

EL USO DEL ALGORITMO DE UZAWA EN PROBLEMAS DE PROGRAMACION LINEAL

R. L. V. Gonzalez

Facultad de Ciencias Exactas e Ingeniería
Universidad Nacional de Rosario
Rosario, Argentina.

OBJETIVO DEL TRABAJO

Se muestra en este trabajo cómo es posible aplicar el algoritmo de Uzawa de cálculo de puntos silla, a la solución de problemas de programación lineal. Este método de solución presenta la ventaja de permitir la solución de problemas de grandes dimensiones en computadoras con memorias pequeñas y de ser fácilmente programable.

1. INTRODUCCION

El algoritmo de Uzawa es uno de los más conocidos para la solución del problema de determinación de puntos silla. En general, para transformar un problema de minimización en uno de búsqueda de un punto silla, se forma el Lagrangiano introduciendo multiplicadores con respecto a las restricciones; sin embargo, el algoritmo de Uzawa converge sólo cuando el problema de minimización original es coercivo, por lo que en nuestro problema de programación lineal procedemos previamente a transformarlo en un problema equivalente coercivo.

Esto lo realizamos de la siguiente forma: definimos una transformación $F: \mathbb{R}^p \rightarrow \mathbb{R}^p$ cuyos únicos puntos fijos son las soluciones del problema de programación lineal y buscamos un puntofijo de F a través de la iteración $y_{n+1} = F(y_n)$. La transformación es la solución de un problema coercivo y lo resolvemos utilizando el algoritmo de Uzawa (previo agregado de los multiplicadores de Lagrange necesarios).

Probamos en este trabajo que la iteración $y_{n+1} = F(y_n)$ converge en un número finito de pasos. Sin embargo, como el algoritmo de Uzawa requiere en principio un número infinito de pasos, para obtener un algoritmo implementable (en el sentido de Polak (4)) construimos una diagonalización del algoritmo de Uzawa y el del punto fijo que resulta implementable y probamos su convergencia.

El algoritmo implementable así obtenido permite resolver problemas de grandes dimensiones en computadoras con memorias pequeñas ya que sólo es necesario tener registrados los elementos no nulos de la matriz que define el problema de programación lineal, las variables independientes x y los multiplicadores de Lagrange p . Además, tiene la ventaja de ser fácilmente programable (en un centenar de instrucciones Fortran) y al ser un algoritmo iterativo, corrige automáticamente los errores de redondeo acumulados al transcurrir el cálculo.

2. ELEMENTOS DEL PROBLEMA ORIGINAL

Se debe minimizar la función.

$$1) J(x) = (c, x), \quad x \in R^V$$

dentro del conjunto W .

$$2) W = \{x \in K / Ax + b \leq 0\}$$

(\cdot, \cdot) designa el producto escalar en R^V ,
 A matriz $\mu \times \nu$

$K \subset R^V$ es un conjunto convexo de la forma siguiente:

$$3) K = \prod_{i=1}^m [m_i, M_i] \quad m_i, M_i \in \bar{R}, \text{ pueden ser finitos o no.}$$

Definimos el conjunto de soluciones óptimas:

$$4) Y = \{\bar{x} \in W / (c, \bar{x}) \leq (c, x) \quad \forall x \in W\}$$

Supondremos en adelante que $Y \neq \emptyset$ y que

$$5) \exists x \in K / (Ax + b)_i < 0 \quad \forall i = 1, \nu$$

3. SOLUCION DEL PROBLEMA P_1 POR MEDIO DEL METODO DEL PUNTO FIJO

Definimos una transformación de $R^V \rightarrow W$ cuyos puntos fijos son los puntos de Y (y viceversa)

$$6) F(x) = P_W(x - \frac{1}{\epsilon} c)$$

siendo P_W la proyección sobre W y $\epsilon > 0$.

$F(x)$ es en consecuencia la única solución de los problemas equivalentes (7), (8) o la inecuación variacional (9):

$$P_2 \left\{ \begin{array}{l} 7) \min_{y \in W} (y - x + \frac{1}{\epsilon} c, y - x + \frac{1}{\epsilon} c) \\ 8) \min_{y \in W} ((c, y) + \frac{\epsilon}{2} (y - x, y - x)) \\ 9) (y - F(x), F(x) - x + \frac{1}{\epsilon} c) \geq 0 \quad \forall y \in W \end{array} \right.$$

Probaremos ahora que los puntos fijos de F pertenecen a Y y que además todos los puntos de Y son invariantes de la transform

mación F . De esta forma el problema P_1 queda reducido a la búsqueda de los puntos fijos de F .

Proposición 1

Si $x = F(x)$, entonces $x \in Y$; asimismo, si $\bar{x} \in Y$ entonces $F(\bar{x}) = \bar{x}$.

Demostración:

Si $x = F(x)$, de (9) obtenemos:
 $(y-x, c) \geq 0 \quad \forall y \in W$, luego $x \in Y$ en virtud de (4).

Si $\bar{x} \in Y$, entonces $(c, \bar{x}) \leq (c, y) \quad \forall y \in W$, por lo que la inecuación variacional (9) se satisface para $F(\bar{x}) = \bar{x}$.

Buscamos entonces las soluciones del problema P_1 o sea los puntos fijos de F utilizando el siguiente algoritmo iterativo A_1 :

- $$\left\{ \begin{array}{l} \text{Paso 0: } m=0, \quad x_m \in R^V \\ \text{Paso 1: } x_{m+1} = F(x_m) \\ \text{Paso 2: Si es } x_{m+1} = x_m \text{ se finaliza el algoritmo; si} \\ \quad x_{m+1} \neq x_m \text{ se toma } m = m+1 \text{ y se regresa al pa} \\ \quad \text{so 1.} \end{array} \right.$$

Este algoritmo tiene la propiedad de que encuentra los puntos de Y en un número finito de iteraciones.

Teorema 1

El algoritmo A_1 converge en un número finito de pasos a un punto de Y .

Demostración:

a) Comprobamos en primer lugar que la sucesión $\{x_m\}$ generada por el algoritmo es acotada. En efecto x_{m+1} satisface la inecuación variacional:

$$10) (x_{m+1} - x_m + \frac{1}{\epsilon} c, x - x_{m+1}) \geq 0 \quad \forall x \in W$$

como además $\forall \bar{x} \in Y (c, \bar{x} - x_{m+1}) \leq 0$ obtenemos:

$$11) (x_{m+1} - x_m, \bar{x} - x_{m+1}) \geq 0 \quad \forall \bar{x} \in Y, \text{ luego}$$

$$(x_{m+1} - \bar{x}, \bar{x} - x_{m+1}) + (\bar{x} - x_m, \bar{x} - x_{m+1}) \geq 0 \quad \forall \bar{x} \in Y$$

por lo que Cauchy-Schwaitz se tiene:

$$\|\bar{x} - x_m\| \cdot \|\bar{x} - x_{m+1}\| \geq \|x_{m+1} - \bar{x}\|^2$$

o sea:

$$12) \|x_{m+1} - \bar{x}\| \leq \|x_m - \bar{x}\| \quad \forall \bar{x} \in Y$$

Luego por inducción resulta que $\{x_m\}$ es acotado ya que:

$$13) \|x_m - \bar{x}\| \leq \|x_0 - \bar{x}\| \quad \forall m \geq 1, \quad \forall \bar{x} \in Y$$

b) Demostramos ahora que $\lim_{m \rightarrow \infty} \|x_{m+1} - x_m\| = 0$ para $m \rightarrow \infty$
 Por definición del problema equivalente (8) se cumple que:

14) $\frac{\epsilon}{2} \|x_{m+1} - x_m\|^2 + (c, x_{m+1}) \leq (c, x_m)$

en consecuencia,

15) $(c, x_{m+1}) \leq (c, x_m)$ Esta sucesión, además de ser monótona, es acotada en virtud de que $\{x_m\}$ es acotada, por lo que será en consecuencia convergente.

De esta forma obtenemos de (14) que $\lim_{m \rightarrow \infty} \|x_{m+1} - x_m\|^2 \leq 0$ lo que implica que:

16) $\lim_{m \rightarrow \infty} \|x_{m+1} - x_m\| = 0$

c) Demostramos ahora que los puntos de acumulación de $\{x_m\}$ pertenecen a Y .

Sea entonces \bar{x} punto de acumulación de $\{x_m\}$ (existe al menos uno en virtud de que $\{x_m\}$ es acotada) y sea $\{x_{m'}\}$ una subsucesión convergente a \bar{x} . Tomando límite $m' \rightarrow \infty$ en (10) y considerando (16), obtenemos: $(c, x - \bar{x}) \geq 0 \forall x \in W$

lo que implica que $\bar{x} \in Y$.

d) Mostramos finalmente que la sucesión $\{x_m\}$ alcanza el conjunto Y para un valor finito de m .

Como $\|x_m - \bar{x}\| \leq \|x_0 - \bar{x}\| \forall \bar{x} \in Y$

17) $\|x_m\| \leq \|x_0\| + R \forall m=1, \dots$; siendo $R = \min_{\bar{x} \in Y} \|x_0 - \bar{x}\|$
 Dado que $\|x\|_\infty = \max_{i=1, \dots, r} |x_i| \leq \|x\| = (\sum_{i=1}^r x_i^2)^{1/2}$
 entonces $x_m \in W' \forall m \geq 1$, donde

$W' = \{x \in W / \|x\|_\infty \leq \|x_0\| + R\}$

y x_{m+1} será entonces la solución de la inecuación

18) $(x_{m+1} - x_m, x - x_{m+1}) + \frac{1}{\epsilon} (c, x - x_{m+1}) \geq 0 \forall x \in W'$

Veremos que esto implica que a partir de un cierto índice, $x_{m+1} \in Y$.

En efecto, si designamos \hat{x}_{m+1} la proyección de x_m sobre el conjunto Y , \hat{x}_{m+1} es la solución de la inecuación variacional:

19) $(\hat{x}_{m+1} - x_m, x - \hat{x}_{m+1}) \geq 0 \forall x \in Y$

Debemos demostrar que a partir de un cierto índice, es

$\hat{x}_{m+1} = x_{m+1}$

Siendo W' un convexo acotado generado por un número finito de restricciones lineales, puede considerarse como la cápsula convexa del conjunto de sus puntos extremales $\{v_1, \dots, v_r\}$. Sea \bar{x} punto de acumulación de $\{x_m\}$ y $\{x_{m'}\}$ sucesión convergente a \bar{x} . Definimos:

$I = \{i / (c, v_i - \bar{x}) > 0\} \quad \delta = \min_{i \in I} (c, v_i - \bar{x})$

Por definición de \hat{x}_{m+1} , tenemos:

$\|\hat{x}_{m+1} - x_m\| \leq \|\bar{x} - x_m\|$

Dado que $\lim_{m' \rightarrow \infty} \|\bar{x} - x_{m'}\| = 0$ entonces $\lim_{m' \rightarrow \infty} \|\hat{x}_{m'+1} - x_{m'}\| = 0$

y existirá en consecuencia

20) $(\hat{x}_{m'+1} - x_{m'}, v_i - \hat{x}_{m'+1}) + \frac{1}{\epsilon} (c, v_i - \hat{x}_{m'+1}) \geq \frac{\delta}{2\epsilon} \forall i \in I$

ya que como $\hat{x}_{m'+1} \in Y$, es $(c, \hat{x}_{m'+1}) = (c, \bar{x})$.

Si $v_i / i \notin I$ entonces

21) $(c, v_i - \hat{x}_{m'+1}) = (c, v_i - \bar{x})$ lo que implica que $v_i \in Y$, por lo que por (19) tendremos:

22) $(\hat{x}_{m'+1} - x_{m'}, v_i - \hat{x}_{m'+1}) \geq 0$

De (20), (21), (22) obtenemos que

23) $(\hat{x}_{m'+1} - x_{m'}, v_i - \hat{x}_{m'+1}) + 1/\epsilon (c, v_i - \hat{x}_{m'+1}) \geq 0 \quad \forall i=1, \dots, r$

La misma desigualdad valdrá para las combinaciones convexas de v_1, \dots, v_r , o sea:

24) $(\hat{x}_{m'+1} - x_{m'}, x - \hat{x}_{m'+1}) + 1/\epsilon (c, x - \hat{x}_{m'+1}) \geq 0 \quad \forall x \in W'$

Siendo esta la inecuación variacional (18) cuya única solución es $x_{m'+1}$, concluimos que $\hat{x}_{m'+1} = x_{m'+1} \quad \forall m' \geq m_1$.

Considerando que $\hat{x}_{m'+1} \in Y$ y que la transformación F tiene como puntos fijos los elementos de Y , obtenemos (llamando m'_1 al primer $m' \geq m_1$) $x_{m'_1+2} = F(x_{m'_1+1}) = x_{m'_1+1} \in Y$. Por inducción resulta $x_{m'_1+h} = x_{m'_1+1} \in Y \quad \forall h=2, \dots$ con lo que el teorema queda demostrado.

4. SOLUCION DEL PROBLEMA P_2 UTILIZANDO EL ALGORITMO DE UZAWA

Para caracterizar a la solución $F(x)$ del problema P_2 , formamos el correspondiente Lagrangiano introduciendo multiplicadores con respecto a las restricciones $Ax+b \leq 0$ es decir

$$25) \begin{cases} L(y, q) = \frac{\epsilon}{2} (y-x, y-x) + (c, y) + (q, Ay+b) \\ y \in K \subset R^v, q \in \Delta = (R^+)^m \end{cases}$$

La solución $F(x)$ del problema P_2 queda caracterizada de la siguiente forma:

Proposición 2:

Existe al menos un $p \in \Delta$ tal que $(F(x), p)$ es un punto silla del Lagrangiano $L(y, q)$ en el conjunto $K \times \Delta$ es decir:

$$26) L(F(x), q) \leq L(F(x), p) \leq L(y, p) \quad \forall y \in K, \forall q \in \Delta$$

Corolario 1

De (26) se observa que son válidas las siguientes propiedades:

$$27) A \cdot F(x) + b \leq 0$$

$$28) (p, A \cdot F(x) + b) = 0$$

$$29) L(F(x), p) = \min_{y \in K} L(y, p)$$

$$30) \max_{q \in \Delta} \min_{y \in K} L(y, q) = \min_{y \in K} \max_{q \in \Delta} L(y, q) = \min_{y \in W} ((c, y) + \frac{\epsilon}{2} (y-x, y-x)) = L(F(x), p)$$

La demostración de las propiedades (26) - (30) está esencialmente contenida en [1], [2], [7].

Utilizamos ahora la equivalencia de los problemas

$$\max_{q \in \Delta} \min_{y \in K} L(y, q) \quad \text{y} \quad \min_{y \in W} ((c, y) + \frac{\epsilon}{2} (y-x, y-x)) \quad (P_2)$$

para solucionar este último resolviendo el primero por medio del método de Uzawa (ver [2]).
 La función $S(q) = \min L(y, q)$ $y \in K$ se calcula muy fácilmente lo mismo que el único elemento (que en lo sucesivo llamaremos $y(q)$) que produce ese mínimo.

En efecto, como

$$31) L(y, q) = \frac{\varepsilon}{2} (y-x, y-x) + (c, y) + (q, Ay+b) = \\ = \frac{\varepsilon}{2} \left\| y-x + \frac{1}{\varepsilon} (c+A^*q) \right\|^2 + (c, x) + (q, Ax+b) - \frac{1}{2\varepsilon} \|c+A^*q\|^2$$

el $y(q)$ que minimiza esta función es

$$32) y(q) = P_K \left(x - \frac{1}{\varepsilon} (c+A^*q) \right)$$

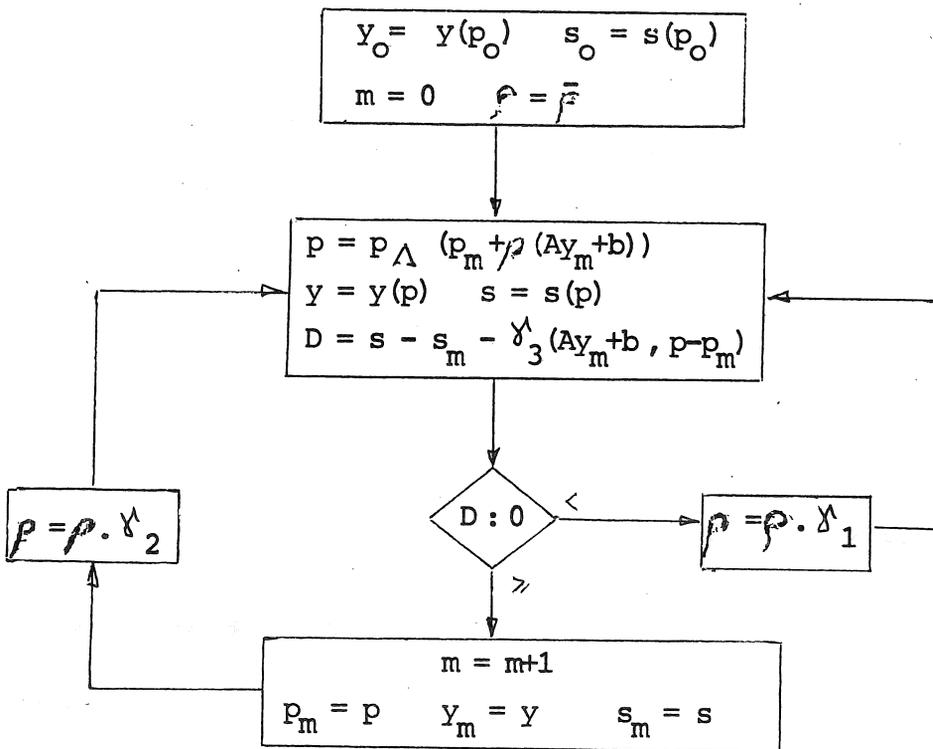
Teniendo en cuenta (3) se calcula en forma explícita las componentes de $y(q)$:

$$33) (y(q))_i = \max \left(\min \left(x_i - \frac{1}{\varepsilon} (c+A^*q)_i, M_i \right), m_i \right) \quad i=1, \nu$$

Daremos ahora un esquema del algoritmo de Uzawa con el agregado de un procedimiento de cálculo automático del paso p . Los parámetros iniciales del algoritmo son:

$$34) p_0 \in \Delta, \bar{\rho} > 0, \gamma_2 > 1, 0 < \gamma_1 < 1, 0 < \gamma_3 < 1$$

Algoritmo A_2



Observamos que el algoritmo es el de Uzawa con paso ρ variable, donde para decidir si un paso ρ es aceptable se usa el test de Armijo (ver E. Polak [4]).

Con respecto a la convergencia de este algoritmo, se tiene el siguiente resultado, cuya demostración está esencialmente contenida en [2]

Teorema 2:

La sucesión de puntos y_m generada por el algoritmo converge a $F(x)$ solución del problema P_1 . Además, $\{p_m\}$ es acotado y todos sus puntos de acumulación \bar{p} satisfacen: $(\bar{p}, A \cdot F(x) + b) = 0$

Observación:

Es posible demostrar que el test de Armijo se satisface (en nuestro algoritmo A_2) $\forall \rho \leq \tilde{\rho} = (1 - \delta_3) \epsilon / (\|A\| \cdot \|A^*\|)$.

En ejemplos de aplicación hemos comprobado, sin embargo, que este valor $\tilde{\rho}$ es muy pequeño y que si se usa una constante $\rho = \tilde{\rho}$ en todo el cálculo se obtiene una convergencia muy lenta. El método adaptativo mostrado (cálculo de ρ con el test de Armijo), mejora este aspecto y el ρ (variable) que utiliza realmente el algoritmo A_2 resulta (en problemas de 10 variables y 20 restricciones) superior en 10 a 100 veces $\tilde{\rho}$.

5. DIAGONALIZACIÓN DEL METODO DEL PUNTO FIJO Y EL ALGORITMO DE UZAWA.

Hemos visto que para solucionar el problema original de programación lineal P_1 , debemos hacer un número finito de iteraciones del algoritmo del punto fijo, $(x_{m+1} = F(x_m))$; como para hallar cada x_m debemos resolver por el método de Uzawa en forma completa un problema dual (que en principio necesita un número infinito de iteraciones), la combinación directa de A_1 y A_2 no resulta implementable. Para hacerlo implementable, se debe "diagonalizar" la combinación A_1, A_2 , es decir se debe dar una regla para "truncar" el proceso de cálculo de A_2 en cada iteración de A_1 . Esto se realiza en el algoritmo que se muestra a continuación, donde se procede a truncar el proceso A_2 cada vez que:

$$35) Ay + b \leq a_1(m)$$

$$36) (p, Ay + b) \geq -a_2(m)$$

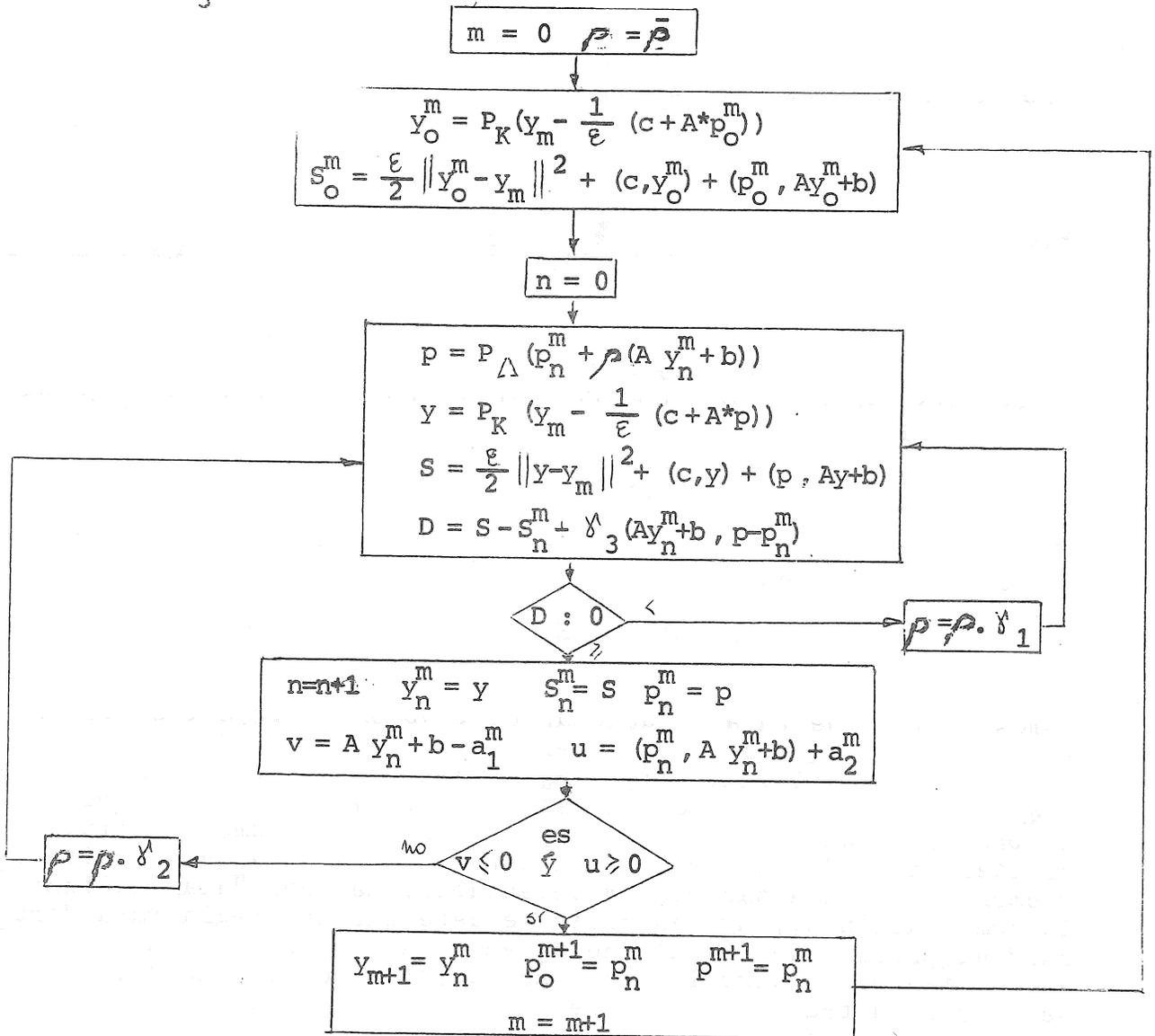
decir $a_1(m) \in (R^+)^{\mu}$, $a_2(m) \in R^+$ son sucesiones sumables, es decir $\sum_{m=1}^{\infty} (a_1(m))_i < +\infty \quad \forall i = 1, \dots, \mu$; $\sum_{m=1}^{\infty} a_2(m) < +\infty$

Esta condición de sumabilidad es la que permite demostrar la convergencia del algoritmo.

Los parámetros y datos iniciales del algoritmo son:

$$37) \epsilon > 0, \bar{p} > 0, 0 < \delta_1 < 1, \delta_2 > 1, 0 < \delta_3 < 1, y_0 \in R^{\nu}, p_0 \in \Delta$$

Algoritmo A₃



Con respecto a la convergencia de este algoritmo A₃ tenemos el siguiente resultado:

Teorema 3:

El algoritmo genera una sucesión de puntos y_m cuyos puntos de acumulación $\bar{y} \in Y$, es decir son soluciones del problema de programación lineal original. Además, todos los puntos de acumulación \bar{p} de $\{p^m\}$ satisfacen:

$$(\bar{p}, A\bar{x} + b) = 0 \quad \forall \bar{x} \in Y$$

6. APLICACIONES

Se ha utilizado este algoritmo A_3 en la resolución de problemas de programación lineal que resultan de discretizar problemas de control óptimo continuo. En particular se ha aplicado a un problema de tiempo de detención óptimo cuya discretización condujo a una programación lineal con 121 variables y 321 restricciones. Se programó el algoritmo A_3 en 120 instrucciones Fortran IV, (incluyendo instrucciones de entrada y salida), utilizándose los siguientes valores de los parámetros

$p_0 = 0$, $\varepsilon = 3$, $\gamma_1 = 0.3$, $\gamma_2 = 1.1$, $\gamma_3 = 0.1$
La sucesión $\{a_1(m), a_2(m)\}$ usada tiene la forma

$$\begin{cases} (a_1(m))_i = r^m & r = 0.75 & (i=1, \mu) \\ a_2(m) = r^m \end{cases}$$

Se obtuvo la convergencia del algoritmo (invariancia de los cinco primeros dígitos) al cabo de 1.500 cálculos de p . En general, el test que toma en cuenta el valor D se satisface en una sola iteración, es decir sin necesidad de reducir el paso P y la convergencia se obtuvo al cabo de 6 iteraciones del último lazo de A_3 .

BIBLIOGRAFIA

- (1) J. Cea, - Optimisation, théorie et algorithmes. Dunod éd, Paris 1971.
- (2) J. L. Lions, R. Glowinski, R. Tremoliers. Methodes numériques de resolution des problemes d'inequations variationnelles en mecanique et en physique. Dunod, Paris 1973.
- (3) R. Glowinski, - La methode de relaxation. Applications a la minimization avec et sans contraintes de fonctionnelles convexes. Quaderni dei Rendiconti - Ist. Mat Univ. Roma 1971.
- (4) E. Polak, Computational methods in Optimization. Academic Press. 1971.
- (5) J. L. Lions, G. Stampacchia, Variational inequalities. Comm. Pure Appl. Math. 20 (1967) 493-519
- (6) R. Glowinski, Methodes iteratives duales pour la minimization de fonctionnelles convexes. Constructive aspects of Functional Analysis. Ed. Cremonese. Roma 1973.
- (7) A. Bensoussan, Optimization in Hilbert space. Course at Singapur, May 1979.