - J(v_k) &= \langle J'(v_k), \alpha_k J'(v_k) \rangle + o(\alpha_k) \\ &= -\alpha_k \|J'(v_k)\|^2 + o(\alpha_k) \\ &= \alpha_k \left[-\|J'(v_k)\|^2 + \frac{o(\alpha_k)}{\alpha_k} \right] < 0 \end{aligned}$$

for $0 < \alpha_k \ll 1$.

2. Gradient Projection Method (GPM)

GPM is used for the approximate solution of the problem

$$J(u) \to \inf, \ u \in U$$

where U is a convex and closed subset of the Hilbert space H. Suppose $J(u) \in J'(U)$. The gradient method itself is not applicable in general, since even if $v_k \in U$, the next iteration $v_k - \alpha_k J'(v_k)$ may be outside of U. The idea of the GPM is to construct the sequence of approximate solutions as

$$v_{k+1} = P_U\Big(v_k - \alpha_k J'(v_k)\Big)$$

where P_U is the projection operator onto U.

Definition 1. $w \in U$ is called the ***projection*** of $u \in H$ onto $U \subset H$ if

$$\|u - w\| = \inf_{v \in U} \|u - v\|$$

We will use the notation $w = P_U(u)$.

Let us establish some properties of the projection operator:

Theorem 1. *Let U be a closed, convex subset of H. Then*

(1) *For arbitrary $u \in H$, there exists a unique $w = P_U(u)$.*
(2) *$w \in U$ is $P_U(u)$ if and only if*

$$\langle w - u, v - w \rangle \geq 0, \ \forall v \in U$$

(3) *P_U satisfies a Lipschitz condition:*

$$\|P_U(u) - P_U(v)\| \leq \|u - v\|, \ \forall u, v \in H$$

Proof. (1) Consider the functional

$$g(v) = \|v - u\|^2 : U \to \mathbb{R}$$

for fixed $u \in H$. g is strongly convex in H, since

$$g\Big(\alpha v_1 + (1-\alpha)v_2\Big) = \alpha g(v_1) + (1-\alpha)g(v_2) - \alpha(1-\alpha)\|v_1 - v_2\|^2$$

Since g is a strongly convex C^1 functional, it takes its minimum on U at a unique point $w \in U$, at which

$$g(v) = \|v - u\|^2 \geq \|w - u\|^2, \ \forall v \in U$$

That is,

$$\|v - u\| \geq \|w - u\|, \ \forall v \in U$$

Equality is only possible when $v = w$.

(2) In order for g to achieve its minimum at w, the necessary and sufficient condition implies that

$$\langle g'(w), v-w\rangle \geq 0,\ \forall v \in U$$

that is,

$$2\langle w-u, v-w\rangle \geq 0,\ \forall v \in U$$

(3) Let us prove the last assertion; we have

$$\langle P_U(u)-u, v-P_U(u)\rangle \geq 0,\ \forall v \in U$$

Consider any $u_1, u_2 \in H$ we have

$$\langle P_U(u_1)-u_1, v-P_U(u_1)\rangle \geq 0,\ \forall v \in U \tag{1}$$

and

$$\langle P_U(u_2)-u_2, v-P_U(u_2)\rangle \geq 0,\ \forall v \in U \tag{2}$$

or

$$\langle u_2-P_U(u_2), P_U(u_2)-v\rangle \geq 0,\ \forall v \in U \tag{3}$$

Adding (1) & (3) implies

$$\langle P_U(u_1)-P_U(u_2)-u_1+u_2, P_U(u_2)-P_U(u_1)\rangle \geq 0$$

so

$$\langle P_U(u_1)-P_U(u_2), P_U(u_2)-P_U(u_1)\rangle \leq \langle u_1-u_2, P_U(u_2)-P_U(u_1)\rangle$$

That is,

$$\|P_U(u_1)-P_U(u_2)\|^2 \leq \|u_1-u_2\|\,\|P_U(u_2)-P_U(u_1)\|$$

so

$$\|P_U(u_1)-P_U(u_2)\| \leq \|P_U(u_2)-P_U(u_1)\|$$ □

Theorem 2. *Let U be a convex and closed subset of H, $J(u) \in C^1(U)$, and suppose $U_* \neq \emptyset$. A necessary condition for an element u_* to be a minimizer of J on U, i.e. $u_* \in U_*$, is*

$$u_* = P_U(u_*-\alpha J'(u_*)),\ \forall \alpha > 0 \tag{4}$$

Moreover, if J is convex on U, then (4) *is sufficient to assert $u \in U_*$.*

Proof. According to Theorem 1, (4) is equivalent to

$$\left\langle u_*-\left(u_*-\alpha J'(u_*)\right), v-u_*\right\rangle \geq 0,\ \forall \alpha > 0,\ \forall v \in U$$

which is equivalent to

$$\langle J'(u_*), v-u_*\rangle \geq 0,\ \forall v \in U$$

which is the necessary condition for arbitrary $J(u) \in C^1(U)$, and is the necessary and sufficient condition for convex $J(u) \in C^1(U)$. □

3. Choosing the step size

GPM consists of building a sequence $\{u_k\}$ as

$$u_{k+1} = P_U\Big(u_k - \alpha_k J'(u_k)\Big),\ k = 1, 2, \dots$$

If for some k $u_{k+1} = u_k$, then the iterative process is interrupted, and according to Theorem 2, u_k satisfies an optimality condition. If in addition $J(u)$ is convex it follows that $u_k \in U_*$. Indeed.

$$u_k = P_U\Big(u_k - \alpha_k J'(u_k)\Big) \ \Leftrightarrow \alpha_k \left\langle J'(u_k), v - u_k \right\rangle \geq 0 \implies u_k \in U_*$$

There are different ways of choosing α:

(1) Choose α_k according to $f_k(\alpha_k) = \inf_{\alpha \geq 0} f_k(\alpha)$, where
$$f_k(\alpha) = P_U\Big(u_k - \alpha J'(u_k)\Big)$$
(2) Choose $\alpha_k = \alpha > 0$ and check if $J(u_{k+1}) < J(u_k)$, dividing α by $2, 4, \dots, 2^n$ until this is achieved.
(3) If $J(u) \in C^{1,1}(U)$ (i.e. J' satisfies a Lipschitz condition) with $\|J'(u) - J'(v)\| \leq L\|u - v\|$ then choose α_k as any number satisfying $0 < \epsilon_0 \leq \alpha_k \leq 2/(1+L)$

4. Explicit Projection Operators

GPM method based on the iteration $u_{k+1} = P_U\Big(u_k - \alpha_k J'(u_k)\Big)$ is convenient if there is an explicit formula to find the projection onto U. Consider some examples.

4.1. **Projection onto a ball.** The projection of $u \in H$ onto the ball

$$U = B(\bar{u}; R) = \{u \in H : \|u - \bar{u}\| \leq R\}$$

is given by

$$P_U(u) = \begin{cases} \bar{u} + R\frac{u - \bar{u}}{\|u - \bar{u}\|} & \text{if } \|u - \bar{u}\| > R \\ u & \text{if } \|u - \bar{u}\| \leq R \end{cases}$$

For a rigorous proof it is enough to check

$$\langle P_U(u) - u, v - P_U(u) \rangle \geq 0,\ v \in U$$

or

$$\left\langle \bar{u} - u + R\frac{u - \bar{u}}{\|u - \bar{u}\|}, v - \bar{u} - R\frac{u - \bar{u}}{\|u - \bar{u}\|} \right\rangle \geq 0,\ v \in U$$

The case $u \in U$ (i.e. $\|u - \bar{u}\| < R$) trivially satisfies the preceding condition; assume that $\|u - \bar{u}\| > R$. We have

$$\left(\frac{R}{\|u - \bar{u}\|} - 1\right)\left\langle u - \bar{u}, v - \bar{u} - R\frac{u - \bar{u}}{\|u - \bar{u}\|} \right\rangle =$$
$$= \left(\frac{R}{\|u - \bar{u}\|} - 1\right)\left(\langle u - \bar{u}, v - \bar{u}\rangle - R\|u - \bar{u}\|\right) \geq 0$$

since $\|u - \bar{u}\| > R$ implies $R/\|u - \bar{u}\| - 1 < 0$,

$$\langle u - \bar{u}, v - \bar{u}\rangle = \leq \|u - \bar{u}\|\,\|v - \bar{u}\| \leq R\|u - \bar{u}\|$$

4.2. **Projection onto a strip.** Let

$$U = \{p \in L_2[0,T] : 0 < p_{\min} \le p(t) \le p_{\max}\}$$

Then

$$P_U(p) = \begin{cases} p_{\min}, & \text{if } p(t) < p_{\min} \\ p(t), & \text{if } p_{\min} \le p(t) \le p_{\max} \\ p_{\max}, & \text{if } p_{\max} < p(t) \end{cases}$$

For a rigorous proof we have to check

$$\langle P_U(p) - p, q - P_U(p) \rangle \ge 0, \ \forall q \in U$$

Indeed,

$$\int_0^T \left(P_U(p) - p\right)\left(q - P_U(p)\right) \, dt \ge 0$$

since the integrand ≥ 0 for a.e. $t \in [0,T]$.

5. GPM applied to Optimal Control for the Heat Equation

Hence, in our optimal control problem for the heat equation, GPM reduces to the following construction:

$$v_{k+1} = P_{V_R}\Big(v_k - \alpha_k J'(v_k)\Big)$$

In particular,

$$p_{k+1} = \begin{cases} p_{\min}, & \text{if } p_k(t) - a^2\nu\alpha_k\psi(l,t;v_k) < p_{\min} \\ p_k(t) - a^2\nu\alpha_k\psi(l,t;v_k), & \text{if } p_{\min} \le p_k(t) - a^2\nu\alpha_k\psi(l,t;v_k) \le p_{\max} \\ p_{\max}, & \text{if } p_{\max} < p_k(t) - a^2\nu\alpha_k\psi(l,t;v_k) \end{cases}$$

and

$$f_{k+1}(x,t) = \begin{cases} R\frac{f_k(x,t) - \alpha_k\psi(x,t;v_k)}{\|f_k(x,t) - \alpha_k\psi(x,t;v_k)\|_{L_2(D)}} & \text{if } \|f_k(x,t) - \alpha_k\psi(x,t;v_k)\|_{L_2(D)} > R \\ f_k(x,t) - \alpha_k\psi(x,t;v_k) & \text{if } \|f_k(x,t) - \alpha_k\psi(x,t;v_k)\|_{L_2(D)} \le R \end{cases}$$

E-mail address: `abdulla@fit.edu`

Mathematical Sciences Department, Florida Institute of Technology, 150 W. University Blvd., Melbourne, FL 32901