

510
B27

UNIVERSIDAD NACIONAL DEL CALLAO

FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA

DEPARTAMENTO ACADÉMICO DE MATEMÁTICA



“MÉTODO DEL PUNTO PROXIMAL PARA DESIGUALDADES
VARIACIONALES EN VARIEDADES RIEMANNIANAS”

Tesis para optar el Título Profesional de Licenciado en Matemática

Nancy Baygorrea Cusihuallpa

CALLAO - PERÚ

Diciembre del 2010

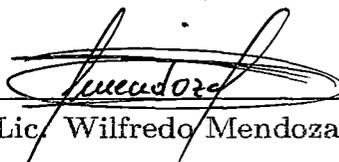
HOJA DE PRESENTACIÓN

Método del Punto Proximal para Desigualdades Variacionales
en Variedades Riemannianas

Nancy Baygorrea Cusihualpa

Tesis presentada a consideración del Cuerpo Docente de la Facultad de
Ciencias Naturales y Matemática de la Universidad Nacional del Callao,
como parte de los requisitos para obtener el Título Profesional de
Licenciado en Matemática.

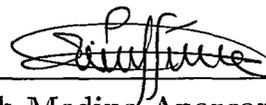
Aprobada por:



Lic. Wilfredo Mendoza Quispe, Lic
(Presidente)



Lic. Ezequiel Fajardo Campos, Lic



Lic. Ruth Medina Aparcana, Lic



Dr. Erik Alex Papa Quiroz, DSc.

Diciembre-2010

Bellavista-Callao, Perú

FICHA CATALOGRÁFICA

BAYGORREA CUSIHUALLPA, NANCY

Método del punto proximal para desigualdades variacionales en variedades riemannianas, Callao [2010].

vii, 100 p. 29,7 cm (UNAC, Licenciado en Matemática, 2010)

Tesis.

Universidad Nacional del Callao.

Facultad de Ciencias Naturales y Matemática.

Matemática

1. UNAC / FCNM II. Título (Serie)

AGRADECIMIENTOS

- A mi familia, Felix, Claudia e Irma, por el apoyo constante, cariño, comprensión, ejemplo y paciencia, dignos de ser admirados, cada uno de ellos, en la vida.
- Al que me orientó en el desarrollo de este trabajo, mi asesor Erik Papa, quién considera que lo más importante en la vida es ser bueno sólo en la parte académica.
- A todos los profesores que hacen parte de mi jurado de Tesis, por la disponibilidad y paciencia que mostraron en la ardua tarea de la revisión del trabajo.
- A todos los que hacen parte de la Escuela Profesional de Matemática de la Universidad Nacional del Callao, mi alma mater.
- Finalmente, a todos los que directa o indirectamente, contribuyeron para la realización de este trabajo.

RESUMEN

MÉTODO DEL PUNTO PROXIMAL PARA DESIGUALDADES VARIACIONALES EN VARIEDADES RIEMANNIANAS

NANCY BAYGORREA CUSIHUALLPA

Diciembre - 2010

Asesor: Erik Alex Papa Quiroz

Título Obtenido: Licenciado en Matemática.

El problema de desigualdad variacional (PDV) fué introducido por los matemáticos Philip Hartman y Guido Stampacchia en 1966. Existe una gran variedad de ejemplos de PDV aplicados a muchos campos de las Ciencias e ingenierías tales como la mecánica, transporte, economía, teoría de juegos, así como también en problemas de optimización.

El problema de desigualdad variacional en una variedad de Hadamard consiste en encontrar un vector $x^* \in K$ tal que

$$\langle F(x^*), \exp_{x^*}^{-1}x \rangle \geq 0, \quad \forall x \in K,$$

donde $F : M \rightarrow TM$ es un campo vectorial y $K \subset M$ es un conjunto no vacío, convexo y cerrado. En esta tesis presentamos una extensión del método de punto proximal con distancias de Bregman para resolver el modelo anterior. Bajo algunas hipótesis naturales sobre F , obtenemos la convergencia de la sucesión generada por el MPP para una solución del problema planteado.

Palabras Claves: Problemas de desigualdad variacional, distancias de Bregman, Método de Punto Proximal, variedades riemannianas y operadores monótonos maximales.

ABSTRACT

PROXIMAL POINT METHODS TO SOLVE VARIATIONAL INEQUALITY PROBLEMS IN RIEMANNIAN MANIFOLDS

NANCY BAYGORREA CUSIHUALLPA

December - 2010

Adviser: Erik Alex Papa Quiroz

Obtained Degree: Mathematician.

The variational inequality problem (VIP) was introduced by Philip Hartman and Guido Stampacchia in 1966. The finite-dimensional VIP provides an unified framework in the study of most models in sciences and engineering, particular cases of these problems are optimization, game theory, traffic assignment, economical equilibrium and optimization problems.

The variational inequality problem in Hadamard manifold consist of finding a vector $x^* \in K$ such that

$$\langle F(x^*), \exp_{x^*}^{-1}x \rangle \geq 0, \quad \forall x \in K,$$

where $F : M \rightarrow TM$ is a monotone vector field and $K \subset M$ to be a nonempty closed and convex set. In this thesis we present an extension of the proximal point methods with Bregman distances to solve the previous model. Some natural assumptions are supposed on F to obtain the convergence result of the sequence generated by the PPM to obtain a solution of previous model.

Keywords: Variational inequality problems, Bregman distances, Proximal Point Methods, riemannian manifolds and maximal monotone operators.

Índice General

Introducción	1
1 Preliminares	6
1.1 Algunos resultados de análisis	6
1.2 Elementos de geometría riemanniana	19
1.3 Elementos de análisis convexo	32
1.4 Funciones y Distancias de Bregman en variedades Riemannianas	44
2 El problema de desigualdad variacional (PDV)	51
2.1 Una formulación para el (PDV)	53
2.2 Resultados de existencia general para (PDV)	55
2.3 Resultados de existencia usando monotonicidad	58
2.4 Método del Punto Proximal (MPP)	61
3 PDV en variedades riemanniannas (PDVV)	64
3.1 Motivaciones	65
3.1.1 Problema de singularidad	66
3.1.2 Problemas de optimización	67
3.1.3 Problema del Equilibrio de Nash	69
3.2 Resultados de existencia de (PDVV)	72
4 Método del punto proximal para resolver PDVV en variedades riemanniannas	77
4.1 MPP en variedades riemannianas para problemas de optimización	78

4.2	MPP para resolver PDVV	80
5	Aplicaciones del MPP para resolver PDVV	86
5.1	Problema	87
5.2	Método	87
5.3	Implementación	88
	Referencias Bibliográficas	97

Introducción

Un problema sencillo del cálculo en \mathbb{R} se da como sigue: sea f una función derivable en los reales, encontrar $x_0 \in I = [a, b]$ tal que:

$$f(x_0) = \min_{x \in I} f(x).$$

Observemos que x_0 satisface una de las siguientes tres posibilidades:

- i. Si $a < x_0 < b$, entonces $f'(x_0) = 0$.
- ii. Si $x_0 = a$, entonces $f'(x_0) \geq 0$.
- iii. Si $x_0 = b$, entonces $f'(x_0) \leq 0$.

Los ítems dados anteriormente pueden ser resumidos de la siguiente forma:

$$f'(x_0)(x - x_0) \geq 0, \quad \forall x \in I.$$

Tal desigualdad puede ser llamada como “problema de desigualdad variacional en \mathbb{R} ”.

Así, podemos generalizar el problema de desigualdad variacional (PDV) dado como sigue: Dado un conjunto $X \subset \mathbb{R}^n$ y una aplicación punto-conjunto $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, nuestro objetivo es encontrar $x^* \in X$ y $p \in T(x^*)$ tal que

$$\langle p, x - x^* \rangle \geq 0, \quad \forall x \in X \tag{1.1}$$

El PDV puede ser considerados como una generalización de problemas de optimización convexa. En efecto, consideremos el problema:

$$(p) \quad \begin{cases} \min & f(x) \\ s.a & \\ & x \in X \end{cases}$$

donde f es una función convexa y diferenciable en \mathbb{R}^n y $X \subset \mathbb{R}^n$ es un conjunto convexo. Estableciendo $T(x) = \nabla f(x)$, entonces (1.1) es la condición de optimalidad necesaria y suficiente para resolver el problema (p).

El problema de desigualdad variacional, PDV, fué introducido por los matemáticos Philip Hartman y Guido Stampacchia [23] en 1966.

Los primeros estudios de PDV fueron establecidos en el contexto del cálculo de variaciones asociados con la minimización de funcionales de dimensión infinita, como una herramienta analítica para estudios de problemas de frontera libre definidos con operadores diferenciales parciales no lineales y como resultado de problemas unilaterales en la teoría de elasticidad y la mecánica. Una introducción a PDV pueden ser encontrado en Kinderlehrer y Stampacchia [31].

El estudio de PDV en dimensión finita ocurrió en 1980 cuando Stella Dafermos [12] descubrió que el problema de equilibrio de tráfico, formulado por Michael J. Smith [45] en 1979, tenía una estructura de desigualdad variacional.

Existe una gran variedad de ejemplos de PDV aplicados a muchos campos de las ciencias e ingenierías tales como la mecánica, transporte, economía, teoría de juegos, etc, que pueden ser encontrados en [12] [28], [31] y [45] y en las notas de lectura de Marcotte [33].

Además, la conexión de PDV con la optimización puede ser visto en Harker y Pang[22].

Uno de los métodos más populares para resolver problemas de optimización irrestricta, convexa y no diferenciable, que es un caso particular de PDV, es el Método del Punto Proximal (MPP). El MPP clásico genera una sucesión de puntos $\{x^k\}$ dado por

$x^0 \in \mathbb{R}^n$ y

$$x^{k+1} \in \arg \min \left\{ f(x) + \frac{\lambda_k}{2} d(x, x^k)^2 : x \in \mathbb{R}^n \right\}, \quad (1.2)$$

donde la función objetivo del problema de optimización, f , es convexa y diferenciable, λ_k es un parámetro positivo y d es la distancia euclidiana, esto es, $d(x, y) = \|x - y\|$.

Este algoritmo fué introducido por Martinet[34] en 1970 y fué muy estudiado por Rocafellar[42] para resolver problemas más generales de encontrar ceros de operadores monótonos maximales y ha sido extendido a diferentes contextos, ver [20] y [26].

Por otro lado, considere el problema:

$$\begin{cases} \min & f(x) \\ \text{s.a.} & \\ & x \in M \end{cases}$$

donde $f : M \rightarrow \mathbb{R}$ y M es una variedad riemanniana. Esas extensiones hacen posible resolver problemas no convexos restringidos en espacios euclidianos transformándolos en convexos desde el punto de vista de la métrica en alguna variedad riemanniana. Referencias básicas sobre variedades riemannianas pueden ser encontrados en [2, 13, 44, 46].

El MPP en variedades riemannianas es dado como: $x^0 \in M$ y

$$x^{k+1} = \arg \min \{ f(y) + \lambda_k \rho_{x^k}(y) \}, \quad (1.3)$$

con $\rho_x : M \rightarrow \mathbb{R}$ definido como $\rho_x(y) = \frac{1}{2} d^2(x, y)$, donde d es la distancia riemanniana y $\{\lambda_k\}$ es una sucesión de números positivos.

En [17] se prueba la buena definición de la sucesión generada por (1.3) y la convergencia a un minimizador de f , cuando existe el mínimo y M es una variedad riemanniana con curvatura seccional no positiva.

El MPP en espacios euclidianos, para resolver PDV, consiste en generar una sucesión de $\{x^k\}$ definida como: $x^0 \in \mathbb{R}^n$, dado x^k , encontrar x^{k+1} tal que

$$0 \in T_k(x^{k+1}),$$

donde $T_k = T(\cdot) + \lambda_k \nabla d(\cdot, x^k)$, λ_k es un parámetro positivo y d es la distancia euclidiana.

Se han logrado muchas extensiones del MPP para resolver PDV por ejemplo en espacios de Hilbert en [6] y en espacios de Banach en [7]. Para resolver PDV sobre un subconjunto de \mathbb{R}^n , el MPP usualmente sustituye la distancia euclidiana por otra función cuyas propiedades se asemejan a la distancia euclidiana.

Una de las funciones más usadas es la llamada “distancia de Bregman” la cuál está definida como

$$D_h(x, y) := h(x) - h(y) - \langle \nabla h(x), y - x \rangle,$$

donde h es una función que satisface algunas condiciones de convexidad y diferenciabilidad. Esta clase de distancias fué introducido por Bregman[3] es su artículo original; el término “función y distancia de Bregman” fué utilizado por primera vez por Censor y Lent[8] en 1981.

El objetivo de esta tesis es generalizar el MPP, para resolver un PDV, en variedades riemannianas desde que existen problemas que no pueden ser resueltos en el espacio euclidiano, sin embargo son solucionables si lo consideramos inmersos en una variedad riemanniana.

La tesis está dividida en 5 capítulos. En el Capítulo 1, presentamos las herramientas necesarias para el desarrollo de este trabajo, resultados de la teoría de operadores monótonos máximos, de geometría Riemanniana y de análisis convexo. En el Capítulo 2 se presenta el problema de PDV en espacios euclidianos, estudiaremos los resultados de existencia del conjunto solución del PDV, introduciremos el MPP clásico en la cuál obviamos la prueba de los resultados de convergencia desde que hay abundante literatura que concierne al tema y que serán citados en el transcurso de este capítulo. En el Capítulo 3, estudiaremos el problema de desigualdad variacional en variedades Riemannianas (PDVV). Presentamos, a manera de motivación, tres problemas impor-

tantes en la ciencia matemática relacionados al problema PDVV tales como el problema de singularidad, problema de optimización y el equilibrio de Nash; así como también los resultados de existencia para el PDVV. En el Capítulo 4 es presentado nuestro principal aporte, el MPP para resolver PDV en variedades Riemannianas. Consideraremos una generalización de este método inexacto (considerando una sucesión de error para cálculos computacionales) donde la distancia euclidiana es sustituida por una distancia más general, llamada distancia de Bregman en variedades riemannianas. Además, nos enfocaremos en los resultados de convergencia de este método utilizando conceptos del análisis convexo y geometría riemanniana dados en el primer capítulo. En el Capítulo 5, presentaremos un problema de desigualdad variacional lo cual resolveremos mediante el MPP. Además, se muestran resultados numéricos para comparar el costo computacional para diferentes parámetros positivos.

Capítulo 1

Preliminares

En este capítulo daremos los elementos matemáticos básicos que serán necesarios para el entendimiento y desarrollo de esta tesis. Iniciamos presentando algunos resultados de análisis, teoría de operadores maximales, conceptos básicos de geometría Riemanniana y del análisis convexo.

1.1 Algunos resultados de análisis

Definición 1.1 *Dado un conjunto X , un subconjunto τ de partes de X se dice que es una topología en X si:*

- i. $\emptyset \in \tau$ y $X \in \tau$.
- ii. Si $A, B \in \tau$, entonces $A \cap B \in \tau$.
- iii. Si $(A_i)_{i \in I}$ es una familia de elementos de τ , con I como un conjunto de índices, entonces $\bigcup_{i \in I} A_i \in \tau$.

Al par (X, τ) se llama espacio topológico y los elementos de τ se llaman los abiertos del espacio topológico (X, τ) .

Definición 1.2 *Sea (X, τ) y (Y, τ') dos espacios topológicos y una función $f : (X, \tau) \rightarrow (Y, \tau')$, se dice que f es:*

- i. continua en $a \in X$ si para todo $V \in \tau'$, $f(a) \in V$ entonces existe $U \in \tau$ tal que $a \in U$ y $f(U) \subseteq V$.
- ii. se dice continua si para todo $V \in \tau'$, $f^{-1}(V) \in \tau$.

Definición 1.3 Sean (X, τ) y (Z, τ') dos espacios topológicos. Una aplicación biyectiva continua $f : (X, \tau) \rightarrow (Z, \tau')$ tal que $f^{-1} : (Z, \tau') \rightarrow (X, \tau)$ es también continua, es llamado homeomorfismo. Además, decimos que (X, τ) y (Z, τ') son espacios homeomorfos, denotado como $(X, \tau) \cong (Z, \tau')$, si existe un homeomorfismo $f : (X, \tau) \rightarrow (Z, \tau')$.

Toda la teoría concerniente a espacios topológicos será de utilidad en nuestro trabajo desde que una variedad riemanniana es considerado un espacio topológico con ciertas propiedades. Para mas detalles se puede ver [10]. En toda esta tesis, \mathbb{R}^n denotará el espacio vectorial euclidiano con norma $\|x\| = \left(\sum_{i=1}^n x_i^2 \right)^{1/2}$

Definición 1.4 El conjunto $S \subset \mathbb{R}^n$ es afín, denotado como $\text{aff } S$, si

$$(1 - \lambda)x + \lambda y \in S, \quad \text{para todo } x, y \in S, \text{ y todo } \lambda \in \mathbb{R}.$$

Definición 1.5 Sea $S \subset \mathbb{R}^n$. El interior relativo de S , denotado como $\text{ri } S$, es el conjunto dado por

$$\text{ri } S := \{x \in S : B(x, \epsilon) \cap \text{aff } S \subseteq S, \quad \text{para algún } \epsilon > 0\}.$$

Definición 1.6 Sea $K \subseteq \mathbb{R}^n$ un subconjunto no vacío. Se dice que K es un conjunto sólido si este tiene un interior topológico, es decir, $\text{int}(K) \neq \emptyset$.

Lema 1.1 Sea $\{a_n\}$ y $\{\psi_n\}$ dos sucesiones en \mathbb{R} tal que $\{a_n\}$ es acotado inferiormente, $\sum_{n=0}^{+\infty} \psi_n$ existe y es finito y se cumple que:

$$a_{n+1} \leq a_n + \psi_n, \quad \forall n \geq 0,$$

entonces $\{a_n\}$ es convergente.

Prueba. Sea $b_n = \sum_{i=1}^n \psi_i$. Desde que $\infty > \sum_{i=1}^n \psi_i = b$, se tiene que

$$b_n \rightarrow b,$$

esto es, para todo $\epsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que para todo $n \geq n_0$, se tiene que $|b_n - b| < \epsilon$, esto es,

$$\epsilon > \left| \sum_{i=1}^n \psi_i - \sum_{i=1}^{\infty} \psi_i \right| = \left| \sum_{i=n+1}^{\infty} \psi_i \right|,$$

y, para todo $\epsilon > 0$, existe n_1 tal que $m, n \geq n_1$, $m > n$, se tiene

$$\begin{aligned} |a_m - a_n| &\leq |a_m - a_{m-1}| + |a_{m-1} - a_{m-2}| + \dots + |a_{m-(p-1)} - a_n| \\ &= \psi_{m-1} + \psi_{m-2} + \dots + \psi_n \\ &= \sum_{i=n}^{m-1} \psi_i. \end{aligned}$$

Desde que $\sum_{i=n}^{m-1} \psi_i < \sum_{i=n}^{\infty} \psi_i < \epsilon$, se tiene que $|a_m - a_n| < \epsilon$, esto es, $\{a_n\}$ es una sucesión de Cauchy en \mathbb{R} . Por lo tanto, $\{a_n\}$ es una sucesión convergente. ■

Definición 1.7 Dada una función $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$. Se dice que f es semicontinua inferior en $\bar{x} \in S$, si para toda sucesión $\{x^k\} \subset S$ convergente a \bar{x} se tiene que:

$$\liminf_{k \rightarrow +\infty} f(x^k) \geq f(\bar{x}). \quad (1.1)$$

Si f es semicontinua inferior para todo $x \in S$, entonces decimos que f es semicontinua inferior en S .

Definición 1.8 Sea $K \subseteq \mathbb{R}^n$, una función $T : K \rightarrow \mathbb{R}^n$ es llamada aplicación gradiente si existe una función $\theta : K \rightarrow \mathbb{R}$ tal que $T(x) = \nabla \theta(x)$, para todo $x \in \text{dom}(T)$.

Definición 1.9 Una función $T : K \rightarrow \mathbb{R}^n$ es simétrico si es una aplicación gradiente. Caso contrario, T es llamado asimétrico.

Definición 1.10 Sea $K \subset \mathbb{R}^n$, una aplicación $T : K \rightarrow \mathbb{R}^n$ es coerciva de orden p con respecto a K , si para cada $q \in T(x)$ existe $p \in (-\infty, 1)$ y $\hat{x} \in K$ tal que

$$\lim_{x \in K, \|x\| \rightarrow +\infty} \frac{\langle q, x - \hat{x} \rangle}{\|x\|^p} = +\infty. \quad (1.2)$$

Observación 1.1 Cuando $p = 1$, (1.2) se reduce a la definición estandar de coercividad de una aplicación.

Definición 1.11 Sea G una matriz simétrica definida positiva. La norma inducida por G es definida como

$$\|x\|_G = \langle x, Gx \rangle^{1/2},$$

donde $\langle \cdot, \cdot \rangle$ denota el producto interno usual (producto interno euclidiano).

Definición 1.12 Sea $\|\cdot\|_G$ una norma inducida por G . Sea $T : K \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$ una aplicación sobre un subconjunto cerrado no vacío. T es una contracción con respecto a la norma $\|\cdot\|_G$ si existe una constante $\alpha \in (0, 1]$ tal que

$$\|T(x) - T(y)\|_G \leq \alpha \|x - y\|_G, \quad \forall x, y \in K.$$

La constante α es llamada constante de contracción o módulo de T .

Definición 1.13 Sea una aplicación $T : K \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$. Se dice que $x^* \in K$ es un punto fijo de T si $T(x^*) = x^*$.

A continuación, un resultado de existencia y unicidad de puntos fijos para aplicaciones contractivas.

Proposición 1.1 Sea $K \subseteq \mathbb{R}^n$ un subconjunto no vacío y cerrado y $T : K \rightarrow K$ una aplicación contractiva para alguna norma $\|\cdot\|_G$ con módulo $\alpha \in [0, 1)$. Entonces, la aplicación T tiene un único punto fijo $x^* \in K$.

Prueba.

a) **Unicidad.** Fijemos $x(0) \in K$ y consideremos una sucesión $\{x(t)\}$ generada por

$$x(t+1) = T(x(t)).$$

Desde que T es una aplicación contractiva, para cada $\alpha \in [0, 1]$, sigue que

$$\|x(t+1) - x(t)\| = \|T(x(t)) - T(x(t-1))\| \leq \alpha \|x(t) - x(t-1)\|, \quad \forall t \geq 1$$

lo que implica que

$$\|x(t+1) - x(t)\| \leq \alpha^t \|x(1) - x(0)\|, \quad \forall t \geq 0. \quad (1.3)$$

Para cada $t \geq 0$, $m \geq 1$ y de la ecuación (1.3) tenemos

$$\begin{aligned}
\|x(t+m) - x(t)\| &\leq \|x(t+m) - x(t+(m-1))\| + \|x(t+(m-1))\| + \dots \\
&\quad + \|x(t+i) - x(t+i-1)\| \\
&= \sum_{i=1}^m \|x(t+i) - x(t+i-1)\| \\
&= \sum_{i=1}^m \alpha^{t+i-1} \|x(1) - x(0)\| \\
&= \|x(1) - x(0)\| \alpha^t (1 + \alpha + \dots + \alpha^{m-1}) \\
&\leq \|x(1) - x(0)\| \frac{\alpha^t}{1 - \alpha}.
\end{aligned}$$

Entonces, $\{x(t)\}$ es una sucesión de Cauchy. Por tanto, la sucesión es convergente, esto es, existe x^* tal que $x(t) \rightarrow x^*$. De la cerradura del conjunto K , se tiene que $x^* \in K$. Además,

$$\begin{aligned}
\|T(x^*) - x^*\| &\leq \|T(x^*) - x(t)\| + \|x(t) - x^*\| \\
&= \|T(x^*) - T(x(t-1))\| + \|x(t) - x^*\| \\
&\leq \alpha \|x^* - x(t-1)\| + \|x(t) - x^*\|,
\end{aligned}$$

y de la convergencia de $\{x(t)\}$ tenemos que $T(x^*) = x^*$.

b) **Unicidad** Supongamos que y^* es otro punto fijo de la aplicación T , entonces tenemos

$$\|x^* - y^*\| = \|T(x^*) - T(y^*)\| \leq \alpha \|x^* - y^*\|,$$

lo cuál implica que $x^* = y^*$. ■

A continuación, el famoso teorema de punto-fijo de Brouwer, aplicado a espacios dimensionalmente finito, lo cuál forma la base para varios teoremas de punto-fijos generales. La prueba de esos teoremas pueden ser encontrados en referencias básicas como [19, 14, 10].

Teorema 1.1 *Sea $\overline{B(0,1)} \subset \mathbb{R}^n$ una bola unitaria cerrada y $T : \overline{B(0,1)} \rightarrow \overline{B(0,1)}$ una aplicación continua. Entonces T tiene un punto fijo.*

Prueba. Ver [19], pp. 65. ■

Definición 1.14 Sea $K \subseteq \mathbb{R}^n$ un subconjunto no vacío y cerrado. Se dice que K es un subconjunto con propiedad de punto fijo si existe una aplicación $T : K \rightarrow K$ tal que T tiene un punto fijo.

El siguiente resultado es llamado el Teorema de Schauder-Tychonoff, dado por Dunford-Schwartz. Este resultado es una extensión al caso infinito dimensional del Teorema de Punto Fijo de Brouwer.

Teorema 1.2 Un subconjunto convexo, compacto y no vacío de un espacio topológico conexo tiene la propiedad de punto fijo.

Prueba. Ver [14], Teorema 5, pp. 456. ■

Definición 1.15 Se dice que T es una aplicación punto-conjunto en el espacio euclidiano n -dimensional si para todo $x \in \mathbb{R}^n$, $T(x) \subseteq \mathbb{R}^n$. Este tipo de aplicación lo denotaremos como $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$.

Observación 1.2 Es habitual identificar $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ como el gráfico de una aplicación punto-conjunto, denotado por $\text{graf}(T)$. Por tanto se define

$$T(x) := \{y : (x, y) \in T\}.$$

Definición 1.16 Se define el dominio de la aplicación punto-conjunto T como

$$\text{dom} T := \{x : (x, y) \in T\} = \{x : T(x) \neq \emptyset\}.$$

Definimos la suma de dos aplicaciones $T, Q \subseteq \mathbb{R}^n \times \mathbb{R}^n$ tal como

$$T + Q = \{(x, y + w) : (x, y) \in T \text{ y } (x, w) \in Q\},$$

de donde, $(T + Q)(x) = T(x) + Q(x)$.

Definición 1.17 Sea $(x, y), (x', y') \in T$. Un conjunto $T \subset \mathbb{R}^n \times \mathbb{R}^n$ es

1. *monótono(estrictamente)*, si $\langle y - y', x - x' \rangle \geq 0$ (> 0 para $x \neq x'$).
2. *pseudo-monótono*, si $\langle y', x - x' \rangle \geq 0$ entonces $\langle y, x - x' \rangle \geq 0$.

3. *paramonótono*, si es monótono y $\langle x - x', y - y' \rangle = 0$ entonces $(x, y), (x', y) \in T$.
4. *monótono maximal*, si es monótono y si para cualquier T' monótono tal que $T(x) \subset T'(x)$ para todo x , se tiene que $T = T'$.

Ejemplos 1.1 a. La aplicación $M : [-2, 2] \Rightarrow \mathbb{R}$ definida como

$$M(x) = \begin{cases} x^2 & , \text{ Si } x \in [-2, 0] \\ 1 + x^2 & , \text{ Si } x \in [0, 2] \end{cases}$$

es pseudo-monótona.

En efecto, sin pérdida de generalidad podemos suponer que $x \in [-2, 0]$ y $x' \in [0, 2]$, esto es, $-x \in [0, 2]$, de donde

$$x' - x \in [0, 4]. \quad (1.4)$$

Además, $0 \leq (1 + (x')^2)(x - x') = (x - x') + (x')^2(x - x')$, de donde

$$x' - x \leq (x')^2(x - x'). \quad (1.5)$$

Comparando (1.4) y (1.5) se tiene que $0 \leq (x')^2(x - x')$. Por lo tanto, M es una aplicación pseudo-monótona.

b. Las aplicaciones $N : [0, +\infty] \rightarrow \mathbb{R}$ y $T : [0, 2] \Rightarrow \mathbb{R}$ definidas como

$$N(x) = \frac{1}{1+x}, \quad y \quad T(x) = \begin{cases} -x+2 & , \text{ Si } x \in [0, 1] \\ x & , \text{ Si } x \in [1, 2] \end{cases}$$

son pseudo-monótonas pero no monótonas.

En efecto, supongamos que es monótona, esto es,

$$\begin{aligned} 0 &\leq \left(\frac{1}{1+x} - \frac{1}{1+x'} \right) (x - x') \\ &= -\frac{1}{(1+x)(1+x')} |x - x'|^2, \end{aligned}$$

lo cual es una contradicción desde que $(1+x)(1+x') > 0$. Por lo tanto, la aplicación punto-punto N no es monótona. Además,

$$0 \leq \frac{1}{(1+x')} (x - x')$$

desde que $(1+x') > 0$, se tiene que $x - x' \geq 0$. Entonces, $\left(\frac{1}{1+x}\right)(x - x') \geq 0$ pues $x \in [0, +\infty]$. Por lo tanto, la aplicación N es pseudo-convexa. Análogamente se tiene para la aplicación punto-conjunto T .

c. Sea $f : \mathbb{R} \rightarrow \mathbb{R}$ una función continuamente diferenciable tal que

$$f'(x) = \begin{cases} f'(x) < 0, & x < a \\ 0, & x \in [a, b] \\ f'(x) > 0, & x > b. \end{cases}$$

Es fácil probar que la función f es pseudo-convexa. El subdiferencial de esta función es el operador pseudo-monótono, esto es, $T(x) = \{f'(x)\}$.

Definición 1.18 Decimos que $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ es sobreyectivo si para todo $y \in \mathbb{R}^n$, existe un $x \in \mathbb{R}^n$ tal que $y \in T(x)$. La aplicación T será inyectivo si para $x \neq y$ se tiene que $T(x) \cap T(y) = \emptyset$.

Teorema 1.3 (Teorema de Minty) Si $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ es monótono maximal y $\lambda_k > 0$, entonces la aplicación $(I + \lambda_k T)$ es inyectiva y sobreyectiva.

Demostración: Ver [35]. ■

Lema 1.2 Si T y Q son aplicaciones monótonas entonces $T + Q$ es monótona.

Prueba. La prueba es trivial por la definición de suma de aplicaciones dadas anteriormente. ■

El siguiente lema, el cuál demostramos, será una herramienta útil para demostrar el Teorema 4.1.

Lema 1.3 Sea $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ una aplicación monótona maximal entonces T es un conjunto cerrado, es decir, $(z^k, y^k) \in T$ tal que $y^k \rightarrow \bar{y}$ y $z^k \rightarrow \bar{z}$, entonces $(\bar{z}, \bar{y}) \in T$.

Prueba. Definamos,

$$\left. \begin{array}{l} (z, y) \in T, \quad z \neq \bar{z} \\ (\bar{z}, \bar{y}) \in T, \quad z = \bar{z} \end{array} \right\} \in \tilde{T}$$

Sea $(z', y') \in T$ y $(z'', y'') \in T$, por la monotonicidad la aplicación T se tiene,

$$\langle z' - z'', y' - y'' \rangle \geq 0. \quad (1.6)$$

Por hipótesis, tenemos que $(z^k, y^k) \in T$.

Si $z^k \neq \bar{z}$, sustituyendo $z'' = z^k$ y $y'' = y^k$ en (1.6),

$$\langle z' - z^k, y' - y^k \rangle \geq 0, \quad (z', y') \in T.$$

Tomando límite cuando $k \rightarrow +\infty$,

$$\langle z' - \bar{z}, y' - \bar{y} \rangle \geq 0, \quad (z', y') \in T.$$

Entonces, $(\bar{z}, \bar{y}) \in T$.

Si $z^k = \bar{z}$, entonces $(z^k, y^k) = (\bar{z}, y^k) \in T$, tomando límite cuando $k \rightarrow +\infty$, se tiene que $(\bar{z}, \bar{y}) \in \tilde{T}$. Por hipótesis, \tilde{T} es maximal, esto es, $T = \tilde{T}$. Por lo tanto, $(\bar{z}, \bar{y}) \in T$. ■

Corolario 1.1 Sea $K \subset \mathbb{R}^n$. Si T_1 y T_2 son aplicaciones paramonótonas sobre K entonces, $T_1 + T_2$ es paramonótono sobre K .

Prueba. La prueba es trivial, sigue de la Definición 1.17, ítem 3. ■

Proposición 1.2 Sea $K \subset \mathbb{R}^n$ y sean $T_1, T_2 : K \rightrightarrows K$, aplicaciones monótonas maximales. Supongamos que cualquiera de las siguientes dos condiciones se cumple:

- i. Si $\text{dom}(T_1) \cap \text{int}(\text{dom}(T_2)) \neq \emptyset$, ó
- ii. Si $K = \mathbb{R}^n$ y $\text{ri}(\text{dom}(T_1)) \cap \text{ri}(\text{dom}(T_2)) \neq \emptyset$.

Entonces, $T_1 + T_2$ es una aplicación monótona maximal.

Prueba. i. a. Supongamos que $\text{dom}(T_2) \subset K$ es un subconjunto limitado. Trasladando T_1 y T_2 si fuera necesario y restando una aplicación constante de T_1 , podemos asumir que

$$0 \in T_1(0), \quad 0 \in \text{int}(\text{dom}(T_2)) \quad (1.7)$$

Sea J una aplicación definida como $x \rightarrow J(x)$ tal que $\langle x, J(x) \rangle = \|x\|^2 = \|J(x)\|^2$.

Por un resultado de análisis convexo, se tiene que si $im(T_1 + T_2 + J) = K$ si y solo si $T_1 + T_2$ es monótono maximal.

Entonces, probaremos que $K \subset im((T_1 + T_2) + J)$.

Restando una aplicación constante de T_2 podemos reducir el argumento para el caso donde $x^* = 0$. Necesitamos mostrar que existe $x \in K$ tal que

$$0 \in (T_1 + T_2 + J)(x), \quad (1.8)$$

esto es, existe $x^* \in K$ tal que

$$-x^* \in (T_1 + 1/2J)(x) \quad (1.9)$$

$$x^* \in (T_1 + 1/2J)(x) \quad (1.10)$$

Definimos las aplicaciones: $S_1, S_2 : K \rightarrow K$ tal que

$$S_1(x^*) = -(T_1 + 1/2J)^{-1}(-x^*) \quad (1.11)$$

$$S_2(x^*) = (T_2 + 1/2J)^{-1}(x^*).$$

Sumando (1.10) y (1.10),

$$\begin{aligned} 0 &\in (T_1 + 1/2J)(x) + (T_2 + 1/2J)(x) \\ &\in S_1(x^*) + S_2(x^*), \end{aligned}$$

esto es,

$$0 \in im(S_1 + S_2) \quad (1.12)$$

Entonces, para probar la existencia de un vector $x \in K$ satisfaciendo (1.8), es suficiente probar (1.12).

Desde que S_1, S_2 son operadores monótonos maximales simples-valuadas y continua tal que $dom(S_1) = K = dom(S_2)$.

Desde que $S_1 + S_2$ es un operador monótono simple-valuado y continuo tal que

$$dom(S_1 + S_2) = dom(S_1) \cap dom(S_2) = K,$$

entonces $S_1 + S_2$ es un operador monótono maximal.

Desde que $J(0) = 0$ y de (1.7) tenemos

$$0 \in T_1(0) = (T_1 + 1/2J)(0) = S_1(0).$$

Por la monotonicidad de S_1 ,

$$\langle S_1(x^*), x^* \rangle = \langle S_1(x^*) - S_1(0), x^* - 0 \rangle \geq 0, \quad \forall x^* \in K,$$

entonces $\langle S_1(x^*), x^* \rangle \geq 0$, para todo $x^* \in K$.

Además, de (1.12)

$$im(S_2) = dom(T_2 + 1/2J) = dom(T_2),$$

donde $im(S_2)$ es la imagen de la aplicación S_2 . Además, $im(S_2)$ es un conjunto limitado, pues por hipótesis $dom(T_2)$ es limitado. De (1.7), se tiene que

$$0 \in int(im(S_2)). \quad (1.13)$$

Afirmación. Para algún $x^*, y^* \in K$ y de la monotonicidad de S_2 tenemos

$$\begin{aligned} 0 &\leq \langle S_2(x^*) - S_2(y^*), x^* - y^* \rangle \\ &= \langle S_2(x^*), x^* - y^* \rangle - \langle S_2(y^*), x^* - y^* \rangle \\ &= \langle S_2(x^*), x^* \rangle + \langle S_2(y^*) - S_2(x^*), y^* \rangle - \langle S_2(y^*), x^* \rangle, \end{aligned}$$

entonces,

$$\langle S_2(x^*), x^* \rangle \geq \langle S_2(y^*), x^* \rangle + \langle S_2(x^*) - S_2(y^*), y^* \rangle. \quad (1.14)$$

Desde que $im(S_2)$ es limitado en K , $im(S_2) \subset B(0, \alpha_1)$ (una bola con centro en el origen y radio $\alpha_1 > 0$).

De la desigualdad de Cauchy Schwartz,

$$\begin{aligned} |\langle S_2(x^*) - S_2(y^*), y^* \rangle| &\leq \|S_2(x^*) - S_2(y^*)\| \|y^*\| \\ &\leq 2\alpha_1 \|y^*\|, \end{aligned} \quad (1.15)$$

pues, $im(S_2) \subset B(0, \alpha_1)$, entonces $S_2(x^*) - S_2(y^*) \in B(0, \alpha_1) - B(0, \alpha_2) = B(0, 2\alpha_1)$ y,

$$\|S_2(x^*) - S_2(y^*)\|_2 = 2\alpha_1.$$

De (1.13), S_2^{-1} es absolutamente acotado en 0. Por tanto, existe $\epsilon > 0$, $\alpha_2 > 0$ tal que

$$\{y : \|y\| \leq \epsilon\} \subset \{S_2(y^*) : \|y^*\| \leq \alpha_2\} \quad (1.16)$$

De (1.14), (1.15) y (1.16),

$$\begin{aligned} \langle S_2(x^*), x^* \rangle &\geq \langle S_2(y^*), x^* \rangle + \langle S_2(x^*) - S_2(y^*), y^* \rangle \\ &\geq \langle S_2(y^*), x^* \rangle - 2\alpha_1 \|y^*\| \\ &= \langle S_2(y^*), x^* \rangle - 2\alpha_1 \alpha_2 \end{aligned}$$

para cada $y^* \in dom(S_2)$ tal que $\|y^*\| \leq \alpha_2$, entonces

$$\begin{aligned} \langle S_2(x^*), x^* \rangle &\geq \sup_{\|y\| \leq \epsilon} \{\langle S_2(y^*), x^* \rangle - 2\alpha_1 \alpha_2\} \\ &= \epsilon \|x^*\| - 2\alpha_1 \alpha_2. \end{aligned}$$

Si $\|x^*\| \geq \frac{2\alpha_1 \alpha_2}{\epsilon}$ entonces $\langle S_2(x^*), x^* \rangle \geq 0$ para $\alpha \geq \frac{2\alpha_1 \alpha_2}{\epsilon}$. ■

Por lo tanto, $0 \in dom(T_1) \cap int(dom(T_2))$, esto es, $dom(T_1) \cap int(dom(T_2)) \neq \emptyset$.

b. Ahora mostraremos que no era necesario la condición de que $dom(T_2)$ sea acotado como fue hecho anteriormente. Sean T_1 y T_2 operadores monótonos maximales tal que

$$dom(T_1) \cap int(dom(T_2)) \neq \emptyset \quad (1.17)$$

donde $dom(T_2)$ no es necesariamente limitado. Trasladando los dominios de T_1 y T_2 si fuera necesario, podemos asumir que el origen pertenece a la intersección en (1.17).

Para cada α , definimos B_α como el subdiferencial de la función indicadora de la bola cerrada U_α , de radio α y alrededor del origen en x . Por tanto, B_α es el operador normal para U_α definido como

$$B_\alpha(x) = \begin{cases} \emptyset, & \|x\| > \alpha \\ \lambda J(x), & \|x\| = \alpha \text{ y } \lambda \geq 0 \\ 0 \in K, & \|x\| < \alpha \end{cases}$$

Donde, B_α es un operador monótono maximal.

Además, se observa que

$$\text{dom}(B_\alpha) = \{x : \|x\| \leq \alpha\} \quad (1.18)$$

es limitado. Entonces, usando el ítem a. tenemos que $T_2 + B_\alpha$ es un operador monótono maximal.

De (1.16) y (1.18), se observa que el origen pertenece a la intersección en (1.16), entonces

$$\text{dom}(T_1) \cap \text{int}(\text{dom}(T_2 + B_\alpha)) \neq \emptyset,$$

donde $\text{dom}(T_2 + B_\alpha)$ es un conjunto limitado. Por tanto, $(T_1 + T_2) + B_\alpha = T_1 + (T_2 + B_\alpha)$ es un operador monótono maximal para cada $\alpha > 0$. Entonces, $T_1 + T_2$ es monótono maximal.

ii. Probaremos el segundo ítem. Trasladando T_1 , T_2 si fuera necesario, podemos suponer que

$$0 \in \text{ri}(\text{dom}(T_1)) \cap \text{ri}(\text{dom}(T_2)) \quad (1.19)$$

tal que la envolvente afin de $\text{dom}(T_1)$ y $\text{dom}(T_2)$ son ciertos subespacios L_1 y L_2 de K respectivamente, esto es, $\text{dom}(T_1) \subset L_1$ y $\text{dom}(T_2) \subset L_2$.

Sea $L_0 = L_1 \cap L_2$ y para $i = 0, 1, 2$,

$$L_i^\perp = \{x^* \in K : \langle x, x^* \rangle = 0, \quad \forall x \in L_i\} \quad (1.20)$$

y,

$$P_i(x) = \begin{cases} L_i^\perp, & x \in L_i \\ \emptyset, & x \notin L_i \end{cases}$$

Cada operador P_i es un operador monótono maximal ($P_i = \partial\delta_{L_i}$), y

$$P_0 = P_1 + P_2 = P_0 + P_1 = P_0 + P_2. \quad (1.21)$$

Dado cualquier $x \in \text{dom}(T_1)$, existe x^* tal que $x^* \in T_1(x)$, $z^* \in L_1^{\text{perp}}$, $y \in \text{dom}(T_1)$ y $y^* \in T_1(y)$,

$$\langle x - y, (x^* - z^*) - y^* \rangle = \langle x - y, x^* - y^* \rangle \geq 0.$$

De la monotonicidad de T_1 , con $\text{dom}(T_1) \subset L_1$, esto implica que la maximalidad de T_1 que $x^* + z^* \in T_1(x)$. Por tanto,

$$T_1 = T_1 + P_1. \quad (1.22)$$

Análogamente, tenemos

$$T_2 = T_2 + P_2, \quad (1.23)$$

y esto sigue que

$$T_1 + T_2 = T_1 + T_2 + P_1 + P_2 = (T_1 + P_0) + T_2.$$

De (1.21) y (1.22), podemos considerar T_1 y P_0 en una forma natural como un operador monótono maximal del espacio L_1 al espacio cociente K/L_1^\perp , lo cuál puede ser identificado con el dual L_1^* de L_1 . Por tanto, $T_1 + P_0$ debe ser un operador monótono maximal.

Análogamente, para el espacio L_2 , de (1.23) y $(T_1 + P_0) + P_2 = T_1 + P_0$, podemos considerar $T_1 + P_0$ y T_2 como operadores monótonos maximales de L_2 a K/L_2^\perp . El interior relativo de $\text{dom}(T_2)$ relativo L_2 pertenece a $\text{dom}(T_1 + P_0) = \text{dom}(T_1) \cap L_2$, lo cuál es un operador monótono maximal. ■

1.2 Elementos de geometría riemanniana

En esta sección se presentarán las notaciones y propiedades fundamentales que serán necesarios para desarrollar el Capítulo 4 de esta tesis y que pueden ser encontrados en referencias básicas como [2, 13, 44, 46].

Definición 1.19 *Una variedad diferenciable de dimensión n es un conjunto M y una familia de aplicaciones biunívocas $x_\alpha : U_\alpha \subset \mathbb{R}^n \rightarrow M$ de abiertos U_α de \mathbb{R}^n en M tales que:*

$$a. \quad \bigcup_\alpha x_\alpha(U_\alpha) = M.$$

- b. Para todo par α, β , con $x_\alpha(U_\alpha) \cap x_\beta(U_\beta) = W \neq \emptyset$, los conjuntos $x_\alpha^{-1}(W)$ y $x_\beta^{-1}(W)$ son abiertos en \mathbb{R}^n y las aplicaciones $x_\beta^{-1} \circ x_\alpha$ son diferenciables.
- c. La familia $\{(U_\alpha, x_\alpha)\}$ es máxima relativamente a las condiciones del ítem a. y b.

Una variedad diferenciable M es conexa si éste no es la unión de dos conjuntos abiertos disjuntos no vacíos. M es simplemente conexa si tiene un camino conexo (esto es, si para cada par de puntos de M puede ser unido por un camino) y cualquier aplicación continua $f : S^1 \rightarrow M$ (donde S^1 denota el círculo unitario en el espacio euclidiano E^2) puede ser contraído a un punto en el siguiente sentido: existe una aplicación continua $F : D^2 \rightarrow M$ (donde D^2 denota el disco unitario en el espacio euclidiano E^2) tal que F restringido a S^1 es la aplicación f .

Sea M una variedad conexa, denotamos por $T_x M$ y $TM = \bigcup_{x \in M} T_x M$ como el espacio tangente de M en x y el fibrado tangente de M respectivamente. Si M es dotada de una métrica riemanniana g , con $G = (g_{ij})$ como la matriz de representación de la métrica g , siendo esta matriz definida positiva, entonces M es una variedad riemanniana. Se define la inversa de la matriz (g_{ij}) como (g^{ij}) . La métrica puede ser usada para definir la longitud de una curva regular por trozos $\alpha : [a, b] \rightarrow M$ tal que $\alpha(a) = x'$ y $\alpha(b) = x$, dado como

$$L(\alpha) = \int_a^b \|\alpha'(t)\|_G dt.$$

Minimizando el funcional sobre el conjunto de todas estas curvas obtenemos una distancia riemanniana $d(x, x')$ que induce una topología natural sobre M .

Un campo de vectores X en una variedad diferenciable M es una correspondencia que a cada punto $p \in M$ asocia un vector $X(p) \in T_p M$. El campo es diferenciable si la aplicación $X : M \rightarrow TM$ es diferenciable.

Sea $\Omega = C^\infty(K) = \{X : K \rightarrow TK \text{ tal que para cada } p \in K, X(p) \in T_p K, X \in C^\infty\}$ el conjunto de los campos vectoriales y $D = D(K) = \{f : K \rightarrow \mathbb{R} : f \in C^\infty\}$ el anillo de las funciones reales de clase C^∞ .

Una conexión afín es una aplicación $\nabla : \Omega \times \Omega \rightarrow \Omega$ donde para cada par de campos $(X, Y) \in \Omega \times \Omega$ se asocia otro campo $\nabla_X Y$ que verifica ciertas condiciones, ver [13]. Los símbolos de Christoffel (o los coeficientes de la conexión afín ∇ en U) son las funciones $\Gamma_{i,j}^m : U \subset M \rightarrow \mathbb{R}$ caracterizadas por

$$\Gamma_{i,j}^m = \frac{1}{2} \sum_{k=1}^n \left\{ \frac{\partial}{\partial x_i} g_{jk} + \frac{\partial}{\partial x_j} g_{ki} - \frac{\partial}{\partial x_k} g_{ij} \right\} g^{km}. \quad (1.24)$$

Sea ∇ una conexión afín asociada a (M, \langle, \rangle) . Por el Teorema de Levi-Civita, ver [13] página 61, existe una única aplicación llamada derivada covariante y es caracterizado en una cierta parametrización como:

$$\frac{DV}{dt} = \sum_{j=1}^n \frac{dv^j}{dt} X_j + \sum_{i,j=1}^n v^j \frac{d\alpha_i}{dt} \nabla_{X_i} X_j. \quad (1.25)$$

Una conexión dada por el Teorema Levi-Civita es denominada conexión de Levi-Civita o conexión riemanniana de M .

Definición 1.20 Sea M una variedad riemanniana unida a su conexión riemanniana. Una curva parametrizada $\alpha : I \rightarrow M$ es una geodésica en $t_0 \in I$ si $\frac{D}{dt} \left(\frac{d\alpha}{dt} \right) = 0$ en el punto t_0 ; si α es geodésica en t , para todo $t \in I$, decimos que α es una geodésica. Si $[a, b] \subset I$ y $\alpha : I \rightarrow M$ es una geodésica, la restricción de α en $[a, b]$ es llamada geodésica ligando $\alpha(a)$ y $\alpha(b)$.

De (1.25), un campo vectorial, a lo largo de α , es llamado paralelo si $\nabla_{\alpha'} V = 0$. Si α' es paralelo entonces α es una geodésica. La ecuación de la geodésica es:

$$\frac{d^2 \alpha_k}{dt^2} + \sum_{i,j=1}^n \Gamma_{i,j}^k \frac{d\alpha_j}{dt} \cdot \frac{d\alpha_i}{dt} = 0, \quad k = 1, 2, \dots, n. \quad (1.26)$$

Sea $U \subset M$, TU es el fibrado tangente de U y

$$\mathcal{U} = \{(q, v) \in TU : q \in V \text{ y } v \in T_q M \text{ con } |v| < \epsilon_1\},$$

donde $V \subset U$ es una vecindad de $p \in M$.

Por un resultado de geometría riemanniana, dado $p \in M$ existe un abierto $V \subset M$, $p \in V$, $\delta > 0$, $\epsilon_1 > 0$ y una aplicación infinitamente diferenciable dado como:

$$\alpha : (-\delta, \delta) \times \mathcal{U} \rightarrow M$$

tal que la curva $t \mapsto \alpha(t, q, v)$, con $t \in (-\delta, \delta)$ es una única geodésica de M en el instante $t = 0$ que pasa por q con velocidad v , para cada $q \in V$ y cada $v \in T_q M$ con $|v| < \epsilon_1$. Por abuso de notación, consideraremos $\alpha(t, q, v) = \alpha(t)$.

Sea $\alpha : [0, 1] \rightarrow M$ una curva geodésica tal que $\alpha(0) = q$ y $\alpha(1) = p$, la aplicación exponencial en \mathcal{U} , denotado como $\exp : T\mathcal{U} \rightarrow M$, es definido por

$$\exp(q, v) = \alpha(1, q, v). \quad (1.27)$$

Es claro que la aplicación \exp es diferenciable. En la mayor parte de las aplicaciones, utilizaremos la restricción de \exp en un abierto del espacio tangente $T_q M$, esto es, definiremos

$$\exp_q : B(0, \epsilon) \subset T_q M \rightarrow M$$

por $\exp_q(v) = \exp(q, v)$.

Definición 1.21 *Una variedad riemanniana M es (geodésicamente) completa si para todo $p \in M$, la aplicación exponencial, \exp_p , está definida para todo $v \in T_p M$, esto es, si las geodésicas $\alpha(t)$ que parten de t están definidas para todo $t \in \mathbb{R}$.*

Se observa también que el Teorema de Hop Rinow asegura la existencia de esta aplicación para todos los puntos de M . La aplicación lineal $P_{\alpha,0,1} : T_{\alpha(0)} M \rightarrow T_{\alpha(1)} M$ es llamado el transporte paralelo a lo largo de α . Se observa que el transporte paralelo $P_{\alpha,0,1}$ preserva el producto interno, esto es, una aplicación isométrica. Por un resultado clásico de geometría riemanniana, se tiene

$$P_{\alpha,0,1} \exp_x^{-1} y = -\exp_y^{-1} x. \quad (1.28)$$

El tensor de curvatura, denotado como R , es definido por

$$R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[Y, X]} Z,$$

donde X, Y y Z son campos vectoriales de M y la curvatura seccional con respecto a X y Y es dado por

$$K(X, Y) = \frac{\langle R(X, Y)Y, X \rangle}{\|X\|^2 \|Y\|^2 - \langle X, Y \rangle^2},$$

donde $\|X\| = \langle X, X \rangle^{1/2}$.

Definición 1.22 Dado un punto $p \in M$ y un subespacio bidimensional $\sigma \subset T_p M$, el número real $K(X, Y) = K(p, \sigma)$, donde $\{X, Y\}$ es una base cualquiera de σ , es llamado curvatura seccional de σ en p .

Corolario 1.2 Sea M una variedad riemanniana, $n = \dim M$, $p \in M$ y $\{e_1, \dots, e_n\}$ una base ortonormal de $T_p M$ y sea $R_{ijkl} = \langle R(e_i, e_j)e_k, e_l \rangle$, $i, j, k, l = 1, \dots, n$. Entonces, $K(p, \sigma) = K_0$ para todo $\sigma \subset T_p M$, si y solo si, $R_{ijij} = -R_{ijji} = K_0$ para todo $i \neq j$, y $R_{ijkl} = 0$ en otros casos.

Prueba. Ver [13], Corolario 3.5, pp.107. ■

Si $K(X, Y) \leq 0$ para todo X y Y , entonces M es llamado una variedad riemanniana de curvatura seccional no positiva.

Definición 1.23 Una variedad de Hadamard es una variedad riemanniana, simplemente conexa, completa y de curvatura seccional no positiva.

Teorema 1.4 Sea M una variedad riemanniana, simplemente conexa y completa con curvatura seccional no positiva. Entonces M es difeomorfa al espacio euclidiano \mathbb{R}^n , $n = \dim M$, esto es, en cualquier punto $x \in M$, la aplicación exponencial $\exp_x : T_x M \rightarrow M$ es un difeomorfismo.

Demostración: Ver [13], Teorema 3.1, pp.165. ■

Ejemplos 1.2 A continuación, se muestra 2 ejemplos de variedades de Hadamard: el plano hiperbólico y el ortante positivo bajo ciertas métricas. Éste último espacio es importante para nuestro trabajo porque será considerado en el capítulo 5.

1. El plano Hiperbólico

Sea el conjunto

$$H = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 > 0\},$$

inducida con la métrica riemanniana con $G = (g_{ij})$ donde

$$g_{ij} = \frac{1}{x_2^2} \delta_{ij}, \quad i, j = 1, 2$$

con

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j. \end{cases}$$

El conjunto H , inducida con tal métrica, es llamado el plano de Poincaré o plano Hiperbólico.

A continuación, demostraremos que el plano hiperbólico H , definido anteriormente, es una variedad de Hadamard, esto es, una variedad riemanniana, simplemente conexa, completa y de curvatura seccional no positiva. Para tal fin, es necesario calcular las componentes de la conexión riemanniana, esto es, los símbolos de Christoffel, las ecuaciones de las geodésicas, la distancia riemanniana y la curvatura seccional.

a. Símbolos de Christoffel.

Las componentes de la métrica están dados como

$$\begin{aligned} g_{11} &= g_{22} = \frac{1}{x_2^2} \\ g_{12} &= g_{21} = 0 \end{aligned} \quad (1.29)$$

Si $k \neq m$, entonces $g_{km} = 0$ entonces, de la ecuación (1.24), se tiene:

$$\Gamma_{ij}^m = \frac{1}{2} g^{mm} \left(\frac{\partial}{\partial x_i} g_{jm} + \frac{\partial}{\partial x_j} g_{mi} - \frac{\partial}{\partial x_k} g_{ij} \right). \quad (1.30)$$

Sustituyendo (1.31) en (1.30), la conexión riemanniana en H tiene las siguientes componentes:

$$\begin{aligned} \Gamma_{11}^1 &= \Gamma_{22}^1 = \Gamma_{12}^2 = \Gamma_{21}^2 = 0 \\ \Gamma_{12}^1 &= \Gamma_{21}^1 = \Gamma_{22}^2 = -\Gamma_{21}^1 = -\frac{1}{x_2} \end{aligned} \quad (1.31)$$

b. Ecuación de la geodésica.

Sea una curva $\alpha(t) = (\alpha_1(t), \alpha_2(t)) \in H$ tal que $\alpha(0) = (y_1, y_2)$ y $\alpha'(0) = (v_1, v_2)$. Entonces sustituyendo las componentes de la conexión afín en (1.26), se tiene el siguiente sistema:

$$\frac{d^2 \alpha_1}{dt^2} - \frac{2}{\alpha_2} \frac{d\alpha_1}{dt} \frac{d\alpha_2}{dt} = 0 \quad (1.32)$$

$$\frac{d^2 \alpha_2}{dt^2} + \frac{1}{\alpha_2} \left(\frac{d\alpha_1}{dt} \right)^2 - \frac{1}{\alpha_2} \left(\frac{d\alpha_2}{dt} \right)^2 = 0. \quad (1.33)$$

Reemplazando $p(t) = \frac{d\alpha_1}{dt}$ en (1.32), se tiene

$$\frac{dp}{dt} - \frac{2}{\alpha_2} \frac{d\alpha_2}{dt} p = 0$$

donde, por un resultado de cálculo diferencial se obtiene la solución:

$$p(t) = k\alpha_2^2(t), \quad k > 0. \quad (1.34)$$

Reemplazando (1.34) en (1.33),

$$\frac{d^2\alpha_2}{dt^2} + k^2\alpha_2^3 - \frac{1}{\alpha_2} \left(\frac{d\alpha_2}{dt} \right)^2 = 0. \quad (1.35)$$

Definimos

$$\frac{d\alpha_2}{dt} := z(t)\alpha_2(t), \quad (1.36)$$

entonces,

$$\frac{d^2\alpha_2}{dt^2} = \alpha_2(t) \left(\frac{dz}{dt} + z(t)^2 \right). \quad (1.37)$$

Sustituyendo (1.36) y (1.37) en (1.35) y desde que $\alpha_2(t) \in H$, esto es, $\alpha_2(t) > 0$ se tiene,

$$z(t) = r - k\alpha_1(t), \quad r \in \mathbb{R}. \quad (1.38)$$

Sustituyendo (1.38) en (1.36),

$$\alpha_2(t) = s e^{\int r - k\alpha_1(t) dt}, \quad s > 0. \quad (1.39)$$

Si $\alpha_1(t) = c$ con c una constante arbitraria, entonces $\alpha_2(t) = se^{ht}$, con $h \in \mathbb{R}$. Por tanto,

$$(\alpha_1(t), \alpha_2(t)) = (c, se^{ht}). \quad (1.40)$$

Si $\alpha_1(t)$ no es constante en (1.39). Definimos:

$$\alpha_2(t) = \frac{s}{k} \operatorname{sech} pt, \quad s > 0. \quad (1.41)$$

Igualando (1.41) en (1.39) y derivando con respecto de t , se tiene

$$\alpha_1(t) - b = \frac{s}{k} \tanh pt, \quad b \in \mathbb{R} \quad (1.42)$$

Por lo tanto,

$$(\alpha_1(t), \alpha_2(t)) = \left(b + \frac{s}{k} \tanh pt, \frac{s}{k} \operatorname{sech} pt \right). \quad (1.43)$$

Además, de (1.41) y (1.42), se tiene

$$(\alpha_1(t) - b)^2 + \alpha_2(t)^2 = R^2,$$

con $0 < R = s/k$, esto es, $\alpha_1, \alpha_2 \in C_{b,R}$ con $\alpha_2(t) > 0$ donde $C_{b,R}$ es la semicircunferencia de centro b y radio r .

Evaluando $\alpha(t) \in H$ en (1.40) y (1.43) tal que $\alpha(0) = (y_1, y_2)$ y $\alpha'(0) = (v_1, v_2)$, se tiene

$$(\alpha_1(t), \alpha_2(t)) = (y_1, y_2 e^{v_2/y_2 t}), \quad v_2 = 0 \quad (1.44)$$

$$(\alpha_1(t), \alpha_2(t)) = \left(y_1 + y_2 \tanh \frac{v_1}{y_2} t, y_2 \operatorname{sech} \frac{v_1}{y_2} t \right), \quad v_1 = 0 \quad (1.45)$$

que son las geodésicas del plano Hiperbólico H .

c. Distancia Riemanniana.

Sea $\alpha(y, v, 1) = (z_1, z_2)$, esto es, $(\alpha_1(t), \alpha_2(t))|_{t=1} = (z_1, z_2)$.

c.1) Si $\alpha(t)$ es la geodésica dada en (1.45), se tiene que $z_1 = y_1$ y $v_2 = y_2 \ln \frac{z_2}{y_2}$. De donde, $\exp_y v = x$, si y solo si, $v = \exp_y^{-1} x$.

Entonces,

$$\begin{aligned} d^2(z, y) &= \|\exp_y^{-1} z\|_G^2 \\ &= \langle v, G(y)v \rangle \\ &= \frac{v_2^2}{y_2^2} \\ &= \ln^2 \frac{z_2}{y_2}. \end{aligned}$$

c.2) Si $\alpha(t)$ es la geodésica dada en (1.44), se tiene que

$$\begin{aligned} z_1 &= y_1 + y_2 \tanh \frac{v_1}{y_2} \\ z_2 &= y_2 \operatorname{sech} \frac{v_1}{y_2}. \end{aligned}$$

Resolviendo el sistema anterior, se obtiene

$$v_1 = \ln \left(\frac{z_1 y_2 + \sqrt{z_1^2 y_2^2 - z_2^2 y_1^2 + z_2^2 y_2^2}}{z_2 (y_1 + y_2)} \right) y_2,$$

de donde,

$$\begin{aligned}
d^2(z, y) &= \|\exp_y^{-1} z\|_G^2 \\
&= \langle v, G(y)v \rangle \\
&= \frac{v_1^2}{y_2^2} \\
&= \ln^2 \left(\frac{z_1 y_2 + \sqrt{z_1^2 y_2^2 - z_2^2 y_1^2 + z_2^2 y_2^2}}{z_2 (y_1 + y_2)} \right)
\end{aligned}$$

Por lo tanto, la distancia hiperbólica entre los puntos $y = (y_1, y_2)$ y $z = (z_1, z_2)$ está dado como:

$$d(z, y) = \begin{cases} \left| \ln \frac{z_1}{y_1} \right|, & x_1 = x_2 \\ \left| \ln \left(\frac{z_1 y_2 + \sqrt{z_1^2 y_2^2 - z_2^2 y_1^2 + z_2^2 y_2^2}}{z_2 (y_1 + y_2)} \right) \right|, & y, z \in C_{b,R}. \end{cases}$$

d. Curvatura seccional.

Sea $\{X_1, X_2\}$ una base local ortonormal del espacio hiperbólico H , lo cuál denotamos por $X_i = \frac{\partial}{\partial x_i}$ con $i = 1, 2$.

Se tiene,

$$\begin{aligned}
\nabla_{X_1} X_1 &= -\nabla_{X_2} X_2 = \frac{1}{x_2} X_2, \\
\nabla_{X_1} X_2 &= -\nabla_{X_2} X_1 = -\frac{1}{x_2} X_1,
\end{aligned}$$

y,

$$\begin{aligned}
\nabla_{X_2} (\nabla_{X_1} X_2) &= \nabla_{X_2} (\nabla_{X_2} X_1) = \nabla_{X_2} (\nabla_{X_2} X_2) = 2 \nabla_{X_1} (\nabla_{X_2} X_2) = -2 \nabla_{X_1} (\nabla_{X_1} X_1) = \frac{2}{x_2^2} X_1 \\
\nabla_{X_2} (\nabla_{X_1} X_1) &= 2 \nabla_{X_1} (\nabla_{X_1} X_2) = -\frac{2}{x_2^2} X_2.
\end{aligned}$$

Además,

$$R(X_i, X_j) X_k = \nabla_{X_j} (\nabla_{X_i} X_k) - \nabla_{X_i} (\nabla_{X_j} X_k),$$

esto es,

$$R(X_1, X_1) X_1 = R(X_2, X_2) X_1 = R(X_1, X_1) X_2 = R(X_2, X_2) X_2 = 0$$

$$R(X_2, X_1) X_1 = -R(X_1, X_2) X_1 = \frac{1}{x_2^2} X_2$$

$$R(X_2, X_1) X_2 = -R(X_1, X_2) X_2 = -\frac{1}{x_2^2} X_1.$$

Desde que $R_{ijkl} = \langle R(X_i, X_j)X_k, X_l \rangle$ y por la ortonormalidad de $\{X_1, X_2\}$, se tiene que

$$R_{1212} = -R_{1221} = -1,$$

$$R_{2121} = -R_{2112} = -1.$$

Luego, del Corolario 1.2 se tiene que la curvatura seccional es -1, esto es, una curvatura seccional no positiva.

Se muestra que, H está dotada de una métrica riemanniana $\langle \cdot, \cdot \rangle$ con $G(x) = \left(\frac{1}{x_2} \delta_{ij}\right)$ para $i, j = 1, 2$, para dos puntos cualquiera dados en H hay uno y sólo un camino conectándolos y que las geodésicas dadas en (1.44) y (1.45) están definidas para todo $t \in \mathbb{R}$, entonces la variedad H es una variedad riemanniana, simplemente conexa y completa respectivamente. Por lo tanto, el espacio hiperbólico H es una variedad de Hadamard.

2. El Ortante Positivo. Sea el conjunto

$$\mathbb{R}_{++}^2 = \{(x_1, x_2) : x_1 > 0, x_2 > 0\}.$$

Si \mathbb{R}_{++}^2 es inducida por una métrica riemanniana diagonal definida como

$$G(x) = \text{diag} \left(\frac{1}{x_1^2}, \frac{1}{x_2^2} \right) = X^{-2},$$

donde $X = \text{diag}(x_1, x_2)$, entonces $(\mathbb{R}_{++}^2, G(x))$ es una variedad riemanniana.

a. Símbolos de Christoffel.

La conexión riemanniana tiene los siguientes coeficientes:

$$\Gamma_{11}^1 = \frac{-1}{x_1}, \quad \Gamma_{22}^2 = \frac{-1}{x_2}, \quad \Gamma_{12}^1 = \Gamma_{21}^1 = \Gamma_{22}^1 = \Gamma_{11}^2 = \Gamma_{12}^2 = \Gamma_{21}^2 = 0.$$

b. Ecuación de la geodésica.

Sea $\alpha(t)$ una curva en \mathbb{R}_{++}^2 , reemplazando todos los coeficientes dados anteriormente en

(1.26), se tiene

$$\begin{aligned}\frac{d^2\alpha_1}{dt^2} - \frac{1}{\alpha_1} \left(\frac{d\alpha_1}{dt}\right)^2 &= 0, \\ \frac{d^2\alpha_2}{dt^2} - \frac{1}{\alpha_2} \left(\frac{d\alpha_2}{dt}\right)^2 &= 0.\end{aligned}$$

Solucionando el sistema anterior, y considerando las condiciones iniciales para $\alpha(t)$ dados como $\alpha(0) = (y_1, y_2)$ y $\alpha'(0) = (v_1, v_2) = v$ se tiene

$$\alpha(t) = (y_1 e^{(v_1/y_1)t}, y_2 e^{(v_2/y_2)t}) \quad i = 1, 2. \quad (1.46)$$

c. Distancia Riemanniana.

Sea $x = (x_1, x_2)$ y $y = (y_1, y_2) \in \mathbb{R}_{++}^2$ tal que $\alpha(1) = x$, se tiene que si $\alpha(y, v, 1) = y e^{v/y}$ entonces

$$(v_1, v_2) = \left(y_1 \ln \left(\frac{x_1}{y_1} \right), y_2 \ln \left(\frac{x_2}{y_2} \right) \right). \quad (1.47)$$

De donde, la ecuación de la geodésica que une los puntos y y x está dado como

$$\alpha(t) = (x_1^t y_1^{1-t}, x_2^t y_2^{1-t}). \quad (1.48)$$

Desde que $v = \exp_y^{-1}x$ entonces

$$\begin{aligned}d^2(x, y) &= \|\exp_x^{-1}y\|_G, \\ &= \ln^2 \left(\frac{y_1}{x_1} \right) + \ln^2 \left(\frac{y_2}{x_2} \right).\end{aligned} \quad (1.49)$$

Por lo tanto, la distancia riemanniana que une los puntos y y x es dada por:

$$d(x, y) = \left(\ln^2 \left(\frac{y_1}{x_1} \right) + \ln^2 \left(\frac{y_2}{x_2} \right) \right)^{1/2}.$$

d. Transporte paralelo.

Consideremos un campo vectorial $V = \sum_{i=1}^2 v_i X_i$ sobre la curva $\alpha : \mathbb{R} \rightarrow \mathbb{R}_{++}^2$. Sustituyendo en (1.25) obtenemos,

$$\begin{aligned}\frac{dV_1}{dt} - \frac{V_1}{\alpha_1} \frac{d\alpha_1}{dt} &= 0 \\ \frac{dV_2}{dt} - \frac{V_2}{\alpha_2} \frac{d\alpha_2}{dt} &= 0,\end{aligned}$$

tal que $V(0) = (V_1(0), V_2(0)) = (v_1, v_2)$.

Solucionando el sistema, para $i = 1, 2$, se tiene:

$$V_i = \alpha_i k, \quad k > 0,$$

de donde, $V_i(t) = \alpha_i(t) k$, $i = 1, 2$.

Evaluable en $t = 0$, $k = \frac{v_i}{y_i}$, para $i = 1, 2$, y así obtenemos

$$V_i(t) = \frac{\alpha_i(t)v_i}{y_i}, \quad i = 1, 2.$$

De (1.48), el transporte paralelo a lo largo de la geodésica $\alpha(t)$ es dado como:

$$\begin{aligned} P_{\alpha,0,t}V(t) = (V_1(t), V_2(t)) &= \left(\frac{\alpha_1(t)v_1}{y_1}, \frac{\alpha_2(t)v_2}{y_2} \right) \\ &= \left(\frac{v_1}{y_1} x_1^t y_1^{1-t}, \frac{v_2}{y_2} x_2^t y_2^{1-t} \right). \end{aligned} \quad (1.50)$$

Para $t = 1$ se tiene,

$$P_{\alpha(t),0,1}V(1) = \left(\frac{v_1}{y_1} x_1, \frac{v_2}{y_2} x_2 \right). \quad (1.51)$$

e. Curvatura seccional.

Si $i \neq j$ entonces $\nabla_{X_i} X_j = 0$.

Si $i = j$,

$$\begin{aligned} \nabla_{X_1} X_1 &= \Gamma_{11}^1 X_1 + \Gamma_{11}^2 X_2 = -\frac{1}{x_1} X_1, \\ \nabla_{X_2} X_2 &= \Gamma_{22}^1 X_1 + \Gamma_{22}^2 X_2 = -\frac{1}{x_2} X_2 \end{aligned}$$

entonces,

$$\begin{aligned} \nabla_{X_2} (\nabla_{X_1} X_1) &= \nabla_{X_2} \left(-\frac{1}{x_1} X_1 \right) = -\frac{1}{x_1} \nabla_{X_2} X_1 + X_2 \left(-\frac{1}{x_1} \right) X_1 = 0, \\ \nabla_{X_1} (\nabla_{X_2} X_2) &= \nabla_{X_1} \left(-\frac{1}{x_2} X_2 \right) = -\frac{1}{x_2} \nabla_{X_1} X_2 + X_1 \left(-\frac{1}{x_2} \right) X_2 = 0, \end{aligned}$$

de donde,

$$R(X_i, X_j) X_k = \nabla_{X_j} (\nabla_{X_i} X_k) - \nabla_{X_i} (\nabla_{X_j} X_k) = 0, \quad \forall i, j, k = 1, 2.$$

Entonces,

$$R(X, Y)Z = \sum_{i,j,k,l} u^i v^j w^k R(X_i, X_j) X_k = 0.$$

Por lo tanto, \mathbb{R}_{++}^2 tiene curvatura cero.

Se observa que el ortante positivo \mathbb{R}_{++}^2 es inducido por una métrica riemanniana diagonal X^{-2} , para dos puntos cualquiera dados en \mathbb{R}_{++}^2 hay uno y sólo un camino conectándolos y que las geodésicas dadas en (1.26) está definidas para todo $t \in \mathbb{R}$, entonces la variedad $(\mathbb{R}_{++}^2, X^{-2})$ es una variedad riemanniana, simplemente conexa y completa respectivamente. Además, es de curvatura seccional cero. Por lo tanto, el espacio $(\mathbb{R}_{++}^2, X^{-2})$ es una variedad de Hadamard.

Definición 1.24 Sea M una variedad diferenciable, G una métrica riemanniana sobre M y $f : M \rightarrow \mathbb{R}$ una función diferenciable. Entonces, el gradiente de la función f sobre M en el punto q es definido como el vector tangente $\text{grad } f^T(q) \in T_q M$ tal que

$$\langle \text{grad } f(q), v \rangle = df(q)v, \quad \forall v \in T_q M,$$

donde $df(q)v$ es la derivada de f en q en la dirección de v .

Si $M = \mathbb{R}^n$, se puede caracterizar el campo vectorial gradiente como:

$$\text{grad } f^T(q) = G^{-1}(q) \nabla f^T(q),$$

donde $G^{-1}(q)$ es la matriz inversa de la matriz simétrica definida positiva $G(q)$ y $\nabla f(q)$ es la gradiente clásica de f en el punto q .

Ejemplo 1.1 Sea $(\mathbb{R}_{++}^2, X^{-2})$ una variedad riemanniana, y una función $f : \mathbb{R}_{++}^2 \rightarrow \mathbb{R}$ tal que

$$f(x) = \ln^2 x_1 + \ln^2 x_2.$$

El campo vectorial gradiente $\text{grad } f$ es dado como

$$\begin{aligned}\text{grad } f(x) &= (X^{-2})^{-1} \nabla f(x) \\ &= X^2 \nabla f(x) \\ &= \begin{pmatrix} 2x_1 \ln x_1 \\ 2x_2 \ln x_2 \end{pmatrix}.\end{aligned}\tag{1.52}$$

Definición 1.25 Sea $f : K \rightarrow \mathbb{R}$ una función de clase C^k , $k \geq 2$. La Hessiana de f , denotada por H^f , es definida como la derivada covariante del campo gradiente, esto es,

$$H^f = \frac{D}{dt}(\text{grad } f)$$

Así, la Hessiana de una función f en el punto p en la dirección de $v \in T_p K$ es dada como:

$$H_p^f(v) = \frac{D}{dt}(\text{grad } f)(p) = \nabla_v \text{grad } f(p).$$

Ejemplo 1.2 En la variedad riemanniana $(\mathbb{R}_{++}^2, X^{-2})$, mediante resultados de geometría riemanniana, la Hessiana de la función f en el punto p es dado como:

$$H_p^f = \nabla^2 f(p) + X^{-1} \nabla f(p),\tag{1.53}$$

1.3 Elementos de análisis convexo

Definición 1.26 Sea M una variedad riemanniana y $K \subset M$ un subconjunto convexo. Suponga que $\text{Front}(K) \neq \emptyset$, $x' \in \text{Front}(K)$, $s \in T_{x'} M$ y $s \neq 0$. El subespacio de $T_{x'} M$ definido por

$$S_{s,x'} := \{v \in T_{x'} M : \langle s, v \rangle = 0\},$$

es llamado el soporte del conjunto K en x' si

$$\langle s, \exp_{x'}^{-1} x \rangle \geq 0, \quad \forall x \in K.$$

Definición 1.27 Sea $K \subset M$ y sea $x \in \text{Front}(K)$. El conjunto

$$\{y \in M : \langle a, \exp_x^{-1} y \rangle = 0, \quad \forall a \in T_x M \setminus \{0\}\},$$

es llamado una hipervariiedad de soporte para K en x si $\langle a, \exp_x^{-1} y \rangle \geq 0$, para todo $y \in K$.

Definición 1.28 Un subconjunto $K \subset \mathbb{R}^n$ es llamado convexo si para todo $x, y \in K$, $t \in [0, 1]$, se tiene que $tx + (1 - t)y \in K$.

Definición 1.29 Un conjunto $K \subset \mathbb{R}^n$ es llamado fuertemente convexo si para todo $x, y \in K$, $y \neq x$, $t \in (0, 1)$, existe $r > 0$ tal que $B((1 - t)x - ty, r) \subseteq K$.

Proposición 1.3 Sean $D_i \subseteq \mathbb{R}^n$, $i \in I$, conjuntos convexos, donde I es un conjunto arbitrario. Entonces la intersección $D = \bigcap_{i \in I} D_i$ también es un conjunto convexo.

Prueba. Ver [27], Proposición 3.2.1, pp. 72. ■

Ejemplo 1.3 Sea K un conjunto convexo. El conjunto definido como

$$\mathcal{S}_y := \{x \in K : \langle F(y), y - x \rangle \geq 0\} \quad (1.54)$$

es convexo. En efecto, sea $x^1, x^2 \in \mathcal{S}_y$ y $\alpha \in [0, 1]$, entonces

$$\langle F(y), \alpha(y - x^1) \rangle \geq 0 \quad (1.55)$$

$$\langle F(y), (1 - \alpha)(y - x^2) \rangle \geq 0. \quad (1.56)$$

Sumando (1.55) y (1.56) se tiene,

$$\langle F(y), y - (\alpha x^1 + (1 - \alpha)x^2) \rangle \geq 0;$$

entonces, $\alpha x^1 + (1 - \alpha)x^2 \in \mathcal{S}_y$. Por lo tanto, \mathcal{S}_y es un conjunto convexo.

Definición 1.30

1. Sea $K \subset \mathbb{R}^n$ un conjunto convexo, $f : K \rightarrow \mathbb{R}$ es una función convexa si para cada $x, y \in K$ y $\alpha \in [0, 1]$, se tiene que $f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y)$.
2. La función f es estrictamente convexa cuando la desigualdad anterior es estricta para todo $x \neq y$ y $\alpha \in (0, 1)$.
3. La función $f : K \subset \mathbb{R}^n \rightarrow \mathbb{R}$ es pseudoconvexa si $f(x) < f(y)$ entonces $\langle \nabla f(y), x - y \rangle < 0$. La función f es pseudocóncava si $-f$ es pseudoconvexa.

Definición 1.31 El epígrafo de la función $f : K \rightarrow \mathbb{R}$ es el conjunto dado por

$$\text{epi}(f) = \{(x, c) \in K \times \mathbb{R} : f(x) \leq c\}.$$

La relación entre convexidad de conjuntos y de funciones es dada por el siguiente teorema.

Teorema 1.5 Sea $K \subset \mathbb{R}^n$ un conjunto convexo. Una función $f : K \rightarrow \mathbb{R}$ es convexa en K si y solo si el conjunto $\text{epi}(f)$ es convexo en $\mathbb{R}^n \times \mathbb{R}$.

Prueba. Ver [27], pp. 67.

Ejemplo 1.4 Sea $K \subset \mathbb{R}^n$ un conjunto cerrado y convexo. La función $\delta_K : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ definida por:

$$\delta_K(x) = \begin{cases} 0, & \text{si } x \in K \\ +\infty, & \text{si } x \notin K \end{cases}$$

es llamada la función indicadora de K .

La función δ_K es convexa si y sólo si K es un conjunto convexo.

En efecto, si $x \in K$ se tiene que $\delta_K(x) = 0$, entonces $\delta_K(x) \leq \beta$, con $\beta \in [0, +\infty)$. Por tanto, $(x, \beta) \in \text{epi}(\delta_K)$ donde $\text{epi}(\delta_K) = K \times [0, +\infty)$.

Sea δ_K una función convexa. Del Teorema (1.5), $\text{epi}(\delta_K)$ es convexo, esto es, $K \times [0, +\infty)$ es un conjunto convexo. Por lo tanto, K es un conjunto convexo.

Recíprocamente, si K es un conjunto convexo, entonces $K \times [0, +\infty) = \text{epi}(\delta_K)$ es convexo. Del Teorema (1.5), se tiene que δ_K es una función convexa.

Si $x \notin K$, entonces $\delta_K(x) = +\infty$. Así, para todo $\beta \in \mathbb{R}$, $\beta < +\infty = \delta_K(x)$, esto es, $(x, \beta) \notin \text{epi}(\delta_K)$.

El problema general de optimización en el espacio euclidiano consiste en resolver el siguiente problema:

$$\begin{cases} \min f(x) \\ \text{s.a.} \\ x \in K \subset \mathbb{R}^n, \end{cases} \quad (1.57)$$

Teorema 1.6 Considerando el problema (1.57), si $K \subset \mathbb{R}^n$ es un conjunto compacto y f es una función continua sobre K . Entonces, existe un mínimo global del problema.

Prueba. Ver [27], Teorema 1.2.1, pp.7. ▀

Teorema 1.7 Sea $K \subset \mathbb{R}^n$ un conjunto convexo y $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función diferenciable en el punto $\bar{x} \in K$. Si \bar{x} es un minimizador local del problema (1.57), entonces

$$\langle \nabla f(\bar{x}), x - \bar{x} \rangle \geq 0, \quad \forall x \in K.$$

Demostración. Ver [27], Teorema 3.1.3, pp.66. ▀

Teorema 1.8 Sea $K \subset \mathbb{R}^n$ un conjunto convexo y $f : K \rightarrow \mathbb{R}$ una función convexa sobre K . Entonces todo minimizador local del problema (1.57) es global y el conjunto de minimizadores es convexo. Además, si f es estrictamente convexa, no puede haber más de un minimizador.

Demostración. Supongamos que $\bar{x} \in K$ es un minimizador local que no es global. Entonces, existe $y \in K$ tal que

$$f(y) < f(\bar{x}). \quad (1.58)$$

Definimos $z(\alpha) = \alpha y + (1 - \alpha)\bar{x}$. Por la convexidad de K se tiene que $z(\alpha) \in K$ para todo $\alpha \in [0, 1]$. Desde que f es una función convexa para $\alpha \in (0, 1]$, y de (1.58) se tiene que

$$\begin{aligned} f(z(\alpha)) &\leq \alpha f(y) + (1 - \alpha)f(\bar{x}) \\ &= f(\bar{x}) + \alpha(f(y) - f(\bar{x})) \\ &< f(\bar{x}). \end{aligned}$$

Tomando $\alpha > 0$ suficientemente pequeño, se garantiza que el punto $z(\alpha)$ está arbitrariamente próximo al punto \bar{x} , y aún se tiene que $f(z(\alpha)) < f(\bar{x})$, para $z(\alpha) \in K$. Lo cuál es una contradicción desde que \bar{x} es un minimizador del problema. Por tanto, \bar{x} es un minimizador global.

Además, sea S el conjunto de minimizadores globales y $\bar{z} \in \mathbb{R}$ es un valor óptimo del

problema, esto es, $f(x) = \bar{z}$ para cualquier $x \in S$. Para cualquier $x \in S$ y $\bar{x} \in S$, $\alpha \in [0, 1]$, por la convexidad de f tenemos:

$$\begin{aligned} f(\alpha x + (1 - \alpha)\bar{x}) &\leq \alpha f(x) + (1 - \alpha)f(\bar{x}) \\ &= \alpha \bar{z} + (1 - \alpha)\bar{z} = \bar{z}, \end{aligned}$$

esto es, $f(\alpha x + (1 - \alpha)\bar{x}) \leq \bar{z}$. Desde que \bar{z} es un valor óptimo del problema, se tiene que $f(\alpha x + (1 - \alpha)\bar{x}) = \bar{z}$. Por lo tanto, $\alpha x + (1 - \alpha)\bar{x} \in S$. Entonces, el conjunto de minimizadores es convexo. ■

Teorema 1.9 Sea K un conjunto convexo y abierto y $f : K \rightarrow \mathbb{R}$ una función diferenciable en K . Entonces las siguientes propiedades son equivalentes:

i. La función f es convexa en K .

ii. Para todo $x, y \in K$,

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle.$$

Demostración. Sea f una función convexa sobre K . Para $x, y \in K$, $\alpha \in (0, 1]$ y definiendo $d = y - x$ tenemos

$$\begin{aligned} f(x + \alpha d) &= f(x + \alpha(y - x)) \\ &= f(\alpha y + (1 - \alpha)x) \\ &\leq \alpha f(y) + (1 - \alpha)f(x), \end{aligned}$$

de donde, $\alpha(f(y) - f(x)) \geq f(x + \alpha d) - f(x)$. Dividiendo por $\alpha > 0$,

$$f(y) - f(x) \geq \frac{f(x + \alpha d) - f(x)}{\alpha}.$$

Desde que f es una función convexa existe el límite. Tomando límite, cuando $\alpha \rightarrow 0^+$ tenemos

$$\begin{aligned} f(y) - f(x) &\geq \lim_{\alpha \rightarrow 0^+} \frac{f(x + \alpha d) - f(x)}{\alpha} \\ &= \langle \nabla f(x), d \rangle \\ &= \langle \nabla f(x), y - x \rangle. \end{aligned}$$

Entonces, $f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle$.

Recíprocamente, sea $x, y \in \text{int}(\text{dom } f)$ y para $\alpha \in [0, 1]$ tal que

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle. \quad (1.59)$$

Sustituyendo $x = x + \alpha(y - x)$ y $y = x$ en (1.59), se tiene

$$\begin{aligned} f(x) &\geq f(x + \alpha(y - x)) + \langle \nabla f(x + \alpha(y - x)), x - (x + \alpha(y - x)) \rangle \\ &= f(x + \alpha(y - x)) + \langle \nabla f(x + \alpha(y - x)), \alpha(x - y) \rangle. \end{aligned}$$

Multiplicando por $(1 - \alpha)$,

$$(1 - \alpha)f(x) \geq (1 - \alpha)f(x + \alpha(y - x)) + (1 - \alpha) \langle \nabla f(x + \alpha(y - x)), \alpha(x - y) \rangle. \quad (1.60)$$

Sustituyendo $x = x + \alpha(y - x)$ en (1.59) se tiene

$$\begin{aligned} f(y) &\geq f(x + \alpha(y - x)) + \langle \nabla f(x + \alpha(y - x)), y - (x + \alpha(y - x)) \rangle \\ &= f(x + \alpha(y - x)) + \langle \nabla f(x + \alpha(y - x)), (1 - \alpha)(y - x) \rangle. \end{aligned}$$

Multiplicando por α ,

$$\alpha f(y) \geq \alpha f(x + \alpha(y - x)) + \alpha \langle \nabla f(x + \alpha(y - x)), (1 - \alpha)(y - x) \rangle. \quad (1.61)$$

Sumando (1.60) y (1.61) tenemos

$$\alpha f(y) + (1 - \alpha)f(x) \geq f(x + \alpha(y - x)) = f(\alpha y + (1 - \alpha)x).$$

Por lo tanto, f es una función convexa. ■

Definición 1.32 Dada una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $s \in \mathbb{R}^n$ es llamado subgradiente de f en x si

$$f(y) \geq f(x) + \langle s, y - x \rangle, \quad \forall y \in \mathbb{R}^n.$$

Observación 1.3 El conjunto de todos los subgradientes de f en x , denotado por $\partial f(x)$, es llamado subdiferencial de f en x .

Ejemplo 1.5 Sea K un conjunto convexo y sea $s \in \partial\delta_K(x)$, entonces

i. Si $x \in K$, se tiene que $\delta_K(x) = 0$, y para todo $y \in \mathbb{R}^n$ se tiene que

$$\begin{aligned}\delta_K(y) &\geq \delta_K(x) + \langle s, y - x \rangle, \\ &= \langle s, y - x \rangle.\end{aligned}\tag{1.62}$$

i.1 Si $y \in K$ se tiene que $\delta_K(y) = 0$. De (1.62),

$$\langle s, y - x \rangle \leq 0.\tag{1.63}$$

i.2 Si $y \notin K$ se tiene que $\delta_K(y) = +\infty$. De (1.62),

$$\langle s, y - x \rangle \leq +\infty.\tag{1.64}$$

Por lo tanto, de (1.63) y (1.64), se tiene que

$$\partial\delta_K(x) = \{s \in \mathbb{R}^n : \langle s, y - x \rangle \leq 0, \text{ para todo } y \in K\}.$$

ii. Si $x \notin K$ se tiene que $\delta_K(x) = +\infty$, entonces

$$\delta_K(y) \geq +\infty, \text{ para todo } y \in \mathbb{R}^n,$$

lo cual es una contradicción por la definición de la función indicadora δ_K . Entonces,

$$\partial\delta_K(x) = \emptyset.$$

Por lo tanto,

$$\partial\delta_K(x) = \begin{cases} \{s \in \mathbb{R}^n : \langle s, y - x \rangle \leq 0, \forall y \in K\}, & \text{si } x \in K \\ \emptyset, & \text{si } x \notin K \end{cases}$$

Proposición 1.4 Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función propia, semicontinua inferior y convexa, entonces ∂f es monótona maximal.

Prueba. Ver [43], Teorema A, pp. 213. ■

Definición 1.33 Sea $K \subseteq \mathbb{R}^n$ un conjunto convexo y cerrado. El operador cono normal $\mathcal{N}_K \subseteq K \times \mathbb{R}^n$ es dado como

$$\mathcal{N}_K(x) := \{(x, y) : x \in K, \langle y, w - x \rangle \leq 0, \forall w \in K\}.\tag{1.65}$$

Se observa que los conjuntos $\mathcal{N}_K(x)$ y $\partial\delta_K(x)$ coinciden.

Lema 1.4 *Si K es un conjunto convexo y cerrado entonces N_K es un operador monótono maximal.*

Prueba. Sigue de la Proposición 1.4. Para mas detalle ver [43]. ■

Definición 1.34 *Sea $K \subseteq \mathbb{R}^n$, el cono dual de K , denotado por K^* , es definido como*

$$K^* = \{y \in \mathbb{R}^n : \langle x, y \rangle \geq 0, \quad \forall x \in K\}.$$

Definición 1.35 *Sea $K \subseteq \mathbb{R}^n$. El cono de recesión del conjunto K , denotado por 0^+K , es definido como el conjunto*

$$0^+K = \{x \in \mathbb{R}^n : y + \delta x \in K, \quad \forall y \in K, \delta \in [0, \infty)\},$$

donde los elementos del conjunto 0^+K son llamados direcciones de recesión.

Ahora, extenderemos algunos resultados a variedades riemannianas.

Definición 1.36 *Sea M una variedad riemanniana completa. Una función $f : M \rightarrow \mathbb{R}$ es*

- i. *convexa, si para toda geodésica $\gamma : \mathbb{R} \rightarrow M$, la composición $f \circ \gamma : \mathbb{R} \rightarrow \mathbb{R}$ es una función convexa, esto es,*

$$f \circ \gamma(ta + (1-t)b) \leq tf(\gamma(a)) + (1-t)f(\gamma(b))$$

para cualquier $a, b \in \mathbb{R}$ y $0 \leq t \leq 1$.

- ii. *es cóncava, si para toda geodésica $\gamma : \mathbb{R} \rightarrow M$, la composición $f \circ \gamma : \mathbb{R} \rightarrow \mathbb{R}$ es una función cóncava, esto es,*

$$f \circ \gamma(ta + (1-t)b) \geq tf(\gamma(a)) + (1-t)f(\gamma(b))$$

para cualquier $a, b \in \mathbb{R}$ y $0 \leq t \leq 1$.

- iii. *es afín, si es cóncava y convexa a la vez, esto es,*

$$f \circ \gamma(ta + (1-t)b) = tf(\gamma(a)) + (1-t)f(\gamma(b))$$

Observación 1.4 Sea M una variedad riemanniana completa y $f : M \rightarrow \mathbb{R}$ una función convexa. Dada una geodésica $\psi : \mathbb{R} \rightarrow M$, consideramos la función real $h : \mathbb{R} \rightarrow \mathbb{R}$ definida como

$$h(t) = f(\psi(t)).$$

Por la regla de la cadena, tenemos que el subdiferencial de h , ∂h , es dado como:

$$\partial h(t) = \{(s, \psi'(t)) : s \in \partial f(\psi(t))\} = \langle \partial f(\psi(t)), \psi'(t) \rangle.$$

Teorema 1.10 Sea M una variedad riemanniana y $f : M \rightarrow \mathbb{R}$ de clase C^1 (continuamente diferenciable). Si x^* es un punto mínimo, entonces $\text{grad } f(x^*) = \Theta$.

Demostración. Sea $v \in T_{x^*}M$ (arbitrario) y $\phi : \mathbb{R} \rightarrow M$ una geodésica tal que $\phi(0) = x^*$ y $\phi'(0) = v$. Entonces $t = 0$ es un punto de mínimo local para la función

$$h(t) = f(\phi(t)).$$

Por lo tanto, $h'(0) = 0$, esto es, $h'(0) = \langle \text{grad } f(x^*), v \rangle$. Desde que $v \in T_{x^*}M$ fué elegido arbitrariamente, se tiene que $\text{grad } f(x^*) = \Theta$. ■

Definición 1.37 Sea M una variedad de Hadamard y $f : M \rightarrow \mathbb{R}$ una función convexa. Dado $x \in M$, un vector $s \in T_x M$ es un subgradiente de f en x si

$$f(y) \geq f(x) + \langle s, \exp_x^{-1} y \rangle, \quad \forall y \in M$$

El conjunto de todos los subgradientes de f en x , denotado por $\partial f(x)$, es llamado el subdiferencial de f en x .

Teorema 1.11 Sea M una variedad de Hadamard y $f : M \rightarrow \mathbb{R}$ una función convexa, entonces $\partial f(x) \neq \emptyset$ para todo $x \in M$.

Demostración. De la convexidad de f se sigue que el grafo de la función, $\text{graf}(f)$, es un conjunto convexo de la variedad producto $M \times \mathbb{R}$. Además, la inversa de la aplicación exponencial de la variedad $M \times \mathbb{R}$ en el punto $(x, f(x)) \in \text{epi}(f)$ (arbitrario) es dado como

$$\exp_{(x, f(x))}^{-1}(p, r) = (\exp_x^{-1} p, r - f(x)).$$

Entonces, sea $S_{((s,\alpha),(x,f(x)))}$ el subespacio soporte de $\text{epi}(f)$ en $(x, f(x))$ dado como

$$\begin{aligned} 0 &\geq \langle (s, \alpha), \exp_{(x,f(x))}^{-1}(p, r) \rangle \\ &= \langle (s, \alpha), (\exp_x^{-1}p, r - f(x)) \rangle \\ &= \langle s, \exp_x^{-1}p \rangle + \alpha(r - f(x)), \end{aligned} \tag{1.66}$$

para todo $(p, r) \in \text{epi}(f)$.

Reemplazando $\tilde{p} = \exp_x s$ y $f(\tilde{p}) < \tilde{r}$ en (1.66) tenemos que $(\tilde{p}, \tilde{r}) \in \text{epi}(f)$ tal que

$$\begin{aligned} 0 &\geq \langle s, \exp_x^{-1}\tilde{p} \rangle + \alpha(\tilde{r} - f(\tilde{p})) \\ &= \langle s, s \rangle + \alpha(\tilde{r} - f(\tilde{p})) \\ &= \|s\|^2 + \alpha(\tilde{r} - f(\tilde{p})), \end{aligned}$$

esto es, $\alpha(f(\tilde{p}) - \tilde{r}) \geq \|s\|^2 \geq 0$.

Si $\alpha = 0$ se tiene que $s = 0$, lo cuál es una contradicción según Definición 1.26. Por lo tanto, $\alpha \neq 0$.

Sin pérdida de generalidad, tomamos $\alpha = -1$ en (1.66),

$$r \geq f(x) + \langle s, \exp_x^{-1}p \rangle.$$

Tomando $f(p) = r$ para todo $p \in M$ tenemos,

$$f(p) \geq f(x) + \langle s, \exp_x^{-1}p \rangle.$$

Por lo tanto; para todo $x \in M$ existe $s \in T_x M$ tal que

$$f(p) \geq f(x) + \langle s, \exp_x^{-1}p \rangle$$

para todo $p \in M$, esto es, $\partial f(x) \neq \emptyset$ para todo $x \in M$. ■

Definición 1.38 Sea M una variedad de Hadamard, un operador monótono T en M es un conjunto $T \subseteq M \times T M$ tal que si $(x, T(x)), (y, T(y)) \in T$ se tiene que

$$\langle P_{\gamma,0,1}T(y) - T(x), \exp_x^{-1}y \rangle \geq 0,$$

donde $\gamma : [0, 1] \rightarrow M$ es la (única) geodésica tal que $\gamma(0) = y$ y $\gamma(1) = x$ y $P_{\gamma,0,1}$ es el transporte paralelo sobre la curva γ .

Definimos la función real punto-conjunto $\phi : \mathbb{R} \rightrightarrows \mathbb{R}$ tal que

$$\phi_{(X,\psi)}(t) = \{\langle \psi'(t), v \rangle : v \in X(\psi(t))\},$$

donde X es un campo vectorial punto-conjunto sobre M y ψ es una geodésica en M .

Teorema 1.12 *Un campo vectorial punto-conjunto X sobre M es monótono si y solo si $\phi_{(X,\psi)}$ es monótono para toda geodésica ψ en M .*

Demostración. Supongamos que X es un campo vectorial monótono. Sea ψ una geodésica en M con $t_1 \neq t_2$ tal que $\psi(t_1) \neq \psi(t_2)$, $r_1 \in \phi_{(X,\psi)}(t_1)$, $r_2 \in \phi_{(X,\psi)}(t_2)$, $v_1 \in X(\psi(t_1))$, $v_2 \in X(\psi(t_2))$ tal que $r_1 = \langle \psi'(t_1), v_1 \rangle$ y $r_2 = \langle \psi'(t_2), v_2 \rangle$. Entonces,

$$\begin{aligned} (t_1 - t_2)(r_2 - r_1) &= (t_1 - t_2) (\langle \psi'(t_2), v_2 \rangle - \langle \psi'(t_1), v_1 \rangle) \\ &= (t_2 - t_1) (\langle P_{\psi^{-1}, t_2, t_1} \psi'(t_2), P_{\psi^{-1}, t_2, t_1} v_2 \rangle - \langle \psi'(t_1), v_1 \rangle) \\ &= (t_2 - t_1) (\langle \psi'(t_1), P_{\psi^{-1}, t_2, t_1} v_2 \rangle - \langle \psi'(t_1), v_1 \rangle) \\ &= (t_2 - t_1) \langle \psi'(t_1), P_{\psi^{-1}, t_2, t_1} v_2 - v_1 \rangle. \end{aligned}$$

Desde que X es un campo vectorial monótono, se tiene que $\langle \psi'(t_1), P_{\psi^{-1}, t_2, t_1} v_2 - v_1 \rangle \geq 0$, $t_2 > t_1$. Por lo tanto, $(t_2 - t_1)(r_2 - r_1) \geq 0$, esto es, $\phi_{(X,\psi)}$ es monótono para toda geodésica $\psi : [t_1, t_2] \rightarrow M$.

Recíprocamente, supongamos que $\phi_{(X,\psi)}$ es monótona. Sea $p, q \in M$, $u \in X(p)$, $v \in X(q)$ y sea la geodésica $\psi : [0, 1] \rightarrow M$ tal que $\psi(0) = p$ y $\psi(1) = q$, $r_1 = \langle \psi'(0), u \rangle \in \phi_{(X,\psi)}(0)$ y $r_2 = \langle \psi'(1), v \rangle \in \phi_{(X,\psi)}(1)$. Entonces,

$$\begin{aligned} \langle \psi'(0), P_{\psi,1,0} v - u \rangle &= \langle \psi'(0), P_{\psi,1,0} v \rangle - \langle \psi'(0), u \rangle \\ &= \langle P_{\psi,1,0} \psi'(1), P_{\psi,1,0} v \rangle - \langle \psi'(0), u \rangle \\ &= \langle \psi'(1), v \rangle - \langle \psi'(0), u \rangle \\ &= r_2 - r_1. \end{aligned}$$

Desde que $\phi_{(X,\psi)}$ es monótona, esto es, $(1 - 0)(r_2 - r_1) \geq 0$ se tiene que $\langle \psi'(0), P_{\psi,1,0} v - u \rangle \geq 0$. Por lo tanto, X es monótono. ■

Proposición 1.5 *Si $f : M \rightarrow \mathbb{R}$ es convexa, entonces ∂f es monótono.*

Prueba. Sea ψ una geodésica en M . Desde que f es convexa, la función real $f \circ \psi$ es convexa y $\partial(f \circ \psi)$ es monótona. De la Observación 1.4, se sigue que

$$\partial(f \circ \psi) = \{ \langle s, \psi'(t) \rangle : s \in \partial(f \circ \psi)(t) \} = \phi_{(\partial f, \psi)}(t)$$

es convexa. Del Teorema 1.12, se tiene que ∂f es monótono. ■

Observación 1.5 *Sea $K \subset M$ un subconjunto convexo y cerrado. La función indicadora sobre K es definido como:*

$$\delta_K(x) = \begin{cases} 0, & x \in K, \\ +\infty, & x \notin K. \end{cases}$$

Desde que K es un conjunto convexo, se tiene que δ_K es una función convexa.

El cono normal de $K \subset M$ en x es definido por

$$\mathcal{N}_K(x) = \partial\delta_K(x),$$

donde $\partial\delta_K(x)$ es el subdiferencial de δ_K en x . Desde que δ_K es una función convexa, su subdiferencial es monótona maximal.

Definición 1.39 *Sea (M, d) un espacio métrico completo. La sucesión $\{x^k\} \subset M$ es Fejér convergente a un conjunto no vacío $U \subset M$ si*

$$d(x^{k+1}, y) \leq d(x^k, y),$$

para todo $y \in U$.

1.4 Funciones y Distancias de Bregman en variedades Riemannianas

Definición 1.40 Sea M una variedad de Hadamard, $S \subset M$ un conjunto convexo, abierto y no vacío con \bar{S} como la cerradura topológica de S . Considere $h : \bar{S} \rightarrow \mathbb{R}$ una función convexa y diferenciable sobre S . La distancia Bregman asociado a h , denotado por D_h , es definido como una función $D_h(\cdot, \cdot) : \bar{S} \times S \rightarrow \mathbb{R}$ tal que:

$$D_h(x, y) := h(x) - h(y) - \langle \text{grad}h(y), \exp_y^{-1}x \rangle. \quad (1.67)$$

Esta clase de distancia fué introducido por Bregman en su paper original[3].

Definición 1.41 Sea $\bar{S} \subset M$. La función $h : \bar{S} \rightarrow \mathbb{R}$ es llamada función Bregman con zona S si:

- i. $S \subset M$ es un conjunto abierto.
- ii. h es finito y continuo sobre \bar{S} .
- iii. h es estrictamente convexo sobre \bar{S} .
- iv. h es continuamente diferenciable sobre S .
- v. Dado $y \in S$, $x \in \bar{S}$ y $\alpha \in \mathbb{R}$, los conjuntos de niveles parciales $L_1(\alpha, y)$ y $L_2(x, \alpha)$ de D_h , para $\alpha \in \mathbb{R}$, definidos por

$$L_1(\alpha, y) := \{x \in \bar{S} : D_h(x, y) \leq \alpha\},$$

$$L_2(x, \alpha) := \{y \in S : D_h(x, y) \leq \alpha\},$$

son acotados.

- vi. Si $\{y^k\} \subset S$ es una sucesión convergente tal que $y_k \rightarrow \bar{y}$, entonces $D_h(\bar{y}, y^k) \rightarrow 0$.
- vii. Si $\{v^k\} \subset \bar{S}$, $\{w^k\} \subset S$ son sucesiones tales que $w^k \rightarrow \bar{w}$ y $\{v^k\}$ es acotada y además $D_h(v^k, w^k) \rightarrow 0$ entonces $v^k \rightarrow \bar{w}$.

Lema 1.5 Sea M una variedad de Hadamard y $h : M \rightarrow \mathbb{R}$ una función diferenciable.

Se define $g : M \rightarrow \mathbb{R}$ tal que

$$g(x) = \langle \text{grad} h(y), \exp_y^{-1} x \rangle, \quad y \in M,$$

para $x \in M$. Entonces se cumple que:

i. $\text{grad} g(x) = P_{\gamma,0,1} \text{grad} h(y)$, donde $\gamma : [0,1] \rightarrow M$ es la curva geodésica tal que $\gamma(0) = y$ y $\gamma(1) = x$.

ii. g es una función afín in M .

Prueba. i. Sea $v \in T_x M$ (arbitrario). Consideremos la variación de la geodésica $\alpha : [0,1] \times (-\epsilon, \epsilon) \rightarrow M$ tal que

$$\alpha(t, s) = \exp_y(tu(s)),$$

donde $u(s) = \exp_y^{-1} x + sP_{\gamma,0,1}v$. Entonces,

$$\begin{aligned} (dg)_x v &= \left. \frac{d}{ds} (g(\alpha(1, s))) \right|_{s=0} \\ &= \left. \frac{d}{ds} \langle \text{grad} u(s) \rangle \right|_{s=0} \\ &= \langle \text{grad} h(y), u'(0) \rangle \\ &= \langle \text{grad} h(y), P_{\gamma,0,1}v \rangle \\ &= \langle P_{\gamma,0,1} \text{grad} h(y), v \rangle \end{aligned}$$

Por lo tanto,

$$\text{grad} g(x) = P_{\gamma,0,1} \text{grad} h(y). \quad (1.68)$$

ii. Desde que el campo vectorial $\text{grad} g(x)$ es un transporte paralelo de $\text{grad} h(y)$, el resultado sigue de i.. ■

Lema 1.6 Sea h una función de Bregman,

i. Sea $x, y \in M$ y $\gamma : [0,1] \rightarrow M$ una curva geodésica tal que $\gamma(1) = x$ y $\gamma(0) = y$, el gradiente de la función D_h es dado como

$$\text{grad} D_h(x, y) = \text{grad} h(x) - P_{\gamma,0,1} \text{grad} h(y) \quad (1.69)$$

ii. $D_h(\cdot, y)$ es estrictamente convexa sobre M para cada $y \in M$.

iii. Para todo $x, y \in M$, $D_h(x, y) \geq 0$ y $D_h(x, y) = 0$ si y solo si $x = y$.

Prueba. i. La gradiente de la función D_h es dado como

$$\begin{aligned} \text{grad } (D_h(\cdot, y))_{(x)} &= \text{grad } (h(\cdot) - h(y) - \langle \text{grad } h(y), \exp_y^{-1}(\cdot) \rangle)_{(x)} \\ &= \text{grad } h(\cdot)_{(x)} - \text{grad } \langle \text{grad } h(y), \exp_y^{-1}(\cdot) \rangle_{(x)}. \end{aligned}$$

Definimos, $g(\cdot)_{(x)} := \langle \text{grad } h(y), \exp_y^{-1} \cdot \rangle_y$. De (1.68), tenemos

$$\text{grad } D_h(x, y) = \text{grad } h(x) - P_{\gamma, 0, 1} \text{grad } h(y).$$

ii. De la definición de D_h , nosotros tenemos

$$D_h(\cdot, y) = h(\cdot) - h(y) - \langle \text{grad } h(y), \exp_y^{-1} \cdot \rangle_y.$$

De la definición 1.41 ítem iii., h es estrictamente convexo sobre M y por Lema 1.5 ítem

ii. $\langle \text{grad } h(y), \exp_y^{-1} \cdot \rangle_y$ es convexa desde que esta es una función lineal afín sobre M .

Entonces, $D_h(\cdot, y)$ es estrictamente convexa.

iii. Por la convexidad de M , para todo $x \in M$ tenemos

$$h(x) \geq h(y) + \langle \text{grad } h(y), \exp_y^{-1} x \rangle_y,$$

esto es, $0 \leq h(x) - h(y) - \langle \text{grad } h(y), \exp_y^{-1} x \rangle_y$.

Entonces, de la definición de $D_h(x, y)$, obtenemos $D_h(x, y) \geq 0$. Además, si $D_h(x, y) = 0$

entonces $h(x) - h(y) = \langle \text{grad } h(y), \exp_y^{-1} x \rangle_y$. Sea $g(x) := h(x) - h(y)$, del Lema 1.5

ítem ii., g es una función lineal afín, esto es,

$$g(\gamma(t)) = tg(x) + (1-t)g(y),$$

donde $\gamma : [0, 1] \rightarrow M$ es la curva geodésica tal que $\gamma(0) = x$ y $\gamma(1) = y$.

Entonces,

$$h(\gamma(t)) - h(y) = t(h(y) - h(y)) + (1-t)(h(x) - h(y))$$

$$h(\gamma(t)) = h(y) + (1-t)h(x) - (1-t)h(y)$$

$$h(\gamma(t)) = th(x) + (1-t)h(y),$$

lo cual es una contradicción desde que h es estrictamente convexa.

Recíprocamente, si $x = y$ entonces $h(x) - h(y) = 0$ desde que $h(x) = h(y)$.

Además, $\exp_y^{-1} x = \exp_y^{-1} y = 0$. Por lo tanto,

$$0 = h(x) - h(y) - \langle \text{grad} h(y), \exp_y^{-1} x \rangle_y = D_h(x, y). \quad \blacksquare$$

Lema 1.7 Si h es una función de Bregman con zona S , $y \in S$ y $z \in S$. Entonces, la función

$$G(x) := D_h(x, y) - D_h(x, z)$$

es afín lineal sobre \bar{S} .

Prueba. De la definición de función de Bregman dada en (1.67),

$$G(x) = h(z) - h(y) + \langle \text{grad} h(z), \exp_z^{-1} x \rangle - \langle \text{grad} h(y), \exp_y^{-1} x \rangle.$$

Del Lema 1.5 ii., $\langle \text{grad} h(z), \exp_z^{-1} x \rangle$ y $\langle \text{grad} h(y), \exp_y^{-1} x \rangle$ son funciones afines lineales.

Por lo tanto, $G(x)$ es afín lineal. \blacksquare

Lema 1.8 Dado $h : \mathbb{R}^n \rightarrow (-\infty, +\infty]$, D_h es definido en (1.67), s, w y $u \in \mathbb{R}^n$ tal que $h(s)$, $h(w)$ y $h(u)$ son finitos y h es una función diferenciable en s y w , entonces

$$D_h(u, s) = D_h(u, w) + D_h(w, s) + \langle \text{grad} D_h(w, s), \exp_w^{-1} u \rangle \quad (1.70)$$

Prueba. Sea $u, w \in M$ y sea $\gamma : [0, 1] \rightarrow M$ una curva geodésica tal que $\gamma(0) = w$ y $\gamma(1) = u$. Del Lema 1.7, $G(u)$ es una función afín lineal, esto es, convexo y cóncavo.

Entonces, de la Definición 1.36 iii., se tiene:

$$G(\gamma(t)) = tG(u) - (1-t)G(w),$$

de donde,

$$\frac{1}{t} (D_h(\gamma(t), s) - D_h(w, s)) - \frac{1}{t} (D_h(\gamma(t), w) - D_h(w, w)) = D_h(u, s) - D_h(u, w) - D_h(w, s).$$

Tomando límite, cuando $t \rightarrow 0$ y reordenando los términos se tiene el resultado. \blacksquare

Ejemplo 1.6 Sea $x = (x_1, x_2) \in \mathbb{R}_{++}^2$. Consideremos una función h definido como

$$h(x) = \sum_{i=1}^2 \ln^2 x_i + x_i^2,$$

donde, aquí por convención, $\ln 0 = 0$.

Además, el campo vectorial gradiente de h en x es dado como:

$$\text{grad } h(x) = X^2 \nabla h(x) = \begin{pmatrix} 2x_1 \ln x_1 + 2x_1^3 \\ 2x_2 \ln x_2 + 2x_2^3 \end{pmatrix}. \quad (1.71)$$

La distancia de Bregman asociado a h es dado como:

$$\begin{aligned} D_h(x, y) &= h(x) - h(y) - \langle \text{grad } h(x), \exp_y^{-1} x \rangle \\ &= \sum_{i=1}^2 \ln^2 x_i + x_i^2 - \ln^2 y_i - y_i^2 - 2x_i y_i \ln \frac{x_i}{y_i} (\ln x_i - x_i^2) \\ &= \sum_{i=1}^2 \ln \frac{x_i}{y_i} (\ln x_i y_i - 2x_i^2 y_i \ln x_i - 2x_i^3 y_i) + x_i^2 - y_i^2 \end{aligned} \quad (1.72)$$

Tomando $V = \text{grad } h(y)$ en (1.51), el transporte paralelo es dado como:

$$\begin{aligned} P_{\alpha, 0, 1} V(t) &= \left(\frac{(2y_1 \ln y_1 + 2y_1^3)}{y_1} x_1, \frac{(2y_2 \ln y_2 + 2y_2^3)}{y_2} x_2 \right) \\ &= (2 \ln y_1 x_1 + 2y_1^2 x_1, 2 \ln y_2 x_2 + 2y_2^2 x_2). \end{aligned} \quad (1.73)$$

Sustituyendo (1.71) y (1.73) en (1.69), se tiene

$$\text{grad } D_h(x, y) = \begin{pmatrix} 2x_1 \ln x_1 + 2x_1^3 - 2 \ln y_1 x_1 - 2y_1^2 x_1 \\ 2x_2 \ln x_2 + 2x_2^3 - 2 \ln y_2 x_2 - 2y_2^2 x_2 \end{pmatrix}. \quad (1.74)$$

Además, se prueba que h es una función Bregman con zona S .

En efecto,

- i. $S = \mathbb{R}_{++}^2$ es un conjunto abierto.
- ii. Por definición de la función h , se observa que h es una función continua sobre \mathbb{R}_{++}^2 , pues es la suma de funciones continuas.

Además, sea $\{x^k\} \subset \mathbb{R}_{++}^2$. De (1.49), la distancia de x^k y 0 está dado por:

$$d(x^k, y) = \ln^2 \frac{y}{x_1^k} + \ln^2 \frac{y}{x_2^k}.$$

Sea $y = \Theta$, y tomando $d(x^k, \Theta) \rightarrow 0$ tenemos

$$\ln \frac{0}{x_i^k} \rightarrow 0, \quad \text{para } i = 1, 2,$$

esto es, $x_i^k \rightarrow 0$, $i = 1, 2$.

De donde,

$$h(x^k) = \ln^2 x^k + (x^k)^2,$$

tomando límite, cuando $k \rightarrow +\infty$, se tiene que $h(0) \rightarrow 0$. Garantizando así la continuidad sobre $\bar{S} = \mathbb{R}_+^2$.

iii. De (1.53), el hessiano de la función h sobre la variedad $(\mathbb{R}_{++}^2, X^{-2})$ es dado como:

$$\begin{aligned} H_x^f &= \nabla^2 f(x_1, x_2) + X^{-1} \nabla f(x_1, x_2) \\ &= \begin{pmatrix} 2 \frac{(1-\ln x_1)}{x_1^2} + 2 & 0 \\ 0 & \frac{(1-\ln x_2)}{x_2^2} + 2 \end{pmatrix} + \begin{pmatrix} 2 \frac{1}{x_1} & 0 \\ 0 & 2 \frac{1}{x_2} \end{pmatrix} \begin{pmatrix} 2 \frac{\ln x_1}{x_1} + 2x_1 & 0 \\ 0 & 2 \frac{\ln x_2}{x_2} + 2x_2 \end{pmatrix} \\ &= \begin{pmatrix} \frac{2}{x_1^2} + 4 & 0 \\ 0 & \frac{2}{x_2^2} + 4 \end{pmatrix}. \end{aligned}$$

Desde que $x_i > 0$, $i = 1, 2$, se tiene que $2/x_i^2 + 4 > 0$. Entonces se tiene que $H_x^h(x)$ es una matriz definida positiva, esto es, $h(x)$ es una función estrictamente convexa en $S = \mathbb{R}_{++}^2$.

Analizaremos la estricta convexidad en el conjunto cerrado \mathbb{R}_+^2 . En efecto, para $k \in \mathbb{N}$ definimos:

$$\tilde{\alpha}(t) = \begin{cases} \alpha(t) = \left(\frac{1}{k}\right)^t y^{1-t}, & 0 \leq t \leq 1 - \frac{1}{k}, \\ 0 & t = 1 \end{cases}$$

donde $\tilde{\alpha}(t)$ es la geodésica extendida que une los puntos $\tilde{\alpha}(0) = y$ y $\tilde{\alpha}(1) = 0$.

Por la continuidad de $\tilde{\alpha}(t)$, es suficiente analizar en $\alpha(t)$ para $t \in (0, 1)$.

$$\begin{aligned} h(\alpha(t)) &= \ln^2(\alpha(t)) + (\alpha(t))^2 \\ &= \ln^2 \left(\left(\frac{1}{k}\right)^t y^{1-t} \right) + \left(\frac{1}{k}\right)^{2t} y^{2(1-t)} \\ &< \ln^2 \left(\frac{1}{k}\right)^t + \ln^2 y^{1-t} + \left(\frac{1}{k}\right)^{2t} y^{2(1-t)}, \end{aligned}$$

tomando límite cuando $k \rightarrow +\infty$, se tiene

$$h(\alpha(t)) < (1-t) \ln^2 y < (1-t) (\ln^2 y + y^2) < t \left(\ln^2 \frac{1}{k} + \left(\frac{1}{k} \right)^2 \right) + (1-t) (\ln^2 y + y^2).$$

Por lo tanto, h es estrictamente convexo sobre $\bar{S} = \mathbb{R}_+^2$.

iv. Desde que $h \in C^2$, se observa que h es continuamente diferenciable, esto es, $h \in C^1$.

v. Dado $y \in \mathbb{R}_{++}^2$, $x \in \mathbb{R}_+^2$ y $\alpha \in \mathbb{R}$, el conjunto de nivel parcial

$$L_1(\alpha, y) = \{x \in \mathbb{R}_+^2 : D_h(x, y) \leq \alpha\}.$$

Desde que $x \in \mathbb{R}_+^2$, se tiene que $x > 0$. Falta obtener una cota superior. Supongamos que $x \in L_1(\alpha, y)$ no sea acotado. Tomando límite en (1.72) se tiene que $D_h(x, y) \rightarrow +\infty$, cuando $x \rightarrow +\infty$, lo cuál es una contradicción pues $x \in L_1(\alpha, y)$, con $\alpha \in \mathbb{R}$ dado.

vi. Tomando $y = y^k$ y $x = \bar{y}$ y en (1.72), y desde que $y^k \rightarrow \bar{y}$, se tiene que

$$\ln \frac{\bar{y}}{y^k} \rightarrow 0.$$

Entonces, $D_h(\bar{y}, y^k) \rightarrow 0$.

vii. Tomando $x = v^k$ y $y = w^k$ en (1.72) tales que $w^k \rightarrow \bar{w}$ y $\{v^k\}$ sea acotada, se tiene lo pedido.

Por lo tanto, $h(x) = \sum_{i=1}^2 \ln^2 x_i + x_i^2$ es una función de Bregman.

Capítulo 2

El problema de desigualdad variacional (PDV)

En este capítulo, estudiaremos el problema de desigualdad variacional en \mathbb{R}^n , denotado como (PDV). Todos los resultados mostrados en este capítulo considera a T como una aplicación punto-punto y que naturalmente puede ser extendido, sin excepción alguna, a aplicaciones punto-conjunto. El capítulo se distribuye como sigue: en la primera sección daremos una formulación para el (PDV), y estudiaremos la existencia de una solución al problema (PDV), esto es, el conjunto solución del problema (PDV).

En la sección 2.2 y 2.3 presentaremos los resultados fundamentales y sus consecuencias, otros resultados pueden ser encontrados en las referencias bibliográficas citadas al final de esta monografía. En la Sección 2.3, estudiaremos los resultados de existencia usando la propiedad de monotonicidad de F . La Proposición 2.5 fué originalmente probado por Minty[35] para el caso de desigualdades variacionales monótonas en espacios de Hilbert y luego fué generalizado por Karamardian[29] para problemas pseudomonótonos. Otra clase de problemas que son permitidos para un conjunto K no acotado es utilizando una herramienta útil del análisis, la coercividad, lo cual trataremos en la Sección 2.3. La Sección 2.4, está enfocado al estudio del Método de Punto Proximal clásico. Este método fue propuesto en 1970 por Martinet [34] y desarrollado por Rockafellar [42].

Definición 2.1 Sea $K \subseteq \mathbb{R}^n$ un conjunto arbitrario y $T : K \rightarrow \mathbb{R}^n$ una aplicación, el

problema de desigualdad variacional consiste en encontrar un vector $x \in K$ tal que

$$(PDV) \quad \langle y - x, T(x) \rangle \geq 0, \quad \forall y \in K,$$

donde $K \subset \text{dom}(T)$. El conjunto de soluciones del problema (PDV) es denotado como $\text{SOL}(K, T)$.

Observación 2.1 El significado geométrico de resolver el problema (PDV) es que un punto $x \in \text{SOL}(K, T)$ si, y solo si, $T(x)$ forma un ángulo no obtuso con cada vector de la forma $y - x$, para todo $y \in K$, esto es, $-T(x)$ pertenece al cono normal de K en el punto x , o equivalentemente

$$\Theta \in T(x) + \mathcal{N}(x, K).$$

Proposición 2.1 Sea $K \subseteq \mathbb{R}^n$ un conjunto convexo y $T : K \rightarrow \mathbb{R}^n$ una aplicación pseudomonótona. Sea el conjunto \mathcal{S}_y definido en (1.54). Entonces

$$\text{SOL}(K, T) = \bigcap_{y \in K} \mathcal{S}_y.$$

Prueba. Sea $x \in \text{SOL}(K, T)$ entonces $\langle T(x), y - x \rangle \geq 0$ para todo $y \in K$. Como T es pseudomonótona, se tiene

$$\langle T(y), y - x \rangle \geq 0, \quad \forall y \in K$$

entonces, $x \in \bigcap_{y \in K} \mathcal{S}_y$.

Recíprocamente, sea $x \in \bigcap \mathcal{S}_y$, entonces existe $z \in K$ tal que $x \in \mathcal{S}_z$, esto es,

$$\langle T(z), z - x \rangle \geq 0.$$

Además, el conjunto K es convexo, entonces para $\tau \in [0, 1]$ se tiene,

$$z = \tau y + (1 - \tau)x \in K.$$

De donde, $\langle T(\tau x + (1 - \tau)z), z - x \rangle \geq 0$. Tomando $\tau \rightarrow 1$, se tiene

$$\langle T(x), y - x \rangle \geq 0, \quad \forall z \in K.$$

Por tanto, $x \in \text{SOL}(K, T)$. ■

2.1 Una formulación para el (PDV)

En el análisis computacional, algunos algoritmos para el caso irrestricto no son aplicables a optimización con restricciones, incluso si el punto inicial está contenido en el conjunto viable, las iteraciones podrían estar fuera del conjunto X . Una simple solución del punto de vista teórico es proyectar los iterados sobre el conjunto X .

Definición 2.2 Sea $K \subseteq \mathbb{R}^n$ un subconjunto convexo, cerrado y no vacío y $G \in \mathbb{R}^{n \times n}$ una matriz simétrica definida positiva. La proyección ortogonal bajo la G -norma de un punto $y \in \mathbb{R}^n$ hacia el conjunto K , denotado por $\Pi_{G,K}(y)$, es definido como la solución del siguiente programa matemático

$$\begin{cases} \min \frac{1}{2} \|y - x\|_G^2 \\ \text{s.a.} \\ x \in K. \end{cases} \quad (2.1)$$

donde $\Pi_{G,K}$ es llamado el operador proyector ortogonal bajo la G -norma sobre K .

Proposición 2.2 Sea $K \subseteq \mathbb{R}^n$ un subconjunto no vacío, cerrado y convexo y $G \in \mathbb{R}^{n \times n}$ una matriz simétrica definida positiva.

- i. Para cada vector $x \in \mathbb{R}^n$, existe un único $\Pi_{G,K}(x)$.
- ii. Para cada vector $x \in \mathbb{R}^n$, $\Pi_{G,K}(x)$ es el único vector satisfaciendo

$$\langle y - \Pi_{G,K}(x), G(\Pi_{G,K}(x) - x) \rangle \geq 0, \quad \forall y \in K.$$

- iii. La aplicación $F : \mathbb{R}^n \rightarrow K$, definida por $F(x) = \Pi_{G,K}(x)$ es continua y no expansiva, esto es, para todo $x, y \in \mathbb{R}^n$,

$$\|\Pi_{G,K}(y) - \Pi_{G,K}(x)\|_G \leq \|y - x\|_G.$$

- iv. La función distancia mínima dada por $\rho(x) = \frac{1}{2} \|x - \Pi_{G,K}(x)\|_G^2$ para $x \in \mathbb{R}^n$ es continuamente diferenciable en x , además $\nabla \rho(x) = x - \Pi_{G,K}(x)$.

Prueba. Ver [15], Teorema 1.5.5, pp. 77. ■

Proposición 2.3 Sea $K \subset \mathbb{R}^n$ un subconjunto convexo cerrado y no vacío, γ algún escalar positivo y $G \in \mathbb{R}^{n \times n}$ una matriz simétrica definida positiva. Entonces $x^* \in \text{SOL}(K, F)$, si y solo si,

$$x^* = \Pi_{G,K}(x^* - \gamma G^{-1}F(x^*));$$

esto es, si y solo si, x^* es un punto fijo de la aplicación $T_G : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definido por

$$T_G(x) = \Pi_{G,K}(x - \gamma G^{-1}F(x)). \quad (2.2)$$

Prueba. Ver [1], Proposición 5.1, pp. 267. ■

La monotonicidad de F no necesariamente implica la monotonicidad de la aplicación $T_G(x)$ definido en (2.2). En efecto, sea

$$K := \{x \in \mathbb{R}^2 : x_1 \geq 0, x_2 = 0\}$$

y,

$$F(x) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix},$$

Es fácil verificar que K es un conjunto convexo y cerrado de \mathbb{R}^2 y que para cualquier $x = (x_1, x_2)$, $y = (y_1, y_2) \in \mathbb{R}^2$ tenemos que

$$\begin{aligned} \langle F(x) - F(y), x - y \rangle &= \langle (-x_2 + y_2, x_1 - y_1), (x_1 - y_1, x_2 - y_2) \rangle \\ &= 0 \geq 0. \end{aligned}$$

Por lo tanto, F es monótona en \mathbb{R}^2 . Sea $\gamma > 0$, $y = (1, \alpha/2)^T$ y $u = (0, 0)^T$, donde se tiene que $T_G(y) = (-\alpha^2/2, 0)^T$, $T_G(u) = (0, 0)^T$ y,

$$\langle T_G(u) - T_G(y), u - y \rangle = -\alpha^2/2 < 0;$$

lo cual implica que T_G no es monótona sobre \mathbb{R}^2 . Por lo tanto, se requiere de una condición más fuerte que la monotonicidad de F para garantizar la monotonicidad de T_G . Si F es coerciva, de Gabay[18], se puede verificar que T_G es monótona. Así, podemos resolver el problema de desigualdad variacional coercivo resolviendo el sistema de ecuaciones monótonos $T_G(x) = 0$.

A manera de motivación del operador proyector sobre K , $\Pi_{G,K}$, sea un conjunto convexo cerrado K y h una función de Bregman definido en la Sección 1.4 para variedades riemannianas, se puede definir un operador proyección $P_{h,K}$, no necesariamente ortogonal, sobre K dado como

$$P_{h,K}(y) := \operatorname{argmin} \{D_h(x, y) : x \in \bar{S} \cap K\}, \quad (2.3)$$

donde $D_h(\cdot, \cdot)$ es definido en (1.67). La unicidad de esta proyección es garantizada por la convexidad estricta de h . Hay una gran cantidad de algoritmos que permanecen válidos cuando el operador proyección $P_{h,K}$ definido en (2.3) es sustituido por un operador proyección ortogonal, por ejemplo $\Pi_{G,K}$ definido en (2.1); esta sustitución es posible porque el operador D_h tiene una cierta propiedad no expansiva semejante a la proximalidad de proyecciones ortogonales.

Una formulación alternativa para el operador $P_{h,K}(y)$ sobre un subconjunto cerrado y convexo $K \subset \mathbb{R}^n$ es dado por

$$(\nabla h + \mathcal{N}_K)^{-1}(\nabla h(y)),$$

donde \mathcal{N}_K denota el operador punto conjunto cono normal sobre K definido en (1.65).

2.2 Resultados de existencia general para (PDV)

Los resultados del Teorema 2.1, Lema 2.1, Teorema 2.2 y Corolario 2.1 resultan bastantes elementales, estas herramientas fueron la base para derivar resultados de existencia para desigualdades variacionales en \mathbb{R}^n .

Teorema 2.1 *Sea $K \subseteq \mathbb{R}^n$ un subconjunto compacto, convexo y no vacío y $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ una aplicación continua. Entonces, $\operatorname{SOL}(K, F)$ es no vacío.*

Demostración. Fijemos un escalar positivo γ y una matriz simétrica definida positiva G . La función $x \mapsto x - \gamma G^{-1}F(x)$ es continua desde que F es continua y T_G , definido en (2.2), también es continua desde que el operador proyección, Proposición 2.2 iii., es

continuo. Por el Teorema 1.2, T_G tiene un punto fijo x^* . Por tanto, por Proposición 2.3, $x^* \in \text{SOL}(K, F)$. ■

Lema 2.1 *Sea $K \subseteq \mathbb{R}^n$ un subconjunto convexo no vacío y $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ una aplicación. El conjunto $\text{SOL}(K, F)$ es no vacío, si y solo si, existe un conjunto $E \subseteq \mathbb{R}^n$ tal que $x^* \in \text{SOL}(K, F)$ con $x^* \in \text{int}(E)$.*

Prueba. Sea $x^* \in \text{SOL}(K, F)$, esto es, $\langle F(x^*), y - x^* \rangle \geq 0$ para todo $y \in K$. Consideremos E una bola euclidiana cerrada y centrada en x^* , esto es, $x^* \in K \cap E$, de donde restringiendo se tiene

$$\langle F(x^*), y - x^* \rangle \geq 0, \quad \forall y \in K \cap E,$$

esto es, $x^* \in \text{SOL}(K \cap E, F)$. Recíprocamente, supongamos que existe un subconjunto E con las propiedades dadas en la hipótesis. Sea $x^* \in \text{SOL}(K \cap E, F) \cap \text{int}(E)$ y sea $y \in K$ arbitrario, $x^* \in K \cap E \cap \text{int}(E) \subset K \cap \text{int}(E)$ tal que

$$\langle F(x^*), y - x^* \rangle \geq 0, \quad \forall y \in K \cap E. \quad (2.4)$$

Además, K es convexo, esto es, para $x^* \in K \cap \text{int}(E)$, $\forall y \in K \cap E$ y $\alpha \in (0, 1]$ se tiene

$$z = x^* + \alpha(y - x^*) \in K \cap E;$$

tomando $y = z$ en (2.4) se tiene

$$\langle F(x^*), \alpha(y - x^*) \rangle \geq 0, \quad \forall y \in K \cap E.$$

Por tanto, $x^* \in \text{SOL}(K, F)$ desde que $K \cap E \subset K$. ■

Teorema 2.2 *Sea $K \subseteq \mathbb{R}^n$ un conjunto no vacío, convexo y cerrado y $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ continuo. Supongamos que existe un conjunto convexo sólido E satisfaciendo*

- i. $K \cap E$ es no vacío y compacto.
- ii. Para cada $u \in K \cap \text{Front}(E)$, existe $v \in K \cap \text{int}(E)$ tal que

$$\langle F(u), u - v \rangle \geq 0 \quad (2.5)$$

entonces, $x^* \in \text{SOL}(K, F)$ tal que $x^* \in \text{int}(E)$.

Demostración. Por la condición i. y Teorema 2.1, existe x^* tal que $x^* \in \text{SOL}(K \cap E, F)$, de donde, $x^* \in K \cap E$. Supongamos que $x^* \in K \cap \text{Front}(E)$, entonces existe $\tilde{y} \in K \cap \text{int}(E)$ tal que $\langle F(x^*), x^* - \tilde{y} \rangle \geq 0$, esto es,

$$\langle F(x^*), \tilde{y} - x^* \rangle \leq 0;$$

lo cual es una contradicción pues $x^* \in \text{SOL}(K \cap E, F)$, de donde, $x^* \in K \cap \text{int}(E)$, esto es, $x^* \in \text{int}(E)$. Finalmente, por Lema 2.1, se tiene que $x^* \in \text{SOL}(K, F)$ tal que $x^* \in \text{int}(E)$. ■

Corolario 2.1 Sea $K \subseteq \mathbb{R}^n$ un subconjunto no vacío, convexo y cerrado y $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ una aplicación continua. Supongamos que existe un subconjunto $D \subseteq K$ no vacío y acotado tal que para cada $x \in K \setminus D$, existe $y \in D$ satisfaciendo

$$\langle F(x), x - y \rangle \geq 0,$$

entonces el conjunto $\text{SOL}(K, F) \neq \emptyset$.

Prueba. Sea E una bola euclidiana cerrada, suficientemente grande tal que $D \subset \text{int}(E)$. Por hipótesis, para cada $x \in K \setminus D$, existe $y \in D \cap K$ tal que se cumple la ecuación (2.5); en particular, para $x \in K \cap \text{Front}(E)$, existe $y \in D \cap K \subset \text{int}(E) \cap K$ tal que $\langle F(x), x - y \rangle \geq 0$. Entonces, por Teorema 2.2, el conjunto $\text{SOL}(K, F)$ tiene solución. ■

Proposición 2.4 (Existencia y unicidad) Supongamos que existe un $\gamma > 0$, una matriz G simétrica definida positiva y $\alpha \in [0, 1)$ tal que la aplicación $R_G(x) = x - \gamma G^{-1}F(x)$ satisface

$$\|R_G(x) - R_G(y)\|_G \leq \alpha \|x - y\|_G, \quad \forall x, y \in K$$

entonces, el conjunto $\text{SOL}(K, F)$ es unitario.

Prueba. Por Proposición 2.2 iii., el operador proyección $\Pi_{G,K}$ es no expansiva con respecto a la norma $\|\cdot\|_G$, esto es,

$$\|T_G(x) - T_G(y)\|_g \leq \|R_G(x) - R_G(y)\|_G.$$

Por hipótesis, $\|R_G(x) - R_G(y)\|_G \leq \alpha \|x - y\|_G$, entonces

$$\|T_G(x) - T_G(y)\|_G \leq \alpha \|x - y\|_G.$$

Así, T_G es una contracción con respecto a la norma $\|\cdot\|_G$. En consecuencia, por Proposición 1.1, se tiene un único punto fijo. Por tanto, de la Proposición 2.3, el conjunto $\text{SOL}(K, F)$ es unitario. ■

2.3 Resultados de existencia usando monotonicidad

En general, una identificación del conjunto D en el Corolario 2.1 es difícil. Si la aplicación F posee algunas propiedades adicionales entonces la existencia de D es obtenida. Las clases de aplicaciones monótonas juegan un rol importante en los problemas de desigualdad variacional tanto como la clase de funciones convexas en optimización. En esta sección introduciremos a la aplicación F las propiedades de monotonicidad y sus derivados para establecer la existencia y unicidad de una solución para el problema (PDV).

Teorema 2.3 *Sea $K \subseteq \mathbb{R}^n$ un conjunto convexo y F una aplicación continua. Si F es estrictamente monótono sobre K , entonces el conjunto $\text{SOL}(K, F)$ tiene a lo más una solución.*

Demostración. Sea $\text{SOL}(K, F) \neq \emptyset$ y supongamos que existen $x_1, x_2 \in \text{SOL}(K, F)$ con $x_1 \neq x_2$, así se tiene que

$$\langle F(x_1), y - x_1 \rangle \geq 0, \quad \forall y \in K \quad (2.6)$$

y,

$$\langle F(x_2), y - x_2 \rangle \geq 0, \quad \forall y \in K; \quad (2.7)$$

sustituyendo $y = x_2$ en (2.6) y $y = x_1$ en (2.7) y sumando los términos, se tiene

$$\langle F(x_1) - F(x_2), x_2 - x_1 \rangle \geq 0,$$

o equivalentemente $\langle F(x_1) - F(x_2), x_1 - x_2 \rangle \leq 0$, lo cual contradice con la hipótesis de la monotonicidad estricta de F . Esta contradicción viene de suponer que $x_1 \neq x_2$, por tanto se tiene que $\text{SOL}(K, F)$ tiene a lo más un elemento. ■

Observación 2.2 Del Teorema 2.3, la estricta monotonidad de F no es condición necesaria para que el conjunto $\text{SOL}(K, F)$ tenga solución. Por ejemplo la función $e^x = 0$ es estrictamente monótona pero no tiene solución en \mathbb{R} .

Proposición 2.5 Sea $K \subseteq \mathbb{R}^n$ un conjunto cerrado y convexo y $F : K \rightarrow \mathbb{R}^n$ una aplicación pseudomonótona continua. Entonces $x^* \in \text{SOL}(K, F)$ si y solo si $x^* \in K$ y

$$\langle F(y), y - x^* \rangle \geq 0, \quad \forall y \in K.$$

Prueba. Sea $x^* \in \text{SOL}(K, F)$, de la Proposición 2.1 se tiene que

$$x^* \in S_y, \quad \forall y \in K.$$

De (1.54), $\langle F(y), y - x^* \rangle \geq 0$ para $x^* \in K$. Recíprocamente, Si $x^* \in K$ y $\langle F(y), y - x^* \rangle \geq 0$ para todo $y \in K$, entonces $x^* \in \bigcap_{y \in K} S_y$. Por Proposición 2.1, se tiene que $x^* \in \text{SOL}(K, F)$. ■

Observación 2.3 Se observa que ni el Teorema 2.3, ni la Proposición 2.5 garantizan la existencia de una solución del problema (PDV). Para lograr este objetivo, es necesario imponer una condición de coercividad, bajo esta propiedad se tiene el siguiente teorema.

Teorema 2.4 Sea $K \subseteq \mathbb{R}^n$ un conjunto convexo, cerrado y no vacío; $F : K \rightarrow \mathbb{R}^n$ una aplicación continua. Si F es coerciva con respecto a K , entonces el conjunto $\text{SOL}(K, F)$ es no vacío y compacto.

Demostración. Por la coercividad de F con respecto a K se tiene que existe $y \in K$ tal que

$$\lim_{x \in K, \|x\| \rightarrow +\infty} \frac{\langle F(x), x - y \rangle}{\|x\|^p} = +\infty,$$

esto es, escogiendo $A > \|x\|$ suficientemente grande tal que para todo $x \in B(0, r)$,

$$\langle F(x), x - y \rangle > A \|x\| > 0, \quad \forall x \in K. \quad (2.8)$$

Ahora, supongamos que el conjunto $\text{SOL}(K, F)$ no tenga solución, esto es, $y \in K$ se tiene:

$$\langle F(x), y - x \rangle < 0,$$

lo cuál, de (2.8), se observa que es una contradicción. Esta contradicción viene de haber supuesto que el conjunto $\text{SOL}(K, F)$ es no vacío.

Además, definimos el conjunto

$$D_y := \{x \in K : \langle F(x), x - y \rangle \leq 0\} \quad (2.9)$$

Es evidente que el conjunto D_y es cerrado. Afirmación: El conjunto D_y es acotado.

En efecto, supongamos que D_y no sea acotado, entonces para todo $R > 0$ se tiene

$$\|x\| \geq \frac{R}{A}, \quad x \in D_y.$$

De (2.8), se tiene

$$\langle F(x), x - y \rangle > A \|x\| > 0,$$

entonces $\langle F(x), x - y \rangle > 0$, lo cuál contradice la definición de D_y dada en (2.9), entonces D_y es acotado. Por lo tanto, D_y es un conjunto compacto.

Ahora probaremos que el conjunto solución $\text{SOL}(K, F)$ es un conjunto compacto. Sea $x \in \text{SOL}(K, F)$, esto es,

$$\langle F(x), y - x \rangle \geq 0, \quad \forall y \in K,$$

o equivalentemente,

$$\langle F(x), x - y \rangle \leq 0, \quad \forall y \in K,$$

entonces $x \in \bigcap_{y \in K} D_y \subseteq D_y$. Por lo tanto, $\text{SOL}(K, F) \subset D_y$. Desde que D_y es un conjunto compacto, entonces el conjunto $\text{SOL}(K, F)$ es también compacto. ■

El siguiente teorema muestra que una condición de pseudomonotonidad de F es suficiente para la existencia de una solución de un PDV. Para obtener este resultado, necesitamos la noción del cono dual K^* asociado a un conjunto arbitrario K y del cono de recesión del conjunto K , denotado como 0^+K .

Teorema 2.5 *Sea $K \subseteq \mathbb{R}^n$ un conjunto convexo y cerrado y $F : K \rightarrow \mathbb{R}^n$ una aplicación continua. Supongamos que F es una aplicación pseudomonótona sobre K , entonces*

- i. *El conjunto $\text{SOL}(K, F)$ es convexo.*

ii. Si existe un vector $x^0 \in K$ tal que $F(x^0) \in \text{int}((0^+K)^*)$, entonces el conjunto $\text{SOL}(K, F)$ es no vacío, convexo y compacto.

Demostración. i. Por la Proposición 2.1 y el Ejemplo 1.3 se tiene el resultado.

ii. Por la pseudoconvexidad de F se tiene,

$$\{x \in K : \langle F(x), x^0 - x \rangle \geq 0\} \subseteq \{x \in K : \langle F(x^0), x^0 - x \rangle \geq 0\} = \mathcal{S}_{x^0}$$

Si \mathcal{S}_{x^0} no es acotado, entonces existe $\Theta \neq x^0 \in K$ una dirección de recesión tal que $\langle F(x^0), x^0 \rangle \leq 0$. Desde que $F(x^0) \in \text{int}((0^+K)^*)$, esto es, para algun escalar $\delta > 0$, $F(x^0) - \delta x^0 \in (0^+K)^*$. Por tanto, para todo $x^0 \in (0^+K)^*$ se tiene

$$0 \leq \langle x^0, F(x^0) - \delta x^0 \rangle = \langle x^0, F(x^0) \rangle - \delta \langle x^0, x^0 \rangle \leq -\delta \langle x^0, x^0 \rangle,$$

lo cual es una contradicción. De donde, \mathcal{S}_{x^0} es acotado. Por tanto,

$$\{x \in K : \langle F(x), x^0 - x \rangle \geq 0\}$$

es un conjunto acotado. Entonces, para todo $x \in K$,

$$\text{SOL}(K, F) = \bigcap_{x \in K} \{x \in K : \langle F(x), x - x^0 \rangle \geq 0\}$$

es acotado. Por lo tanto, es compacto. ■

2.4 Método del Punto Proximal (MPP)

Consideremos el problema de optimización sin restricciones:

$$\begin{cases} \min f(x) \\ \text{s.a.} \\ x \in \mathbb{R}^n \end{cases} \quad (2.10)$$

donde f es una función convexa y limitada inferiormente.

El MPP clásico fué propuesto por Martinet[34] en 1970 como un método de regularización en el contexto de optimización convexa en espacios de Hilbert y desarrollado posteriormente por Rockafellar[42] para resolver el problema de optimización convexa

(2.10) y para obtener ceros de operadores monótonos maximales en espacios de Hilbert, respectivamente.

El método proximal para resolver (2.10) genera una sucesión $\{x^k\} \subseteq \mathbb{R}^n$, a partir de un punto arbitrario $x^0 \in \mathbb{R}^n$, dado como

$$x^{k+1} = \operatorname{argmin} \left\{ f(x) + \frac{\lambda_k}{2} \|x - x^k\|^2 : x \in \mathbb{R}^n \right\} \quad (2.11)$$

donde λ_k es un parámetro positivo y $\|\cdot\|$ es la norma euclidiana.

Se puede probar que si f es acotada inferiormente y semicontinua inferior entonces la sucesión $\{x^k\}$ generada por (2.11), está bien definida.

En el análisis de convergencia del MPP clásico, se observa que los requerimientos más importantes para obtener la convergencia de la sucesión $\{x^k\}$ a la solución es que la condición de que $\{x^k\}$, generada por (2.11), sea una sucesión Fejér convergente y que todos los puntos de acumulación pertenecen a cierto conjunto cerrado y convexo. Para más detalles del MPP clásico y su prueba de convergencia, ver [42].

Por la condición de mínimo, de (2.11) se tiene,

$$\Theta \in \partial f(x^{k+1}) + \lambda_k (x^{k+1} - x^k). \quad (2.12)$$

De la Proposición 1.4, ∂f es un operador maximal; sustituyendo $T = \partial f$ y reordenando (2.12) se tiene

$$x^k \in \left(I + \frac{1}{\lambda_k} T \right) (x^{k+1}), \quad \forall k \geq 0.$$

Por cuestiones computacionales, Rockafellar[42] consideró una sucesión de error $\{e^k\}$ que está relacionado con la sucesión $\{x^k\}$ de la siguiente ‘versión inexacta’:

$$x^k + e^{k+1} \in \left(I + \frac{1}{\lambda_k} T \right) (x^{k+1}), \quad \forall k \geq 0.$$

El modelo anterior muestra que el MPP clásico es una versión particular del método proximal para encontrar ceros de operadores monótonos maximales, el cuál genera una sucesión $\{x^k\} \subseteq \mathbb{R}^n$, a partir de un punto arbitrario $x^0 \in \mathbb{R}^n$, dado como

$$x^{k+1} \in \left(I + \frac{1}{\lambda_k} T \right)^{-1} (x^k + e^{k+1}).$$

Del Teorema 1.3, la existencia y unicidad de la sucesión $\{x^k\}$ es garantizado desde que $(I + 1/\lambda_k T)$ es biyectivo, (la existencia por la sobreyectividad y la unicidad por la inyectividad).

Trabajos anteriores a Rockafellar[42] basaron sus estudios del MPP en la ‘forma exacta’ en la cual $e^k = 0$ para todo k . Rockafellar[42] mostró que si e^k converge a 0 suficientemente rápido tal que $\sum_{k=1}^{+\infty} \|e^k\| < \infty$, entonces $x^k \rightarrow z \in \mathbb{R}^n$, donde $0 \in T(z)$. Los resultados básicos de Rockafellar sobre la convergencia del MPP para resolver (PDV) con operadores monótonos maximales fueron generalizados por Luque [32] concerniente al grado de convergencia. Podemos decir que el análisis de convergencia del MPP para operadores maximales es análogo a lo hecho en el modelo clásico con la salvedad de sustituir la condición de convexidad por la monotonicidad de T .

Capítulo 3

PDV en variedades riemanniannas (PDVV)

En esta sección, el problema (PDV) en espacios euclidianos estudiado en la sección anterior es extendido a variedades riemanniannas, este nuevo problema será denotado como (PDVV). Consideraremos $(M, \langle \cdot, \cdot \rangle)$ una variedad de Hadamard y $K \subset M$ como un conjunto convexo geodésicamente cerrado.

En la primera sección de este capítulo trataremos, a manera de motivación, de tres problemas interesantes en la ciencia e ingeniería que están estrechamente relacionados con un problema (PDV) y que han sido clásicamente estudiados en espacios euclidianos. Luego nos enfocaremos en la extensión de esos problemas a variedades riemanniannas. En el primer problema se muestra, mediante la Proposición 3.1, que la singularidad de una aplicación es condición necesaria para que el punto singular sea una solución del problema (PDV), y un resultado de extensión de \mathbb{R}^n a variedades riemanniannas de este problema es dado en el Corolario 3.1. Segundo, el problema de optimización con restricciones (p). La Proposición 3.4 muestra que una condición necesaria para que un vector sea solución del problema (PDV) es que el vector solucione el problema (p), siempre que el conjunto viable sea convexo, y la generalización a variedades riemanniannas está dado por la Proposición 3.5. Tercero, el problema del equilibrio de Nash en variedades riemanniannas, comenzaremos dando una breve

introducción de la teoría de juegos para luego plantear el objetivo del problema con sus respectivas notaciones. En la Proposición 3.7 se dá condiciones necesarias y suficientes para que una solución del equilibrio de Nash sea una solución de un problema (PDV).

En la segunda sección, se presenta los resultados de existencia para el problema (PDV) en variedades riemannianas. Teorema 3.1 es presentado cuando K es compacto. La no compacidad del conjunto K será sustituido por la condición de coercividad para algún $x_0 \in K$.

Definición 3.1 *Sea M una variedad riemanniana, $K \subset M$ y sea F un campo vectorial sobre K . Entonces el problema de desigualdad variacional en variedades riemannianas consiste en encontrar un $x \in K$ tal que*

$$(PDVV) \quad \langle F(x), \exp_x^{-1}y \rangle \geq 0, \quad \forall y \in K.$$

El conjunto de soluciones del problema (PDVV) es denotado como $SOL(K,M,F)$.

Observación 3.1 *Si $K = M$ entonces el conjunto $SOL(K,M,F)$ se reduce a $SOL(M,F)$. Si $M = \mathbb{R}^n$, $SOL(K,M,F)$ se reduce a $SOL(K,F)$.*

3.1 Motivaciones

En la actualidad existe una gran cantidad de problemas prácticos que pueden ser formulados como un problema de desigualdad variacional (PDV) tales como en sistemas de ecuaciones, optimización restringida e irrestringida, teoría de juegos, asignación del tráfico urbano, problemas de complementariedad, de punto fijo, de punto silla y los tan conocidos y estudiados problemas de equilibrio que involucran el tráfico de redes, el precio espacial, mercado oligopolístico y financiero, entre otros.

En esta sección introduciremos formulaciones de estos problemas pero ahora en variedades riemannianas. El motivo de esta nueva formulación es que existen problemas que no pueden ser resueltos en el espacio euclidiano, sin embargo son solucionables si lo consideramos inmersos (dentro) de una variedad riemanniana.

3.1.1 Problema de singularidad

Las desigualdades variacionales incluyen muchos casos especiales que son importantes en las ciencias e ingeniería, entre ellos está el más simple que es de resolver sistemas de ecuaciones no lineales.

Proposición 3.1 Si $x \in \text{SOL}(K, F)$ y $x \in \text{int}(K)$ entonces $F(x) = \Theta$.

Prueba. Como $x \in \text{int}(K)$, existe $\tau > 0$ (suficientemente pequeño) tal que $y = x + \tau(-F(x)) \in K$, entonces

$$\begin{aligned} 0 &\leq \langle F(x), y - x \rangle \\ &\leq \langle F(x), (x + \tau(-F(x))) - x \rangle \\ &\leq -\tau \langle F(x), F(x) \rangle \end{aligned}$$

de donde, $0 \geq \tau \langle F(x), F(x) \rangle = \tau \|F(x)\|^2$. Por tanto, $F(x) = 0$. ■

Si $K = \mathbb{R}^n$, tenemos la siguiente proposición:

Proposición 3.2 $x \in \text{SOL}(\mathbb{R}^n, F)$ si y solo si $F(x) = \Theta$.

Prueba. Sea $x \in \text{SOL}(\mathbb{R}^n, F)$, entonces de la Proposición 3.1 se tiene que $F(x) = \Theta$. Recíprocamente, sea $x \in K$ y $F(x) = \Theta$, es inmediato verificar que $\langle F(x), y - x \rangle \geq 0$ para todo $y \in K$, en particular para $K = \mathbb{R}^n$, esto es, $x \in \text{SOL}(\mathbb{R}^n, F)$. ■

Extendiendo la singularidad de un campo vectorial a variedades se tiene:

Corolario 3.1 Sea M una variedad de Hadamard, $K \subseteq M$ un subconjunto convexo cerrado. Si $x \in \text{SOL}(K, M, F)$ y $x \in \text{int}(K)$, entonces $F(x) = \Theta$.

Prueba. Como $x \in \text{int}(K)$, existe $\epsilon > 0$ tal que $z = \exp_x(\epsilon \exp_x^{-1}y) \in K$ para todo $y \in K$, entonces

$$\langle F(x), \exp_x^{-1}(\exp_x(\epsilon \exp_x^{-1}y)) \rangle = \langle F(x), \epsilon \exp_x^{-1}y \rangle = \epsilon \langle F(x), \exp_x^{-1}y \rangle \geq 0.$$

Desde que $x \in \text{SOL}(K, M, F)$, $\langle F(x), \epsilon \exp_x^{-1}y \rangle \geq 0, \forall y \in K$. Por lo tanto, $F(x) = \Theta$. ■

Corolario 3.2 *Sea $x \in \text{SOL}(K, M, F)$ y $x \in \text{Front}(K)$. Entonces $F(x)$ determina una hipervariiedad de soporte para K en x siempre que $F(x) \neq \Theta$.*

Prueba. Desde que $x \in \text{SOL}(K, M, F)$, $\langle F(x), \exp_x^{-1}y \rangle \geq 0$ para todo $y \in K$. Por hipótesis, $x \in \text{Front}(K)$ y por Definición 1.27 se tiene la hipervariiedad de soporte K en x como el conjunto

$$\{y \in M : \langle F(x), \exp_x^{-1}y \rangle = 0, \quad \forall F(x) \in T_x M \setminus \{0\}\} \quad \blacksquare$$

Observación 3.2 *Es fácil notar que si $x \in K$ y $F(x) = \Theta$ entonces $x \in \text{SOL}(K, M, F)$. Por tanto, si $K = M$ se tiene que $x \in \text{SOL}(M, F)$ si y solo si $F(x) = \Theta$.*

3.1.2 Problemas de optimización

En los años 90, Fukushima y otros establecieron la equivalencia entre el problema (PDV) y un problema de optimización diferenciable con condiciones suficientes de optimalidad global.

Consideremos el problema de optimización con restricciones:

$$(p_1) \begin{cases} \min f(x) \\ \text{s.a. :} \\ x \in K \end{cases}$$

donde $f \in C^1$ y $K \subseteq \mathbb{R}^n$ un subconjunto convexo y cerrado.

Proposición 3.3 *Supongamos que existe $x \in K$ que es solución del problema (p_1) , entonces $x \in \text{SOL}(K, \nabla f)$.*

Prueba. Sea $x \in K$ un mínimo del problema (p_1) . Por Lema 1.7, todo minimizador local debe satisfacer que $\langle \nabla f(x), y - x \rangle \geq 0$, para todo $y \in K$, esto es, $x \in \text{SOL}(K, \nabla f)$.

■

Proposición 3.4 *Si $f : K \rightarrow \mathbb{R}^n$ es una función convexa, la matriz jacobiana es simétrica y $x \in \text{SOL}(K, \nabla f)$, entonces x es una solución del problema de optimización (p_1) .*

Prueba. Como f es convexa, por definición se tiene

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle, \quad \forall y \in K. \quad (3.1)$$

Desde que $x \in \text{SOL}(K, \nabla f)$,

$$\langle \nabla f(x), y - x \rangle \geq 0. \quad (3.2)$$

De (3.1) y (3.2),

$$f(y) \geq f(x), \quad \text{para todo } y \in K.$$

Por lo tanto, x es solución del problema (p_1) . ■

Observación 3.3 Sea $K \subseteq \mathbb{R}^n$ un subconjunto convexo y cerrado, una función convexa $f : K \subseteq \mathbb{R}^n$ y $T = \nabla f$, el PDV con respecto a T y K es equivalente al problema de optimización (p_1) .

Considerando M una variedad de Hadamard, T un campo vectorial con $T = \text{grad } f$ y $K \subseteq M$ un subconjunto geodésicamente convexo y cerrado, extendiendo el problema de optimización a variedades se tiene los siguientes resultados:

Proposición 3.5 Supongamos que existe $x \in K \subset M$ tal que x es una solución del problema (p_1) , entonces $x \in \text{SOL}(K, M, T)$.

Prueba. Definimos, $z(t) = \exp_x(t \exp_x^{-1} y)$ para todo $t \in [0, 1]$. Por la convexidad geodésica de K , se tiene que $z(t) \in K$.

Definimos,

$$\phi(t) = f(z(t)), \quad t \in [0, 1].$$

Como $f(x) \leq f(y)$ para todo $y \in K$, entonces $f : M \rightarrow \mathbb{R}$ toma su mínimo valor cuando $t = 0$ (pues $\exp_x 0 = x$). De la condición necesaria de mínimos se tiene $\phi'(t)|_{t=0} = 0$, entonces

$$0 = \langle \text{grad } f(z(t)), z'(t) \rangle|_{t=0}. \quad (3.3)$$

Además,

$$z(0) = \exp_x 0 = x \quad (3.4)$$

y,

$$z'(t)|_{t=0} = \exp_x^{-1}y. \quad (3.5)$$

Reemplazando (3.4) y (3.5) en (3.3) se tiene

$$0 = \phi'(t)|_{t=0} = \langle \text{grad } f(x), \exp_x^{-1}y \rangle.$$

Entonces $\langle \text{grad } f(x), \exp_x^{-1}y \rangle \geq 0$ para todo $y \in K$. Por hipótesis, $\text{grad } f(x) = T(x)$.

Por lo tanto, $\langle T(x), \exp_x^{-1}y \rangle \geq 0$. ■

Proposición 3.6. *Supongamos que f es una función convexa geodésica y $x \in \text{SOL}(K, M, T)$ entonces x soluciona el problema de optimización (p_1) .*

Prueba. Desde que K es un conjunto convexo y f es una función convexa, entonces para todo $x, y \in K$ y $\gamma: [0, l] \rightarrow K$ una curva geodésica con $\gamma(0) = x$ y $\gamma(l) = y$ se tiene $f \circ \gamma$ es convexa, esto es,

$$\begin{aligned} (f \circ \gamma)(l) &\geq (f \circ \gamma)(0) + (f \circ \gamma)'(0); \\ f(\gamma(l)) &\geq f(\gamma(0)) + \langle \text{grad } f(\gamma(0)), \gamma'(0) \rangle; \\ f(y) &\geq f(x) + \langle \text{grad } f(x), \exp_x^{-1}y \rangle \\ &= f(x) + \langle T(x), \exp_x^{-1}y \rangle. \end{aligned} \quad (3.6)$$

Por hipótesis, $\langle T(x), \exp_x^{-1}y \rangle \geq 0$ para todo $y \in K$, entonces (3.6) queda como

$$f(y) \geq f(x), \quad y \in K.$$

Por tanto, x resuelve el problema de optimización (p_1) . ■

3.1.3 Problema del Equilibrio de Nash

La teoría de juegos es una teoría de toma de decisiones racionales, fundada por John Von Neumann y Morgenstern [47] en 1944. En la teoría de juegos, si hay un solo participante con un solo objetivo se trata de un problema de investigación de operaciones, si ese único participante tiene dos o más objetivos entonces es un problema de decisión multicriterio; los juegos no cooperativos son donde participan 2 o más jugadores que tienen dos o más

objetivos a diferencia de los juegos cooperativos que sólo tienen un objetivo.

Nash[37] formalmente definió el equilibrio de un juego no cooperativo, comúnmente llamado equilibrio de Nash, como un perfil de estrategias para cada uno de los jugadores, tal que la estrategia de cada jugador maximiza la utilidad esperada en contra de las estrategias dadas de los otros jugadores.

La primera aplicación clara del equilibrio de Nash, en un modelo matemático preciso, fué dado en el trabajo de Cournot[9], quién en 1838, construyó una teoría de firmas oligopolísticas que incluían competidores monopolistas y perfectos, lo cuál analizó por la metodología del equilibrio de Nash muy independiente de que esto fuese un siglo antes de Nash, así nos deberíamos preguntar si Cournot debería obtener el crédito para el equilibrio no cooperativo.

Planteamiento del problema del equilibrio de Nash: En un juego no cooperativo general, hay n jugadores cada uno de ellos tiene asociado un conjunto de estrategias (independiente de la estrategia de los otros) y una función utilidad que depende de la estrategia de todos los jugadores.

Las notaciones que usaremos para expresar el modelo son:

- $N = \{i : 1 \leq i \leq n\}$ el conjunto de n jugadores.
- $X_i = \{x_i : i \in N\} \subseteq M^{n_i}$ el conjunto de estrategias del jugador i con M^{n_i} como variedad riemanianna de dimensión n_i .
- $u_i : X \rightarrow \mathbb{R}$ con $X = \prod_{i \in N} X_i$, función de utilidad del jugador i .
- $u = (u_1, u_2, \dots, u_n)^T$.
- $x_{N \setminus \{i\}} = (x_j : j \in N, j \neq i, i \in N)$.

El objetivo es determinar para cada jugador i con la estrategia fijada x_j de su oponente, con $i \neq j$, la estrategia optimal x_i que resuelve el problema de maximización de la utilidad

en la variable x_i , esto es,

$$\begin{cases} \max u_i(x_i, x_{N \setminus \{i\}}) \\ \text{s.a. :} \\ x_i \in X_i \end{cases}$$

o equivalentemente,

$$(p_2) \begin{cases} \min -u_i(x_i, x_{N \setminus \{i\}}) \\ \text{s.a. :} \\ x_i \in X_i \end{cases}$$

Denotamos el conjunto solución del problema (p_2) como $S_i(x_{N \setminus \{i\}})$.

Un equilibrio de Nash, denotado como $NE(X, u)$, es el conjunto dado como:

$$NE(X, u) = \{x^* = \{x_i^* : i = 1, \dots, N\}, x_i^* \in S_i(x_{N \setminus \{i\}})\},$$

donde x^* es un vector de estrategias donde ningun jugador puede incrementar su utilidad unilateralmente apartándose de su estrategia de equilibrio x_i^* .

Proposición 3.7 *Sea $X_i \subseteq M$ un subconjunto no vacío convexo y cerrado, $u_i(x_i, x_{N \setminus \{i\}})$ una función continuamente diferenciable y pseudoconvexa con respecto a x_i para cada $x_{N \setminus \{i\}}$ fijado, entonces $x^* \in NE(X, u)$ si y solo si $x^* \in SOL(X, M, F)$ donde*

$$X = \prod_{i \in N} X_i$$

y, $F : \prod_{i \in N} M^{n_i} \rightarrow \prod_{i \in N} M^{n_i}$ es dado como

$$F(x^*) = (-\text{grad } u_i(x^*))_{i=1}^n.$$

Prueba. Sea $x^* \in NE(X, u)$, entonces $x_i^* \in S_i(x_{N \setminus \{i\}}), \forall i \in N$, esto es, x_i^* satisface el problema (p_2) , por Teorema 1.10 se tiene que $\text{grad}(-u_i(x_i^*, x_{N \setminus \{i\}})) = \Theta$.

Sea $\psi : [0, 1] \rightarrow M$ una curva geodésica tal que $\psi(0) = x_i^*$ y $\psi(1) = x_i$, de donde se tiene

$$\langle \text{grad}(-u_i(x_i^*, x_{N \setminus \{i\}})), \psi'(0) \rangle \geq 0, \quad \forall x_i \in X_i$$

esto es, $\langle F(x^*), \psi'(0) \rangle \geq 0$. Por lo tanto, $x^* \in SOL(K, M, F)$.

Recíprocamente, si $x^* \in SOL(X, M, F)$ entonces

$$\langle \text{grad}(-u_i(x_i^*, x_{N \setminus \{i\}})), \psi'(0) \rangle \geq 0, \quad \forall i \in N.$$

Además, $u_i(\cdot, x_{N \setminus \{i\}})$ es pseudocóncava, o equivalentemente, $-u_i(\cdot, x_{N \setminus \{i\}})$ es pseudoconvexa, entonces se tiene,

$$-u_i(x_i, x_{N \setminus \{i\}}) \geq -u_i(x_i^*, x_{N \setminus \{i\}}), \forall x_i \in X_i.$$

De donde x_i resuelve (p_2) , esto es, $x_i^* \in S_i(x_{N \setminus \{i\}})$. Por tanto, $x^* \in \text{NE}(X, u)$. ■

El problema de equilibrio de Nash generalizado (PENG) es una generalización del problema de equilibrio de Nash estandar y puede ser formulado como un problema de desigualdad cuasi-variacional (PDCV), ver [21, 40]. Sin embargo, a diferencia del PDV hay pocos métodos disponibles para resolver efectivamente un (PDCV) [40].

El PENG de N jugadores es encontrar un perfil de estrategias tal que la estrategia de cada jugador es una respuesta óptima a la estrategia del jugador rival, donde el conjunto de estrategias de cada jugador puede depender de la estrategia del jugador rival.

3.2 Resultados de existencia de (PDVV)

Definición 3.2 *Un conjunto A de una variedad riemanniana es llamado geodésico ‘en forma de estrella’ con respecto a $\sigma \in A$ si para cada $x \in A$ hay un segmento geodésico en A uniendo σ y x .*

Análogo al Teorema punto fijo de Brouwer se tiene el siguiente lema:

Lema 3.1 *Sea M una variedad riemanniana, si $K \subset M$ es un conjunto compacto, entonces K tiene la propiedad de punto fijo.*

Prueba. Definimos:

$$\hat{K} := \exp^{-1}(K)$$

tal que $\hat{K} \subset \mathbb{R}^n$. Por el Teorema 1.4, la aplicación $\exp_0 : TM \rightarrow M$ es un difeomorfismo, esto es, un homeomorfismo; de donde se tiene

$$K \cong \hat{K} \tag{3.7}$$

Afirmación: $\hat{K} \cong \bar{B}(0, 1)$; donde $\bar{B}(0, 1)$ es una bola unitaria cerrada centrada en el origen del punto.

En efecto, consideremos la función:

$$\begin{aligned} g : \hat{K} &\rightarrow \bar{B}(0, 1) \\ x &\mapsto \frac{x}{\|x\|} \end{aligned}$$

La función g es continua desde que es el producto de dos funciones continuas. De donde, $g^{-1} \circ g(x) = I_{id}(x)$, esto es, $g^{-1}(x) = x \|x\|$. Además, $g^{-1} : \bar{B}(0, 1) \rightarrow \hat{K}$ es una función continua pues es el producto de dos funciones continuas.

Entonces, $g : \hat{K} \rightarrow \bar{B}(0, 1)$ es un homeomorfismo, esto es, $\hat{K} \cong \bar{B}(0, 1)$. Del Teorema 1.1 se tiene, $h : \hat{K} \rightarrow \hat{K}$ continua, entonces h tiene un punto fijo, esto es,

$$\exists x \in \hat{K} \text{ tal que } h(x) = x. \quad (3.8)$$

Por hipótesis, K es geodésicamente convexo, entonces es un conjunto 'en forma de estrella'.

Reemplazando (3.7) en (3.8) se tiene,

$$\exists \bar{x} \in K \text{ al que } h(\bar{x}) = \bar{x},$$

esto es, h tiene un punto fijo. Entonces, K es de propiedad punto fijo. ■

Definición 3.3 La aplicación proyección Π_K sobre $K \subset M$ es caracterizado por

$$\langle \exp_{\Pi_K} z, \exp_{\Pi_K} y \rangle \geq 0, \quad \forall z \in M, \forall y \in K. \quad (3.9)$$

Teorema 3.1 Si V es un campo vectorial continuo y K es compacto, entonces el problema de desigualdad variacional en variedades (PDVV) tiene una solución.

Demostración. Consideremos la aplicación continua $\Pi_K(\exp(-V)) : K \rightarrow K$. Por hipótesis, K es compacto, entonces por Lema 3.1 tiene un punto fijo, esto es, existe $x \in K$ tal que $\Pi_K(\exp(-V))(x) = x$.

Definimos, $z = \exp_x(-V)$; de donde se tiene

$$\Pi_K z(x) = x. \quad (3.10)$$

Reemplazando (3.10) en (3.9) se tiene para todo $y \in K$,

$$\begin{aligned} 0 &\geq \left\langle \exp_{\Pi_K(\exp_x)^{-1}(-V(x))}^{-1} \exp_x(-V(x)), \exp_{\Pi_K(\exp_x)^{-1}(-V(x))}^{-1} y \right\rangle \\ &= \left\langle \exp_x^{-1} \exp_x(-V(x)), \exp_x^{-1} y \right\rangle \\ &= \left\langle -V(x), \exp_x^{-1} y \right\rangle \end{aligned}$$

Entonces, existe $x \in K$ tal que $\langle V(x), \exp_x^{-1} y \rangle \geq 0$ para todo $y \in K$, resolviendo así el problema (PDVV). ■

Observación 3.4 En el caso que el conjunto K no sea acotado, entonces definimos \sum_R como una bola geodésica cerrada tal que $K_R := K \cap \sum_R$; de donde K_R es no vacía pues $o \in K$ por hipótesis y $o \in \sum_R$ pues es una bola centrada, entonces el problema (PDVV) es encontrar $x_R \in K_R$ tal que

$$\langle V_{x_R}, \exp_{x_R}^{-1} y \rangle \geq 0, \quad y \in K_R \quad (3.11)$$

Teorema 3.2 Sea V un campo vectorial continuo sobre K , una condición necesaria y suficiente para la existencia del problema (PDVV) es que exista $R > 0$ tal que $x_R \in K_R$ definido en (3.11), satisfaciendo $d(o, x_R) < R$. Con d como función distancia geodésica.

Demostración. Si x soluciona el problema (PDVV), entonces $\langle V(x), \exp_x^{-1} y \rangle \geq 0$ para todo $y \in K$. En particular, tomando $x = x_R$ tenemos,

$$\langle V(x_R), \exp_{x_R}^{-1} y \rangle \geq 0,$$

para todo $y \in K$. En particular para $y \in K \cap \sum_R$, con $R > 0$, donde \sum_R es una bola geodésica, entonces existe $x_R \in \sum_R$ que soluciona (3.11).

Recíprocamente, supongamos que existe $R > 0$ tal que una solución $x_R \in K_R$ definido en (3.11) que satisface $d(o, x_R) < R$.

Como $x_R \in \text{int } K_R$ entonces existe $\epsilon \geq 0$ suficientemente pequeño, $y \in K$ tal que

$$\exp_{x_R}(\epsilon \exp_{x_R}^{-1} y) \in K_R.$$

Definimos, $w := \exp_{x_R}(\epsilon \exp_{x_R}^{-1} y) \in K_R$. Desde que x_R es solución para el problema (PDVV),

$$\{x_R \in K_R : \langle V(x_R), \exp_{x_R}^{-1} w \rangle \geq 0, \quad w \in K_R.\}$$

Como $K_R \subset K$, se tiene

$$\{x_R \in K : \langle V(x_R), \exp_{x_R}^{-1}y \rangle \geq 0, \quad y \in K.\}$$

Entonces, x_R es solución para el problema (PDVV). ■

De este teorema podemos deducir una condición suficiente para la existencia de solución del problema (PDVV),

Corolario 3.3 *Sea V un campo vectorial continuo sobre K satisfaciendo la condición de coercividad para algun $x_0 \in K$. Entonces, existe una solución para el problema (PDVV).*

Prueba. Usando la coercividad, por definición de límites divergentes, para todo $|A| > 0$ se tiene,

$$\langle V(x), \exp_x^{-1}x_0 \rangle + \langle V(x_0), \exp_{x_0}^{-1}x \rangle \leq Ad(x_0, x). \quad (3.12)$$

también, $d(o, x) \rightarrow +\infty$ si y solo si para todo $R > 0$, existe $d(o, x_0)$ tal que $d(o, x) > d(o, x_0)$, entonces $d(o, x) > R$, para todo $x \in K$; en particular para $d(o, x_0) < R < d(x, o)$.

Además, como V es un campo continuo, entonces existe $H > 0$ tal que $\|V(x_0)\| < H$, de donde

$$\begin{aligned} -\langle V(x_0), \exp_{x_0}^{-1}x \rangle &\leq |\langle V(x_0), \exp_{x_0}^{-1}x \rangle| \leq \|V(x_0)\| \|\exp_{x_0}^{-1}x\| - \langle V(x_0), \exp_{x_0}^{-1}x \rangle \\ &\leq \|V(x_0)\| d(x_0, x). \end{aligned} \quad (3.13)$$

Reemplazando (3.13) en (3.12) se obtiene,

$$\langle V(x), \exp_x^{-1}x_0 \rangle \leq -A d(x_0, x) + \|V(x_0)\| d(x_0, x).$$

De donde, $\langle V(x), \exp_x^{-1}x_0 \rangle \leq (\|V(x_0)\| - A) d(x_0, x)$.

En particular, para $A = H$ y por la desigualdad triangular se tiene

$$\langle V(x), \exp_x^{-1}x_0 \rangle \leq (\|V(x_0)\| - H) (d(x_0, o) - d(o, x)). \quad (3.14)$$

Désde que $d(o, x_0) < R$ entonces $x_0 \in \sum_R$ y por la hipótesis, $x_0 \in K$, de donde, $x_0 \in \sum_R \cap K \equiv K_R$; entonces $x_0 \in K_R$. Supongamos que $d(o, x) = R$, entonces de (3.14) se tiene

$$\langle V(x), \exp_x^{-1}x_0 \rangle < 0. \quad (3.15)$$

Por (3.11),

$$x_R \in K_R : \langle V(x_R), \exp_{x_R}^{-1}y \rangle \geq 0, \quad \forall y \in K_R.$$

En particular para $y = x_0$, se tiene

$$x_R \in K_R : \langle V(x_R), \exp_{x_R}^{-1}x_0 \rangle \geq 0, \quad x_0 \in K_R.$$

Así, lo cual contradice la ecuación (3.15); de donde se tiene que $d(o, x_R) \neq R$. Entonces $d(o, x_R) < 0$. ■

Capítulo 4

Método del punto proximal para resolver PDVV en variedades riemannianas

El método del Punto Proximal (MPP) es un método que fué introducido por Martinet[34] para resolver problemas de optimización convexa y que ha motivado una serie de algoritmos para resolver (PDV): algoritmos de programación convexa relacionados con métodos multiplicadores, modificaciones de la técnica de penalidad, inversas parciales de operadores monótonos así como también algoritmos de descomposición.

En este capítulo estudiaremos el MPP extendido a espacios más generales como son las variedades riemannianas.

En la primera sección de este capítulo, se presenta el MPP para problemas de optimización. En la segunda sección, se plantea el (MPP) como un método para resolver el problema de desigualdad variacional en variedades riemannianas. Este problema puede ser caracterizado como el problema de encontrar ceros de operadores monótonos maximales, los cuales fueron estudiados por Brezis[4], Bruck[5], Kamimura[30], Martinet[34] y Rockafellar[42] para espacios de Hilbert, generalizados a variedades riemannianas.

En este capítulo, usando las herramientas del análisis convexo y suponiendo algunas hipótesis naturales obtendremos la convergencia del método a una solución del problema de desigualdad variacional, que constituye nuestro principal resultado de esta tesis.

4.1 MPP en variedades riemannianas para problemas de optimización

El MPP clásico en espacios euclidianos, vistos en la Sección 2.4 del Capítulo 2, fué naturalmente extendido a variedades riemannianas por Ferreira y Oliveira[17] en el 2002, donde la distancia euclidiana en (2.11) fué sustituido por la distancia riemanniana.

Sea (M, d) una variedad riemanniana completa y una función convexa $f : M \rightarrow \mathbb{R}$. Consideremos el problema de optimización

$$\begin{cases} \min & f(x) \\ \text{s.a.} & \\ & x \in M \end{cases}$$

El MPP en variedades riemannianas genera, a partir de un punto inicial $x^0 \in M$, una sucesión $\{x^k\} \subset M$ definido como:

$$x^{k+1} := \operatorname{argmin} \{f(y) + \lambda_k \rho_{x^k}(y) : y \in M\}, \quad (4.1)$$

donde $\rho_x : M \rightarrow \mathbb{R}$ es definido por $\rho_x(y) = (1/2)d^2(x, y)$ siendo d la distancia riemanniana y λ_k una sucesión de numeros positivos.

Ferreira y Oliveira[17] mostraron que la sucesión generada por (4.1) es Fejér convergente y asumiendo la hipótesis de que la sucesión $\{\lambda_k\}$ es tal que $\sum_{k=0}^{\infty} (1/\lambda_k) = +\infty$ y que el conjunto de minimizadores U^* es no vacío entonces la sucesión $\{x^k\}$ converge a un punto de U^* .

Consideremos ahora el problema de optimización restringida:

$$\begin{cases} \min & f(x) \\ \text{s.a. :} & \\ & x \in \bar{S} \end{cases} \quad (4.2)$$

donde $S \subset M$ es un conjunto abierto y convexo y $f : M \rightarrow \mathbb{R}$ es una función limitada inferiormente.

Las distancias de Bregman, ver Sección 1.4 del Capítulo 1, fué introducido por Bregman es su paper original[3]; el término "función y distancia de Bregman" fué utilizado por primera vez por Censor y Lent[8] en 1981. Esta teoría será utilizada para resolver, en particular, el problema (4.2).

El MPP en variedades riemannianas con distancias de Bregman fué estudiada por Papa Quiroz[41], a partir de un punto inicial $x^0 \in S$, genera una sucesión $\{x^k\} \subset M$ tal que

$$x^{k+1} \in \operatorname{argmin} \{f(x) + \lambda_k D_h(x, x^k) : x \in \bar{S}\}, \quad (4.3)$$

donde h es una función de Bregman con zona S y $D_h(\cdot, \cdot)$ es una distancia de Bregman según Definición 1.41 y ecuación (1.67) respectivamente y, $\lambda_k \in (0, \bar{\lambda})$ es una sucesión de números positivos con $\lambda_k > 0$.

La convergencia del método fué garantizado para el caso irrestricto, esto es, $S = M$, ver Capítulo 4 de Papa Quiroz[41]. Considerando $S = M$ y por una condición necesaria de mínimos, y sustituyendo $T = \partial f$ se tiene de (4.3) que

$$\begin{aligned} \Theta &\in \partial (f + \lambda_k D_h(x, x^k))(x^{k+1}), \\ &= T(x^{k+1}) + \lambda_k \operatorname{grad} D_h(x^{k+1}, x^k). \end{aligned} \quad (4.4)$$

De (1.69) y reordenando se tiene:

$$P_{\gamma,0,1} \operatorname{grad} h(x^k) \in \operatorname{grad} h(x^{k+1}) + \lambda_k T(x^{k+1}). \quad (4.5)$$

donde $\gamma : [0, 1] \rightarrow M$ es una geodésica tal que $\gamma(0) = x^k$ y $\gamma(1) = x^{k+1}$.

4.2 MPP para resolver PDVV

Existen muchos métodos para resolver PDVV, entre ellos se tiene el MPP en variedades riemannianas con una aplicación monótona punto-conjunto T , que genera a partir de un punto inicial arbitrario $x^0 \in S$, una sucesión de puntos $\{x^k\} \in S$ como sigue

dado $x^k \in M$, y $e^k \in T_{x^k}M$ encontrar $x^{k+1} \in M$ y $e^{k+1} \in T_{x^{k+1}}M$ tal que

$$e^{k+1} \in T_{k+1}(x^{k+1})$$

donde $T_k(x) = T(x) + \lambda_k \text{grad } D_h(x, x^k)$ y $\lambda_k \in (0, \bar{\lambda})$ para algun $\bar{\lambda} > 0$.

Para obtener resultados de convergencia asumiremos las siguientes hipótesis:

Hipótesis A: $S \cap \text{ri}(\text{dom}(T)) \neq \emptyset$.

Hipótesis B: Tenemos las siguientes condiciones para la sucesión del error $\{e^k\}$

- i. Existe x^k .
- ii. $\sum_{k=1}^{+\infty} \|e^k\| < +\infty$ (esto implica que $\|e^k\| \rightarrow 0$) tal que $e^k \in T_{x^k}M$;
- iii. $\sum_{k=1}^{+\infty} \langle e^k, \exp_{x^k}^{-1}z \rangle$ existe y es finito.

Hipótesis C: Supongamos que cualquiera de las siguientes condiciones se cumple:

- i. $\overline{\text{dom}(T)} \subseteq S$;
- ii. $T = \partial f$, la aplicación subdiferencial de una función convexa propia $f : M \rightarrow (-\infty, +\infty]$ convexa, propia y semicontinua inferior;
- iii. T cumple con dos propiedades:
 - iii.1 Si $\{(x^k, y^k)\} \subset T$, $\{x^k\} \subset S$ y $\{x^k\}$ es convergente, entonces $\{y^k\}$ tiene un punto límite.
 - iii.2 T es paramonótona.

Lema 4.1 *Supongamos que $\{x^k\} \subset S$, $\{e^k\} \subset T_{x^k}M$ y $\{\lambda^k\} \subset \mathbb{R}$ una sucesión de números positivos que conforman la recursión (4.5), S y h satisfacen la Definición 1.40. Suponga que $\Theta \in \hat{T}(z)$, donde $\hat{T} = T + \mathcal{N}_{\bar{S}}$, entonces para todo $k \geq 0$ se tiene*

$$D_h(z, x^{k+1}) \leq D_h(z, x^k) - D_h(x^{k+1}, x^k) - \langle e^{k+1}, \exp_{x^{k+1}}^{-1} z \rangle. \quad (4.6)$$

Prueba: Sea $\varphi : [0, 1] \rightarrow M$ una curva geodésica tal que $\varphi(0) = x^{k+1}$ y $\varphi(1) = z$. Sustituyendo $s = x^k$, $w = x^{k+1}$ y $u = z$ en (1.70) tenemos

$$D_h(z, x^k) = D_h(z, x^{k+1}) + D_h(x^{k+1}, x^k) + \langle \text{grad } D_h(x^{k+1}, x^k), \exp_{x^{k+1}}^{-1} z \rangle. \quad (4.7)$$

Además,

$$\langle \text{grad } D_h(x^{k+1}, x^k), \exp_{x^{k+1}}^{-1} z \rangle = \langle \text{grad } D_h(x^{k+1}, x^k), \exp_{x^{k+1}}^{-1} z \rangle - \langle e^{k+1}, \exp_{x^{k+1}}^{-1} z \rangle + \langle e^{k+1}, \exp_{x^{k+1}}^{-1} z \rangle. \quad (4.8)$$

Sustituyendo (4.8) en (4.7),

$$D_h(z, x^k) = D_h(z, x^{k+1}) + D_h(x^{k+1}, x^k) + \langle \text{grad } D_h(x^{k+1}, x^k) - e^{k+1}, \exp_{x^{k+1}}^{-1} z \rangle + \langle e^{k+1}, \exp_{x^{k+1}}^{-1} z \rangle. \quad (4.9)$$

Desde que $\{x^k\}$ cumple con (4.5) y de (1.69), reordenando se tiene

$$\begin{aligned} T(x^{k+1}) &\ni \frac{1}{\lambda_k} (P_{\gamma,0,1} \text{grad } h(x^k) - \text{grad } h(x^{k+1}) + e^{k+1}) \\ &= \frac{1}{\lambda_k} (-\text{grad } D_h(x^{k+1}, x^k) + e^{k+1}), \end{aligned}$$

donde $\psi : [0, 1] \rightarrow M$ es una geodésica tal que $\psi(0) = x^k$ y $\psi(1) = x^{k+1}$ y desde que $x^{k+1} \in S \subset \bar{S}$ tal que $\Theta \in \mathcal{N}_{\bar{S}}(x^{k+1})$,

$$\frac{1}{\lambda_k} (-\text{grad } D_h(x^{k+1}, x^k) + e^{k+1}) \in T(x^{k+1}) + \mathcal{N}_{\bar{S}}(x^{k+1}) = \hat{T}(x^{k+1}). \quad (4.10)$$

Por hipótesis $\Theta \in \hat{T}(z)$ y de la monotonidad de \hat{T} se tiene que

$$\begin{aligned} 0 &\leq \langle \exp_z^{-1} x^{k+1}, P_{\varphi,0,1} \hat{T}(x^{k+1}) - \Theta \rangle, \\ &= \langle -P_{\varphi,0,1} \exp_{x^{k+1}}^{-1} z, P_{\varphi,0,1} \hat{T}(x^{k+1}) \rangle. \end{aligned}$$

con $\varphi : [0, 1] \rightarrow M$ tal que $\varphi(0) = x^{k+1}$ y $\varphi(1) = z$.

Por la isometría lineal del transporte paralelo $P_{\varphi,0,1}$,

$$\left\langle \exp_{x^{k+1}}^{-1} z, -\hat{T}(x^{k+1}) \right\rangle \geq 0. \quad (4.11)$$

Sustituyendo (4.10) en (4.11) se tiene:

$$0 \leq \left\langle \exp_{x^{k+1}}^{-1} z, \frac{1}{\lambda_k} (\text{grad } D_h(x^{k+1}, x^k) - e^{k+1}) \right\rangle,$$

Sustituyendo (4.12) en (4.9) y reordenando se tiene:

$$D_h(z, x^{k+1}) \leq D_h(z, x^k) - D_h(x^{k+1}, x^k) - \langle e^{k+1}, \exp_{x^{k+1}}^{-1} z \rangle. \quad \blacksquare$$

Lema 4.2 *Supongamos que se cumple la hipótesis del Lema 4.1, h es una función Bregman y que la serie*

$$E(z) = \sum_{k=0}^{+\infty} \langle e^{k+1}, \exp_{x^{k+1}}^{-1} z \rangle$$

existe y es finito, entonces $\{x^k\}$ es acotado y $D_h(x^{k+1}, x^k) \rightarrow 0$.

Prueba. Definimos $E_l(z) = \sum_{k=0}^{l-1} \langle e^{k+1}, \exp_{x^{k+1}}^{-1} z \rangle$. Por hipótesis, $E(z)$ existe y es finito, esto es, $E_l(z) \rightarrow E(z)$ cuando $l \rightarrow +\infty$, donde $E(z) < \infty$ (por hipótesis).

De la ecuación (4.6), para alguna iteración finita con $l \geq 0$,

$$D_h(z, x^l) \leq D_h(z, x^0) - \sum_{k=0}^{l-1} D_h(x^{k+1}, x^k) - \sum_{k=0}^{l-1} \langle e^{k+1}, \exp_{x^{k+1}}^{-1} z \rangle. \quad (4.12)$$

Desde que $D_h(x^{k+1}, x^k) \geq 0$, se tiene

$$\begin{aligned} D_h(z, x^l) &\leq D_h(z, x^0) - \sum_{k=0}^{l-1} \langle e^{k+1}, \exp_{x^{k+1}}^{-1} z \rangle, \\ &= D_h(z, x^0) - E_l(z) < \infty. \end{aligned} \quad (4.13)$$

Entonces, $D_h(z, x^l) < \infty$ con $l \geq 0$ desde que $D_h(z, x^0)$ y $E(z)$ son finitos.

De la Definición 1.41 v., del conjunto de nivel de D_h para $\alpha \in \mathbb{R}_0^+$ y algún $z \in \bar{S}$ y de la ecuación (4.13) se tiene que $L_2(z, \alpha) := \{x^l : D_h(z, x^l) \leq \alpha, l \geq 0\}$ es acotado. Por tanto, $\{x^k\}$ es una sucesión acotada.

Desde que $D_h(z, x^l) \geq 0$ para todo $l \geq 0$, (4.12) queda como

$$0 \leq D_h(z, x^0) - \sum_{k=0}^{l-1} D_h(x^{k+1}, x^k) - E_l(z).$$

Tomando límite cuando $l \rightarrow +\infty$,

$$\lim_{j \rightarrow +\infty} \sum_{k=0}^{l-1} D_h(x^{k+1}, x^k) \leq D_h(z, x^0) - E(z).$$

Desde que $D_h(x^{k+1}, x^k) \geq 0$ para todo $k \geq 0$, se tiene que $\sum_{k=0}^{+\infty} D_h(x^{k+1}, x^k) < +\infty$. Por lo tanto, $D_h(x^{k+1}, x^k) \rightarrow 0$. ■

Teorema 4.1 *Suponga que la sucesión $\{x^k\} \subset S$, $\{e^k\} \subset T_{x^k}M$ conforman (4.5) donde $\{\lambda_k\}$ es una sucesión de positivos escalares, T es un operador maximal sobre M y h es una función Bregman con zona S , donde $S \cap \text{ri}(\text{dom}T) \neq \emptyset$. Si $\hat{T} = T + N_{\hat{S}}$ tiene algún cero, se cumple la hipótesis B sobre $\{e^k\}$ y se cumple también cualquiera de las hipótesis C, i., ii. ó iii.. Entonces, $\{x^k\} \rightarrow \bar{x}$ con $\Theta \in \hat{T}(\bar{x})$.*

Demostración. Consideremos Z el conjunto de ceros de \hat{T} y sea $z \in Z$. De la hipótesis B iii., se tiene que $\sum_{k=0}^{+\infty} \langle e^{k+1}, \exp_{x^{k+1}}^{-1} z \rangle$ existe y es finito. Por un resultado de análisis se tiene que

$$\langle e^{k+1}, \exp_{x^{k+1}}^{-1} z \rangle \rightarrow 0. \quad (4.14)$$

Del Lema 4.2, $D_h(x^{k+1}, x^k) \rightarrow 0$ y $\{x^k\}$ es acotado, esto es, existe una subsucesión $\{x^{k_j}\} \subset \{x^k\}$ tal que $x^{k_j} \rightarrow \bar{x}$ y $D_h(x^{k_j+1}, x^{k_j}) \rightarrow 0$. Por la consistencia de convergencia de la función D_h dado por Definición 1.41 vii., $x^{k_j+1} \rightarrow \bar{x}$ tal que $\bar{x} \in \bar{S}$ para todo $k \geq 0$. Definimos

$$y^{k+1} := \frac{1}{\lambda_k} (-\text{grad} D_h(x^{k+1}, x^k) + e^{k+1}). \quad (4.15)$$

De (4.10),

$$y^{k+1} \in \hat{T}(x^{k+1}) \quad \text{para todo } k \geq 0. \quad (4.16)$$

Del Lema 4.1 y desde que $D_h(x^{k+1}, x^k) \geq 0$,

$$\begin{aligned} 0 \leq D_h(z, x^{k+1}) &\leq D_h(z, x^k) - D_h(x^{k+1}, x^k) - \langle e^{k+1}, \exp_{x^{k+1}}^{-1} z \rangle, \\ &\leq D_h(z, x^k) - \langle e^{k+1}, \exp_{x^{k+1}}^{-1} z \rangle. \end{aligned}$$

Del Lema 1.1, para $a_{k+1} = D_h(z, x^k)$, $\psi_k = \langle e^{k+1}, \exp_{x^{k+1}}^{-1} z \rangle$ y $\{D_h(z, x^k)\}$ acotado inferiormente se tiene que $\{D_h(z, x^k)\}$ es convergente, esto es,

$$D_h(z, x^k) \rightarrow d(z) \quad \text{donde } d(z) \in [0, +\infty). \quad (4.17)$$

Sustituyendo (4.15) en (4.9) y reordenando,

$$\langle -\lambda_k y^{k+1}, \exp_{x^{k+1}}^{-1} z \rangle = D_h(z, x^k) - D_h(x^{k+1}, x^k) - D_h(z, x^{k+1}) - \langle e^{k+1}, \exp_{x^{k+1}}^{-1} z \rangle. \quad (4.18)$$

Tomando límite, cuando $k \rightarrow +\infty$, de (4.14) y (4.17) se tiene que

$$\langle y^{k+1}, \exp_{x^{k+1}}^{-1} z \rangle \rightarrow 0. \quad (4.19)$$

Ahora, analizaremos el teorema cumpliéndose cualquiera de las condiciones dadas en la Hipótesis C:

i. Primero, supongamos que se cumple la Hipótesis C i.. De (4.15) y de (1.69),

$$\begin{aligned} y^{k_j+1} &:= \frac{1}{\lambda_k} (-\text{grad } D_h(x^{k_j+1}, x^{k_j}) + e^{k_j+1}); \\ &= \frac{1}{\lambda_{k_j}} (P_{\gamma,1,0} \text{grad } h(x^{k_j}) - \text{grad } h(x^{k_j+1}) + e^{k_j+1}). \end{aligned}$$

Por la continuidad de $\text{grad } h(\cdot)$ en $\bar{x} \in S$, $e^k \rightarrow \Theta$ y $\lambda_k \in (0, \bar{\lambda})$ para algun $\bar{\lambda} \in \mathbb{R}$, se tiene que $\lim_{j \rightarrow +\infty} y^{k_j+1} = \Theta$.

T es un conjunto cerrado en $M \times TM$ por la maximalidad de T , esto es, $(x^{k_j+1}, y^{k_j+1}) \rightarrow (\bar{x}, \Theta) \in T$. Como $\bar{x} \in \bar{S}$ entonces $(\bar{x}, \Theta) \in \hat{T}$. Por tanto, $\Theta \in \hat{T}(\bar{x})$.

ii. Segundo, supongamos si se cumple la Hipótesis C ii. De (4.16) para $T = \partial f$ se tiene

$$f(z) \geq f(x^{k_j+1}) + \langle y^{k_j+1}, \exp_{x^{k_j+1}}^{-1} z \rangle$$

donde $\varphi : [0, 1] \rightarrow M$ es una geodésica con $\varphi(0) = x^{k+1}$ y $\varphi(1) = z$.

Tomando límite cuando $j \rightarrow +\infty$,

$$f(z) \geq \liminf_{j \rightarrow +\infty} f(x^{k_j+1}) + \liminf_{j \rightarrow +\infty} \langle y^{k_j+1}, \exp_{x^{k_j+1}}^{-1} z \rangle. \quad (4.20)$$

Sustituyendo (4.19) en (4.20) para una subsucesión $\{x^{k_j}\} \subset \{x^k\}$, tal que $x^{k_j} \rightarrow \bar{x}$,

$$f(z) \geq \liminf_{j \rightarrow +\infty} f(x^{k_j+1}). \quad (4.21)$$

Desde que $f : M \rightarrow (-\infty, +\infty)$ es una función semicontinua inferior, de (1.1) se tiene que,

$$f(\bar{x}) \leq \liminf_{j \rightarrow +\infty} f(x^{k_j}). \quad (4.22)$$

De (4.21) y (4.22) se tiene que $f(z) \geq f(\bar{x})$. Por Hipótesis, $\Theta \in \hat{T}(z)$, esto es,

$$z \in \operatorname{argmin}\{f(x) : x \in \bar{S}\} \in Z.$$

Por tanto, $\bar{x} \in Z$, esto es, $\Theta \in T(\bar{x})$.

- iii. Tercero, supongamos que se cumple la Hipótesis C iii. Sea $\{(x^{k_j+1}, y^{k_j+1})\} \subset T$, $\{x^{k_j+1}\} \subset S$ tal que $x^{k_j+1} \rightarrow \bar{x}$. Entonces, existe \bar{y} tal que $y^{k_j+1} \rightarrow \bar{y}$. Por la maximalidad de T , $(\bar{x}, \bar{y}) \in T$.

De (4.19) para las subsucesiones $\{y^{k_j}\} \subset \{y^k\}$ y $\{x^{k_j}\} \subset \{x^k\}$ tal que $x^{k_j} \rightarrow \bar{x}$ se tiene:

$$\langle y^{k_j+1}, \exp_{x^{k_j+1}}^{-1} z \rangle \rightarrow 0.$$

Tomando límite cuando $j \rightarrow +\infty$,

$$\langle \bar{y}, \exp_{\bar{x}}^{-1} z \rangle = 0. \quad (4.23)$$

Por Hipótesis, T es paramonótona y como resultado del Corolario 1.1 se tiene que \hat{T} es paramonótono. Sea $(\bar{x}, \bar{y}) \in \hat{T}$, $(z, \Theta) \in \hat{T}$. De (4.23) y de la paramonotonidad de \hat{T} se tiene que $\Theta \in \hat{T}(\bar{x})$.

Bajo cualquier hipótesis C, i, ii. ó iii., se ha demostrado que $\Theta \in \hat{T}(\bar{x})$, esto es, $\bar{x} \in Z$ (conjunto de ceros de \hat{T}). Reemplazando $z = \bar{x}$ en (4.17) se tiene,

$$D_h(\bar{x}, x^k) \rightarrow d(\bar{x}); \quad \text{donde } d(\bar{x}) \in [0, +\infty). \quad (4.24)$$

Desde que $\{x^{k_j}\} \subset \{x^k\}$, la función D_h , por definición, para $\bar{x} \in \bar{S}$ es dado como

$$D_h(\bar{x}, x^{k_j}) = h(\bar{x}) - h(x^{k_j}) - \langle \operatorname{grad} h(x^{k_j}), \exp_{x^{k_j}}^{-1} \bar{x} \rangle.$$

Por la continuidad de la función h y $\operatorname{grad} h$ en \bar{x} y tomando límite cuando $j \rightarrow +\infty$ se tiene que $D_h(\bar{x}, x^{k_j}) \rightarrow 0$. De (4.24), $d(\bar{x}) = 0$ entonces $D_h(\bar{x}, x^k) \rightarrow 0$.

Sea \tilde{x} un punto límite de $\{x^{k'_j}\}$ tal que $\{x^{k'_j}\} \subset \{x^k\}$, esto es, $x^{k'_j} \rightarrow \tilde{x}$. De donde, $D_h(\bar{x}, x^{k'_j}) \rightarrow 0$. Por la propiedad de la convergencia consistente según Definición 1.41 vii., $x^{k'_j} \rightarrow \bar{x}$, entonces $\bar{x} = \tilde{x}$ desde que $x^{k'_j} \rightarrow \tilde{x}$. Por lo tanto $x^k \rightarrow \bar{x} \in Z$. ■

Capítulo 5

Aplicaciones del MPP para resolver PDVV

Sea el conjunto

$$\mathbb{R}_{++}^2 := \{x = (x_1, x_2) : x_1 > 0, x_2 > 0\}.$$

Consideremos el siguiente problema de optimización:

$$(\text{Prob}) \begin{cases} \min \ln^2 x_1 + \ln^2 x_2 \\ \text{s.a. :} \\ x \in \mathbb{R}_{++}^2, \end{cases}$$

Se observa que la función objetivo definida arriba, $f(x) = \ln^2 x_1 + \ln^2 x_2$ con $x \in \mathbb{R}_{++}^2$, no es convexa en el sentido clásico (con la métrica euclidiana).

En efecto, existen $(x_1, x_2) = (e, e)$, $(y_1, y_2) = (2e, 2e) \in \mathbb{R}_{++}^2$ tal que

$$5.733494 = f(2e, 2e) < f(e, e) + \langle \text{grad } f(e, e), (e, e) \rangle = 29.5562244.$$

Entonces, por el Teorema 1.9 se tiene que $f(x_1, x_2)$ no es convexo en el sentido clásico. Así, con el objetivo de que $f(x_1, x_2)$ sea convexo en \mathbb{R}_{++}^2 se hace un cambio de métrica apropiada. Para tal fin, consideremos la métrica de Dikin, X^{-2} .

De (1.53), el hessiano de la función f sobre la variedad $(\mathbb{R}_{++}^2, X^{-2})$ es dado como:

$$\begin{aligned}
 H_x^f &= \nabla^2 f(x_1, x_2) + X^{-1} \nabla f(x_1, x_2) \\
 &= \begin{pmatrix} 2 \frac{(1 - \ln x_1)}{x_1^2} & 0 \\ 0 & 2 \frac{(1 - \ln x_2)}{x_2^2} \end{pmatrix} + \begin{pmatrix} 2 \frac{1}{x_1} & 0 \\ 0 & 2 \frac{1}{x_2} \end{pmatrix} \begin{pmatrix} 2 \frac{\ln x_1}{x_1} & 0 \\ 0 & 2 \frac{\ln x_2}{x_2} \end{pmatrix} \\
 &= \begin{pmatrix} 2 \frac{(1 - \ln x_1)}{x_1^2} & 0 \\ 0 & 2 \frac{(1 - \ln x_2)}{x_2^2} \end{pmatrix} + \begin{pmatrix} 2 \frac{\ln x_1}{x_1^2} & 0 \\ 0 & 2 \frac{\ln x_2}{x_2^2} \end{pmatrix} \\
 &= \begin{pmatrix} \frac{2}{x_1^2} & 0 \\ 0 & \frac{2}{x_2^2} \end{pmatrix}.
 \end{aligned}$$

Desde que $x_1, x_2 > 0$ se tiene que $2/x_1^2 > 0$, $2/x_2^2 > 0$ y $(2/x_1^2)(2/x_2^2) = 4/(x_1 x_2)^2 > 0$.

Por lo tanto, $f(x)$ es estrictamente convexa en la variedad riemanniana $(\mathbb{R}_{++}^2, X^{-2})$.

5.1 Problema

Encontrar $\bar{x} \in (\mathbb{R}_{++}^2, X^{-2})$ tal que

$$\text{grad } f(x) = \Theta,$$

5.2 Método

El MPP, para resolver PDV, consiste en generar, a partir de un punto inicial arbitrario

$x^0 \in \mathbb{R}_{++}^2$, una sucesión de $\{x^k\}$ como sigue:

Dado x^k , encontrar x^{k+1} tal que

$$\Theta \in T_k(x^{k+1}),$$

donde $T_k(\cdot) = T(\cdot) + \lambda_k \text{grad } D_h(\cdot, x^k)$, $T(\cdot) = \text{grad } f(\cdot)$ y $\lambda_k \in (0, \bar{\lambda})$ con $\lambda > 0$.

De (1.74) y (1.52), se tiene

$$T_k(x^{k+1}) = T(x^{k+1}) + \lambda_k \text{grad } D_h(x^{k+1}, x^k)$$

$$= \begin{pmatrix} 2x_1^{k+1} \ln x_1^{k+1} \\ 2x_2^{k+1} \ln x_2^{k+1} \end{pmatrix} + \lambda_k \begin{pmatrix} 2(x_1^{k+1} \ln x_1^{k+1} + (x_1^{k+1})^3 - x_1^{k+1} \ln x_1^k - (x_1^k)^2 x_1^{k+1}) \\ 2(x_2^{k+1} \ln x_2^{k+1} + (x_2^{k+1})^3 - x_2^{k+1} \ln x_2^k - (x_2^k)^2 x_2^{k+1}) \end{pmatrix}$$

5.3 Implementación

En esta sección, para problemas de implementación, usaremos el software 'Matlab'. Considerando un punto inicial arbitrario $x^0 = (x_1^0, x_2^0) = (2, 3) \in \mathbb{R}_{++}^2$, la implementación del algoritmo del método de Punto Proximal, para resolver el problema (Prob), es dado como:

```
% Algoritmo para resolver el problema, en particular para lambda = 1/ M,   M ∈ ℕ
x=[2;3];
epsilon=0.000001;
Nmax=100;
Mmax=100;
lambda_k=1./M;
fprintf('\ n Punto Inicial: %f',x)
M=1;
% % while M <= Mmax
fprintf('\ n Iteración N: %d',M)
T=[2*x(1)*log(x(1)); 2*x(2)*log(x(2))];
normT=sqrt(T(1,1)^2+T(2,1)^2);
% % %
if normT >= epsilon
k=1;
z=[1.5;2.8];
```

```

%% % Continuación del algoritmo

%% % % INICIO DEL SUBPROBLEMA % % % %

while k <= Nmax,

TK=[2*z(1)*log(z(1))+2*lambda_k*(z(1)*log(z(1))+z(1)^3-z(1)*(log(x(1))+
x(1)^2));2*z(2)*log(z(2))+2*lambda_k*(z(2)*log(z(2))+z(2)^3-z(2)*(log(x(2))+x(2)^2))]

DT1=2*log(z(1))+2+ 2*lambda_k*(log(z(1))+1+3*z(1)^2-log(x(1))-(x(1))^2);
DT2=2*log(z(2))+2+ 2*lambda_k*(log(z(2))+1+3*z(2)^2-log(x(2))-(x(2))^2);

d = [-TK(1,1)/DT1; -TK(2,1)/DT2];

normTK=sqrt(TK(1,1)^2 + TK(2,1)^2);

% if normTK >= epsilon

z=z+ d;

k=k+1;

else

x=z;

fprintf('\ n criterio de error del subproblema satisfecho =: %f',normTK)

fprintf('\ n =====')

fprintf('\ n el punto x1 aproximado es: %f',x(1))

fprintf('\ n el punto x2 aproximado es: %f',x(2))

fprintf('\ n =====')

fprintf('\ n numero de iteraciones del subproblema es: %d',k-1)

k=1000000;

end

end

%% % % % FIN DEL SUBPROBLEMA % % % %

```

```

M=M+1;

else
fprintf('\ n criterio de error satisfecho Normgradiente=: %f',normT)
fprintf('\ n ')
fprintf('\ n =====')
fprintf('\ n el punto x1 aproximado es: %f',x(1))
fprintf('\ n el punto x2 aproximado es: %f',x(2))
fprintf('\ n =====')
fprintf('\ n ')
fprintf('\ n numero de iteraciones es: %d',M-1)

M =1000000;
end

end

% % %
% % % % % if M > 999999
fprintf('\ n resolvimos el problema aproximadamente : %f')
else
fprintf('\ n : numero de iteraciones excedido a %d',M-1)
end

% % % % FIN % % % %

```

En el Capítulo anterior, estudiamos los resultados de convergencia del método de punto proximal con distancias de Bregman. A continuación, mostraremos resultados numéricos computacionales para la resolución de nuestro problema (Prob). Notaremos que la rapidez de convergencia o el costo computacional va a depender de la elección apropiada del parámetro λ_k , esto es, si elegimos un parámetro λ_k tal que $\lambda_k \rightarrow 0$ se encontrará el punto óptimo para realizar la menor cantidad de pasos iterativos, lo que implica un menor costo computacional.

a). Para $\lambda_k = 1$.

```
Criterio de error satisfecho Normagradiente=: 0.000001
=====
el punto  $x_1$  aproximado es: 1.000000
el punto  $x_2$  aproximado es: 1.000000
=====

número de iteraciones es: 60
resolvimos el problema aproximadamente :
```

b). Para $\lambda_k = 1/k$.

```
Criterio de error satisfecho Normagradiente=: 0.000000
=====
el punto  $x_1$  aproximado es: 1.000000
el punto  $x_2$  aproximado es: 1.000000
=====

número de iteraciones es: 15
resolvimos el problema aproximadamente .
```

c). Para $\lambda_k = k$.

Criterio de error satisfecho Normagradiente=: 0.000000

=====

el punto x_1 aproximado es: 1.161588

el punto x_2 aproximado es: 1.689004

=====

número de iteraciones excedido a 2000

Materiales y Métodos.

Para la elaboración de este trabajo de tesis, se recopiló una gran variedad de papers relacionados a temas de desigualdad variacional, método de punto proximal, operadores monótonos, distancias de Bregman, sobretodo se trabajó con los artículos de Iusem[26], Ferreira y Oliveira[17], Németh[39], Facchinei y Pang[15] Vol.I, Burachik y Iusem[6] y los materiales indispensables para las herramientas de geometría riemanniana se obtuvieron de Udriste[46], Papa Quiroz[41], Ferreira[16] y Do Carmo[13]. Los demás materiales, que no dejan de ser importantes, pueden ser encontrados en la bibliografía. Mucho de todos ellos fueron proporcionados por el asesor.

Fué de gran utilidad para la digitación de este trabajo el editor científico Latex2X, recomendado para este tipo de textos científicos por su elegancia y la gran variedad de recursos y simbologías matemáticas que posee.

Resultados.

- Se presenta un estudio básico de conceptos de operadores monótonos punto-conjunto en espacios euclidianos. Es importante resaltar la importancia de la Proposición 1.2, donde se dá las hipótesis necesarias para que la suma de dos operadores monótonos maximales resulte otro operador monótono maximal.
- Mediante herramientas de geometría riemanniana, se prueba que el plano Hiperbólico H y el ortante positivo \mathbb{R}_{++}^2 son variedades de Hadamard, esto es, una variedad riemanniana, simplemente conexa, completa y de curvatura seccional no positiva.
- Se presenta condiciones necesarias y suficientes sobre resultados de existencia general de una solución al problema de desigualdad variacional en espacios euclidianos, se estudia también estos resultados usando la propiedad de monotonicidad y sus variantes como son la pseudomonotonicidad, paramonotonicidad, entre otros; así también, la extensión de los resultados de existencia a variedades riemannianas.
- Se presenta 3 problemas particulares de PDV de gran importancia en las ciencias e ingenierías, tales como son el problema de singularidad, el problema de optimización y el equilibrio de Nash, todas ellas en variedades riemannianas.
- Se plantea el problema de encontrar ceros de operadores monótonos maximales, estudiados por Brezis[4], Bruck[5], Kamimura[30], Martinet[34] y Rockafellar[42] para espacios de Hilbert, generalizados a variedades riemannianas.
- Se plantea un problema no convexo en el sentido clásico y que mediante una elección adecuada de métrica, el problema no convexo se vuelve convexo en el mismo espacio.

Discusiones.

- Hay una gran variedad de artículos que trabajan con distancias generalizadas (por ejemplo de Bregman) en espacios de Hilbert, por ejemplo Iusem, Burachik, Kiwiel y Censor. Cabe resaltar que la bibliografía existente que trata sobre distancias de Bregman en variedades riemannianas es muy escasa porque el estudio de la geometría riemanniana relacionado con la optimización es reciente.
- Cuando la función objetivo es convexa, el estudio de aquella función encaja en la teoría del análisis convexo. La dificultad radica cuando la función no es convexa, por tal razón se hace necesario tener un conocimiento analítico de geometría riemanniana, pues con una elección adecuada de métrica riemanniana la función se vuelve convexa.

Conclusiones.

- La extensión de conceptos y técnicas en espacios euclidianos a espacios topológicos convexos tales como las variedades riemannianas es casi siempre natural y no trivial. Muchos problemas tales como ecuaciones no lineales, problemas de minimización y PDV en espacios de Hilbert son extendidos a variedades riemannianas.
- Introducimos una nueva clase de distancias generalizadas, llamadas distancias de Bregman, lo cuál juega un rol importante en la convergencia y el análisis de error del algoritmo iterativo para resolver PDV.
- Las propiedades de coercividad han jugado un importante rol en la teoría de existencia de (PDV), ver [36, 24, 25]. Hartman y Stampacchia[23] y Moré[36] muestran que si la aplicación es coerciva, entonces PDV tiene por lo menos una solución.
- Es indispensable hacer un estudio analítico de resultados de existencia en variedades riemannianas. Al igual que en espacios euclidianos, la no compacidad del conjunto K es sustituido por la propiedad de coercividad para algún punto de K .
- Con ciertas condiciones impuestas para la sucesión del error $\{e^k\}$, y considerando la maximalidad de la suma de dos operadores monótonos maximales bajo cierta condición es que se logra el principal resultado de convergencia del MPP, para resolver problemas de desigualdad variacional, en variedades riemannianas, Teorema 4.1.
- Es importante anotar que un dominio de las herramientas de geometría riemanniana puede ayudar a facilitar la resolución de problemas en optimización.

Bibliografia

- [1] BERTSEKAS T. - *Parallel and distributed computation: Numerical methods*. SIAM J. OPTIM. Vol 8 Nro1 pp. 197 - 216. February 1998.
- [2] BISHOP, R.L., O'NEILL, B. - *Manifolds of Negative Curvature*. Trans. Amer. math. Soc., 145, pp1-49. 1969.
- [3] BREGMAN, L.M.- *The Relaxation Method of Finding the Common Point of Convex Sets and its Application to the Solution of Problems in Convex Programming*. USSR Comput. Math. and Math. Phys. 7,3 200-217. 1967.
- [4] BREZIS,H., LIONS,P.L.-*Produits infinis de resolvents*.J.Math.29:329-345.Isr.1970.
- [5] BRUCK, R.E.- *A Strongly Convergent Iterative Solution of $0 \in U(x)$ for a Maximal Monotone Operator in Hilbert Space*. J. Math. Anal Appl.48,114-126.1974.
- [6] BURACHIK R., IUSEM A.- *A Generalized Proximal Point Algorithm for the Variational Inequality Problem in a Hilbert Space*. SIAM J. OPTIM. 8, 197-216. 1998.
- [7] BURACHIK R. , SCHEIMBERG S. - *A Proximal Point Method for the Variational Inequality Problem in Banach Spaces*. SIAM J. Control Optim. 39, 1633-1649. 2000.
- [8] CENSOR, Y., LENT, A.- *An Iterative Row-Action Method for Interval Convex Programming*. J. Optim. Theory Appl. 34, 3 321-353. 1981.
- [9] COURNOT, A.- *Resecherches Sur les Principes Mathématiques de la Théorie des Richesses*. Paris.1838. (English) N.Y.MacMillan. 1927.
- [10] DUGUNDJI, J.- *Topology* McGraw-Hill Companies. 1966.

- [11] DA CRUZ NETO, J.X.- *Métodos Geodésicos na Programação Matemática*. Tesis de doutorado. Programa de engenharia de sist. e comp. COOPE, UFRJ.-Brasil. 1995.
- [12] DAFERMOS S.C. - *Traffic Equilibrium and Variational Inequalities*. Transportation Science, 14, 42-54. 1980.
- [13] DO CARMO, M.P. - *Riemannian Geometry*. Boston, Birkhäuser. 1992.
- [14] DUNFORD N., SCHWARTZ J. - *Linear Operators Part I: General Theory*. A serie of text and monographs. Volume VII. Interscience Publishers, INC., N.Y. 1957.
- [15] FACCHINEI F., PANG J.S.- *Finite dimensional variational inequalities and nonlinear complementarity problems*. VOL.1-2, Springer-Verlag. 2003.
- [16] FERREIRA O. P.- *Programação Matemática em Variedades Riemannianas: Algoritmos Subgradiente e Ponto Proximal*. Tesis de Doutorado.R.J.-Brasil.1997.
- [17] FERREIRA O.P., OLIVEIRA P. *Proximal Point Algorithm on Riemannian Manifolds*. Optimization. 2002, Vol. 51(2), pp 257-270.
- [18] GABAY, D.- *Applications of the Method of Multipliers to Variational Inequalities*. Augmented Lagrangian Methods,229-332. Amstersam.1983.
- [19] GUILLEMIN V., POLLACK ALAN *Differential topology*. Massachusetts Institute of Technology. Prentice Hall. Inc., Englewood Chiffs, N.Y., 1974.
- [20] GÜLER O. *On the convergence of the Proximal Point Algotihm for Convex Minimization*. SIAM J. Control and Optimization, 29(2), 403-419. 1991.
- [21] HARKER, P.T. - *Generalized Nash Games and Quasi-variational Inequality*. Eur. J. Oper. Res. 54, 81-94. 1991.
- [22] HARKER P.T., PANG J.S. - *Finite dimensional variational inequalities and nonlinear complementarity problems*. SIAM J. Control Optim., 29 (1991), pp. 403-419.
- [23] HARTMAN P., STAMPACCHIA G. - *One Some Nonlinear Elliptic Differential Functional Equations*. Acta Math, 115, 153-188. 1966.

- [24] ISAC, G., BULAVASKI, V., KALASHNIKOV, V. *Exceptional families, Topological Degree and Complementarity Problems*. J. Global Optim., 10(2) pp. 207. 1997.
- [25] ISAC, G., OBUCHOWSKA, W.T. *Functions without Exceptional Family of Elements and Complementarity Problems*. J. Optim. Theory Appl., 99(1):147. 1998.
- [26] IUSEM A. N. *Métodos de ponto proximal em otimização, 20 Colóquio Brasileiro de Matemática*. IMPA.
- [27] IZMAILOV A.F., SOLODOV M.V. *Otimização - Volume 1. Condições de Otimalidade, Elementos de Análise Convexa e de Dualidade*. Rio de Janeiro. IMPA, 2005.
- [28] JOFRÉ A., ROCAFELLAR T. *Variational Inequalities and Economic Equilibrium*. Center for Math. Modelling and Dept. of Math. Engineering, Univ. of Chile.
- [29] KARAMARDIAN S. - *An Existence Theorem for the Complementarity Problem*. Journal of Optimization Theory and Applications 18, pp. 445-454. 1976.
- [30] KAMIMURA, S., TAKAHASHI, W.- *Approximating solutions of Maximal Monotone Operators in Hilbert Spaces*. J. Approx. Theory 13, 226-240. 2000.
- [31] KINDERLEHRER D., STAMPACCHIA G. - *An Introduction to Variational Inequalities and Their Applications*. Academic Press, New York, 1980.
- [32] LUQUE, F.J. - *Asymptotic Convergence Analysis of the Proximal Point Algorithms*. SIAM J. Control and Optimization 22, 277-293. 1984.
- [33] MARCOTTE P. - *Inéquations variationnelles: motivation, algorithmes de résolution et quelques applications*. Cours donné à Zinal, Suisse en 1997.
- [34] MARTINET B. - *Régularisation d'inéquations variationnelles par approximations successives*. Revue Française d'Informatique et de Recherche Operationell, 154-159. 1970.
- [35] MINTY, B.- *Monotone nonlinear Operators in Hilbert Space*. Duke Mathematical Journal 29, 341-346. 1978.

- [36] MORÉ, J.J.- *Coercivity Conditions in Nonlinear Complementarity Problems*. SIAM. Rev., 16(1):1. 1974.
- [37] NASH J.F. - *Equilibrium Points in n-person Games* - Proceedings of the National Academy of Sciences 36, pp.48-49.1950.
- [38] NASH J.F. - *Non-cooperative Games* - Annals of Mathematics 54, pp.286-295. 1951.
- [39] NEMETH. Z. - *Variational Inequalities on Hadamard Manifolds*. Nonlinear Anal. 52. pp. 1491-1498. 2003.
- [40] PANG, J.S., FUKUSHIMA, M. - *Quasi-variational Inequalities, Generalized Nash Equilibria, and Multi-leader-follower Games*. Comput. Manag. Sc.2, 21-56. 2005.
- [41] PAPA QUIROZ, E.A.-*Algumas aplicações da geometria Riemanniana à otimização*. Tesis de doutorado. Prog. de Eng. de Sist. e Comp., COOPE, UFRJ. R.J.-Brasil. 2007.
- [42] ROCKAFELLAR R.T. - *Monotone Operators and the Proximal Point Algorithm*. SIAM J. Control Optim., 14(1976), pp. 877-898.
- [43] ROCKAFELLAR R.T. - *On the Maximal Monotonicity of Subdifferential Mapping*. Pacific Journal of Mathematics. Vol 33, Nro 1. 1970.
- [44] SAKAI, T. - *Riemannian Geometry, Translations of Mathematical Monographs*. American Mathematical Soc., 149, Providence, R.I. 1996.
- [45] SMITH M.J. - *The Existence, Uniqueness and Stability of Traffic Equilibria*. Transportation Research 13B, 295-304.1979.
- [46] UDRISTE, C. - *Convex Functions and Optimization Methods on Riemannian Manifolds*. Mathematics and Its Applications. Kluwer Academic Publishers. 1994.
- [47] VON NEWMANN J., MORGENSTERN O. - *Theory of Games and Economic Behavior*. Princeton University Press. Princeton, N.J. 1944.