UNIVERSIDAD NACIONAL DE INGENIERÍA

FACULTAD DE CIENCIAS

UNIDAD DE POSGRADO

“MÉTODOS ALGORÍTMICOS LAGRANGIANO AUMENTADO − PROXIMAL − HACES PARA PROBLEMAS DE DESIGUALDAD VARIACIONAL”

TESIS

PARA OPTAR EL GRADO ACADÉMICO DE MAESTRO EN CIENCIAS CON MENCIÓN EN MATEMÁTICA APLICADA

PRESENTADO POR:

ERNESTO ORÉ ALBORNOZ

ASESOR:

Dr. ELADIO OCAÑA ANAYA

LIMA – PERÚ
2015

Dedicatoria

A toda mi familia,
en especial a mi madre.

AGRADECIMIENTOS

Agradesco a mi asesor, Dr. Eladio Ocaña, por su paciencia, consejos y asesoramiento durante la elaboración de este trabajo. Asimismo agradesco al Instituto de Matemática y Ciencias Afines (IMCA), por el apoyo científico, económico e inmobiliario brindado.

RESUMEN

En este trabajo se extienden los métodos del Lagrangiano Aumentado, Proximal y Haces desarrollados en la reciente monografía del IMCA N^o 55 "Métodos Algorítmicos Haces-Proximal-Lagrangiano Aumentado para Problemas de Optimización Matemática"(y referencias citadas en esta) al caso general de problemas de desigualdad variacional monótono con restricciones de igualdades lineales y desigualdades convexas.

Índice general

Introducción

Consideremos el siguiente problema de desigualdad variacional:

$$\text{Determinar } x \in C \text{ tal que existe } x^* \in T(x) \text{ satisfaciendo}$$

$$\langle x^*, y - x\rangle \geq 0 \ \text{ para todo } \ y \in C \tag{V}$$

donde $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ es un operador y C un conjunto definido por

$$C := \{x \in \mathbb{R}^n : Ax = a,\ g_j(x) \leq 0, j = 1, \cdots, q\}$$

con $g_j : \mathbb{R}^n \to \mathbb{R}$ función convexa, $A \in \mathbb{R}^{p\times n}$ y $a \in \mathbb{R}^p$.

El objetivo de este trabajo es desarrollar algoritmos que nos permitan aproximar una solución (si existe) del problema variacional (V). Al ser el problema de optimización convexa un caso especial del problema de desigualdad variacional monótono, proponemos extender el algoritmo Lagrangiano Aumentado y el algoritmo Lagrangiano Aumentado–Proximal desarrollado en [5] y [8] respectivamente. El trabajo está organizado del siguiente modo:

El primer capítulo muestra cómo deducir el marco de dualidad para problemas de desigualdad variacional a partir del marco de dualidad para problemas de optimización. Obteniendo así problemas equivalentes al problema (V) denominados problema variacional dual y minimax.

El segundo capítulo desarrolla el método de Lagrangiano Aumentado para el problema variacional (V) a partir de su problema variacional dual. Se considera dos casos para una mejor comprensión: el caso con restricciones de igualdades lineales y el caso con restricciones de desigualdades convexas.

El tercer capítulo desarrolla el método Lagrangiano Aumentado–Proximal para el problema variacional (V), en el cual combinamos el método obtenido en el Capítulo 2 con el método Proximal. Esto nos lleva a un algoritmo en el cual debemos resolver subproblemas variacionales irrestrictos donde sus operadores asociados son fuertemente monótonos.

El cuarto capítulo desarrolla el método de Haces (en inglés, bundle) para resolver el problema variacional irrestricto de un operador fuertemente monótono. Obteniendo un algoritmo implementable, en el cual sólo se exige que el operador sea maximal fuertemente monótono de coeficiente conocido $\rho > 0$ y cuyo dominio sea $\mathbb{R}^n$.

Capítulo 1

Dualidad para problemas de desigualdad variacional

Comencemos recordando la dualidad para problemas de optimización para luego extenderla a problemas de desigualdad variacional.

1.1. Dualidad para problemas de optimización

Dado $f : \mathbb{R}^n \to \bar{\mathbb{R}} := [-\infty, +\infty]$, consideremos el problema de optimización irrestricta:

$$V_P := \text{ínf}\,[f(x) :\ x \in \mathbb{R}^n], \tag{P}$$

el cual incluye los problemas del tipo de optimización restricta:

$$\text{ínf}\,[g(x) :\ x \in C]. \tag{P_C}$$

donde $g : \mathbb{R}^n \to \mathbb{R}$ y C es un subconjunto no vacío de $\mathbb{R}^n$, pues este último problema es equivalente al problema (P) considerando $f = g + \delta_C$.

Asociado al problema (P) consideremos una función $\varphi : \mathbb{R}^n \times \mathbb{R}^p \to \bar{\mathbb{R}}$ de manera que $\varphi(\cdot, 0) = f(\cdot)$, llamada función de perturbación. Este nos lleva a considerar para cada $u \in \mathbb{R}^p$ (variable de perturbación) los siguientes problemas perturbados de (P):

$$h(u) := \text{ínf}\,[\varphi(x,u) :\ x \in \mathbb{R}^n], \tag{P_u}$$

donde claramente $h(0) = V_P$. Utilizando la conjugada de frenchel, se obtiene una regularización de h por medio de su bidual (h^{**}). En general $h(0) \geq h^{**}(0)$ y bajo ciertas hipótesis sobre h se da la igualdad. Esto nos conduce a determinar $h^{**}(0)$. Primero al calcular la conjugada de h se obtiene:

$$h^*(u^*) = \sup_{x,u}[\langle u^*, u\rangle - \varphi(x,u)] = \varphi^*(0,u^*), \tag{1.1}$$

luego, definiendo la función $\beta(u^*) := -\varphi^*(0,u^*)$, se tiene que $h^{**}(0) = \sup_{u^*} \beta(u^*)$. Entonces obtenemos un nuevo problema denominado problema dual:

$$V_D := \sup\,[\beta(u^*) :\ u^* \in \mathbb{R}^p], \tag{D}$$

el cual por construcción es un problema convexo (maximar una función concava), cuyo conjunto solución es $\partial h^{**}(0)$ y $V_P \geq V_D$. En resumen se tiene el siguiente esquema de dualidad:

$$\left.\begin{array}{rcl} h(0) & = & \text{ínf}_x\, f(x) \\ h^{**}(0) & = & \sup_{u^*} \beta(u^*) \end{array}\right\} \text{ donde } \left\{\begin{array}{rcl} h(u) & = & \text{ínf}_x\, \varphi(x,u) \\ f(x) & = & \varphi(x,0) \\ \beta(u^*) & = & -\varphi^*(0,u^*) \end{array}\right.$$

Análogamente se desarrolla un esquema de dualidad para (D), pues como $-V_D = \text{ínf}_y\, \varphi^*(0,y)$, entonces considerando φ^* como la función de perturbación, definimos los problemas duales perturbados:

$$k(x^*) := \text{ínf}\,[\varphi^*(x^*,u^*) :\ u^* \in \mathbb{R}^p], \tag{D_{x^*}}$$

luego como $k^*(x^*) = \varphi^{**}(x^*,0)$, consideramos el problema bidual como:

$$V_{DD} := \text{ínf}\,[\varphi^{**}(x,0) :\ x \in \mathbb{R}^n], \tag{DD}$$

Por construcción se cumple que el conjunto solución de (DD) es $\partial k^{**}(0)$ y $V_D \leq V_{DD}$. En resumen se tiene el siguiente esquema de dualidad:

$$\left.\begin{array}{rcl} -k(0) & = & \sup_{u^*} \beta(u^*) \\ -k^{**}(0) & = & \text{ínf}_x\, \tilde{f}(x) \end{array}\right\} \text{ donde } \left\{\begin{array}{rcl} -k(x^*) & = & \sup_{u^*} -\varphi^*(x^*,u^*) \\ \beta(u^*) & = & -\varphi^*(0,u^*) \\ \tilde{f}(x) & = & \varphi^{**}(x,0) \end{array}\right.$$

Cuando φ es convexo sci y propia, entonces $\varphi^{**} = \varphi$, luego se tiene que el problema bidual (DD) es el problema primal (P).

Análogamente, desarrollando un esquema de dualidad para el problema bidual (DD), resulta que su dual es (D).

Relacionemos ahora los problemas primal y dual con el problema de minimax.

1.1.1. Problema Minimax

Dado $l : X \times Y \to \bar{\mathbb{R}}$, donde X y Y son dos conjuntos no vacíos de $\mathbb{R}^n$ y $\mathbb{R}^m$ respectivamente, definimos las funciones:

$$\phi_l(x) := \sup_{y \in Y} l(x,y) \quad \text{y} \quad \psi_l(y) := \inf_{x \in X} l(x,y).$$

Asociados a estas funciones se definen los siguientes problemas

$$\bar{\alpha} := \inf\,[\phi_l(x) : \; x \in X], \tag{P_l}$$

$$\bar{\beta} := \sup\,[\psi_l(y) : \; y \in Y]. \tag{D_l}$$

El problema (P_l) es llamado problema primal, y el problema (D_l) es llamado problema dual, ambos asociados a la función l. Los términos asignados a estos problemas se debe a que en el caso particular cuando l es la lagrangiana asociado a una función de perturbación del problema (P) y considerando hipótesis no tan restringidas, estos problemas son exactamente los problemas (P) y (D) respectivamente, veremos con más detalle en la siguiente sección.

Siempre se cumple que:

$$\bar{\beta} = \sup_{y \in Y} \inf_{x \in X} l(x,y) \leq \inf_{x \in X} \sup_{y \in Y} l(x,y) = \bar{\alpha}, \tag{1.2}$$

pero en general no siempre se da la igualdad. El problema de minimax nos conduce al concepto de punto silla.

Definición 1.1 (Punto silla) *Un punto* $(\bar{x},\bar{y}) \in X \times Y$ *es llamado punto silla de* l *si*

$$l(\bar{x},y) \leq l(\bar{x},\bar{y}) \leq l(x,\bar{y}) \text{ para todo } (x,y) \in X \times Y$$

o equivalentemente si

$$\inf_{x\in X} l(x,\bar{y}) = l(\bar{x},\bar{y}) = \sup_{y\in Y} l(\bar{x},y).$$

El conjunto de todos los puntos silla de l *será denotado por* $\operatorname{argminmax}_{X,Y} l$.

Una propiedad importante de los puntos silla es:

$$(\bar{x},\bar{y}) \in \operatorname{argminmax}_{X,Y} l \text{ si y solo si } \begin{cases} \bar{x} \in \operatorname{argmin}_{x\in X} \phi_l(x); \\ \bar{y} \in \operatorname{argmax}_{y\in Y} \psi_l(y); \\ \sup_{y\in Y} \psi_l(y) = \inf_{x\in X} \phi_l(x). \end{cases}$$

En consecuencia se obtiene la siguiente relación

$$\operatorname{argmin}_{x\in X} \phi_l(x) \times \operatorname{argmax}_{y\in Y} \psi_l(y) \supseteq \operatorname{argminmax}_{X,Y} l\,.$$

Si no hay salto de dualidad ($\bar{\alpha} = \bar{\beta}$), esta inclusión es una igualdad.

Otra caracterización de los puntos silla es atraves de la subdiferencial:

$$(\bar{x},\bar{y}) \in \operatorname{argminmax}_{X,Y} l \text{ si y solo si } \begin{cases} 0 \in \partial_x l(\bar{x},\bar{y}); \\ 0 \in \partial_y[-l](\bar{x},\bar{y}). \end{cases}$$

1.1.2. Relaciones entre $(P) - (P_l)$ y $(D) - (D_l)$

En esta parte justificaremos los términos primal y dual asignados a los problemas (P_l) y (D_l).

De la relación (1.1) el problema (D) tiene la siguiente expresión

$$V_D = \sup_{u^*}[\beta(u^*)] = \sup_{u^*\in\mathbb{R}^p} \inf_{x\in\mathbb{R}^n} \left[\inf_{u\in\mathbb{R}^p} [\varphi(x,u) - \langle u^*, u\rangle] \right].$$

Considerando la función lagrangiana $l : \mathbb{R}^n \times \mathbb{R}^p \to \overline{\mathbb{R}}$ definida por

$$l(x, u^*) = \inf[\varphi(x, u) - \langle u^*, u\rangle : \ u \in \mathbb{R}^p]$$

se obtiene, por la relación (1.1), que

$$\psi_l(u^*) = \inf_x l(x, u^*) = \beta(u^*),$$

por lo tanto el problema (D) coincide con el problema (D_l).

De otro lado, la función objetivo del problema (P_l) es

$$\phi_l(x) = \sup_y \inf_u [\varphi_x(u) - \langle u, y\rangle] = (\varphi_x)^{**}(0) \leq \varphi_x(0) = f(x),$$

donde $\varphi_x(u) = \varphi(x, u)$.

Si para cada x, la función φ_x, donde $\varphi_x(u) = \varphi(x, u)$, es propia convexa sobre $\mathbb{R}^p$ y sci en 0, entonces el problema (P_l) es igual al problema (P). Bajo estas condiciones se observa que $(\bar{x}, \bar{u}^*)$ es un punto silla de l si y solo si $\bar{x}$ es una solución óptima de (P), $\bar{u}^*$ es una solución óptima de (D) y $V_P = V_D$.

1.1.3. Reformulación de los problemas de optimización como problemas de Desigualdad Variacional

Las formulaciones variacionales de los problemas primal, dual y minimax son respectivamente:

$$0 \in \partial f(\bar{x}), \quad 0 \in \partial\beta(\bar{u}^*) \quad \text{ y } \quad (0, 0) \in T_l(\bar{x}, \bar{u}^*),$$

donde

$$T_l(x, u^*) := \partial_x l(x, u^*) \times \partial_u[-l](x, u^*).$$

A continuación analizaremos la relación de los operadores de estos problemas con respecto a $\partial\varphi$. Para ello definimos las funciones $\pi_1 : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^n$ y $\pi_2 : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^p$, como:

$$\pi_1(x, y) = x, \quad \pi_2(x, y) = y\,.$$

Al cumplirse que $\pi_1(\partial\varphi(x,0)) \subseteq \partial_x\varphi(x,0) = \partial f(x)$, consideramos el siguiente problema variacional respecto del operador $\pi_1(\partial\varphi(\cdot,0))$:

$$\text{Determinar } x \in \mathbb{R}^n \text{ tal que } 0 \in T_P(x) \tag{P_V}$$

donde $T_P : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ se define como

$$T_P(x) := \{x^* \in \mathbb{R}^n : \exists u^* \in \mathbb{R}^p, \text{ tal que } (x^*, u^*) \in \partial\varphi(x,0)\}.$$

Se observa que x es una solución de (P_V) si y solo si

$$\exists\, u^* \in \mathbb{R}^p \ \text{ tal que } \ (0, u^*) \in \partial\varphi(x,0),$$

el cual es equivalente a

$$x \in \operatorname{argmin} f\,, \quad u^* \in \operatorname{argmax} \beta \quad \text{ y } \quad \mathrm{V_P} = \mathrm{V_D}. \tag{1.3}$$

Si φ es convexa propia y $0 \in \operatorname{ri}(\pi_2(\operatorname{dom}\varphi))$ entonces $T_P(x) = \partial_x\varphi(x,0)$, por lo tanto el problema (P_V) es equivalente a la formulación variacional de (P). Además bajo estas condiciones el problema dual (D) admite solución y su valor optimal coincide con el del problema (P).

Análogamente, como $\pi_2(\partial\varphi^*(0,u^*)) \subseteq \partial_{u^*}\varphi^*(0,u^*) = \partial\beta(u^*)$, consideramos el siguiente problema variacional:

$$\text{Determinar } u^* \in \mathbb{R}^p \text{ tal que } 0 \in T_D(u^*) \tag{D_V}$$

donde $T_D : \mathbb{R}^p \rightrightarrows \mathbb{R}^p$ se define como

$$T_D(u^*) := \{u \in \mathbb{R}^p : \exists x \in \mathbb{R}^n, \text{ tal que } (x,u) \in \partial\varphi^*(0,u^*)\}.$$

Se observa que u^* es una solución de (D_V) si y solo si

$$\exists\, x \in \mathbb{R}^n \ \text{ tal que } \ (x,0) \in \partial\varphi^*(0,u^*),$$

el cual es equivalente a

$$u^* \in \operatorname{argmax} \beta\,, \quad x \in \operatorname{argmin} \tilde{f} \quad \text{ y } \quad V_D = V_{DD}. \tag{1.4}$$

Si φ es propia y $0 \in \mathrm{ri}\,(\pi_1(\mathrm{dom}\,\varphi^*))$ entonces $T_D(u^*) = \partial_y \varphi^*(0, u^*)$, por lo tanto el problema (D_V) es equivalente a la formulación variacional de (D). Además bajo estas condiciones el problema bidual (DD) admite solución y su valor optimal coincide con el del problema (D).

Los operadores T_D y T_l, bajo hipóteis no restrictivas, se pueden reformular por medio de proyecciones de φ:

Si suponemos φ convexa sci y propia, entonces $\partial\varphi^* = (\partial\varphi)^{-1}$, luego se obtiene una nueva representacion de T_D:

$$T_D(u^*) = \{u \in \mathbb{R}^p : \exists x \in \mathbb{R}^n, \text{ tal que } (0, u^*) \in \partial\varphi(x, u)\}.$$

Respecto al problema de encontrar un punto silla de l, se cumple que si para cada x la función φ_x es propia convexa y sci sobre $\mathbb{R}^p$, entonces

$$(x^*, u^*) \in \partial\varphi(x, u) \text{ si y solo si } x^* \in \partial_x l(x, u^*),\ u \in \partial_u[-l](x, u^*)$$

luego se obtiene una nueva representación de T_l:

$$T_l(x, u^*) := \{(x^*, u) \in \mathbb{R}^n \times \mathbb{R}^p : (x^*, u^*) \in \partial\varphi(x, u)\}.$$

En ambas reformulaciones encontrar una solución de sus problemas, es equivalente a encontrar $(x, u^*) \in \mathbb{R}^n \times \mathbb{R}^p$ tal que $(0, u^*) \in \partial\varphi(x, 0)$, y esta condición se cumple si y solo si $x \in \mathrm{argmin}\, f$; $u^* \in \mathrm{argmax}\, \beta$ y $V_P = V_D$.

Los resultados anteriores, nos muestra que dado el operador $\partial\varphi$, es posible obtener las formulaciones variacionales del problema primal, dual y lagrangiano, por medio de proyecciones de $\partial\varphi$. Esto nos inspira a desarrolar un marco de dualidad para el problema de desigualdad variacional.

1.2. Dualidad para el problema de Desigualdad Variacional

Dado $F_0 : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ una multifunción, consideremos el problema de desigualdad variacional (inclusión diferencial):

$$\text{Determinar } x \in \mathbb{R}^n \text{ tal que } 0 \in F_0(x). \tag{V}$$

Asociado al problema (V), decimos que un operador $F : \mathbb{R}^n \times \mathbb{R}^p \rightrightarrows \mathbb{R}^n \times \mathbb{R}^p$ es una perturbación de F_0, si cumple la siguiente propiedad:

$$x^* \in F_0(x) \text{ si y solo si } \exists\, u^* \in \mathbb{R}^p \text{ tal que } (x^*, u^*) \in F(x, 0).$$

Consideremos $F : \mathbb{R}^n \times \mathbb{R}^p \rightrightarrows \mathbb{R}^n \times \mathbb{R}^p$ un operador de perturbación de F_0. Se deduce que el problema (V) es equivalente a:

$$\text{Determinar } x \in \mathbb{R}^n \text{ tal que } \exists\, u^* \in \mathbb{R}^p \text{ con } (0, u^*) \in F(x, 0) \tag{V_p}$$

Este problema es equivalente al problema llamado **problema minimax**:

$$\text{Determinar } (x, u^*) \in \mathbb{R}^n \times \mathbb{R}^p \text{ tal que } (0, 0) \in L(x, u^*) \tag{V_l}$$

donde $L : \mathbb{R}^n \times \mathbb{R}^p \rightrightarrows \mathbb{R}^n \times \mathbb{R}^p$ se define como

$$L(x, u^*) := \{(x^*, u) \in \mathbb{R}^n \times \mathbb{R}^p : (x^*, u^*) \in F(x, u)\}.$$

Esto nos lleva a considerar el siguiente problema llamado **problema dual**:

$$\text{Determinar } u^* \in \mathbb{R}^p \text{ tal que } 0 \in G_0(u^*) \tag{V_d}$$

donde $G_0 : \mathbb{R}^p \rightrightarrows \mathbb{R}^p$ se define como

$$G_0(u^*) := \{u \in \mathbb{R}^p : \exists x \in \mathbb{R}^n, \text{ tal que } (0, u^*) \in F(x, u)\}.$$

Capítulo 2

Lagrangiano Aumentado para Problemas de Desigualdad Variacional

Similar al capítulo anterior comencemos desarrollando este método para problemas de optimización; para luego extenderlo a problemas de desigualdad variacional. Comencemos asumiendo que el conjunto admisible está determinado por un sistema de igualdades lineales, las tecnicas desarrolladas se extienden inmediatamente al caso de sistema de desigualdades lineales . Finalmente en la sección 2,2 trataremos el caso donde el conjunto admisible está formado por un sistema de desigualdades convexas, no necesariamente lineales.

2.1. El caso con restricciones de igualdades lineales

Comencemos recordando este método para problemas de minimización sobre un subespacio afín

$$V_P = \inf_x \left[f(x) : Ax = a \right] \tag{P}$$

donde $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ es una función convexa sci y propia, $A \in \mathbb{R}^{p \times n}$ y $a \in \mathbb{R}^p$.

Usaremos la hipótesis siguiente sobre la función f:

(H1) existe $\gamma \geq 0$ tal que la función $\tilde{f} : \mathbb{R}^n \to \bar{\mathbb{R}}$ definida por

$$\tilde{f}(x) := f(x) - \frac{\gamma}{2}\|x\|^2,$$

es conveva. Este es sci y propia debido a que f lo es.

Esta hipótesis se cumple para $\gamma > 0$ si y sólo si f es fuertemente convexa con módulo $\gamma > 0$. Sin embargo la hipótesis siempre se satisface para $\gamma = 0$.

Dado $r > 0$, consideramos la siguiente función de perturbacion para (P):

$$\varphi(x,u) = \begin{cases} f(x) + \frac{1}{2r}\|u\|^2 & \text{si } Ax + u = a, \\ +\infty & \text{caso contrario.} \end{cases} \tag{2.1}$$

El problema dual asociado es:

$$V_D := \sup_{u^*} [\beta(u^*)] \tag{D}$$

donde

$$\beta(u^*) := \inf_x \left[f(x) + \frac{1}{2r}\|Ax - a\|^2 + \langle Ax - a, u^* \rangle \right]. \tag{M_{u^*}}$$

La función lagrangiana asociada a esta perturbación, llamada función lagrangiana aumentada de (P), es:

$$l(x, u^*) = f(x) + \langle Ax - a, u^* \rangle + \frac{1}{2r}\|Ax - a\|^2.$$

Se verifica que los problemas (P) y (P_l) son equivalentes. También son equivalentes los problemas (D) y (D_l).

Respecto a la existencia de solución del problema M_{u^*}, Usaremos la siguiente hipótesis:

(H2) para todo $u^* \in \mathbb{R}^p$ existe una solución x_{u^*} de (M_{u^*}).

Observe que si $\gamma > 0$, entonces f es fuertemente convexa sci propia, y por lo tanto **(H2)** se cumple, en este caso, la solución óptima de (M_{u^*}) es única.

Bajo la hipótesis **(H2)** se demuestra (ver [5]) que β es diferenciable sobre todo $\mathbb{R}^p$ con

$$\nabla\beta(u^*) = Ax_{u^*} - a.$$

Respecto a la solución del problema (D), haremos uso de las siguientes hipótesis:

(H3) $A \in \mathbb{R}^{p\times n}$ es de rango p

(H4) $V_P > -\infty$ y existe $\tilde{x} \in \operatorname{int}(\operatorname{dom}(f))$ tal que $A\tilde{x} = a$.

Bajo la hipótesis **(H4)** el conjunto solución de (D) es cerrado no vacío y $V_P = V_D$ finito. Estas propiedades se siguen cumpliendo si consideramos interior relativo en lugar de interior en **(H4)**. El conjunto solución de (D) es acotado (y por lo tanto compacto) si además de la condición **(H4)**, la condición **(H3)** se cumple.

Tenemos la siguiente relación entre las soluciones de (P) y (D).

Proposición 2.1 *Asumiendo la hipótesis $(H2)$ y la condicion $V_P = V_D$, se cumple que $\bar{u} \in \operatorname{sol}(D)$ si y sólo si $\operatorname{sol}(M_{\bar{u}}) = \operatorname{sol}(P)$.*

Prueba. Sea $\bar{u} \in \operatorname{sol}(D)$, entonces dado $x_{\bar{u}} \in \operatorname{sol}(M_{\bar{u}})$ se cumple que $\nabla\beta(\bar{u}) = Ax_{\bar{u}} - a = 0$, luego se tiene que

$$V_P = \beta(\bar{u}) = f(x_{\bar{u}}) + \frac{1}{2r}\|Ax_{\bar{u}} - a\|^2 + \langle Ax_{\bar{u}} - a, \bar{u}\rangle = f(x_{\bar{u}})$$

por lo tanto $x_{\bar{u}} \in \operatorname{sol}(P)$. Recíprocamente, si $\bar{x} \in \operatorname{sol}(P)$, entonces $A\bar{x}-a = 0$, luego se tiene que

$$\beta(\bar{u}) = V_P = f(\bar{x}) = f(\bar{x}) + \frac{1}{2r}\|A\bar{x} - a\|^2 + \langle A\bar{x} - a, \bar{u}\rangle$$

por lo tanto $\bar{x} \in \operatorname{sol}(M_{\bar{u}})$.

Por otro lado, la igualdad $\text{sol}\,(M_{\bar{u}}) = \text{sol}\,(P)$, implica $x_{\bar{u}} \in \text{sol}\,(P) \cap \text{sol}\,(M_{\bar{u}})$ y por lo tanto $\nabla\beta(\bar{u}) = Ax_{\bar{u}} - a = 0$. Se deduce que $\bar{u} \in \text{sol}\,(D)$. ■

Observe que las siguientes dos condiciones juntas: $\gamma > 0$ en $(H1)$ y $\text{ri}\,(\text{dom}\,(f)) \cap \{x : Ax = a\} \neq \emptyset$, implica que (P) admite una única solución y satisface $V_P = V_D$, siendo este valor finito.

2.1.1. Desigualdad variacional asociada al problema de optimización

Denotamos por $C := \{x \in \mathbb{R}^n : Ax = a\}$ al conjunto admisible de (P) el cual es convexo cerrado.

Usaremos la siguiente hipótesis:

(H5) $\text{ri}\,(\text{dom}\,(f)) \cap C \neq \emptyset$.

Asumiendo $(H5)$, la formulación variacional del problema (P) equivale a:

$$\text{Determinar } x \in \mathbb{R}^n \text{ tal que } 0 \in \partial f(x) + N_C(x) \qquad (V_{op})$$

siendo $N_C(x)$ el cono normal de C en x, el cual posee la siguiente expresión:

$$N_C(x) = \begin{cases} \text{Im}\,(\text{A}^{\text{t}}) & \text{si } Ax = a, \\ \emptyset & \text{caso contrario.} \end{cases} \tag{2.2}$$

Por lo tanto, bajo la condición $(H5)$, el problema de encontrar una solución de (P) equivale a:

$$\text{Determinar } x \in \mathbb{R}^n \text{ tal que } \exists\, u \in \mathbb{R}^p \text{ satisfaciendo}$$

$$\begin{cases} 0 \in \partial f(x) + A^t u, \\ Ax = a. \end{cases}$$

Denotemos por $D := \{(x,u) \in \mathbb{R}^n \times \mathbb{R}^p : Ax + u = a\}$, entonces el subdiferencial de la función φ definida en (2.1),

$$\varphi(x,u) = f(x) + \frac{1}{2r}\|u\|^2 + \delta_D(x,u),$$

es

$$\partial\varphi(x,u) = \left[\partial f(x) \times \{\frac{1}{r}u\}\right] + N_D(x,u),$$

el cual es equivalente a:

$$\partial\varphi(x,u) = \begin{cases} \left[\partial f(x) \times \{\frac{1}{r}u\}\right] + \bigcup_{w\in\mathbb{R}^p} \{(A^t w, w)\} & \text{si } Ax + u = a, \\ \emptyset & \text{caso contrario.} \end{cases}$$

De ese modo la relación entre $F_0 = \partial f + N_C$ y $\partial\varphi$ es:

$$x^* \in F_0(x) \;\text{ si y solo si }\; \exists\, u^* \in \mathbb{R}^p \text{ tal que } \;(x^*, u^*) \in \partial\varphi(x,0).$$

Entonces $\partial\varphi$ es un operador de perturbación, del cual obtenemos los problemas dual (V_{od}) y lagrangiana (V_{ol}), equivalentes al problema (V_{op}), donde sus respectivos operadores se deducen que son:

$$G_0(u^*) = \{a - Ax : 0 \in \partial f(x) + A^t u^* + \frac{1}{r}A^t(Ax - a)\},$$

$$L(x,u^*) = \{\partial f(x) + A^t u^* + \frac{1}{r}A^t(Ax - a)\} \times \{a - Ax\}.$$

2.1.2. El problema de desigualdad variacional

Asociado a un operador $T : \mathbb{R}^n \to \mathbb{R}^n$ y al conjunto $C = \{x : Ax = a\}$, consideremos el siguiente problema de desigualdad variacional:

$$\text{Determinar } x \in C \text{ tal que existe } \; x^* \in T(x) \text{ satisfaciendo}$$

$$\langle x^*, y - x\rangle \geq 0 \;\text{ para todo }\; y \in C.$$

Este problema es equivalente a

$$\text{Determinar } x \in \mathbb{R}^n \text{ tal que } 0 \in T(x) + N_C(x) \tag{V_p}$$

el cual incluye al problema (V_{op}) al considerar $T = \partial f$. A continuación extenderemos los resultados anteriores obtenidos del problema (V_{op}).

Debido a (2.2), el problema (V_p) es equivalente a

$$\text{Determinar } x \in \mathbb{R}^n \text{ tal que } \exists\, u \in \mathbb{R}^p \text{ satisfaciendo}$$

$$-A^t u \in T(x) \ \text{ y } \ Ax = a.$$

Análogo al operador $\partial\varphi$, consideremos el operador $F : \mathbb{R}^n \times \mathbb{R}^p \rightrightarrows \mathbb{R}^n \times \mathbb{R}^p$ definido por

$$F(x,u) = \begin{cases} \left[T(x) \times \{\frac{1}{r}u\}\right] + \bigcup\limits_{w \in \mathbb{R}^p} \{(A^t w, w)\} & \text{si } Ax + u = a, \\ \emptyset & \text{caso contrario.} \end{cases}$$

De ese modo la relación entre $F_0 = T + N_C$ y F es:

$$x^* \in F_0(x) \ \text{ si y solo si } \ \exists\, u^* \in \mathbb{R}^p \text{ tal que } \ (x^*, u^*) \in F(x, 0).$$

Entonces F es un operador de perturbación, del cual deducimos los siguientes problemas equivalentes de (V):

El problema de desigualdad variacional dual

$$\text{Determinar } u^* \in \mathbb{R}^p \text{ tal que } \ 0 \in G_0(u^*) \tag{V_d}$$

donde

$$G_0(u^*) = \{a - Ax : 0 \in T(x) + A^t u^* + \frac{1}{r} A^t (Ax - a)\}.$$

El problema tipo lagrangiano:

$$\text{Determinar } (x, u^*) \in \mathbb{R}^n \times \mathbb{R}^p \text{ tal que } \ (0,0) \in L(x, u^*) \tag{V_l}$$

donde

$$L(x, u^*) = \{T(x) + A^t u^* + \frac{1}{r} A^t (Ax - a)\} \times \{a - Ax\}.$$

Haremos uso de las siguientes hipótesis:

$(\mathbf{H_1^\gamma})$ Existe $\gamma \geq 0$ tal que $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ satisface:

$$\langle z - z', w - w' \rangle \geq \gamma \|z - z'\|^2 \text{ cuando } w \in T(z), w' \in T(z').$$

$(\mathbf{H_2^v})$ Dado $u^* \in \mathbb{R}^p$ arbitrario, existe $x_{u^*} \in \mathbb{R}^n$ tal que:

$$0 \in T(x_{u^*}) + A^t u^* + \frac{1}{r} A^t (A x_{u^*} - a),$$

o equivalentemente se cumple que $G_0(u^*) \neq \emptyset$.

La hipótesis (H_1^v) se cumple para $\gamma > 0$ si y sólo si T es fuertemente monótono con módulo $\gamma > 0$, resultando en este caso que x_{u^*} en (H_2^v) es único. De otro lado, si T es monótono, esta hipótesis se satisface para $\gamma = 0$.

La hipótesis (H_2^v) se cumple si $\gamma > 0$ y T es monótono maximal, pues al ser $A^t(A(\cdot) - a)$ una función continua monótono en $\mathbb{R}^n$, se deduce que $T(\cdot) + \frac{1}{r} A^t(A(\cdot) - a)$ es fuertemente monótono maximal. Luego dado $u^* \in \mathbb{R}^p$ existe un único $x \in \mathbb{R}^n$ tal que $-A^t u^* \in T(x) + \frac{1}{r} A^t(Ax - a)$.

La siguiente proposición muestra que G_0 es univaluada, esto es, una función.

Proposición 2.2 *Asumiendo las hipótesis* (H_1^v) *y* (H_2^v) *se cumple que* G_0 *es un operador univaluado, cuyo dominio es* $\mathbb{R}^p$ *y satisface:*

$$\langle G_0(u) - G_0(c), u - c \rangle \geq \frac{1}{r} \|G_0(u) - G_0(c)\|^2 + \gamma \|x_u - x_c\|^2 \tag{2.3}$$

En particular se cumple que G_0 *es una función monótono y Lipschitz.*

Prueba. Dado $u, c \in \mathbb{R}^p$, por (H_2^v), se tiene que existen $x_u, x_c \in \mathbb{R}^n$ tal que

$$-A^t u - \frac{1}{r} A^t (A x_u - a) \in T(x_u) \ \text{ y } \ -A^t c - \frac{1}{r} A^t (A x_c - a) \in T(x_c)$$

Luego por (H_1^v), se tiene

$$\langle A^t c - A^t u + \frac{1}{r} A^t A (x_c - x_u), x_u - x_c \rangle \geq \gamma \|x_u - x_c\|^2$$

$$\langle c - u, A x_u - A x_c \rangle - \frac{1}{r} \|A(x_c - x_u)\|^2 \geq \gamma \|x_u - x_c\|^2$$

$$\langle c - u, (a - A x_c) - (a - A x_u) \rangle \geq \frac{1}{r} \|(a - A x_c) - (a - A x_u)\|^2 + \gamma \|x_u - x_c\|^2 \tag{2.4}$$

por la desigualdad de cauchy schwarz, se tiene

$$\|c-u\|\|(a-Ax_c)-(a-Ax_u)\| \geq \frac{1}{r}\|(a-Ax_c)-(a-Ax_u)\|^2$$

$$\|(a-Ax_c)-(a-Ax_u)\| \leq r\|c-u\|$$

Por lo tanto G_0 es una función Lipschitz sobre $\mathbb{R}^p$, en particular univaluado, luego (2.3) se deduce de (2.4). ■

Debido a que G_0 es una función monótono y Lipschitz, consideramos el método de la subgradiente para resolver el problema (V_d), entonces obtenemos el siguiente algoritmo.

Algoritmo (LACI)

1. **Inicialización:** Escoger $u_0 \in \mathbb{R}^p$ arbitrario, luego hacer $k = 0$.

2. **Etapa k:** Buscar $x_k \in \mathbb{R}^n$ tal que

$$0 \in T(x_k) + A^t u_k + \frac{1}{r}A^t(Ax_k - a)$$

   - Si $Ax_k - a = 0$, PARAR.
   - Caso contrario, hacer $u_{k+1} = u_k + \frac{1}{r}(Ax_k - a)$, $k = k + 1$. Retornar al ítem 2.

Si el algoritmo se detiene en la etapa k, entonces $G_0(u_k) = Ax_k - a = 0$, y por lo tanto u_k es una solución del problema (V_d) y x_k es una solución del problema (V_p).

Theorem 2.1.1 *Bajo las hipótesis anteriores. Supongamos que existe una solución de (V_p). La sucesión generada $\{u_k\}$ converge hacia u', donde u' es una solución de (V_d). Si $\gamma > 0$, la sucesión $\{x_k\}$ converge a $x_{u'}$ la única solución de (V_p). Si T es maximal, dado x' un valor de adherencia de $\{x_k\}$ entonces x' es una solución de (V_p), más aún (x', u') es una solución de (V_l).*

Prueba. Al ser (V_p) equivalente a (V_d), entonces existe $\bar{u}$ solución de (V_d), luego $G(\bar{u}) = 0$. Por otro lado se observa que $x_{u_k} = x_k$. Luego considerando $u = \bar{u}$ y $p = u_k$ en (2.3), se obtiene

$$\langle u_k - \bar{u}, G_0(u_k)\rangle \geq \frac{1}{r}\|G_0(u_k)\|^2 + \gamma\|x_k - x_{\bar{u}}\|^2$$

$$\langle u_{k+1} - \bar{u}, G_0(u_k)\rangle \geq \gamma\|x_k - x_{\bar{u}}\|^2 \tag{2.5}$$

Como $u_{k+1} - \bar{u} + \frac{1}{r}G_0(u_k) = u_k - \bar{u}$, se tiene que

$$\|u_k - \bar{u}\|^2 = \|u_{k+1} - \bar{u}\|^2 + \|\frac{1}{r}G_0(u_k)\|^2 + \frac{2}{r}\langle u_{k+1} - \bar{u}, G_0(x_k)\rangle$$

Luego usando (3.5), se tiene

$$\|u_k - \bar{u}\|^2 \geq \|u_{k+1} - \bar{u}\|^2 + \frac{1}{r^2}\|G_0(u_k)\|^2 + \frac{2\gamma}{r}\|x_k - x_{\bar{u}}\|^2$$

Entonces $\{\|u_k - \bar{u}\|\}$ es decreciente y al estar acotado converge, luego se tiene que $\{G_0(u_k)\}$ tiende a 0 y si $\gamma > 0$, $\{x_k\}$ tiende a $x_{\bar{u}}$.

Como $\{u_k\}$ es acotada, existe u' valor de adherencia de $\{u_k\}$, por la continuidad de G_0 se tiene que $G_0(u') = 0$, Luego tomando $\bar{u} = u'$ en el principio, se obtiene que $\{\|u_k - u'\|\}$ es decreciente, por lo tanto se deduce que $\{u_k\}$ converge a u' una solución de (V_d).

Si x' es un valor de adherencia de $\{x_k\}$, como $u_{k+1} = u_k + \dfrac{1}{r}(Ax_k - a)$, entonces $Ax' = a$, además si T es maximal (el grafico de T es cerrado), como $-A^t(u_{k+1}) \in T(x_k)$ se deduce que $-A^t(u') \in T(x')$, por lo tanto (x', u') es una solución de (V_l). ■

Theorem 2.1.2 *Asumamos T monótono maximal y (H_1^v) con $\gamma > 0$. Entonces*

1. *Si* $\mathrm{ri}\,(\mathrm{dom}\,(T)) \cap C \neq \emptyset$ *entonces existe una única solución de (V_p).*

2. *Si $A \in \mathbb{R}^{p\times n}$ es de rango p e* $\mathrm{int}\,(\mathrm{dom}\,(T)) \cap C \neq \emptyset$, *entonces el conjunto solución del problema (V_d) es compacto no vacío.*

Prueba.

1. T es fuertemente monótono maximal, N_C es monótono maximal y como $\mathrm{ri}\,(\mathrm{dom}\,(T)) \cap \mathrm{ri}\, C = \mathrm{ri}\,(\mathrm{dom}\,(T)) \cap C \neq \emptyset$ se cumple que $T + N_C$ es fuertemente monótono maximal, por lo tanto existe una única solución de (V_p).

2. existe $\tilde{x} \in \mathrm{int}\,(\mathrm{dom}\,(T))$ tal que $A\tilde{x} = a$. Luego existe V una vencidad de $\tilde{x}$ tal que $V \subseteq \mathrm{int}\,(\mathrm{dom}\,(T))$, por ser A de rango p, se deduce que $U := a - A(V)$ es una vecindad de 0.

   Dado $u^* \in U$, existe $x^* \in V$ tal que $a - Ax^* = u^*$, denotando $C_{u^*} := \{x : Ax = a - u^*\}$ se tiene que $x^* \in \mathrm{int}\,(\mathrm{dom}\,(T)) \cap C_{u^*}$, luego por el item anterior, se tiene que existe x_{u^*} tal que $0 \in (T + N_{C_{u^*}})(x_{u^*})$ lo cual es equivale a que exista $u' \in \mathbb{R}^p$ tal que $-A^t u' \in T(x_{u^*})$ y $Ax_{u^*} = a - u^*$. Denotando $\hat{u} = u' + \frac{1}{r}u^*$, se tiene que

   $$-A^t[\hat{u} + \frac{1}{r}(Ax_{u^*} - a)] \in T(x_{u^*})$$

   Entonces $G_0(\hat{u}) = a - Ax_{u^*} = u^*$.

   Por lo tanto se obtiene que $U \subseteq \mathrm{dom}\,(G_0^{-1})$, es decir $0 \in \mathrm{int}\,(\mathrm{dom}\,(G_0^{-1}))$. Entonces se cumple que $G_0^{-1}(0)$ es compacto no vacío. ■

2.2. El caso con restricciones de desigualdades convexas

Comencemos recordando este método para problemas de optimización convexa de la forma

$$V_P = \inf_x \,[f(x) : g_j(x) \leq 0,\ j = 1, \ldots, q] \qquad (P)$$

donde $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ es una función convexa sci y propia, y las funciones $g_j : \mathbb{R}^n \to \mathbb{R}$ convexas.

Nuevamente, usaremos la hipótesis siguiente sobre la función f:

($\hat{\mathbf{H}}\mathbf{1}$) Existe $\gamma \geq 0$ tal que la función $\tilde{f} : \mathbb{R}^n \to \bar{\mathbb{R}}$ definida por

$$\tilde{f}(x) := f(x) - \frac{\gamma}{2}\|x\|^2,$$

es conveva.

Dado $r > 0$, consideremos la función de perturbación para (P),

$$\varphi(x, v) = \begin{cases} f(x) + \frac{1}{2r}\|v\|^2 & \text{si } g(x) + v \leq 0, \\ +\infty & \text{caso contrario.} \end{cases} \tag{2.6}$$

El problema dual asociado es

$$V_D := \sup_{v^*} [\beta(v^*)] \tag{D}$$

donde

$$\beta(v^*) := \inf_x \left[f(x) + \frac{1}{2r}\|[g(x) + rv^*]^+\|^2 - \frac{r}{2}\|v^*\|^2 \right]. \tag{M_{v^*}}$$

donde $[\alpha]^+$ es un vector cuyas componentes son las partes positivas de las componentes de α.

La función lagrangiana asociada a esta perturbación, llamada *función lagrangiana aumentada de* (P), es:

$$l(x, v^*) = f(x) + \frac{1}{2r}\|[g(x) + rv^*]^+\|^2 - \frac{r}{2}\|v^*\|^2.$$

Se deduce por lo tanto que los problemas (P) y (P_l) son equivalentes, también son equivalentes los problemas (D) y (D_l).

Haremos también uso de la hipótesis siguiente:

$(\mathbf{\hat{H}2})$ para todo $v^* \in \mathbb{R}^q$ existe una solución x_v^* de (M_{v^*}).

Observe que si $\gamma > 0$, entonces f es fuertemente convexa sci propia, y por lo tanto $(\hat{H}2)$ se cumple, en este caso la solución óptima de (M_{v^*}) es única.

Bajo la hipótesis $(\hat{H}2)$ se demuestra (ver [5]) que β es diferenciable sobre todo $\mathbb{R}^q$ con

$$\nabla\beta(v^*) = [g(x_{v^*}) + rv^*]^+ - rv^*.$$

Se deduce que $\bar{v}$ es solución del problema (D) si y solo si

$$[g(x_{\bar{v}}) + r\bar{v}]^+ - r\bar{v} = 0, \tag{2.7}$$

que equivalente a la relación:

$$\bar{v} \geq 0,\ \ g(x_{\bar{v}}) \leq 0 \ \text{ y } \ \langle \bar{v}, g(x_{\bar{v}})\rangle = 0. \tag{2.8}$$

Respecto a la solución del problema (D), haremos uso de la siguiente hipótesis:

$(\tilde{H}3)$ $V_P > -\infty$ y existe $\tilde{x} \in \operatorname{dom}(f)$ tal que $g(\tilde{x}) < 0$.

Bajo la hipótesis $(\tilde{H}3)$ el conjunto solución de (D) es compacto no vacío y $V_P = V_D$, siendo este un valor finito.

Se cumple la siguiente relación entre las soluciones de (P) y (D):

Proposición 2.3 *Asumiendo la hipótesis $(\hat{H}2)$ y la condición $V_P = V_D$. Se cumple que dado $\bar{v} \in \operatorname{sol}(D)$, entonces $\operatorname{sol}(M_{\bar{v}}) = \operatorname{sol}(P)$.*

Prueba. Dado $\bar{v} \in \operatorname{sol}(D)$. Consideremos $x_{\bar{v}} \in \operatorname{sol}(M_{\bar{v}})$ entonces $\nabla\beta(\bar{v}) = [g(x_{\bar{v}}) + r\bar{v}]^+ - r\bar{v} = 0$, de donde $\bar{v} \geq 0$ y $g(x_{\bar{v}}) \leq 0$, luego se tiene que

$$V_P = V_D = \beta(\bar{v}) = f(x_{\bar{v}}) + \frac{1}{2r}\|[g(x_{\bar{v}}) + r\bar{v}]^+\|^2 - \frac{r}{2}\|\bar{v}\|^2 = f(x_{\bar{v}}),$$

y por lo tanto $x_{\bar{v}} \in \operatorname{sol}(P)$. Para mostrar la inclusión contraria, consideremos $\bar{x} \in \operatorname{sol}(P)$, entonces $g(\bar{x}) \leq 0$, y por lo tanto

$$V_P = \beta(\bar{v}) \leq f(\bar{x}) + \frac{1}{2r}\|[g(\bar{x}) + r\bar{v}]^+\|^2 - \frac{r}{2}\|\bar{v}\|^2 \leq f(\bar{x}) = V_P$$

de donde $\bar{x} \in \operatorname{sol}(M_{\bar{v}})$. ■

Si $\operatorname{dom}(f) \cap \{x : g(x) < 0\} \neq \emptyset$ y $\gamma > 0$ en $(\hat{H}1)$, entonces el problema (P) admite una única solución, el conjunto solución de (D) es compacto no vacío, y $V_P = V_D$, siendo este valor finito.

2.2.1. Desigualdad variacional asociado al problema de optimización

Denotamos por $C := \{x \in \mathbb{R}^n : g(x) \leq 0\}$ al conjunto admisible del problema (P) el cual es convexo cerrado.

Usaremos la siguiente hipótesis de regularizacón (condición de Slater):

($\hat{\mathbf{H}}\mathbf{4}$) Existe $\tilde{x} \in \operatorname{ri}(\operatorname{dom}(f))$ tal que $g(\tilde{x}) < 0$.

Asumiendo $(\hat{H}4)$, la formulación varacional del problema (P) equivale a:

$$\text{Determinar } x \in \mathbb{R}^n \text{ tal que } 0 \in \partial f(x) + N_C(x), \qquad (V_{op})$$

siendo $N_C(x)$ el cono normal de C en x. Si existe $\tilde{x} \in \mathbb{R}^n$ tal que $g(\tilde{x}) < 0$ (en particular, si $(\hat{H}4)$ se cumple), la expresión del cono normal $N_C(x)$ es:

$$N_C(x) = \begin{cases} \displaystyle\sum_{j \in J(x)} N_j(x) & \text{si } x \in C \text{ y } J(x) \neq \emptyset, \\ \{0\} & \text{si } x \in C \text{ y } J(x) = \emptyset, \\ \emptyset & \text{caso contrario }, \end{cases} \qquad (2.9)$$

donde $J(x) := \{j \in 1, \cdots, q : g_j(x) = 0\}$ y

$$N_j(x) := \mathbb{R}_+ \partial g_j(x) = \bigcup_{\alpha \geq 0} \alpha \partial g_j(x).$$

Por lo tanto, bajo la condición $(\hat{H}4)$, el problema de encontrar una solución de (P) equivale a:

Determinar $x \in \mathbb{R}^n$ tal que $\exists\, v \in \mathbb{R}^q_+$ satisfaciendo

$$\begin{cases} 0 \in \partial f(x) + \sum_{i=1}^{p} v_i \partial g_i(x), \\ \\ g(x) \leq 0 \text{ y } \langle v, g(x) \rangle = 0. \end{cases}$$

Denotemos por $D := \{(x, v) \in \mathbb{R}^n \times \mathbb{R}^q : g(x) + v \leq 0\}$, entonces

$$N_D(x,v) = \begin{cases} \displaystyle\sum_{j \in J(x,v)} N_j(x,v) & \text{si } \ (x,v) \in D \ \text{ y } \ J(x,v) \neq \emptyset, \\ \{0\} & \text{si } \ (x,v) \in D \ \text{ y } \ J(x,v) = \emptyset, \\ \\ \emptyset & \text{caso contrario }, \end{cases}$$

donde $J(x,v) := \{j \in 1, \cdots, q : g_j(x) - v_j = 0\}$ y

$$N_j(x,v) := \mathbb{R}_+[\partial g_j(x) \times \{e_j\}] = \bigcup_{\alpha \geq 0} \alpha[\partial g_j(x) \times \{e_j\}].$$

De este modo, el subdiferencial de la función φ definida en (2.6),

$$\varphi(x,v) = f(x) + \frac{1}{2r}\|v\|^2 + \delta_D(x,v),$$

es

$$\partial\varphi(x,v) = \left[\partial f(x) \times \{\frac{1}{r}v\}\right] + N_D(x,v),$$

el cual es equivalente a:

$$\begin{cases} \left[\partial f(x) \times \{\frac{1}{r}v\}\right] + \displaystyle\sum_{j \in J(x,v)} N_j(x,v) & \text{si } \ (x,v) \in D \ \text{ y } \ J(x,v) \neq \emptyset, \\ \partial f(x) \times \{\frac{1}{r}v\} & \text{si } \ (x,v) \in D \ \text{ y } \ J(x,v) = \emptyset, \\ \\ \emptyset & \text{caso contrario }. \end{cases}$$

De este modo, la relación entre $F_0 := \partial f + N_C$ y $\partial\varphi$ es:

$$x^* \in F_0(x) \text{ si y solo si } \exists\, v^* \in \mathbb{R}^p \text{ tal que } (x^*, v^*) \in \partial\varphi(x, 0).$$

Entonces $\partial\varphi$ es un operador de perturbación. Asociado a este operador de perturbación, obtenemos los problemas dual (V_{od}) y lagrangiana (V_{ol}), equivalentes al problema (V_{op}), donde sus respectivos operadores se deducen que son:

$$G_0(v^*) = \left\{ rv^* - [g(x_{v^*}) + rv^*]^+ : x_{v^*} \in \operatorname{argmin}_x \left[f(x) + \frac{1}{2r} \|[g(x) + rv^*]^+\|^2 \right] \right\},$$

$$L(x, v^*) = \left\{ \partial f(x) + \frac{1}{r} \sum_{j=1}^{q} [g_j(x) + rv_j^*]^+ \partial g_j(x) \} \times \{ rv^* - [g(x) + rv^*]^+ \right\}.$$

2.2.2. El problema de desigualdad variacional

Asociado a un operador $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ y al conjunto $C = \{x : g(x) \leq 0\}$, consideremos el siguiente problema de desigualdad variacional:

$$\text{Determinar } x \in C \text{ tal que existe } x^* \in T(x) \text{ satisfaciendo}$$

$$\langle x^*, y - x \rangle \geq 0 \text{ para todo } y \in C.$$

Este problema es equivalente a

$$\text{Determinar } x \in \mathbb{R}^n \text{ tal que } 0 \in T(x) + N_C(x) \qquad (V_p)$$

el cual incluye al problema (V_{op}) al considerar $T = \partial f$. A continuación extenderemos los resultados obtenidos para el caso $T = \partial f$ analizados en la sección anterior.

Si existe $\tilde{x} \in \mathbb{R}^n$ tal que $g(\tilde{x}) < 0$, entonces por (2.9), el problema (V_p) es equivalente a:

$$\text{Determinar } v \in \mathbb{R}^n \text{ tal que } \exists\, u \in \mathbb{R}^p_+ \text{ satisfaciendo}$$

$$\begin{cases} 0 \in T(v) + \sum_{i=1}^{p} u_i \partial g_i(v), \\ g(v) \leq 0 \text{ y } \langle u, g(v) \rangle = 0. \end{cases}$$

Análogo al operador $\partial\varphi$, consideremos el siguiente operador $F : \mathbb{R}^n \times \mathbb{R}^p \rightrightarrows \mathbb{R}^n \times \mathbb{R}^p$ definido por:

$$F(x,v) = \begin{cases} \left[T(x) \times \{\frac{1}{r}v\}\right] + \displaystyle\sum_{j \in J(x,v)} N_j(x,v) & \text{si } (x,v) \in D \text{ y } J(x,v) \neq \emptyset, \\ T(x) \times \{\frac{1}{r}v\} & \text{si } (x,v) \in D \text{ y } J(x,v) = \emptyset, \\ \emptyset & \text{caso contrario .} \end{cases}$$

De ese modo la relación entre $F_0 = T + N_C$ y F es:

$$x^* \in F_0(x) \text{ si y solo si } \exists\, v^* \in \mathbb{R}^q \text{ tal que } (x^*, v^*) \in F(x, 0).$$

Entonces F es un operador de perturbación, del cual deducimos los siguientes problemas equivalentes de (V_p):

El problema de desigualdad variacional dual

$$\text{Determinar } v^* \in \mathbb{R}^q \text{ tal que } 0 \in G_0(v^*) \qquad (V_d)$$

donde

$$G_0(v^*) = \left\{ rv^* - [g(x_{v^*}) + rv^*]^+ : 0 \in T(x_{v^*}) + \frac{1}{r}\sum_{j=1}^{q}[g_j(x_{v^*}) + rv_j^*]^+ \partial g_j(x_{v^*}) \right\},$$

El problema tipo lagrangiano:

$$\text{Determinar } (x, v^*) \in \mathbb{R}^n \times \mathbb{R}^q \text{ tal que } (0,0) \in L(x, v^*) \qquad (V_l)$$

donde

$$L(x, v^*) = \{T(x) + \frac{1}{r}\sum_{j=1}^{q}[g_j(x) + rv_j^*]^+ \partial g_j(x)\} \times \{rv^* - [g(x) + rv^*]^+\}.$$

Asumiremos en adelante las siguientes hipótesis:

$(\mathbf{\hat{H}_1^v})$ Existe $\gamma \geq 0$ tal que $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ satisface:

$$\langle z - z', w - w' \rangle \geq \gamma \|z - z'\|^2 \text{ cuando } w \in T(z), w' \in T(z').$$

$(\mathbf{\hat{H}_2^v})$ Dado $v^* \in \mathbb{R}^p$ arbitrario, existe $x_{v^*} \in \mathbb{R}^n$ tal que:

$$0 \in T(x_{v^*}) + \frac{1}{r} \sum_{j=1}^{q} [g_j(x_{v^*}) + rv_i^*]^+ \partial g_j(x_{v^*}).$$

o equivalentemente se cumple que $G_0(v^*) \neq \emptyset$.

La hipótesis $(\hat{H}_1^v)$ se cumple para $\gamma > 0$ si y sólo si T es fuertemente monótono con módulo $\gamma > 0$, resultando en este caso que x_{v^*} en $(\hat{H}_2^v)$ es único. De otro lado, si T es monótono, esta hipótesis se satisface para $\gamma = 0$.

La hipótesis $(\tilde{H}_2^v)$ se cumple si $\gamma > 0$ y T es monótono maximal, pues dado $v^* \in \mathbb{R}^q$ al ser $2\sum_{j=1}^{q}[g_j(\cdot)+rv_j^*]^+\partial g_j(\cdot) = \partial_x \|[g(\cdot)+rv^*]^+\|^2$ monótono maximal sobre $\mathbb{R}^n$, se deduce que $T(\cdot) + \frac{1}{r}\sum_{j=1}^{q}[g_j(\cdot)+rv_j^*]^+\partial g_j(\cdot)$ es fuertemente monótono maximal. Por lo tanto dado $v^* \in \mathbb{R}^q$ existe un único $x \in \mathbb{R}^n$ tal que $0 \in T(x) + \frac{1}{r}\sum_{j=1}^{q}[g_j(x) + rv_i^*]^+\partial g_j(x)$.

La siguiente proposición muestra que el operador G_0 es univaluada, esto es, una función.

Proposición 2.4 *Asumiendo las hipótesis $(H1)$ y $(H2)$ se cumple que G_0 es un operador univaluado, cuyo dominio es $\mathbb{R}^q$ y satisface:*

$$\langle G_0(v) - G_0(d), v - d \rangle \geq \frac{1}{r}\|G_0(v) - G_0(d)\|^2 + \gamma \|x_v - x_d\|^2 \tag{2.10}$$

En particular se cumple que G_0 es una función monótono y Lipschitz.

Prueba. Dados $v, d \in \mathbb{R}^p$, por $(H2)$ existen x_v, $x_d \in \mathbb{R}^n$ tales que existen $s^{j,v} \in \partial g_j(x_v)$, $s^{j,d} \in \partial g_j(x_d)$ para $j = 1, \ldots, q$, satisfaciendo

$$0 \in T(x_v) + \frac{1}{r}\sum_{j=1}^{q}[g_j(x_v) + rv_j]^+ s^{j,v} \text{ y } 0 \in T(x_d) + \frac{1}{r}\sum_{j=1}^{q}[g_j(x_d) + rd_j]^+ s^{j,d}.$$

Por $(H1)$ y multiplicando por r, se tiene que

$$\langle\sum_{j=1}^{q}[g_j(x_d)+rd_j]^+s^{j,d}-\sum_{j=1}^{q}[g_j(x_v)+rv_j]^+s^{j,v},x_v-x_d\rangle \geq r\gamma\|x_v-x_d\|^2 \quad (2.11)$$

Se cumple que dados $b,e \geq 0$ y $f:\mathbb{R}^n \to \mathbb{R}$ tal que $s_1 \in \partial f(x)$, $s_2 \in f(y)$

$$(b-e)(f(x)-f(y)) \geq \langle x-y, bs_2-es_1\rangle$$

Luego aplicando esta desigualdad a cada función g_j, considerando $b_j = [g_j(x_d)+rd_j]^+$ y $e_j = [g_j(x_v)+rv_j]^+$, y sumandolos juntamente con (3.1), se deduce

$$\langle[g(x_d)+rd]^+ - [g(x_v)+rv]^+, g(x_v)-g(x_d)\rangle \geq r\gamma\|x_v-x_d\|^2 \quad (2.12)$$

Por otro lado, como para todo $y \in \mathbb{R}^q$ se cumple $y-y^+ \in N_{\mathbb{R}^q_+}(y^+)$. Denotando $\lambda_v := rv-[g(x_v)+rv]^+$ y $\lambda_d := rd-[g(x_d)+rd]^+$, se deduce por la monotonia de $N_{\mathbb{R}^q_+}$ y sumandolo con (3.2)

$$\langle[g(x_d)+rd]^+ - [g(x_u)+rv]^+, \lambda_d-\lambda_v\rangle \geq r\gamma\|x_v-x_d\|^2 \quad (2.13)$$

Al cumplirse que

$$\langle\lambda_v-\lambda_d, [g(x_v)+rv]^+ - [g(x_d)+rd]^+\rangle = r\langle\lambda_v-\lambda_d, v-d\rangle - \|\lambda_v-\lambda_d\|^2$$

entonces por (3.3) y multiplicando por $\frac{1}{r}$, se tiene

$$\langle\lambda_v-\lambda_d, v-d\rangle - \frac{1}{r}\|\lambda_v-\lambda_d\|^2 \geq \gamma\|x_v-x_d\|^2 \quad (2.14)$$

Entonces por la desigualdad de cauchy schwarz se deduce que G_0 es una función lipschitz, en particular es univaluado, luego (2.10) se deduce de (3.4) ■

Debido a que G_0 es una función monótono y Lipschitz, consideramos el método de la subgradiente para resolver el problema (V_d), entonces obtenemos el siguiente algoritmo

Algoritmo (LACD)

1. **Inicialización:** Escoger $v^0 \in \mathbb{R}^q$ arbitrario, luego hacer $k = 0$.

2. **Etapa k:** Buscar $x^k \in \mathbb{R}^n$ tal que

$$0 \in T(x^k) + \frac{1}{r}\sum_{j=1}^{q}[g_j(x^k) + rv_j^k]^+ \partial g_j(x^k)$$

   - Si $[g(x^k) + rv^k]^+ = rv^k$, PARAR.
   - Caso contrario, hacer $v^{k+1} = \dfrac{1}{r}[g(x^k) + rv^k]^+$, $k = k + 1$. Retornar al ítem 2.

Si el algoritmo se detiene en la etapa k, entonces $G_0(v^k) = [g(x^k)+rv^k]^+ = rv^k$, y por lo tanto v^k es una solución del problema (V_d) y x^k es una solución del problema (V_p).

Theorem 2.2.1 *Bajo las hipótesis anteriores. Supongamos que existe una solución de (V_p). La sucesión generada $\{v^k\}$ converge hacia v', donde v' es una solución de (V_d). Si $\gamma > 0$, la sucesión $\{x^k\}$ converge a $x_{v'}$ la única solución de (V_p). Si T es maximal, dado x' un valor de adherencia de $\{x^k\}$ entonces x' es una solución de (V_p), más aún (x', u') es una solución de (V_l).*

Prueba. Al cumplir G_0 las mismas propiedades de la proposición (2.2), la demostración de este teorema es análoga a la del teorema (2.1.1). ■

Capítulo 3

Lagrangiano Aumentado–Proximal para Problemas de Desigualdad Variacional

En este capítulo se desarrollan dos algoritmos que permiten encontrar una solución aproximada del problema variacional monótono con restriciones de igualdades lineales y desigualdades convexas.

Asociado a un operador $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ consideramos el siguiente problema de desigualdad variacional:

$$\text{Determinar } x \in \mathbb{R}^n \text{ tal que } 0 \in T(x) + N_C(x). \tag{V_p}$$

En el cual el conjunto C es definido como

$$C := \{x \in \mathbb{R}^n : Ax = a,\ g_j(x) \leq 0, j = 1,\ \ldots, q\}$$

donde $g_j : \mathbb{R}^n \to \mathbb{R}$ son convexas para todo $j = 1,\ \ldots, q$, $A \in \mathbb{R}^{p\times n}$ y $a \in \mathbb{R}^p$.

En adelante asumiremos la siguiente hipótesis:

(H1) Existe $\tilde{x} \in \mathbb{R}^n$ tal que $A\tilde{x} = a$ y $g(\tilde{x}) < 0$.

Con esta hipótesis la expresión del cono normal N_C es

$$N_C(x) = \begin{cases} \bigcup_{w \in \mathbb{R}^p} \{A^t w\} + \sum_{j \in J(x)} N_j(x) & \text{si } \; x \in C \; \text{ y } \; J(x) \neq \emptyset, \\ \bigcup_{w \in \mathbb{R}^p} \{A^t w\} & \text{si } \; x \in C \; \text{ y } \; J(x) = \emptyset, \\ \emptyset & \text{caso contrario .} \end{cases}$$

donde $J(x) = \{j \in 1, \cdots, q : g_j(x) = 0\}$ y

$$N_j(x) = \bigcup_{\alpha \geq 0} \alpha[\partial g_j(x)].$$

De ese modo, por $(H1)$, el problema (V_p) es equivalente a:

$$\text{Determinar } x \in \mathbb{R}^n \text{ tal que } \exists\, (u, v) \in \mathbb{R}^p \times \mathbb{R}^q_+ \text{ satisfaciendo}$$

$$\begin{cases} 0 \in T(x) + A^t u + \sum_{i=1}^p v_i \partial g_i(x), \\ Ax = a, \; g(x) \leq 0 \text{ y } \langle v, g(x) \rangle = 0. \end{cases}$$

3.1. Método Teórico

Dado $\rho > 0$, $s > 0$ y $r > 0$, inspirados en el problema dual asociado a (V_p) y el método proximal, definimos el siguiente operador:

$$D(z, u, v) := \left\{ (\rho(z - x), a - Ax, rv - [g(x) + rv]^+) : 0 \in W_{(z,u,v)}(x) \right\},$$

donde el operador $W_{(z,u,v)} : \mathbb{R}^n \to \mathbb{R}^n$ es definido como:

$$W_{(z,u,v)}(x) := T(x) + \rho(x - z) + A^t[u + \frac{1}{s}(Ax - a)] + \frac{1}{r} \sum_{j=1}^p [g_j(x) + rv_j]^+ \partial g_j(x).$$

Se cumple que el siguiente problema

$$\text{Determinar } (z, u, v) \in \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^q \text{ tal que } 0 \in D(z, u, v) \qquad (V_D).$$

es equivalente al problema (V_p), en el sentido que z^* es solución de (V_p) si y solo si existe $(u^*, v^*) \in \mathbb{R}^p \times \mathbb{R}^q$ tal que (z^*, u^*, v^*) es solución de (V_D).

En adelante asumiremos la siguiente hipótesis

(H2) $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ es un operador monótono maximal.

Con esta hipótesis, se muestra en la siguiente proposición, que D satisface importantes propiedades. Para una mejor notación, dados los parametros $(\gamma, \beta, \alpha) > 0$, introducimos el siguiente producto interno y respectiva norma inducida sobre $\mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^q$:

$$\langle (z,u,v),(y,c,d)\rangle_{(\gamma,\beta,\alpha)} := \gamma\langle z,y\rangle + \beta\langle u,c\rangle + \alpha\langle v,d\rangle$$

$$\|(z,u,v)\|^2_{(\gamma,\beta,\alpha)} := \gamma\|z\|^2 + \beta\|u\|^2 + \alpha\|v\|^2.$$

Proposición 3.1 *El operador D es univaluado, cuyo dominio es $\mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^q$. Además dados (z,u,v), $(y,c,d) \in \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^q$ se cumple que:*

$$\langle D(z,u,v) - D(y,c,d), (z,u,v) - (y,c,d)\rangle \geq \|D(z,u,v) - D(y,c,d)\|^2_{(\frac{1}{\rho},\frac{1}{s},\frac{1}{r})}$$

En particular se cumple que D es una función monótono y Lipschitz.

Prueba. Devido a la hipóteis **(H2)**, se observa que dado (z,u,v), el operador $W_{(z,u,v)}$ es fuertemente monótono maximal, por lo tanto se deduce que el dominio de S es $\mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^q$.

Dados $(z,u,v),(y,c,d) \in \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^q$, sean $x,\ x^* \in \mathbb{R}^n$ tales que existen $s^{j,x} \in \partial g_j(x)$, $s^{j,x^*} \in \partial g_j(x^*)$ para $j = 1,\ldots,q$, satisfaciendo

$$0 \in T(x) + \rho(x-z) + A^t[u + \frac{1}{s}(Ax-a)] + \frac{1}{r}\sum_{j=1}^{q}[g_j(x) + rv_j]^+ s^{j,x},$$

$$0 \in T(x^*) + \rho(x^*-y) + A^t[c + \frac{1}{s}(Ax^*-a)] + \frac{1}{r}\sum_{j=1}^{q}[g_j(x^*) + rd_j]^+ s^{j,x^*}.$$

Denotemos l y m como:

$$l := \langle -\rho(x-z) + \rho(x^* - y), x - x^* \rangle$$

$$m := \langle -A^t[u + \frac{1}{s}(Ax - a)] + A^t[c + \frac{1}{s}(Ax^* - a)], x - x^* \rangle$$

se cumple que

$$l = \langle \rho(z-x) - \rho(y - x^*), z - y \rangle - \frac{1}{\rho}\|\rho(z-x) - \rho(y-x^*)\|^2$$

$$m = \langle u - c, (a - Ax) - (a - Ax^*) \rangle - \frac{1}{s}\|(a - Ax) - (a - Ax^*)\|^2$$

Por la monotonia de T y multiplicando por r, se tiene que

$$rl + rm + \langle \sum_{j=1}^{q} [g_j(x^*) + rd_j]^+ s^{j,x} - \sum_{j=1}^{q} [g_j(x) + rv_j]^+ s^{j,x}, x - x^* \rangle \geq 0 \quad (3.1)$$

Se cumple que dados $b, e \geq 0$ y $f : \mathbb{R}^n \to \mathbb{R}$ tal que $s_1 \in \partial f(x)$, $s_2 \in \partial f(y)$

$$(b - e)(f(x) - f(y)) \geq \langle x - y, bs_2 - es_1 \rangle$$

Luego aplicando esta desigualdad a cada función g_j, considerando $b_j = [g_j(x^*) + rd_j]^+$ y $e_j = [g_j(x) + rv_j]^+$, y sumandolos juntamente con (3.1), se deduce que

$$rl + rm + \langle [g(x^*) + rd]^+ - [g(x) + rv]^+, g(x) - g(x^*) \rangle \geq 0 \quad (3.2)$$

Por otro lado, como para todo $y \in \mathbb{R}^q$ se cumple $y - y^+ \in N_{\mathbb{R}^q_+}(y^+)$. Denotando $\lambda_v := rv - [g(x) + rv]^+$ y $\lambda_d := rd - [g(x^*) + rd]^+$, se deduce por la monotonia de $N_{\mathbb{R}^q_+}$ y sumandolo con (3.2)

$$rl + rm + \langle [g(x^*) + rd]^+ - [g(x) + rv]^+, \lambda_d - \lambda_v \rangle \geq 0 \quad (3.3)$$

Al cumplirse que

$$\langle \lambda_v - \lambda_d, [g(x) + rv]^+ - [g(x^*) + rd]^+ \rangle = r\langle \lambda_v - \lambda_d, v - d \rangle - \|\lambda_v - \lambda_d\|^2$$

entonces por (3.3) y multiplicando por $\frac{1}{r}$, se tiene

$$l + m + \langle \lambda_v - \lambda_d, v - d \rangle - \frac{1}{r}\|\lambda_v - \lambda_d\|^2 \geq 0 \tag{3.4}$$

Luego reemplazando los valores de l y m, (3.3) se reescribe como

$$\langle D(z,u,v) - D(y,c,d), (z,u,v) - (y,c,d) \rangle \geq \|D(z,u,v) - D(y,c,d)\|^2_{\frac{1}{\rho},\frac{1}{s},\frac{1}{r}}$$

Entonces se cumple que D es monótono univaluado y por la desigualdad de cauchy schwarz y equivalencias de normas en $\mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^n$, se deduce que D es una función Lipschitz ■

Definamos los siguientes operadores D_1, D_2 y D_3, cuyos rangos estan en $\mathbb{R}^n$, $\mathbb{R}^p$ y $\mathbb{R}^q$ respectivamente, tales que:

$$D = (D_1, D_2, D_3).$$

Considerando el método de la gradiente en el problema (V_D), obtenemos el siguiente algoritmo: Dado $(u^0, v^0, z^0) \in \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^n$ arbitrario,

$$z^{k+1} = z^k - \frac{1}{\rho} D_1(z^k, u^k, v^k),$$

$$u^{k+1} = u^k - \frac{1}{s} D_2(z^k, u^k, v^k),$$

$$v^{k+1} = v^k - \frac{1}{r} D_3(z^k, u^k, v^k),$$

el cual se rescribe como:

Algoritmo (LACI)

1. **Inicialización:** Escoger $(u^0, v^0, z^0) \in \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^n$ arbitrario, luego hacer $k = 0$.

2. **Etapa k:** Buscar x^k satisfaciendo

$$0 \in T(x^k) + \rho(x^k - z^k) + A^t[u^k + \frac{1}{s}(Ax^k - a)] + \frac{1}{r}\sum_{j=1}^{p}[g_j(x^k) + rv_j^k]^+ \partial g_j(x^k)$$

- Si $z_k = x_k$, $Ax^k = a$ y $[g(x^k) + rv^k]^+ = rv^k$, PARAR.
- Caso contrario, hacer

$$u^{k+1} = u^k + \frac{1}{s}(Ax^k - a),$$

$$v^{k+1} = \frac{1}{r}[g(x^k) + rv^k]^+,$$

$$z^{k+1} = x_k,$$

  Hacer $k = k+1$, luego retornar a la Etapa k.

Si el algoritmo se detiene en la etapa k, se tiene que $D(z^k, u^k, v^k) = 0$, entonces x^k es una solución óptima del problema (V_P).

Theorem 3.1.1 *Bajo las hipótesis anteriores. Supongamos que existe solución de* (V_p). *La secuencia generada* $\{z^k, u^k, v^k\}$ *converge hacia* (z', u', v'), *donde* z' *es una solución de* (V_p).

Prueba. Denotemos por $w^k := (z^k, u^k, v^k)$ para todo $k \in \mathbb{N}$. Sea $\bar{z}$ una solución de (V_p), entonces existe $(\bar{u}, \bar{v})$ tal que denotando $\bar{w} = (\bar{z}, \bar{u}, \bar{v})$, se cumple que $D(\bar{w}) = 0$. Luego por la proposición (3.1) se obtiene:

$$\langle w^k - \bar{w}, D(w^k)\rangle \geq \frac{1}{\rho}\|D_1(w^k)\|^2 + \frac{1}{s}\|D_2(w^k)\|^2 + \frac{1}{r}\|D_3(w^k)\|^2$$

$$\langle w^{k+1} - \bar{w}, D(w^k)\rangle \geq 0 \tag{3.5}$$

Como se cumple

$$w^{k+1} - \bar{w} + (\frac{1}{\rho}D_1(w^k), \frac{1}{s}D_2(w^k), \frac{1}{r}D_3(w^k)) = w^k - \bar{w}$$

denotando $c = (\rho, s, r)$ y usando (3.5) se tiene

$$\|w^k - \bar{w}\|_c^2 \geq \|w^{k+1} - \bar{w}\|_c^2 + \|(\frac{1}{\rho}D_1(w^k), \frac{1}{s}D_2(w^k), \frac{1}{r}D_3(w^k))\|_c^2$$

Entonces $\{\|w^k - \bar{w}\|_c\}$ es decreciente y al estar acotado es converge, luego se tiene que $\{D_1(w^k)\}$, $\{D_2(w^k)\}$ y $\{D_3(w^k)\}$ tienden a 0.

Como $\{w^k\}$ es acotada, existe (z', u', v') valor de adherencia de $\{w^k\}$, por la continuidad de D se tiene que $D(z', u', v') = 0$. Luego tomando $(\bar{z}, \bar{u}, \bar{v}) = (z', u', v')$ en el principio, se obtiene que $\{\|(z^k, u^k, v^k) - (\bar{z}, \bar{u}, \bar{v})\|_c\}$ es decreciente, por lo tanto se deduce que $\{(z^k, u^k, v^k)\}$ converge a (z', u', v') una solución de (V_D). ■

3.2. Método Híbrido Teórico

En la práctica, la resolución de los subproblemas planteados en el algortimo anterior, tienen la misma dificultad de resolución del problema original. Por ello teniendo en cuenta los resultados obtenidos por Solodov en [6], en esta sección se desarrolla un algoritmo donde solo es necesario la resolución aproximada de los subproblemas, para ello usaremos el operador ϵ-enlargement definido como:

Definición 3.1 *Dado $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ un operador monotono y $\epsilon \geq 0$, el ϵ-enlargement de T es definido como $T^\epsilon : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ tal que*

$$T^\epsilon(x) = \{u \in \mathbb{R}^n : \langle v - u, y - x\rangle \leq -\epsilon, \forall y \in R^n,\ v \in T(y)\}.$$

Considerando una perturbación del operador D definido en la sección anterior, dado $\epsilon \geq 0$, definimos el siguiente operador

$$D_\epsilon(z, u, v) := \left\{(\rho(z - x), a - Ax, rv - [g(x) + rv]^+) : 0 \in \tilde{W}^\epsilon_{(z,u,v)}(x)\right\},$$

donde el operador $\tilde{W}^\epsilon_{(z,u,v)} : \mathbb{R}^n \to \mathbb{R}^n$ es defindo como:

$$\tilde{W}^\epsilon_{(z,u,v)}(x) := T^\epsilon(x) + \rho(x - z) + A^t[u + \frac{1}{s}(Ax - a)] + \frac{1}{r}\sum_{j=1}^{p}[g_j(x) + rv_j]^+ \partial g_j(x).$$

Observando la prueba de la proposición 3.1, se deduce con facilidad la siguiente proposición.

Proposición 3.2 *Dados* $(z,u,v),(y,c,d)$ *en* $\mathbb{R}^n\times\mathbb{R}^p\times\mathbb{R}^q$ *y* $g\in D_\epsilon(z,u,v)$ *se cumple que:*

$$\langle g-D(y,c,d),(z,u,v)-(y,c,d)\rangle\geq\|g-D(y,c,d)\|^2_{(\frac{1}{\rho},\frac{1}{s},\frac{1}{r})}-\epsilon.$$

En particular se tiene que $D_\epsilon\subseteq D^\epsilon$.

Considerando una perturbación en el método de la gradiente en el problema (V_D), obtenemos el siguiente algoritmo: Dados $(u^0,v^0,z^0)\in\mathbb{R}^p\times\mathbb{R}^q\times\mathbb{R}^n$ y $\sigma\in[0,1)$ arbitrarios,

$$z^{k+1}=z^k-\frac{1}{\rho}g_1^k,$$
$$u^{k+1}=u^k-\frac{1}{s}g_2^k,$$
$$v^{k+1}=v^k-\frac{1}{r}g_3^k,$$

donde $(g_1^k,g_2^k,g_3^k)\in D_{\epsilon_k}(z^k+r^k,u^k,v^k)$ con $g_1^k\in\mathbb{R}^n$, $g_2^k\in\mathbb{R}^p$ y $g_3^k\in\mathbb{R}^q$. En el cual r^k y $\epsilon_k\geq 0$ son escogidos tales que, denotando $d:=(\frac{1}{\rho},\frac{1}{s},\frac{1}{r})$, satisfacen:

$$\|\rho r^k\|^2+2\rho\epsilon_k\leq\sigma\rho\|(g_1^k-\rho r^k,g_2^k,g_3^k)\|_d^2 \tag{3.6}$$

el cual es equivalente a lo siguiente:

$$2\langle r^k,g_1^k\rangle+2\epsilon_k\leq(\sigma-1)\|(g_1^k-\rho r^k,g_2^k,g_3^k)\|_d^2+\|(g_1^k,g_2^k,g_3^k)\|_d^2 \tag{3.7}$$

este algoritmo se rescribe como:

Algoritmo (ILACI)

1. **Inicialización:** Escoger $(u^0,v^0,z^0)\in\mathbb{R}^p\times\mathbb{R}^q\times\mathbb{R}^n$, $\sigma\in[0,1)$ y $c:=(\rho,s,r)>0$ arbitrarios, luego hacer $k=0$.

2. **Etapa k:** Buscar x^k satisfaciendo

   $$\rho r^k\in T^{\epsilon_k}(x^k)+\rho(x^k-z^k)+A^t[u^k+\frac{1}{s}(Ax^k-a)]+\frac{1}{r}\sum_{j=1}^p[g_j(x^k)+rv_j^k]^+\partial g_j(x^k)$$

   donde $r^k\in\mathbb{R}^n$ y $\epsilon_k\geq 0$ cumplen que

   $$\|\rho r^k\|^2+2\rho\epsilon_k\leq\sigma\rho[\|(x^k-z^k,u^{k+1}-u^k,v^{k+1}-v^k)\|_c^2]$$

- Si $z_k + r^k = x_k$, $Ax^k = a$ y $[g(x^k) + rv^k]^+ = rv^k$, PARAR.
- Caso contrario, hacer

$$u^{k+1} = u^k + \frac{1}{s}(Ax^k - a),$$

$$v^{k+1} = \frac{1}{r}[g(x^k) + rv^k]^+,$$
$$z^{k+1} = x_k - r^k,$$

Hacer $k = k + 1$, luego retornar a la Etapa k.

Si el algoritmo se detiene en la etapa k, se tiene que $D(z^k + r^k, u^k, v^k) = 0$, entonces x^k es una solución óptima del problema (V_P).

Theorem 3.2.1 *Bajo las hipótesis anteriores. Supongamos que existe solución de (V_p). La secuencia generada $\{z^k, u^k, v^k\}$ converge hacia (z', u', v'), donde z' es una solución de (V_p).*

Prueba. Para todo $k \in \mathbb{N}$, denotemos $w^k := (z^k, u^k, v^k)$, $q^k := (r^k, 0, 0)$ y $(g_1^k, g_2^k, g_3^k) \in D_{\epsilon_k}(z^k + r^k, u^k, v^k)$ con $g_1^k \in \mathbb{R}^n$, $g_2^k \in \mathbb{R}^p$ y $g_3^k \in \mathbb{R}^q$.
Sea $\bar{z}$ una solución de (V_p), entonces existe $(\bar{u}, \bar{v})$ tal que denotando $\bar{w} = (\bar{z}, \bar{u}, \bar{v})$, se cumple que $D(\bar{w}) = 0$. Luego por la proposición (3.2) se obtiene:

$$\langle w^k + q^k - \bar{w}, (g_1^k, g_2^k, g_3^k)\rangle \geq \frac{1}{\rho}\|g_1^k\|^2 + \frac{1}{s}\|g_2^k\|^2 + \frac{1}{r}\|g_3^k\|^2 - \epsilon_k$$

$$\langle w^{k+1} - \bar{w}, (g_1^k, g_2^k, g_3^k)\rangle \geq -\langle r^k, g_1^k\rangle - \epsilon_k \tag{3.8}$$

Como se cumple que

$$w^{k+1} - \bar{w} + (\frac{1}{\rho}g_1^k, \frac{1}{s}g_2^k, \frac{1}{r}g_3^k) = w^k - \bar{w}$$

denotando $c = (\rho, s, r)$ y usando (3.8) y (3.7) se tiene que

$$\|w^k - \bar{w}\|_c^2 \geq \|w^{k+1} - \bar{w}\|_c^2 + (1 - \sigma)\|(g_1^k - \rho r^k, g_2^k, g_3^k)\|_d^2$$

Entonces se deduce que $\{\|w^k - \bar{w}\|_c\}$ es converge y $\{\|(g_1^k - \rho r^k, g_2^k, g_3^k)\|_d^2\}$ converge a 0, luego de (3.6) se tiene que $\{r^k\}$ y $\{\epsilon^k\}$ tienden a 0, por lo tanto

$\{(g_1^k, g_2^k, g_3^k)\}$ tiende a 0.

Como $\{w^k\}$ es acotada, existe (z', u', v') valor de adherencia de $\{w^k\}$ y al cumplirse que $(g_1^k, g_2^k, g_3^k) \in D^{\epsilon_k}(w^k + (r^k, 0, 0))$, por la maximalidad del operador univaluado D, se deduce que $D(z', u', v') = 0$. Luego tomando $(\bar{z}, \bar{u}, \bar{v}) = (z', u', v')$ en el principio, se obtiene que $\{\|(z^k, u^k, v^k) - (\bar{z}, \bar{u}, \bar{v})\|_c\}$ es decreciente, por lo tanto se deduce que $\{(z^k, u^k, v^k)\}$ converge a (z', u', v') una solución de (V_D). ∎

En particular si consideramos $\rho = s = r$, $\epsilon_k = 0$ y solo las restricciones de desigualdad convexa, obtenemos que el algoritmo anterior es exactamente el algoritmo propuesto en [2], en el cual se restringe que T sea univaluado.

Capítulo 4

Método de Haces para Operadores Fuertemente Monótonos

En este capítulo desarrollamos un algoritmo que permite aproximar la solución el problema variacional asociado a un operador $T : \mathbb{R}^n \rightrightarrows \mathbb{R}$ maximal fuertemente monótono de coeficiente ρ, cuyo dominio es $\mathbb{R}^n$.

En los trabajos [7] y [4] se desorrollan algoritmos ejecutables para aproximar la solución el problema variacional asociado a un operador T monótono maximal, siendo la idea principal de estos algoritmos que dado un punto, buscan un hiperplano que cumpla ciertas restriciones y separe el conjunto solución del problema variacional de tal punto. En este trabajo usaremos ideas análogas pero aprovechando la fuerte monotonia.

Dado $v \in T(y)$, la monoticidad de T implica que $T^{-1}(0)$ este contenida en el siguiente hiperplano:

$$H_{v,y} := \{x \in \mathbb{R}^n : \langle y - x, v\rangle \geq 0\}$$

mas aún si T es fuertemente monótono de coeficiente ρ, se cumple que $T^{-1}(0)$

pertenece a la siguiente bola:

$$B_{v,y} := \{x \in \mathbb{R}^n : \|x - y + \frac{v}{2\rho}\|^2 \leq \|\frac{v}{2\rho}\|^2\}.$$

Por lo tanto usamos estos tipos de conjuntos para lograr una separación de un punto y $T^{-1}(0)$. Se obtiene que en el caso fuertemente monótono, dado un punto x', la busqueda de un hiperplano $H_{v,y}$ (donde $v \in T(y)$), que lo separe de $T^{-1}(0)$, debe satisfacer que dado $u \in T(x)$, se cumple que y pertenese al interior de $B_{u,x}$. Por otro lado, en el caso fuertemente monótono, se obtiene la siguiente proposición análogo a lo obtenido en [7].

Proposición 4.1 *Sea $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ un operador maximal fuertemente monotono con coeficiente $\rho > 0$. Dado $\tilde{x} \in \mathbb{R}^n$, $r > 0$ y el conjunto de duplas $\{z^i, w^i \in T(z^i)\}_{i\in 1,\cdots,m}$, tal que $\|z^i - \tilde{x}\| \leq r$ para todo $i \in 1, \cdots, m$. Sea*

$$(\hat{x}, \hat{s}) := (\sum_{i=1}^{m} \lambda_i z^i, \sum_{i=1}^{m} \lambda_i w^i)$$

con $\lambda_i \geq 0$ y $\sum \lambda_i = 1$. Entonces $\|\hat{x} - \tilde{x}\| \leq r$ y

$$\hat{s} \in [T - \rho(I - \tilde{x})]^{\epsilon}(\hat{x}) + \rho(\hat{x} - \tilde{x})$$

para algún $\epsilon \leq 2rM$, siendo $M := \max\{\|w^i\| : i \in 1, \cdots, m\}$.

Prueba. Al ser T fuertemente monotono con coeficiente $\rho > 0$ se tiene que $T - \rho(I - \tilde{x})$ es monotono maximal. Luego considerando el conjunto de duplas $\{z^i, w^i - \rho(z^i - \tilde{x}) \in [T - \rho(I - \tilde{x})](z^i)\}_{i\in 1,\ldots,m}$ en el corolario 2,3 de [7] se obtiene que $\hat{s} \in [T - \rho(I - \tilde{x})]^{\epsilon}(\hat{x}) + \rho(\hat{x} - \tilde{x})$ con $\|\hat{x} - \tilde{x}\| \leq r$ y $\epsilon = \sum_{i=1}^{m} \lambda_i \langle z^i - \tilde{x}, w^i - \rho(z^i - \tilde{x}) - \hat{s}\rangle \leq 2rM$. ■

4.1. Algoritmo

El algoritmo propuesto a continuación realiza lo siguiente: dado x' se busca un hiperplano $H_{v,y}$ (donde $v \in T(y)$) que lo separe de $T^{-1}(0)$, si x' es $T^{-1}(0)$ el algoritmo realizara una subrutina infinita, de lo contrario se encuentra dicho hiperplano $H_{v',y'}$ (donde $v' \in T(y')$) en un tiempo finito por medio de una subrutina y proyectamos x' sobre $B_{v',y'}$ que esta incluido en $H_{v',y'}$. Luego realizamos el mismo proceso anterior con ese nuevo punto hallado.

Algoritmo (PrFm)

Inicialización: Escoger $x^0 \in \mathbb{R}^n$ arbitrario, $\sigma \in (0,1)$, $\{\alpha_j\} \in]0,1[$ y $\{r_j\} \in]0,1[$ sucesiones decrecientes que convergen a 0, luego hacer $k = 0$.

Etapa k:

paso 0:

0.1 Tomar $u^k \in T(x_k)$. Si $u^k = 0$ entonces PARAR

0.2 De lo contrario definir $s^{k,-1,0} = w^0 = u^k$, $m = 0$ y $j = 0$

paso 1:

1.1 Calcular $d^{k,m,j} = \text{proj}_{\,\text{co}\,\{\text{s}^{\text{k,m}-1,\text{j}},\text{w}^{\text{m}},\text{u}^{\text{k}}\}}(0)$.

1.2 Si $\|d^{k,m,j}\| \leq \alpha_j \|u^k\|$ entonces $s^{k,m,j} = u^k$, $t_j = m$ y $j = j+1$.

1.3 De lo contrario definir $s^{k,m,j} = d^{k,m,j}$.

paso 2:

2.1 Calcular $\beta_j = \text{mín}\{\dfrac{r_j\|u^k\|}{\|s^{k,m,j}\|}, \dfrac{\sigma\langle s^{k,m,j}, u^k\rangle}{\|s^{k,m,j}\|^2}\}$.

2.2 Definir $y^{k,m,j} = x^k - \dfrac{\beta_j}{\rho} s^{k,m,j}$ y tomar $v^{k,m,j} \in T(y^{k,m,j})$.

paso 3:

3.1 Si $\langle v^{k,m,j}, s^{k,m,j}\rangle \leq 0$ entonces
hacer $w^{m+1} = v^{k,m,j}$, $m = m+1$ e ir al paso 1.

3.2 De lo contrario
definir $y^k = y^{k,m,j}$, $v^k = v^{k,m,j}$, $l_k = j$,

$$x^{k+1} = x^k - (1 - \frac{\|v^k\|}{\|2\rho(x^k - y^k) + v^k\|})(x^k - y^k + \frac{v^k}{2\rho}),$$

hacer $k = k+1$ e ir a la etapa k.

4.2. Convergencia

La siguiente proposición es un resultado análogo a lo obtenido en [4].

Proposición 4.2 *Dados p^m y q^m sucesiones en $\mathbb{R}^n$ y $v \in \mathbb{R}^n$ tales que para todo $m \in \mathbb{N} \cup \{0\}$ se cumple que:*

$$p^0 = q^0\,, p^{m+1} = \mathrm{proj}_{\,\mathrm{co}\,\{\mathrm{p^m},\mathrm{q^{m+1}},\mathrm{v}\}}(0) \;\; y \;\; \langle \mathrm{q^{m+1}}, \mathrm{p^m} \rangle \leq 0,$$

entonces para todo $m \in \mathbb{N} \cup \{0\}$ se cumple $\|p^m\| \leq \dfrac{\max\limits_{i \in 0,..,m} \{\|q^i\|\}}{\sqrt{m+1}}$

Prueba. Mostraremos la inecuación por inducción. Para $m = 0$ la inecuación se cumple. Supongamos que la inecuación se cumple para m, mostraremos que tambien se cumple para $m+1$.

$$\|p^{m+1}\|^2 \leq \min_{r \in [0,1]} \|rp^m + (1-r)q^{m+1}\|^2$$

$$\leq \min_{r \in [0,1]} [r^2\|p^m\|^2 + (1-r)^2\|q^{m+1}\|^2]$$

$$= (\frac{\|q^{m+1}\|^2}{\|p^m\|^2 + \|q^{m+1}\|^2})^2\|p^m\|^2 + (\frac{\|p^m\|^2}{\|p^m\|^2 + \|q^{m+1}\|^2})^2\|q^{m+1}\|^2$$

$$= \frac{\|q^{m+1}\|^2\|p^m\|^2}{\|p^m\|^2 + \|q^{m+1}\|^2} \leq (\frac{1}{\|q^{m+1}\|^2} + \frac{1}{\|p^m\|^2})^{-1} \leq \frac{\max\limits_{i \in 0,..,m+1} \{\|q^i\|\}^2}{m+2}$$

■

La proposición anterior nos indica que si $\|d^{k,m,j}\| > \alpha_j \|u^k\|$ en el paso 1.2, entonces despues de un incremento de m maximo de c tal que

$$\frac{c+1}{\max\limits_{i \in 0,..,c} \{\|q^i\|^2\}} \geq \frac{1}{\alpha_j},$$

se logra que se ingrese al paso 1.2 y se aumente j.

La siguiente proposición nos muestra que la subrutina de 3.1 a 1 resulta ser finito si el punto tomado en la etapa k no es $T^{-1}(0)$.

Proposición 4.3 *Sea x^k un iterado del algoritmo. Entonces $x^k \neq T^{-1}(0)$ si y solo si en la etapa k, el ciclo de 3.1 a 1 resulta ser finito.*

Prueba. Si el el ciclo de 3.1 a 1 resulta ser finito entonces se cumple que $\langle v^k, x^k - y^k\rangle > 0$ entonces por la monotonia de T se cumple que $x^k \neq T^{-1}(0)$.

Si $x^k \neq T^{-1}(0)$, dado k fijo, si suponemos por el contraio que el ciclo de 3.1 a 1 es infinito, por la proposición anterior se deduce que j incrementa hacia el infinito creando la sucesion $\{d^{k,n_j,j}\}$ cumpliendo que:

$$\|d^{k,n_j,j}\| \leq \alpha_j,\ d^{k,n_j,j} \in \mathrm{co}\,\{\cup_{\|\mathrm{y}-\mathrm{x}^\mathrm{k}\| \leq \frac{\mathrm{r_j}\|\mathrm{u}^\mathrm{k}\|}{\rho}} \mathrm{T(y)}\}$$

por lo tanto como $\{r_j\}$ y $\{\alpha_j\}$ tienden a 0 y por la maximalidad de T, se cumple que $0 \in T(x^k)$ lo cual es una contradicción. ■

La siguiente proposición muestra como el algoritmo se acerca a la solución durante el ciclo de 3.1 a 1.

Proposición 4.4 *Dado la etapa k en el algoritmo. Si en el ciclo 3.1 a 1, se tiene que $j \geq 1$, entonces se cumple que:*

$$\|x^k - T^{-1}(0)\| \leq \frac{2r_j\|u^k\| + \alpha_j\|u^k\| + \sqrt{\alpha_j^2\|u^k\|^2 + 8r_j\|u^k\|M}}{2\rho}$$

donde $M := \max\{T(y) : \|y - x^k\| \leq \frac{r_j\|u^k\|}{\rho}\}$

Prueba. Si $j \geq 1$, al entrar al paso 1.2, se tiene que:

$$\|d^{k,n_j,j}\| \leq \alpha_j\|u^k\|,\ d^{k,n_j,j} \in \mathrm{co}\,\{\cup_{\|\mathrm{y}-\mathrm{x}^\mathrm{k}\| \leq \frac{\mathrm{r_j}\|\mathrm{u}^\mathrm{k}\|}{\rho}} \mathrm{T(y)}\}$$

denotando $s_j := d^{k,n_j,j}$ y $c_j := \frac{r_j\|u^k\|}{\rho}$ por la proposición 4.1 se tiene:

$$s_j \in [T - \rho(I - x^k)]^{\epsilon_j}(\hat{x}_j) + \rho(\hat{x}_j - x^k) \text{ con } \|\hat{x}_j - x^k\| \leq c_j \text{ y } \epsilon_j \leq 2c_jM,$$

luego por la definición del operador ϵ-enlargement, se tiene:

$$\langle s_j - \rho(\hat{x}_j - x^k) + \rho(x^* - x^k), \hat{x}_j - x^* \rangle \geq -\epsilon_j$$

$$\langle s_j, \hat{x}_j - x^* \rangle \geq \rho \|\hat{x}_j - x^*\|^2 - \epsilon_j$$

por la desigual de cauchy y como $\|s_j\| \leq \alpha_j \|u^k\|$ se obtiene

$$0 \geq \rho \|\hat{x}_j - x^*\|^2 - \alpha_j \|u^k\| \|\hat{x}_j - x^*\| - \epsilon_j$$

lo que implica que

$$\|\hat{x}_j - x^*\| \leq (\alpha_j \|u^k\| + \sqrt{\alpha_j^2 \|u^k\|^2 + 4\rho\epsilon_j})/(2\rho)$$

usando la desigualdad que satisface ϵ_j se tiene

$$\|\hat{x}_j - x^*\| \leq \frac{\alpha_j \|u^k\| + \sqrt{\alpha_j^2 \|u^k\|^2 + 8\rho c_j M}}{2\rho}$$

finalmente reemplazando c_j y por la desigualdad triangular

$$\|x^k - x^*\| \leq \frac{2r_j \|u^k\| + \alpha_j \|u^k\| + \sqrt{\alpha_j^2 \|u^k\|^2 + 8r_j \|u^k\| M}}{2\rho}$$

■

Al pasar de una etapa a otra se logra una mejor aproximación a $T^{-1}(0)$ en el siguiente sentido

Proposición 4.5 *Se cumple las siguientes desigualdades:*

$$\|x^k - T^{-1}(0)\|^2 \geq \|x^{k+1} - T^{-1}(0)\|^2 + \|x^{k+1} - x^k\|^2.$$

$$\|x^k - x^{k+1}\|^2 \geq \|\mathrm{proj}_{\mathrm{H}_{\mathrm{v^k,y^k}}}(\mathrm{x^k}) - \mathrm{x^{k+1}}\|^2 + \|\mathrm{proj}_{\mathrm{H}_{\mathrm{v^k,y^k}}}(\mathrm{x^k}) - \mathrm{x^k}\|^2.$$

Prueba. La primera desigualdad se logra desde que $x^{k+1} = \mathrm{proj}_{\mathrm{B}_{\mathrm{v^k,y^k}}}$. La segunda desigualdad se deduce del hecho que B_{v^k,y^k} esta incluido en H_{v^k,y^k}.■

Con los proposiciones anteriores, se obtiene el siguiente teorema de convergencia para el algoritmo propuesto.

Theorem 4.2.1 *Dado T un operador maximal fuertemente monótono de coeficiente ρ, cuyo dominio es $\mathbb{R}^n$. Sea (x_k) la secuencia generada por el algoritmo, entonces la secuencia es finita con el ultimo elemento igual a $T^{-1}(0)$ o de lo contrario es una sucesión que converge a $T^{-1}(0)$.*

Prueba. Si la secuencia es finita, por la proposición 4.3, se tiene que el ultimo elemento es igual a $T^{-1}(0)$. De lo contrario el algoritmo genera una sucesión, que por la proposicion anterior, solo falta mostrar que $T^{-1}(0)$ es un punto de acumulación de $\{x^k\}$, para mostrar que converge a $T^{-1}(0)$.

Como $\langle v^k, x^k - y^k\rangle > 0$ se tiene que:

$$\|x^k - y^k + \frac{v^k}{2\rho}\|^2 > \|x^k - y^k\|^2 + \|\frac{v^k}{2\rho}\|^2$$

$$\|x^k - y^k + \frac{v^k}{2\rho}\| - \|\frac{v^k}{2\rho}\| > \frac{\|x^k - y^k\|^2}{\|x^k - y^k + \frac{v^k}{2\rho}\| + \|\frac{v^k}{2\rho}\|} \tag{4.1}$$

Al ser x^k acotado entonces u^k es acotado, luego como $\|y^k - x^k + \frac{u^k}{2\rho}\|^2 < \|\frac{u^k}{2\rho}\|^2$ se deduce que y^k es acotado, entonces v^k es acotado. Por lo tanto existe $M > 0$ tal que $[\|x^k - y^k + \frac{v^k}{2\rho}\| + \|\frac{v^k}{2\rho}\|]^{-1} \geq M$. Además al cumplirse que $\langle s^{k,m,l_k}, u^k\rangle \geq \|s^{k,m,l_k}\|^2$ y $\|s^{k,m,l_k}\| > \alpha_{l_k}\|u^k\|$, se deduce por (4.1):

$$\|x^{k+1} - x^k\| = \|x^k - y^k + \frac{v^k}{2\rho}\| - \|\frac{v^k}{2\rho}\| > \frac{M}{\rho^2}\|u^k\|^2 \min\{r_{l_k}\alpha_{l_k}, \sigma\alpha_{l_k}^2\}.$$

Si existe una subsucesión de $\{\|u^k\|\}$ que converge a 0 entonces $T^{-1}(0)$ es un punto de acumulación de $\{x^k\}$. De lo contrario $\{u_k\}$ es acotado inferiormente por un número positivo, luego como $\{\|x^{k+1} - x^k\|\}$ converge a 0 entonces l_k tiende a infinito cuando k tiende a infinito. Por otro lado se tiene que

$$\|d^{k,t_{l_k},l_k}\| \leq \alpha_{l_k}\|u^k\|,\ d^{k,t_{l_k},l_k} \in \mathrm{co}\{\cup_{\|\mathrm{y}-\mathrm{x}^\mathrm{k}\| \leq \frac{\mathrm{r}_{\mathrm{l}_\mathrm{k}}\|\mathrm{u}^\mathrm{k}\|}{\rho}} \mathrm{T(y)}\}$$

Entonces como $\{r_j\}$ y $\{\alpha_j\}$ converge a 0, se deduce que dado un punto de acumulación de $\{x^k\}$ esta es $T^{-1}(0)$. ∎

Conclusiones

Análogamente a los problemas de optimización, en el Capítulo 2, se obtuvo los problemas variacionales equivalentes (denominados dual o minimax), respecto de los problema de desigualdad variacional iniciales considerados en dicho Capítulo. Además en los problemas variacionales duales equivalentes, se obtiene que bajo ciertas hipótesis su operador asociado es una función Lipschitz, lo cual nos conduce a definir nuevos algoritmos para resolver el problema variacional primal, los cuales resultan ser generalizaciones del algoritmo de Lagrangiana Aumentada para el caso de problemas de optimización.

Análogo al algoritmo proximal de multiplicadores para problemas de optimización, en el Capítulo 3, se obtuvo dos algoritmos para problemas variacionales con restricciones de igualdad lineal y desigualdad convexa. El primer algoritmo nos conduce a resolver subproblemas variacionales irrestrictos de operadores fuertemente monotonos. El segundo algoritmo mejora el anterior al considerar sólo soluciones aproximadas de cada subproblema.

En el último Capítulo se desarrollo un algoritmo para encontrar una solución aproximada del problema variacional irrestricto asociado a un operador fuertemente monótono.

Perspectivas

Se espera en un trabajo futuro combinar el último método obtenido para operadores fuertemente monótonos, con el método Proximal-Lagrangiano Aumentado para asi lograr un algoritmo implementable que aproxime la solución (si existe) del problema de desigualdad variacional monótono con restricciones clásicas.

Bibliografía

[1] A. Auslender, M. Teboulle *Asymptotic Cones and Functions in Optimization and Variational Inequalities*, Springer, 2003.

[2] A. N. Iusen, Mostafa Nasri *Augmented Lagrangian methods for variational inequality problems*, RAIRO-Operations Research 44.01 (2010)

[3] D. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, NY, 1982.

[4] I. Konnov *Combined Relaxation Methods for Variational Inequalities*, Vol. 495. Springer, 2001.

[5] J.-P. Crouzeix, E. Ocaña, *Métodos Algorítmicos Haces-Proximal-Lagrangiano Aumentado para Problemas de Optimización Matemática*, Monografías del IMCA N^o 55, Lima, Perú.

[6] M. V. Solodov, B. F. Svaiter *A unified framework for some inexact proximal point algorithms*, IMPA, 2001.

[7] R. S. Burachik, C. Sagastizábal, B. F. Swaiter, *Bundle methods for maximal monotone operators*, Springer Berlin Heidelberg, 1999.

[8] R. T. Rockafellar, *Augmented Lagrangians and application of the proximal point algorithm in convex programming*, Math. Oper. Res., 1 (1976), pp. 97-116.

[9] T. Pennanen, *Dualization of generalized equation of maximal monotone type*, SIAM J. Optimization, 10, 2000, 809-835.