PUBL. Mat. UAB
Nº 18 Abril 1980
Actas II Congreso de Ecuaciones Diferenciales y
Aplicaciones. Valldoreix, Mayo 1979

RESOLUCION NUMERICA DE PROBLEMAS DE CONTROL OPTIMO

CARLOS MORENO

COLEGIO UNIVERSITARIO DE VIGO

1.- Introducción Sea V un espacio de Hilbert real. Una aplicación multívoca $T:V\to V$ se llama operador monótono si

(1.1) $\qquad (u_1-u_2,\ u_1^*-u_2^*)\geq 0 \quad$ si $\quad u_1^*\in T(u_1) \quad$ y $\quad u_2^*\in T(u_2)$

Si el grafo de T no está contenido propiamente en el grafo de ningun otro operador monótono, se dice que T es maximal monótono. Se llama regularización Yosida de T al operador unívoco T_s definido por

(1.2) $\qquad T_s = \dfrac{I-R_s}{s} \quad , \ s>0$

siendo $R_s = (I+sT)^{-1}$ el resolvente de T. Ambos operadores permiten caracterizar la imagen de un operador maximal monotono mediante una igualdad, como se establece en el siguiente lema

Lema 1.1 (ver Bermudez-Moreno [1]). Las siguientes afirmaciones son equivalentes

a) $\quad u_1^*\in T(u_1)$

b) $\quad u_1^*=T_s(u_1+s\ u_1^*) \quad ,s>0$

c) $\quad u_1=R_s(u_1+s\ u_1^*) \quad ,s>0$

Si f es funcional convexo semicontínuo inferiormente (s.c.i.) propio en V, su subdiferencial ∂f es un operador maximal monótono, que permite distinguir a los elementos que realizan el mínimo de f

(1.3) $\qquad f(u)=\underset{v\in V}{\mathrm{Min}} f(v) \quad$ si y solo si $\quad 0\in\partial f(u)$

Puesto que $\partial f_s(u)$ es la diferencial Frechet (ver Brezis [2]) en u del funcional f_s definido por

$$(1.4) \qquad f_s(z)= \underset{v\in V}{\mathrm{Min}}\left\{f(v)+(1/2s)\|v-z\|^2\right\},$$

usando el lema 1.1 se deduce

$$(1.5) \qquad f(u)=\underset{v\in V}{\mathrm{Min}}\ f(v) \quad \text{si y solo si} \quad f_s(u)= \underset{v\in V}{\mathrm{Min}}\ f_s(v)$$

Por consiguiente, el problema de optimizacion no diferenciable puede ser transformado en otro problema equivalente correspondiente a un funcional diferenciable. Es razonable utilizar el método del gradiente para aproximar los elementos que realizan el mínimo de f (o de f_s)

$$(1.6) \qquad u^{k+1}= u^k-r_k\ \partial f_s(u^k)$$

Tomando $s=r_k$, se obtiene

$$(1.7) \qquad u^{k+1}= B_{r_k}(u^k)= \mathrm{Arg.min.}\left\{f(v)+(1/2r_k)\|v-u^k\|^2\right\}'$$

La sucesión $\{u^k\}$ definida por (1.7) coincide con la que genera el algorítmo del punto próximo (ver Rockafellar [6]).

La convergencia fuerte de este algoritmo está garantizada si ∂f es un operador fuertemente monótono, es decir, si f verifica la desigualdad siguiente

$$(1.8) \qquad (u_1-u_2,\ u_1^*-u_2^*) \geq \alpha\|u_1-u_2\|^2 \qquad \text{si } u_1^*\in\partial f(u_1) \text{ y } u_2^*\in \partial f(u_2),\ \alpha>0$$

En el caso general, si existe al menos un elemento que realice el minimo de f, se tiene asegurada la convergencia debil de la sucesión definida -- por (1.7) a uno de tales elementos. Permanece abierta la cuestión de si - la convergencia es también fuerte.

De particular interés es el caso, en el que f= g+h, siendo g y h - funcionales convexos s.c.i. y propios en V, tales que existe $w\in \mathrm{dom}g\wedge\mathrm{dom}h$, en donde g es contínua. El algorítmo del punto próximo genera la sucesión definida por

$$(1.9)\qquad u^{k+1} = \text{Arg. min}\left\{g(v)+h(v)+(1/2s_k)\|v-u^k\|^2\right\}.$$

Utilizando la equivalencia (1.3) se obtiene

$$(1.10)\qquad 0\in\partial g(u^{k+1})+\partial h(u^{k+1})+(1/s_k)\,(u^{k+1}-u^k)\ .$$

Representado por R_s el operador resolvente de g, la relación (1.10) puede transformarse en

$$(1.11)\qquad \begin{cases} u^{k+1} = R_{s_k}(u^k - s_k z^{k+1}) \\ z^k \in \partial g(u^k) \end{cases}$$

Esta formulación del algorítmo del punto próximo está motivada por numerosos ejemplos en los que el funcional a minimizar puede ser descompuesto en dos sumandos, uno de los cuales posee una cierta regularidad y el otro es tal que el operador resolvente de su subdiferencial es simple desde el punto computacional. Ya que la fórmula (1.11) es de tipo implícito parece natural plantearse el estudio de algorítmos análogos de tipo explícito, tales como (se supone que g es G-diferenciable).

$$(1.12)\qquad u^{k+1} = R_{s_k}(u^k - s_k g'(u^k))$$

Si $h=\chi_K$ es la función indicadora del conjunto convexo y cerrado K el algorítmo definido por (1.12) coincide con el clásico algorítmo del gradiente con proyección.

Si g' es un operador fuertemente monótono y $\{s_k\}$ está acotada superiormente por una cota convenientemente pequeña, la sucesión $\{u^k\}$ converge fuertemente al único elemento u que realiza el mínimo de f.

Un segundo tipo de aplicaciones del algorítmo del punto próximo, corresponde al método de Uzawa para la aproximación de puntos de silla de las lagrangianas asociadas al problema de optimización considerado.
Para desarrollar estos métodos se considera un cuadro funcional ligeramente distinto del anterior. Sean E y V dos espacios de Hilbert reales

(de normas $|\ |$ y $\|\ \|$ respectivamente), g (resp.h) un funcional convexo s.c.i. propio en V (resp. en E) y L un operador lineal y continuo de V en E. Se considera el siguiente problema

Hallar $u\in V$ tal que $f(u)\leq f(v)$ para todo $v\in V$, siendo f el funcional definido por

$$f(v)=g(v)+h(Lv) \tag{1.13}$$

El problema dual según el formalismo de Fenchel y Rockafellar consiste en hallar $w\in E$ (que se identifica con su dual) que realiza el mínimo del funcional

$$\tilde{f}(w)=g^*(-L^*w)+h^*(w) \tag{1.14}$$

siendo g^* (resp.h^*) la función conjugada de g (resp.h), es decir

$$g^*(w)=\sup_{v\in V}\left\{(w,v)-g(v)\right\} \tag{1.15}$$

Formalmente (se puede justificar imponiendo hipótesis poco restrictivas sobre g y h) el algorítmo de punto próximo aplicado al problema dual genera la sucesión $\{w^k\}$ definida por

$$\left\{\begin{array}{l} w^{k+1}=\partial h_{s_k}(Lv^{k+1}+s_k w^k) \\ v^k=\text{Arg. Min}\left\{g(v)+(Lv,w^k)\right\} \end{array}\right. \tag{1.16}$$

En las siguientes secciones se considerará también el algorítmo de tipo explícito definido por

$$\left\{\begin{array}{l} w^{k+1}=\partial h_{s_k}(Lv^k+s_k w^k) \\ v^k=\text{Arg. Min}\left\{g(v)+(Lv,w^k)\right\} \end{array}\right. \tag{1.17}$$

En la sección 2 se establecen resultados de convergencia de los algorítmos que se han considerado. En las restantes secciones se dan aplicaciones a la resolución numérica de problemas de control óptimo.

2.- Convergencia de los métodos duales El siguiente teorema proporciona un resultado de convergencia para el algorítmo definido por (1.16)

Teorema 2.1. Bajo las siguientes hipótesis

1) g es fuertemente convexo

2) Existe $w \in \mathrm{dom}\, g \cap \mathrm{dom}\, h$, en donde h es contínua

3) $o < s_k < s$

se tiene

a) $\{v^k\} \to u$, u es la solución del problema original

b) $\{w^k\} \to \bar{w}$, $\bar{w}$ es una solución del problema dual

Demostración. Usando la equivalencia (1.3) y puesto que gracias a la hipótesis 2) las reglas de la suma y de la cadena del cálculo subdiferencial se verifican (ver Ekeland-Temam [3]), se obtiene

$$(2.1) \qquad o \in \partial g(u) + L^* \partial h(Lu)$$

Si $\bar{w} \in \partial h(Lu)$, del lema 1.1 se deduce

$$(2.2) \qquad \bar{w} = \partial h_s(Lu + s\lambda\bar{w}) \quad , \ s > o$$

además

$$(2.3) \qquad u = \mathrm{Arg.min}\, \{g(v) + (Lv, \bar{w})\}$$

Restando la primera igualdad de (1.16) de (2.2) se obtiene

$$(2.4) \qquad |\bar{w}-w^{k+1}|^2 \leqslant (1/s_k)\ (Lu-Lv^{k+1}, \bar{w}-w^{k+1}) + (\bar{w}-w^k, \bar{w}-w^{k+1})$$

. De la relación de $o \in \partial g(u) + L^*\bar{w}$ y la análoga para v^{k+1}, usando la hipótesis 1) (es decir, ∂g es fuertemente monótono de módulo α) se deduce

$$(2.5) \qquad (\bar{w}-w^{k+1}, Lu-Lv^{k+1}) \leqslant -\alpha \|u-v^{k+1}\|^2$$

Combinando (2.4) y (2.5) se obtiene

$$(2.6) \qquad |\bar{w}-w^{k+1}|^2 + (\alpha/s_k)\ \|u-v^{k+1}\|^2 \leqslant (\bar{w}-w^k, \bar{w}-w^{k+1})$$

$$(2.7) \qquad |\bar{w}-w^{k+1}|^2 + (2\alpha/s_k)\ \|u-v^{k+1}\|^2 + |w^{k+1}-w^k|^2 \leqslant |\bar{w}-w^k|^2$$

Consecuentemente

$$(2.8)\qquad \{v^k\}\to u, \quad \|u-v^{k+1}\| = o(\sqrt{s_k})$$

Para realizar la iteración definida por (1.16) es preciso resolver una ecuación no lineal, de la que hasta el momento se ignora si posee solución y en caso afirmativo, si esta es única (lo que es imprescindible para que el algorítmo esté bien definido). No obstante, por un razonamiento similar al anterior se puede probar que la transformación definida por

$$(2.9)\qquad \begin{cases} Tw = \partial h_{s_k+s}\,(Lz+s_k w^k+sw) \\ z = \quad \text{Arg.min}\,\{g(v)+(Lv,w)\} \end{cases}$$

es una contracción estricta que tiene a w^{k+1} como único punto fijo. Además, la existencia de solución de la ecuación (1.16) garantiza que la sucesión $\{w^k\}$ coincide con la generada por el algorítmo del punto próximo aplicado al problema dual. En consecuencia b) se verifica.

El siguiente teorema proporciona un resultado de convergencia para el algorítmo definido por (1.17).

Teorema 2.2. Bajo las siguientes hipótesis

1) g es fuertemente convexo

2) Existe $w \in \text{dom } g \cap \text{dom } h$, en donde h es contínua

3) $(1/2\alpha) < s_k < s$

se tiene que $\{v^k\}\to u$ y la sucesión $\{w^k\}$ es acotada.

Demostración. Las relaciones (2.1), (2.2) y (2.3) son aún válidas, gracias a las hipótesis 1) y 2). Además

$$(2.10)\qquad |\bar{w}-w^{k+1}|^2 \leq (1/s_k^2)\,\|L\|\,|u-v^k|^2 + |\bar{w}-w^k|^2 + (2/s_k)(Lu-Lv^k, \bar{w}-w^k)$$

$$\leq |\bar{w}-w^k|^2 + ((1/s_k^2)-(2\alpha/s_k))\,\|u-u^k\|^2$$

La desigualdad (2.10) permite concluir la tesis del teorema.

3.- Problemas de control óptimo Sea W el espacio de Hilbert de los estados; H_1 un espacio de Hilbert en el que el estado es observado mediante el operador $C_1 \in L(W,H_1)$; U el espacio de Hilbert de los controles (que se identifica a su dual) y K_1 el subconjunto convexo y cerrado de U, constituido por los controles admisibles.

Se supone que el estado es una función afín contínua del control

$$(3.1) \qquad y(v)= Dv+g \quad , D\in L(U,W).$$

Sea $z_d \in H_1$ la observación del estado deseada y J el funcional coste definido por

$$(3.2) \qquad J(v)= \| C_1 y(v)-zd \|^2_{H_1} + r \| v \|^2_U$$

El problema de control óptimo se formula del siguiente modo

Hallar $u\in K_1$, tal que $J(u)\leq J(v)$ para todo $v\in K_1$

Los siguiente resultados de existencia y unicidad son clásicos (ver Lions [4]).

1) Si $r>o$, existe un control óptimo único

2) Si $r=o$ y K_1 es acotado, existe un conjunto convexo cerrado y no vacío de controles.

El problema de control considerado puede incluirse en el cuadro funcional de la sección 1 tomando $g=(1/2)J$ y $h=\chi_{K_1}$

En este caso, la fórmula (1.12) adopta la forma particular

$$(3.3) \qquad u^{k+1}= P_{K_1}(u^K - s_k(D^* C_1^* (C_1 y(u^k)-zd)+ru^k))$$

siendo P_{K_1} el operador de proyección sobre K_1.

Si $r>o$, la sucesión definida por (3.3) es convergente al único control óptimo. Por el contrario, si $r=o$, $\{u^k\}$ no es en general convergente y en todo caso, razones de inestabilidad numérica desaconsejan su uso. En esta situación, resulta adecuado utilizar el algorítmo de punto próximo

$$(3.4) \qquad u^{k+1} = \text{Arg.min} \left\{ (1/2)J(v) + \chi_{K_1}(v) + (1/2s) \| v - u^k \|^2_U \right\}$$

La realización de cada iteración requiere la resolución de un problema de control óptimo del tipo r>o, para lo cual es aconsejable el uso del algorítmo definido por (3.3).

4.- Problemas de control óptimo con restricciones sobre el estado Sea H_2 un espacio de Hilbert real en el que el estado es observado por un segundo operador $C_2 \in L(W, H_2)$. Se formula el siguiente problema de control

Hallar $u \in U_{ad} = \{ v \in K_1 : C_2 y(v) \in K_2 \}$ tal que $J(v) \leq J(u)$ para todo $v \in U_{ad}$

Tomando

$$(4.1) \begin{cases} E = H_2 \quad , \; V = U \\ g(v) = (1/2)J(v) + \chi_{K_1}(v) \quad , \; h(w) = \chi_{K_2}(w + C_2 g) \\ L = D \end{cases}$$

el problema de control puede incluirse en el cuadro funcional de la sección 1. El algorítmo definido por (1.16) adopta en este caso la siguiente forma

$$(4.2) \quad \begin{cases} v^{k+1} = \text{Arg.min} \left\{ (1/2)J(v) + \chi_{K_1}(v) + G_{s_k}(C_2 y(u) + s_k w^k),\; C_2 Dv) \right\} \\ w^{k+1} = G_{s_k}(C_2 y\,(v^{k+1}) + s_k w^k) \end{cases}$$

siendo G_s la regularización Yosida de χ_{K_2}

$$(4.3) \qquad G_s = (1/s)(I - P_{K_2})$$

Similarmente el algorítmo (1.17) es

$$(4.4)\quad \begin{cases} v^{k+1} = \text{Arg.min}\left\{(1/2)J(v)+\chi_{K_1}(v)+(w^k,C_2Dv)\right\} \\ w^{k+1} = G_{s_k}(C_2y(v^k)+s_kw^k) \end{cases}$$

Si $r>o$, el funcional g es fuertemente convexo. Desafortunadamente, la condición 2) de los teoremas de convergencia (2.1) y (2.2) se verifica si y solo si el interior de K_2 es no vacío, lo cual es una fuerte limitación en el caso infinito dimensional. No obstante, en la siguiente sección, se pondrá de manifiesto en ejemplos, como la discretización del problema contínuo, levanta generalmente esta limitación. Si el problema es previamente penalizado, es decir, el operador multívoco $\partial\chi_{K_2}$ se sustituye por $(1/s)(I-P_{K_2})$, la transformación $w^k \rightarrow w^{k+1}$ es contractiva. Un algorítmo próximo al definido por (4.4), ha sido propuesto por Mossino [5] para resolver el problema penalizado

5.- Ejemplo Sea Ω un abierto acotado de $\mathbb{R}^n$, de frontera Γ suficientemente regular. El estado es la solución de

$$(5.1)\quad \begin{cases} y-\Delta y=f+v & \text{en } \Omega \\ y = o & \text{en } \Gamma \end{cases}$$

con $f\in L^2(\Omega)$

El problema de control es

$$\underset{v\in U_{ad}}{\text{Min}} \left\{\int_\Omega (y(v)-z_d)^2dx + r\int_\Omega v^2dx\right\} \quad , \quad z_d\in L^2(\Omega) \quad , \quad r\geq o$$

siendo $U_{ad} = \left\{v\in H_o^1(\Omega) : |\text{grad } v| \leq 1 \quad \text{c.p.d.}\ \Omega\right\}$. A la fórmula (4.4) le corresponde en este caso

$$(5.2)\quad \begin{cases} v^{k+1} = \text{Arg.min}\left\{(1/2)J(v)+\int_\Omega w^k \text{grad } y(v)dx\right\} \\ w^{k+1} = (1/s_k)(I-P_{K_2})(\text{grad } y(v^k)+s_kw^k) \end{cases}$$

siendo $K_2 = \left\{ q \in L^2(\Omega)^n : |q| \leq 1 \quad \text{c.p.d.}\ \Omega \right\}$ De modo similar se describe el algorítmo definido por (4.2)

Puesto que el interior de K_2 es vacío, la hipótesis 2) de los teoremas 2.1 y 2.2 no se verifica y consecuentemente la convergencia estos algorítmos no está garantizada. No obstante, el algorítmo (5.2) para poder ser utilizado requiere una discretización previa del problema, gracias a lo cual esta limitación desaparece.

Sea $L^2(\Omega)_h$ el espacio de las funciones escalonadas

(5.3) $$v_h = \sum_{M \in \mathring{\Omega}_h} v_h(M)\,\omega_{hM}$$

(consultar en Termam [7] las notaciones usuales del método de las diferencias finitas), provisto del producto escalar

(5.4) $$(v_h, w_h)_h = \sum_{M \in \mathring{\Omega}_h} v_h(M) w_h(M)$$

y $H_o^1(\Omega)_h$ el espacio de las funciones escalonadas

(5.5) $$y_h = \sum_{M \in \mathring{\Omega}_h^1} u_h(M)\,\omega_{hM}$$

provisto del producto interior

(5.6) $$((y_h, z_h))_h = \sum_{M \in \mathring{\Omega}_h^1} y_h(M) z_h(M) + \sum_{i=1}^{n} \sum_{M \in \mathring{\Omega}_h^1} \delta_i y_h(M) \delta_i z_h(M)$$

Sea $f_h \in L^2(\Omega)_h$ definido por

(5.7) $$f_h(M) = (h_1 h_2 \ldots h_n)^{-1} \int_{\sigma_h(M)} f(x)\,dx$$

y $z_{dh} \in L^2(\Omega)_h$ definida de modo similar.

El operador grad_h de $H_o^1(\Omega)_h$ en $L^2(\Omega)_h^n$ está definido por sus componentes

(5.8) $$\delta_i q_h(M) = (1/h_i)(y_h(M + h_i e_i) - y_h(M))$$

El operador transpuesto de grad_h es

(5.9) $\qquad -\mathrm{div}_h q_h = \mathrm{grad}_h^{*} q_h - \sum_{i=1}^{n} (1/h_i)(q_h(M-h_i e_i) - q_h(M))$

Se considera

(5.10) $\qquad K_{2h} = \left\{ q_h \in L^2(\Omega)_h^n \; : \; |q_h| \leq 1 \right\}$

El problema de control discretizado es

$$\underset{v \in Uad_h}{\mathrm{Min}} \left\{ (y_h(v_h) - z_{d_h}, \; y_h(v_h) - z_{d_h})_h + r(v_h, v_h)_h \right\}$$

siendo $Uad_h = \left\{ v_h \in L^2(\Omega)_h : |\mathrm{grad}_h v_h| \leq 1 \right\}$ e $y_h(v_h)$ la única solución en $H_0^1(\Omega)_h$ de

(5.11) $\qquad y_h - \mathrm{div}_h \mathrm{grad}_h y_h = f_h$

Resultados de convergencia del problema descretizado al problema exacto se puede encontrar en lemam [7] . Además el algorítmo (5.2) es convergente en este caso, ya que el interior de K_{2h} es no vacío.

REFERENCIAS BIBLIOGRAFICAS

[1] A. BERMUDEZ-C.MORENO, Duality methods for solving variational inequalities, próxima aparición.

[2] H. BREZIS, Opérateurs maximaux monotones et semigroupes de contractions dans les espaces de Hilbert, North-Holland, 1973.

[3] I. EKELAND-R. TEMAM, Analyse convexe et problémes variationnels, Dunod, 1974.

[4] J.L. LIONS, Controle optimal de systemes gouvernés par der equations aux dérivées partielles, Dunod 1.968.

[5] J. MOSSINO, Approximation numérique de problémes de controle optimal avec contrainte sur le controle et sur l'etat, Calcolo, vol XIII, fasc. I, 1976.

[6] R.T. ROCKAFELLAR, Monotone operators and the proximal point algorithm, SIAM J. Control and Optimization, vol 14, no 5, agosto, 1976

[7] R. TEMAM, Analyse numérique, Presses Universitaires de France, Paris 1970.