Universidad Nacional Mayor de San Marcos
Universidad del Perú. Decana de América

Dirección General de Estudios de Posgrado
Facultad de Ciencias Matemáticas
Unidad de Posgrado

Método del punto proximal y su aplicación a modelos económicos

TESIS

Para optar el Grado Académico de Magíster en Investigación de Operaciones y Sistemas

AUTOR

Lucy Haydee DE LA CRUZ CUADROS

ASESOR

Esther BERGER VIDAL

Lima, Perú

2022



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

https://creativecommons.org/licenses/by-nc-sa/4.0/

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

De la Cruz, L. (2022). *Método del punto proximal y su aplicación a modelos económicos.* [Tesis de maestría, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Matemáticas, Unidad de Posgrado]. Repositorio institucional Cybertesis UNMSM.

Metadatos complementarios

Datos de autor	
Nombres y apellidos	LUCY HAYDEE DE LA CRUZ CUADROS
Tipo de documento de identidad	DNI
Número de documento de identidad	06554824
URL de ORCID	https://orcid.org/0000-0002-9788-1531
Datos de asesor	
Nombres y apellidos	ESTHER BERGER VIDAL
Tipo de documento de identidad	DNI
Número de documento de identidad	08766040
URL de ORCID	https://orcid.org/0000-0001-5282-6793
Datos del jurado	
Presidente del jurado	
Nombres y apellidos	CARMELA CATALINA VELÁSQUEZ PINO.
Tipo de documento	DNI
Número de documento de identidad	09078631
Miembro del jurado 1	
Nombres y apellidos	JOSÉ CARLOS ORÉ LUJÁN
Tipo de documento	DNI
Número de documento de identidad	06119405
Miembro del jurado 2	
Nombres y apellidos	LUIS ERNESTO TORRES GUARDIA
Tipo de documento	Cédula de Identidad
Número de documento de identidad	BR /12.166.374-4

Datos de investigación	
Línea de investigación	No aplica.
Grupo de investigación	No aplica.
Agencia de financiamiento	Sin financiamiento. .
Ubicación geográfica de la investigación	País: Perú Departamento: Lima Provincia: Lima Distrito: San Juan de Lurigancho Latitud: -12.16979 Longitud: -76.96222
Año o rango de años en que se realizó la investigación	2012 –2022
URL de disciplinas OCDE	Matemáticas aplicadas https://purl.org/pe-repo/ocde/ford#1.01.02

UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
Universidad del Perú, DECANA DE AMÉRICA

FACULTAD DE CIENCIAS MATEMÁTICAS
VICEDECANATO DE INVESTIGACIÓN Y POSGRADO
UNIDAD DE POSGRADO

ACTA DE SUSTENTACIÓN VIRTUAL DE TESIS PARA OPTAR EL GRADO ACADÉMICO DE MAGÍSTER EN INVESTIGACIÓN DE OPERACIONES Y SISTEMAS

Siendo las, 14:14 horas del día Viernes cuatro de marzo de dos mil veintidos, en la sala virtual http://meet.google.com/ixp-qjme-mha, el Jurado de Tesis conformado por los siguientes docentes:

PRESIDENTA	:	Mg. Carmela Catalina Velásquez Pino
MIEMBRO	:	Mg. José Carlos Oré Luján
MIEMBRO EXTERNO	:	Dr. Luis Ernesto Torres Guardia
MIEMBRO ASESORA	:	Dra. Esther Berger Vidal

se reunieron para la sustentación de la tesis titulada: **«MÉTODO DEL PUNTO PROXIMAL Y SU APLICACIÓN A MODELOS ECONÓMICOS»** presentada por la Señora Bachiller Lucy Haydee De la Cruz Cuadros, egresada de la Maestría en Investigación de Operaciones y Sistemas.

Concluida la exposición, los miembros del Jurado de Tesis procedieron a formular sus preguntas que fueron absueltas por la graduanda; acto seguido se procedió a la evaluación correspondiente, según tabla adjunta, habiendo obtenido la Señora Bachiller **Lucy Haydee De la Cruz Cuadros**, el calificativo de …MUY BUENO… **(18)**.

Habiendo sido aprobada la sustentación de la Tesis, el Jurado Evaluador recomienda para que el Consejo de Facultad apruebe el otorgamiento del Grado Académico de **Magíster en Investigación de Operaciones y Sistemas** a la **Bachiller Lucy Haydee De la Cruz Cuadros**.

Siendo las 15:37 horas, se levantó la sesión, firmando para constancia la presente Acta:

Mg. Carmela Catalina Velásquez Pino
PRESIDENTA

Mg. José Carlos Oré Luján
MIEMBRO

Dr. Luis Ernesto Torres Guardia
MIEMBRO EXTERNO

Dra. Esther Berger Vidal
MIEMBRO ASESORA

Agradecimientos

Al concluir con este trabajo, quiero agradecer a Dios por ser mi fuente de luz en este camino y por bendecirme con la materialización de este sueño anhelado. A mis hijas por su apoyo incondicional y mi compañero de vida Fernando por darme el tiempo necesario para realizarme profesionalmente y por alentarme hasta el final.

A mi asesora de tesis, Dra. Esther Berger Vidal por confiar en mí, por guiarme acertadamente en este arduo proceso, por sus consejos, por su paciencia, por motivarme a seguir creciendo profesionalmente.

RESUMEN

Método del Punto Proximal y su Aplicación a Modelos Económicos

Se considera el siguiente problema de optimización:

$$(P)\ min\{f(\boldsymbol{x}) : \boldsymbol{x} \geq \mathbf{0}\}$$

donde $f: M \to \mathbb{R} \cup \{+\infty\}$ es una función propia semicontinua inferior y $\boldsymbol{x} \geq 0$ denota que todas las componentes del vector $\boldsymbol{x}$ son no negativas.

Se analiza la convergencia de un método de optimización denominado *método de punto proximal* para resolver el problema dado para el caso no convexo y aplica los resultados de convergencia a la solución de problemas económicos en particular de microeconomía. El método considera un punto inicial $x_0 \in \mathbb{R}^n{}_{++}$ (arbitrario) y genera una sucesión de puntos $\{x^k\}$ dada por:

$$0 \in \hat{\partial}(f(\cdot) + (\lambda_k/2)d(.,x^{k-1}))(x^k),$$

donde $\hat{\partial}$ es el subdiferencial de Fréchet, donde d es un tipo de distancia que asegura que las iteraciones se mantengan el interior de las restricciones y λ_k es un parámetro positivo para cada k, donde λ_k cumple una función de regularización, esto es, no puede ser tan pequeño ni tan grande. Con respecto a la aplicación del método a la microeconomía, presentamos un ejemplo cuando la función objetivo es una función de producción e implementamos el algoritmo para resolver el problema. También presentamos adicionalmente algunos experimentos numéricos de la implementación del algoritmo que dan confiabilidad a los resultados teóricos.

Palabras Claves: método de punto proximal, funciones no convexas, convergencia global, función de producción.

ABSTRACT

PROXIMAL POINT METHOD AND ITS APPLICATION TO ECONOMIC MODEL

LUCY HAYDEE DE LA CRUZ CUADROS

ENERO - 2022

Adviser: Dra. Esther Berger Vidal

Obtained Degree: Master in Operational Research.

We consider the following optimization problem:

$$(P)\ min\{f(x) : x \geq 0\}$$

where $f: M \to \mathbb{R} \cup \{+\infty\}$ is a proper lower semicontinuous function and $x \geq 0$ denotes that all the components of the vector x are nonnegative.

La present thesis analise the convergence of an optimization method called proximal point method to solve the above problem for the non convex case and applied the convergence results to the solution of economic problems in particular in microeconomy. The method considers the initial point $x_0 \in \mathbb{R}^n{}_{++}$(arbitrary) and generates a sequence of points $\{x^k\}$ given by:

$$0 \in \hat{\partial}(f(\cdot) + (\lambda_k/2)d(.,x^{k-1}))(x^k),$$

where $\hat{\partial}$ is the Fréchet subdifferential, d is a certain distance which assure the iterations restaint in the interior of the constraints and λ_k is a positive parameter for each k.

With respect to the applications of the proposed method in microeconomy, we present an example when the objective function is a production function and we implemented the algorithm to solve the problem. Also, we additionaly present some numerical experiments to obtain confiability of the theory results.

Keywords: Proximal point method, quasiconvex functions, global convergence, production function.

ÍNDICE GENERAL

ÍNDICE GENERAL vi
ÍNDICE DE TABLAS viii
ÍNDICE DE FIGURAS. ix
ÍNDICE DE ANEXO. x
1.INTRODUCCIÓN……..1
1.1 Situación Problemática. 1
1.2 Formulación del Problema. 2
1.2.1 Problema General. 2
1.2.2 Problemas específicos. 2
1.3 Justificación Teórica. 3
1.4 Justificación Práctica. 3
1.5 Objetivos. 3
1.5.1 Objetivo general. 3
1.5.2 Objetivos específicos. 4
1.6 Hipótesis. 4
2. MARCO TEÓRICO……..5
2.1 Marco Filosófico o Epistemológico de la Investigación. 5
2.2 Antedecedentes de la Investigación. 5
2.3 Bases Teóricas. . . . . .. .11
2.3.1 Existencia de Puntos de Mínimo Global. 11
2.3.2 Condiciones de Optimalidad de Problemas sin Restricciones. 15
2.3.2.1 Algoritmo general de descenso. 18
2.3.2.2 Método de Newton 21
2.3.3 Condiciones de Optimalidad para Restricciones de Igualdad. 24
2.3.4 Condiciones de optimalidad para Restricciones de Desigualdad 29
2.3.5 Elementos de Convexidad. 33
2.3.6 Optimización Convexa. 38

2.3.7 Optimización Cuasi-convexa. 40

2.3.8 Métodos de optimización restringida. 43

2.4. El Método de los Multiplicadores como Caso Particular del APP. 48

2.5 Distancia Generalizada. 54

2.6 Variables utilizadas en el modelo de Cobb Douglas 58

3. METODOLOGÍA. 61

3.1 El universo. 61

3.2 Técnicas de recopilación de datos. 61

3.3 Técnicas descriptivas para la contrastación o demostración de las hipótesis. 61

4 RESULTADOS Y DISCUSIÓN. 63

4.1 Análisis, interpretación y discusión de resultados. 63

4.2 Pruebas de hipótesis. 63

4.3 Presentación de resultados. 64

4.4 Algoritmo Proximal para Minimización. 66

4.4.1 El Algoritmo. . . . . …66

4.4.2 Resultados de Convergencia . . . . . …67

4.5 Aplicaciones a Economía 68

4.5.1 Aplicación del Método a la Teoría Económica de la decisión. . . . . …68

4.5.2. Pruebas econométricas para obtener el valor de α y β 75

4.5.3. Resultados para la función Cobb-Douglas….. 75

4.6 Experimentos Computacionales . . . . . …81

5 IMPACTOS . . . . . ….99

6.1 CONCLUSIONES 101

6.2 RECOMENDACIONES 102

7..REFERENCIAS BIBLIOGRÁFICAS. 103

ÍNDICE DE TABLAS

Tabla 1. Método del Descenso más Rápido. 20

Tabla 2: Algoritmo de Newton. 24

Tabla 3. Crecimiento de la Economía Estadounidense durante el período 1899 – 1922 73

Tabla 4. Distancia proximal de Kullback-Leibler. 78

Tabla 5. Distancia φ – divergencia. 79

Tabla 6. Distancia Logaritmo-cuadrático. 88

Tabla 7. Distancia de Bregman (Kullback-Leibler). 89

Tabla 8. Distancia φ -divergencia con $\lambda_k = 1$. 97

Tabla 9. Distancia φ -divergencia con $\lambda_k = 1/M$. 98

ÍNDICE DE FIGURAS

Figura 2.1. (Grafica de un polinomio con mínimos locales) 13

Figura 2.2: Método del Descenso más Rápido. 20

Figura 2.3.Gráfica de un modelo no lineal y dominio de factibilidad del problema 42

Figura 2.4. Gráfica de un modelo no lineal con dominio de factibilidad del problema 43

.

Figura 4.1. PBI-Millones de Soles-2007 64

Figura 4.2. Modelo de producción con paquete econométrico Eviews 8 65

Figura 4.3. Los estadísticos Cusum y Cusum cuadrado 66

Figura 4.4. Distancia proximal de Kullback-Leibler: 77

Figura 4.5. Distancia φ – divergencia. 79

Figura 4.6. Distancia Logaritmo cuadrático, ...80

ÍNDICE DE ANEXOS

Anexo 1. Demanda interna -Inversión bruta interna . . . . ……... . . 107

Anexo 2. PBI (S/. 2007) 108

Anexo 3. PBI per cápita (S/. 2007)..... 109

Anexo 4. Perú: Producto Bruto Interno total y por habitante,
1994-2019. . . . . .…….. . .. . .. .110

Anexo 5. Perú: Formación bruta de capital fijo público y privado,
1994-2019 . . . . . .……. 111

Anexo 6. Población económicamente activa ocupada, según ámbito
geográfico, 2007 – 2019(miles de personas). .. . ……. . ..112

Anexo 7. Data de Cobb Douglas 113

Capítulo 1
INTRODUCCIÓN

1.1 Situación Problemática

Las aplicaciones de la optimización matemática son muy importantes y de gran interés en la comunidad científica y tecnológica, por ejemplo, en los negocios, en la administración, en la economía, las ingenierías, entre otros, motivo por el cual su estudio se ha intensificado en los últimos 20 años mediante publicaciones científicas, eventos internacionales, libros de especialidad, creación de nuevos softwares de optimización, entre otros.

Los modelos de optimización matemática consisten en minimizar o maximizar una determinada función limitado con recursos escasos en su dominio. Uno de estos modelos en el campo de la Teoría Económica es la toma de decisiones del consumidor, que consiste en maximizar la utilidad entre todos los bienes que se encuentran a su disposición, es decir, se trata de escoger las mejores opciones entre aquellas que sean posibles.

Recientes avances en estos modelos económicos consideran problemas de optimización no convexa como modelos más cercanos a la realidad y por ende se necesita de un software para poder resolver esta clase de problemas. Los métodos de optimización existentes en la actualidad, en general, no garantizan encontrar la solución óptima después de una serie de iteraciones, por lo tanto, es necesario realizar más investigaciones para verificar si el método o la extensión de este es capaz de encontrar estos puntos óptimos.

El "Método del Punto Proximal" , introducido en los años 70 por Martinet (1970) es el más atractivo para estudiar una clase de Métodos de Optimización No Convexa, por algunas razones: una de estas es la íntima relación con la Teoría de Convergencia de los métodos de multiplicadores que son utilizados en muchos softwares para resolver problemas de optimización con restricciones; otro motivo es que son considerados los mejores métodos para resolver problemas difíciles de optimización; otra razón es que las iteraciones del método funcionan de una manera totalmente natural para encontrar el siguiente punto que se aproxime a la solución pues la función que se va a minimizar en la

iteración actual tiene mejores propiedades de convexidad que la función original y, por lo tanto, es de fácil implementación.

En esta tesis extendemos el "Método del Punto Proximal", en la resolución de problemas de optimización no convexa y lo aplicamos a una clase de modelos económicos en teoría de decisión. Este método como cualquiera de los métodos de optimización genera una sucesión de puntos y bajo algunas condiciones de convexidad converge al óptimo del problema y si quitamos la condición de convexidad a la función objetivo convergirá a un punto crítico del problema.

Este trabajo aporta al progreso de una eficiente solución de problemas con funciones objetivos son cuasi-convexas usando una extensión del método de punto proximal. Anteriores trabajos resultan ser casos particulares de nuestro enfoque general.

El procedimiento del método es el siguiente: considerando el problema en el ortante no negativo y dado un punto inicial en el interior de este conjunto el algoritmo en la iteración k-ésima tiene un punto ya conocido x^{k-1} cuyas componentes son estrictamente positivas y encuentra un punto crítico x^k de la función $f(\cdot) + (\lambda_k /2)d(., x^{k-1})$, sin considerar ninguna restricción, observe que la función depende del término conocido x^{k-1}, por este motivo es llamado método proximal. A diferencia de otros métodos para funciones no convexas este método no es un método heurístico ni probabilístico.

1.2 Formulación del Problema

1.2.1 Problema General

¿De qué forma será posible utilizar el Método del Punto Proximal , para resolver modelos económicos de optimización para la toma de decisiones, ¿con funciones no convexas?

1.2.2 Problemas específicos

¿De qué manera será posible utilizar el Método del Punto Proximal, para problemas de maximización o minimización no convexa?

¿Cómo sera posible extender el método de punto Proximal propuesto a modelos de decisión en microeconomía?

1.3 Justificación Teórica

No existe una teoría de convergencia de un algoritmo que obtenga óptimos globales de problemas de optimización no convexa.

Muchos algoritmos de optimización que funcionan para problemas convexos no siempre funcionan o no pueden ser extendidos para resolver problemas no convexos.

Todos los softwares de optimización están basados en algoritmos de optimización y un buen funcionamiento del algoritmo se basa en sus propiedades teóricas de convergencia y que debe ser probado matemáticamente. Otro método de optimización puede existir para resolver el mismo problema y las técnicas para saber qué método es mejor está basado en sus propiedades de convergencia y su velocidad de convergencia. Es por eso que es importante analizar las propiedades de convergencia de un método de optimización.

1.4 Justificación Práctica

Los modelos de optimización que aparecen en economía, por ejemplo, en el área de teoría de decisión, están relacionados naturalmente a la no convexidad de la función objetivo, por lo tanto, es de interés estudiar algoritmos que puedan resolver esta clase de problemas.

Actualmente no existen softwares eficientes de optimización para resolver problemas de optimización no convexa.

Daríamos una opción a los especialistas para construir prototipos o software de optimización que consideren los casos no convexos usando el “Método del Punto Proximal” y sus variantes.

1.5 Objetivos

1.5.1 Objetivo general

Aplicar el “Método del Punto Proximal” para resolver modelos económicos no lineales en funciones no convexas.

1.5.2 Objetivos específicos

Proponer una extensión del " “Método del Punto Proximal” " para resolver problemas no convexos de optimización.

Verificar si la extensión propuesta del “Método del Punto Proximal” es aplicable en la resolución de problemas de decisión en microeconomía.

1.6 Hipótesis

Hipótesis general

Usando un concepto de subdiferencial de Fréchet es posible extender el “Método del Punto Proximal” para solucionar modelos económicos de optimización para la toma de decisiones, con funciones no convexas.

Hipótesis específicas

- Usando un concepto de subdiferencial de Fréchet es posible extender el “Método del Punto Proximal” para solucionar problemas de optimización no convexa.
- Usando un concepto de subdiferencial de Fréchet es posible extender el “Método del Punto Proximal” para solucionar problemas de decisión en microeconomía usando Cobb Douglas.

Capítulo 2

MARCO TEÓRICO

2.1 Marco Filosófico o Epistemológico de la Investigación

Esta tesis parte de un marco filosófico matemático de la Investigación de Operaciones, basado en la extensión de un algoritmo que genera a partir de una sucesión de puntos que basados en demostraciones lógicas darán como resultado la convergencia hacia un punto crítico en el caso no convexo y al óptimo en el caso convexo, luego, basándonos en el mismo marco, aplicamos el algoritmo a la solución de una clase de modelos económicos.

Cuando tenemos un problema cuya formulación matemática tiene una función objetivo no lineal con recursos escasos de un problema es una función no lineal, entonces decimos que pertenece al área de la Programación No Lineal.

2.2 Antedecedentes de la Investigación

El Método del Punto Proximal para minimizar funciones convexas sin restricciones fue introducido por Martinet (3) en1970 y fue muy estudiado por Rockafellar (1976 i) para resolver problemas más generales de encontrar ceros de operadores monótonos maximales, De la Cruz,L, Papa E. (2013) , ver definición 2.3.20.

El algoritmo es el siguiente:

Sea $f: \mathbb{R}^n \longrightarrow \mathbb{R}\cup\{+\infty\}$ una función propia dada la definición 2.3.5, semicontinua inferior, ver definición 2.3.3 y consideremos resolver el problema de minimización

$$min\{f(x) : x \in \mathbb{R}^n\} \tag{2.1}$$

donde f es convexa, esto es cuando f satisface:

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y), \quad x, y \, \epsilon \, dom \, f$$

Se da una sucesión de puntos $\{x^k\}$, con un punto inicial x^0, y un cierto $k = 1,2,\ldots.$ si $0 \in \partial f(x^{k-1})$, donde

$$\partial f(x) = \{s: f(y) \geq f(x) + s^T(y-x), \forall y \epsilon \mathbb{R}^n\}$$

es el subdiferencial de f para funciones convexas (que es la generalización del vector gradiente ∇f para el caso donde f no es diferenciable) entonces el algoritmo finaliza. Caso contrario, encuentra un $x^k \in \mathbb{R}^n$ tal que

$$x^{\mathrm{k}} \in \operatorname{argmin}\left\{\mathrm{f}(\cdot) + \left(\frac{\lambda_k}{2}\right)\|\bullet - x^{k-1}\|^2\right\},$$

que significa que x^{k} es un punto de mínimo de la función $\mathrm{f}(\cdot) + \left(\frac{\lambda_k}{2}\right)\|\bullet - x^{k-1}\|^2$ y λ_k es un parámetro positivo para cada k. Se demuestra en Kaplan y Tichatsche (1998), que se cumple que:

$$\sum_{\mathrm{k}=1}^{+\infty} \frac{1}{\lambda_k} = +\infty$$

decimos,que la sucesión $\{f(x^k)\}$ tiende al ínfimo de f , además el conjunto no vacío de soluciones óptimas, en general los métodos proximales cuando el parámetro tiende a cero la convergencia es superlineal, luego, $\{x^k\}$ converge a una solución del problema.

La función que cumple el parámetro λ_k es asegurar la convergencia de la sucesión $\{x^k\}$, este parámetro es independiente en el sentido que es un dato inicial del algoritmo y en un caso práctico puede tomarse como $\lambda_k = 1/k,$, $\lambda_k = 1,$ donde k es un número entero positivo que contabiliza las iteraciones del algoritmo.

Posteriormente Kaplan y Tichatsche (2004) estudiaron el método para una clase de funciones no convexas cuando la función regularizada $f(\cdot) + \left(\frac{\lambda_k}{2}\right)\|\bullet - x^{k-1}\|^2$ que significa que a la función principal le adicionamos a una función fuertemente convexa y diferenciable, (una función g es fuertemente convexa si $g(\lambda x + (1-\lambda)y) \leq \lambda g(x) + (1-\lambda)g(y) - \mu\lambda(1-\lambda)\|x-y\|^2$ para todo $x, y \epsilon\, dom\, g)$ es fuertemente convexa en un conjunto convexo bajo una adecuada elección de λ_k.

Los autores probaron que la sucesión finaliza en un número finito de iteraciones en un punto estacionario o todo punto cúmulo de $\{x^k\}$ es llamado punto estacionario de la función f, ver Definición 2.3.27.

Cuando f es propia (el dominio no es vacío y todo punto del dominio no es -∞) ver Definición 2.3.5, localmente Lipschitzana (que significa que la función en una vecindad es Lipchitziana), no diferenciable y sobre todo cuasiconvexa, ver Pan S, Chen JS (2007) es,

$$f(\lambda x + (1-\lambda)y) \leq \max\{f(x), f(y)\}, \quad x, y \in dom\, f,$$

Papa Quiroz y Oliveira (2012ii) introdujeron el siguiente algoritmo: dada una sucesión de parámetros positivos $\{\lambda_k\}$ y un punto inicial $x^0 \in \mathbb{R}^n$, para cada $k = 1,2, \ldots$, Si $0 \in \hat{\partial}\,(f(x^{k-1})$, entonces, el método finaliza. De otro modo buscamos un $x^k \in \mathbb{R}^n$ tal que

$$0 \in \hat{\partial}\,(f(\cdot) + (\lambda_k/2)\|.-x^{k-1}\|^2)(x^k),$$

donde $\hat{\partial}$ es el subdiferencial de Fréchet esto es:

$$\hat{\partial} f(x) = \left\{ s \epsilon \mathbb{R}^n : \lim\ inf_{y \to x} \frac{f(y) - f(x) - \langle s, y - x \rangle}{\|y - x\|} \geq 0 \right\}$$

y λ_k es un parámetro positivo para cada k.

Observemos que este método es una extensión del "Método del Punto Proximal" con funciones no convexas, en efecto si f es convexa entonces la iteración se convierte en:

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

Se observa también que esta metodología solo necesita encontrar un punto estacionario, ver definición 2.3.27. $0 \in \hat{\partial}\,(f(\cdot) + (\lambda_k/2)\,\|\,.\,- x^{k-1}\|^2)(x^k)$ (no necesariamente un mínimo global) de la función regularizada, entonces, el algoritmo puede ser usado satisfactoriamente en cada iteración, ya que al ser un problema no convexo esta puede tener puntos críticos que no son mínimos.

Si f es semicontinua inferior, ver definición 2.3.3 y acotada inferiormente entonces las iteraciones $\{x^k\}$ son bien definidas y si f es cuasi-convexa, entonces $\{f(x^k)\}$ es no creciente, $\{x^k\}$ converge a un punto estacionario del problema (P). Definiendo un conjunto $X = \{x : f(x) \leq f(x^k), \forall k\}$ que contiene el conjunto de soluciones óptimas del problema podemos obtener que, si el conjunto de X es vacio, entonces;

$$\lim_{k\to+\infty} f(x^k) = \inf_{x\in\mathbb{R}^n} f(x).$$

Si el conjunto de X es no vacío y $\{\lambda_k\}$ son acotados, entonces $\{x^k\}$ converge a un punto crítico. Además, si cada x^k es un mínimo global de la función regularizada dado por $(f(\cdot) + (\lambda_k/2)\|\,.\,, x^{k-1}\|^2$ y $\lambda_k \to 0$ se probó que $\{x^k\}$ converge a un punto mínimo del problema. Como caso particular, se obtuvo la convergencia del método a un punto mínimo global para funciones pseudoconvexas y convexas. Recordemos que una función es Pseudoconvexa si para todo $x, y \in dom\, f$: $\nabla f(x)^T(y-x) \geq 0$, entonces $f(x) \leq f(y)$.

Tomamos la resolución de un modelo de optimización en general de la forma:

(P) Min f(x)

s.a. $x \in D \subseteq \mathbb{R}^n\ Donde\ f\colon \mathbb{R}^n \to \mathbb{R} \to$ es una función, comúnmente continua y diferenciable, y D es el dominio de factibilidad del problema, generalmente dado por:

$D = \{x \in \mathbb{R}^n / g_i(x) = b_i\ \ i = 1, \dots, m;\ h_r(x) \leq d_r, r = 1, \dots, l\}$

Decimos que $x^* \in D$ es un mínimo global o solución óptima del problema P) ssi:

$f(x^*) \leq f(x)\ para\ todo\ x \in D$

Por otra parte, decimos que $x^{\wedge} \in D$ es un mínimo local del problema P)

ssi: $f(x^{\wedge}) \leq f(x)\ para\ todo\ x\ en\ una\ vecindad\ (x \in D \cap (x^{\wedge}, \in))$

Ahora, consideremos el problema restricto al ortante no negativo

$$min\{f(x) : x \geq 0\}, \tag{2.2}$$

Donde $f\colon \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ es una función propia ver, Definición 2.3.5, acotada inferiormente y semicontinua inferior, ver definición 2.3.3.

y $x \geq 0$ significa que $x_i \geq 0, \forall i = 1, \dots, n$; el método de punto proximal para esta clase de problemas genera una sucesión $\{x^k\}$ dado por $x^0 \in \mathbb{R}^n + +$ (punto arbitrario)

$$x^k \in argmin_{x\geq 0}\{f(x) + \lambda_k d(x, x^{k-1})\}, \quad (2.3)$$

Donde la función d cumple con algunas propiedades aceptables pues se parecen al de una distancia euclidiana con la condición de que las iteraciones esten dentro del interior de un conjunto de restricciones y λ_k es un parámetro positivo. Mostraremos algunas funciones basadas en distancias con esta metodología:

Definición 2.2.1 Distancias de Bregman, *Kullback-Leibler).* Ver Kiwiel 1997:

$$D_{\bar{h}_1}(x, y) = \sum_{i=1}^{n} \left(x_i \, log \left(\frac{x_i}{y_i} \right) + y_i - x_i \right).$$

donde $h : \mathbb{R}^n \to \mathbb{R}$ es una función de Bregman ver Eckstein J (1998) ,(ver la definición formal de función de Bregman, Solodov MV, Svaiter BF (2000)

Definición 2.2.2 Distancias φ –divergentes:

$$d_{\varphi}(x, y) = \sum_{i=1}^{n} y_i \varphi \left(\frac{x_i}{y_i} \right),$$

$$d(x, y) = \sum_{i=1}^{n} \left(y_i^2 \log \left(\frac{x_i}{y_i} \right) + x_i - y_i \right)$$

donde $\varphi: \mathbb{R}+ \to (-\infty, +\infty)$ es una función que cumple con algunas propiedades de diferenciabilidad y convexidad tal que la distancia se transforma en estrictamente convexa en la primera variable, ver Auslender y Teboulle (1999).

Asumiendo que f es una función propia ver definición 2.3.5, semicontinua inferior, convexa,

$domf \cap \mathbb{R}^n + + \neq \emptyset$ y $\{\lambda_k\}$ satisface

$$\sum_{k=1}^{+\infty} \frac{1}{\lambda_k} = +\infty \quad (2.4)$$

Ver Auslender y Teboulle (2006), demostraron que usando la distancia proximal

$$lim_{k\to+\infty} f(x^k) = inf\{f(x) : x \geq 0\}$$

Donde el conjunto de óptimas soluciones es no vacío, dado el problema (2.2), vemos que $\{x^k\}$ tiende a una solución óptima.

Definición 2.2.3 Distancia homogénea de segundo orden definida por:,

$$d_\theta(x,y) = \sum_{i=1}^{n} y_i^2 \theta\left(\frac{x_i}{y_i}\right),$$

$$d(x,y) = \sum_{i=1}^{n} \left(x_i^2 - x_i y_i - y_i^2 \log\left(\frac{x_i}{y_i}\right)\right)$$

la cual es llamada distancia Logaritmo cuadrático.

Si f es una función cuasi-convexa sobre el ortante no negativo hay algunos desarrollos recientes en la literatura.

Por ejemplo, podemos mencionar:

Attouch y Teboulle (2004), con un regularizado Lotka, Volterra, han demostrado la convergencia del método continuo con el punto que pertenece a cierto conjunto de puntos óptimos; ver también Alvarez et al. (2004), que trata una clase general de sistemas dinámicos que incluye la de Attouch y Teboulle (2004). Cunha et al. (2010) y Chen y Pan (2008), con una distancia φ –divergente en particular, han probado la convergencia completa del método proximal al punto KKT del problema cuando el parámetro λ_k está limitado y la convergencia a una solución óptima cuando $\lambda_k \to 0$. Chen y Pan (2008), que incluye el Logarítmo-Cuadrático Proximal, y Souza et al. (2010) con una clase de Bregman separado de distancias, han demostrado el mismo resultado de convergencia de Cunha et al. (2010) y Chen y Pan (2008), probaron la convergencia a un cierto conjunto que pertenece al conjunto de soluciones óptimas y convergencia a un punto de KKT cuando la función es continuamente diferenciable.

Como la tesis está motivada a resolver el problema (2.2), cuando f es no convexa usando el método proximal, Attouch H, Bolte J, Svaiter B (2010), el problema (2.3) no es necesariamente convexo debido a la no convexidad de f, por eso, que el problema (2.2), el resolver que el problema original puede ser tedioso si no se toman algunas consideraciones sobre el algoritmo y también

sobre los parámetros λ_k, ahí radica la importancia del parámetro de regularización.

De esta forma proponemos, similarmente a los trabajos de Papa Quiroz y Oliveira (2012i) la iteración:

$$0 \in \hat{\partial}\{ \mathrm{f}(\cdot) + \lambda_k \, d(., x^{k-1}) \}(x^k) \tag{2.5}$$

donde $\hat{\partial}$ denota el subdiferencial de Fréchet, esto es:

$$\hat{\partial} f(x) = \left\{ s \in \mathbb{R}^n : \lim inf_{y \to x} \frac{f(y) - f(x) - \langle s, y - x \rangle}{\|y - x\|} \geq 0 \right\},$$

Donde $\lim inf_{y \to x} g(y)$ es el límite inferior de g en el punto x. Observemos que (2.5) es mas comprensible resolver un problema teórico y práctico que hallar la función minimizada en el punto $\{x^k\}$.

$$f(\cdot) + \lambda_k \, d(., x^{k-1}),$$

Basta hallar un punto crítico generalizado. Así bajaríamos el costo computacional por las iteraciones en referencia al "Método del Punto Proximal clásico".

Del punto de vista práctico, damos una aplicación del método a la microeconomía, presentando un ejemplo cuando la función objetivo es una función de producción e implementamos el algoritmo para resolver el problema. Además, presentamos adicionalmente algunos experimentos numéricos que dan confiabilidad a los resultados teóricos.

Del punto de vista académico, debemos resaltar que este trabajo será de mucha utilidad también en la construcción de software de optimización.

La diferencia con respecto a los trabajos de Papa Quiroz y Oliveira (2012ii, 2012iii) es que ellos consideraron el caso cuasi-convexo, Langenberg N, Tichatschke R (2012) y en este trabajo estamos considerando el caso no convexo que es más general.

2.3 Bases Teóricas

En esta sección presentamos un resumen de la teoría de optimización diferenciable y optimización convexa, temas que serán útiles para el desarrollo de la tesis. Las demostraciones se pueden encontrar para mas detalle y profundización de los temas en las referencias de Bazaara MS, Sherali HD, Shetty CM (1993), Barbolla et al. (2000), Luenberger and Ye (2008), Minoux (1986).

2.3.1 Existencia de Puntos de Mínimo Global

En esta subsección mostramos soluciones de problemas de optimización dadas las condiciones suficientes que nos garantizen dicha solución.

Debemos observar que el espacio euclidiano en el que se da este estudio para formular problemas de optimización es un espacio natural en problemas de optimización, pues el problema de optimización se da en un espacio de dimensión infinita con diferentes métodos de discretización que nos ayuda a cambiar el problema original en un problema dado en un espacio euclidiano.

Definición 2.3.1 (Mínimo local y mínimo global)

Si definimos la función $f: x \subseteq \mathbb{R}^n \to \mathbb{R}$ sobre x, llamado punto de mínimo local de f sobre X con un punto $\bar{x} \in X$,si existe $\xi > 0$ de manera que la función $f(\bar{x}) \leq f(x)$, para todo $x \in X \cap B(\bar{x}, \xi)$.

El punto $\bar{x} \in X$ decimos que el punto de mínimo global de f sobre X si $f(\bar{x}) \leq f(x)$, para todo $x \in X$.

Observación 2.3.1

Teniendo en cuenta estas definiciones dadas decimos que un mínimo global es también mínimo local, pero el recíproco no iempre se cumple. Veremos en el ejemplo 2.1.

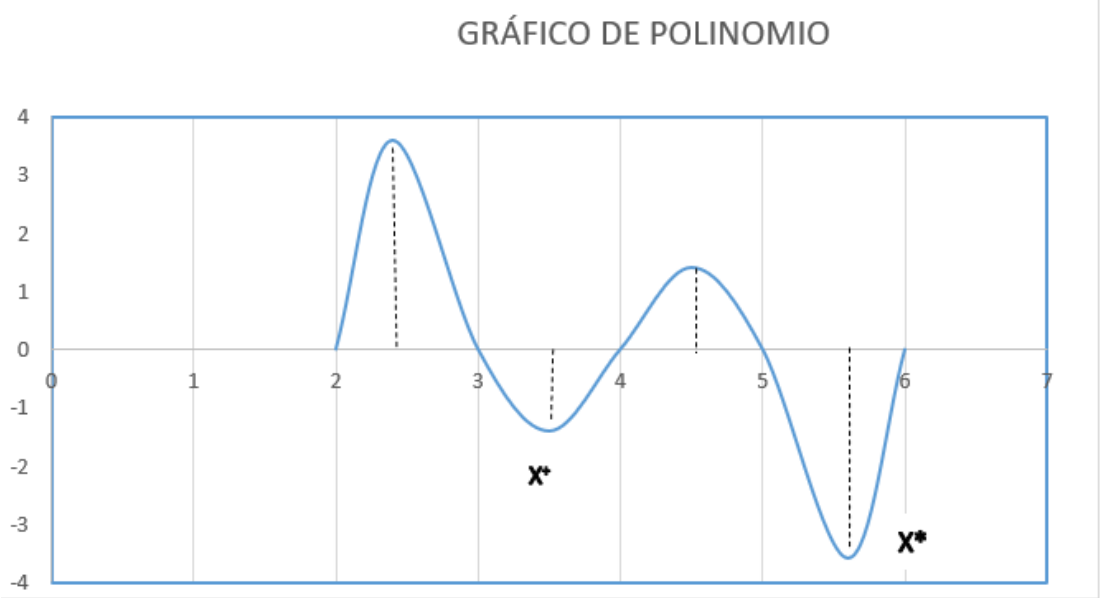
Ejemplo 2.1

Min $\quad f(x) = (x-2)(x-3)(x-4)(x-5)(x-6)$

$s.a. \quad 2 \leq x \leq 6$

Figura 2.1

Grafica de un polinomio con mínimos locales



Nota: En la grafica de la *Figura 1,* observamos que hay dos mínimos locales $x^+ \; y \; x^*$

Donde: Mínimo local es x^+

$Minimo\ global\ es\ x^*$, ver Definición 2.3.1

$Dados\ los\ puntos\ x^+ \; y \; x^*$, la solución óptima o mínimo global es x^*

Definición 2.3.2 (Máximo local y máximo global)

Si definimos $\bar{x} \in X$ un punto de máximo de una función $f: X \subseteq \mathbb{R}^n \to \mathbb{R}$ sobre X. es llamado un de f sobre X donde $\xi > 0$ tal que $f(\bar{x}) \geq f(x)$, para todo $x \in X \cap B(\bar{x}, \xi)$.

Llamamos el punto $\bar{x} \in X$ máximo global de f sobre X es, si $f(\bar{x}) \geq f(x)$, para todo $x \in X$.

Observación 2.3.2

Dado un problema de maximización se puede trabajar con otro problema de minimización, por eso trataremos el análisis de minimización en los problemas.

Consideremos el problema:

$$min\{f(x) : x \in X\},$$

Tenemos una función definida sobre X donde $f: X \subseteq \mathbb{R}^n \to \mathbb{R}$. queremos hallar un punto que resuelva el problema formulado. Por eso será necesario trabajar varias condiciones de f y sobre X.

Definición 2.3.3

Decimos que una función f es semicontinua inferior en $\bar{x},\ Donde\ f: X \subseteq \mathbb{R}^n \to \mathbb{R}. \in X$, si para toda sucesión $\{x^l\}$ de X tiende a $\bar{x}$.

$$Lim\ inf_{k\to+\infty} f(x^k) \geq f(\bar{x}),$$

Si f es semicontinua inferior donde el $Lim\ inf_{k\to+\infty} f(x^k) = sup_{n\in\mathbb{N}}\ inf_{k\geq n} f(x^k)$. para todo $x \in X$, entonces diremos que f es semicontinua inferior en X.

Teorema 2.3.1

Decimos que un conjunto no vacío y compacto X, existe un punto de mínimo global de una función $f: X \subseteq \mathbb{R}^n \to \mathbb{R}$. Donde f es semicontinua inferior, ver definición 2.3.3.

Una condición compacidad de X es muy fuerte para garantizar la existencia de un mínimo global. En efecto, considere el problema

$$min\{x^2 : x \in \mathbb{R}\} \tag{2.6}$$

Aquí $X = \mathbb{R}$ no es compacto, pero existe el mínimo global.

Definamos, para $\alpha \in \mathbb{R}$, el conjunto de compacidad $Lf(\alpha): \{x \in X: f(x) \leq \alpha\}$ para debilitar la hipótesis de compacidad de X, , al que se conoce como conjunto de nivel inferior de f.

Definición 2.3.4 Identificamos la sucesión $\{x^k\}$ en X como crítica (en conexión al conjunto X) si: $limk_{\to+\infty} \| x^k \| = +\infty \vee limk_{\to+\infty}\ x^k = \bar{x} \in \bar{X}\backslash X$, donde la notación $\bar{x} \in \bar{X}\backslash X$ significa que $\bar{x} \in \bar{X}\ y\ \bar{x} \notin X$. La función $f: X \subset \mathbb{R}^n \to \mathbb{R}$ al que llamamos *coerciva* en X si en toda sucesión crítica $\{x^k\}$ tenemos:

$Lim\ sup_{k\to+\infty} f(x^k) = +\infty$, donde $limsup_{k\to+\infty} f(x^k) := inf_{n\in\mathbb{N}}\ sup_{k\geq n} f(x^k)$.

Corolario 2.3.1

Decimos que una función f es coerciva y semicontinua inferior ver definición 2.3.3. Si la función $f: X \subseteq \mathbb{R}^n \to \mathbb{R}$. de un conjunto no vacío X,hay un punto X. mínimo global en f .

Corolario 2.3.2

Sea un conjunto $A \subset \mathbb{R}^n, donde\ A\ es\ cerrado \iff FrA \subset A$

Observación 2.3.3 En el corolario 2.3.2, la condición de cerradura de X es fundamental. En efecto, consideremos:

Ejemplo 2.2, $X = (-1,1]$ y $f(x) = (x+1)^2$. Donde $\alpha \in \mathbb{R}++$ con $Lf(\alpha) = (-1, \sqrt{\alpha - 1}]$ es acotado no vacio, si f es continua pero no tiene un mínimo global de f en X.

Una de las condiciones de existencia de mínimos globales se usa para debilitar dichas condiciones por esto definiremos una función coerciva.

Definición 2.3.5 Decimos que $f: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ es una función propia tal que $dom\ f \cap \mathbb{R}^n_+ \neq \emptyset\ y\ x \geq 0,$ donde cada componente es no negativo en x.

2.3.2 Problemas sin Restricciones con condiciones de Optimalidad

La mayoria de problemas en la práctica se dan con restricciones pero el estudio de técnicas de optimización sin restricciones es fundamental por muchas razones, una de estas razones es que se resuelven problemas con restricción transformandolos en algoritmos con una sucesión de problemas sin restricciones, otra de las razones es que los algoritmos de optimización sin restricciones pueden ser usadas de manera natural en la resolución de problemas restrictos.

Consideramos en esta subsección, los problemas de optimización sin restricciones

$$min\{f(x) : x \in \mathbb{R}^n\} \tag{2.7}$$

donde f es una función definida en el espacio euclideano $\mathbb{R}^n$.

Teorema 2.3.2 (Condición Necesaria de 1er Orden)

Si una función diferenciable $f: \mathbb{R}^n \to \mathbb{R}$ en $x*$. Donde decimos que $x*$ punto de mínimo local, $\nabla f(x*) = 0$, ver problema (2.7).

Teorema 2.3.3 (Condición Necesaria de 2do Orden)

Si una función diferenciable dos veces $f : \mathbb{R}^n \to \mathbb{R}$ en $x*$. Decimos que $x*$ es mínimo local del problema (2.7), por tanto $\nabla f(x*) = 0$ y $\nabla^2 f(x*)$ es semidefinida positiva.

Teorema 2.3.4 (Condición Suficiente de una función de 2do Orden)

Si una función f es diferenciable dos veces, donde $f : \mathbb{R}^n \to \mathbb{R}$ en $x*$. Si $\nabla f(x*) = 0$ y $\nabla^2 f(x*)$ es definida positiva, decimos que $x*$ es un mínimo local estricto ver (2.7).

Observación 2.3.4 Un teorema recíproco de 2.3.4 es falso. En efecto, considere:

Ejemplo 2.3, $f(x) = x^4$ y $\bar{x} = 0$. Dado mínimo global estricto en $\bar{x}$ de una función f con $f''(\bar{x}) = 0$.

Observación 2.3.5 Consideremos ahora el problema de maximización

$max\{f(x) : x \in \mathbb{R}^n\}$.

Vemos que en el ejemplo 2.3 que es un problema semejante a

$min\{-f(x) : x \in \mathbb{R}^n\}$.

Donde los valores obtenidos del problema minimizado teniendo las condiciones en el caso de maximización son:

1. Condición necesaria en 1er orden de: Cuando $f : \mathbb{R}^n \to \mathbb{R}$ es diferenciable en $x*$ como punto local máximo punto de, donde $\nabla f(x*) = 0$.

2. Condición necesaria de segundo orden: decimos que $f : \mathbb{R}^n \to \mathbb{R}$ es diferenciable dos veces en un máximo local $con\ punto\ x*$, con $\nabla^2 f(x*) \preccurlyeq 0$.

3. Condición suficiente de 2do orden: Cuando $\bar{x}$ cumple con $\nabla f(\bar{x}) = 0$ y también $\nabla^2 f(\bar{x}) \prec 0$, entonces, $\bar{x}$ es un punto de máximo local estricto de f.

Consideraremos un problema

P) $Min\ f(x), con\ x \in \mathbb{R}^n$

En este tipo de problemas podemos ver por ejemplo el problema de aproximación y ajuste de curvas. Pero una de las causas para del estudio consiste en la extensión de casos y algoritmos sobre este tipo de problemas de optimización restricta.

Podemos ver un resumen de ciertos resultados teóricos en este tipo de problemas:

Teorema 2.3.5 (condiciones necesarias de 1er orden). Dado una función continuamente diferenciable f con $x^* \in \mathbb{R}^n$ mínimo local de P) donde: $\nabla f(x^*) = 0$.

Teorema 2.3.6 (condiciones necesarias de 2do orden). Dado una función f dos veces diferenciable y continua, donde $x^* \in \mathbb{R}^n$ es mínimo local de P), donde:

$\nabla f(x^*) = 0 \ y \ D^2 f(x^*)$ es definida semipositiva.

Del Teorema 2.3.6, podemos decir que no siempre los puntos de $x \in \mathbb{R}^n$ cumplen con las propiedades mencionadas son mínimos locales de la función, a veces hay resultados que se dan las condiciones suficientes y necesarias para x^* sea mínimo local.

Teorema 2.3.7 (condiciones necesarias y suficientes de 2do orden).

Una función f es dos veces diferenciable y continua en $x^* \in \mathbb{R}^n. Si \ \nabla f(x^+) = 0 \ y \ D^2 f(x^*)$ definida positiva, decimos que x^* es punto estricto mínimo local.

Teorema 2.3.8 Decimos que función convexa f es continuamente diferenciable, donde x* será un mínimo global si y solo si $\nabla f(x^*) = 0$.

Ejemplo 2.4

Consideremos:

$$f(x_1, x_2) = 3x_1^2 + x_2^3 - 1.5x_2^2$$

su gradiente y matriz Hessiana corresponden a:

$$\nabla f(x) = \begin{bmatrix} 6x_1 \\ 3x_2^2 - 3x_2 \end{bmatrix}$$

$$D^2 f(x) = \begin{bmatrix} 6 & 0 \\ 0 & 6x_2 - 3 \end{bmatrix}$$

Donde se dan dos posibles puntos, $x^+ = (0, 0)^T$ y $x^* = (0, 1)^T$, los cuales cumplen con las condiciones necesarias de 1er orden.

$$D^2 f(x) = \begin{bmatrix} 6 & 0 \\ 0 & 6x_2 - 3 \end{bmatrix}$$

Pero en el problema es definida positiva en $x^* = (0,1)$, de manera que x^* es mínimo local.

Los algoritmos generales de descenso en optimización nos ayuda a resolver este tipo de problemas que pertenecen a estos métodos ,ver Humes Jr C, Silva PJS (2005) , estos nos ayudan a hallar un mínimo local y asi reducir el el cálculo de esta secuencia para problemas con búsqueda unidimensional.

Denotamos d como dirección del descenso de una función f,donde $d \in \mathbb{R}^n \rightarrow$ es de la en el punto x^* si y solo si la derivada direccional de f en x^* con dirección negativa de d .

Consideremos además la función unidimensional

$$g(\alpha) = f(x^* + \alpha d)$$

Tomamos α es el tamaño del paso que es un escalar real. La función nos da el resultado de la función f cuando este se desplaza partiendo de un punto x^+ hacia la dirección d de determinado α.

Verificamos, si $g'(0) = \nabla f^T(x^*)d < 0$,seria posible tomar un paso α de manera:

$$g(\alpha) = f(x^* + \alpha d) < f(x^*) = g(0)$$

Donde $g(\alpha)$ reduce el valor en referencia al valor actual evaluado en x^*.

2.3.2.1 Algoritmo general de descenso

1) Considerar como semilla $x = x^0$. $Haciendo\ k = 0$

2) Veremos la dirección de descenso d^k que se toma.

3) Buscaremos un paso α_k lineal, de manera:

$$g_k(\alpha_k) = f(x^k + \alpha_k d^k) < f(x^k) = g_k(0)$$

4) Hacer $x^{k+1} = x^k + \alpha_k d^k$

5) Hacer una prueba de convergencia.

Converge $\begin{cases} si, & parar \\ no, hacer\ k = k+1, volver\ al\ paso\ 2 \end{cases}$

De 5), verificaremos que se cumplan los criterios mas usados de convergencia:

$\nabla f(x^k) < \varepsilon$

$|f(x^{k+1}) - f(x^k)| \ / \ (1 + |f(x^k)|) < \varepsilon$

Para un determinado número L, con números consecutivos k, y tomando un cierto ε como error, en este caso tomaremos como error de tolerancia $\varepsilon = 10^{-4}$.

Hay una variedad de algoritmos para escoger con el método de dirección de descenso.

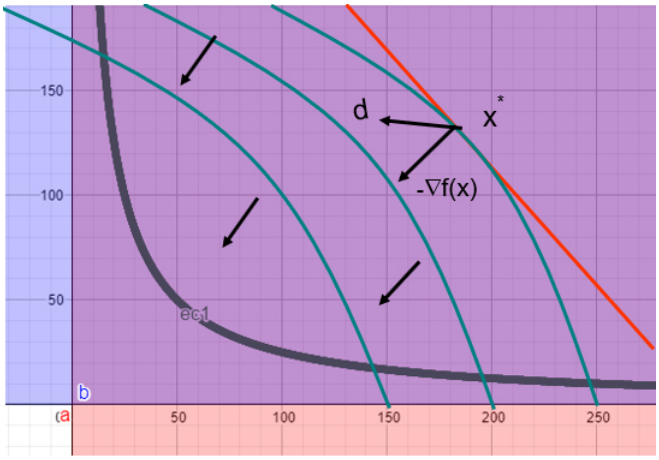
Método del Descenso más Rápido

También conocido como Método de Cauchy o Gradiente, Attouch H, Bolte J, Svaiter B (2010), x^k, la dirección de descenso se escoge como:

$d^k = -\nabla f(x^k)$ (2.8)

Figura 2.2

Método del Descenso más Rápido



Ejemplo 2.5

Dada la función de minimización:

Min $(x_1 - 2)^4 + (x_1 - 2x_2)^2$

S.a:

$x \in \mathbb{R}^2$

Aplicaremos el método del descenso iniciando con los puntos $x_1{}^0 = 0, x_2{}^0 = 3$

$$\nabla f(x) = \begin{bmatrix} 4(x_1 - 2)^3 + 2(x_1 - 2x_2) \\ 2(x_1 - 2x_2)(-2) \end{bmatrix} = \begin{bmatrix} 4(x_1 - 2)^3 + 2(x_1 - 2x_2) \\ -4(x_1 - 2x_2) \end{bmatrix}$$

Evaluando en $x_1{}^0 = 0, x_2{}^0 = 3$ y

$x^{k+1} = x^k + \alpha_k d^k \rightarrow x^1 = x^0 + \alpha_0 d^0$

Verificando $\mathbf{f}(\boldsymbol{x_1}) = (x_1 - 2)^4 + (x_1 - 2x_2)^2 = (0 - 2)^4 + (0 - 2(3))^2 = 52$

$$\nabla f(\boldsymbol{x_1}) = \begin{bmatrix} 4(x_1 - 2)^3 + 2(x_1 - 2x_2) \\ -4(x_1 - 2x_2) \end{bmatrix} = \begin{bmatrix} 4(0 - 2)^3 + 2(0 - 2(3)) \\ -4(0 - 2(3)) \end{bmatrix} = \begin{bmatrix} -44 \\ 24 \end{bmatrix}$$

Tabla 1

Método del Descenso más Rápido

Iteración k	x_k	$f(x_k)$	$\nabla f(x_k)$	α_k
1	(0.00,3.00)	52.00	(-44.00,24.00)	0.062
2	(2.70,1.51)	0.34	(0.73, 1.28)	0.24
3	(2.52,1.20)	0.09	(0.80, -0.48)	0.11
4	(2.43,1.25)	0.04	(0.18, 0.28)	0.31
5	(2.37,1.16)	0.02	(0.30, -0.20)	0.12
6	(2.33,1.18)	0.01	(0.08, 0.12)	0.36
7	(2.30,1.14)	0.009	(0.15, -0.08)	0.13
8	(2.28,1.15)	0.007	(0.05, 0.08)	

Autor: J. Moré and S. Wright (1993)

2.3.2.2 Método de Newton

El algoritmo de Newton es uno de los que utilizan derivadas, no siempre encontraremos su solución como óptimo global, como consecuencia, por lo que se tomará diferentes puntos para ver si el problema tiene varias soluciones óptimas.

Cuando un problema tiene funciones diferenciables, entonces podemos aplicar programación no lineal podemos usar varios algoritmos de soluciónen los que se hallan derivadas. Por ello debemos tener conceptos básicos como el Hessiano el cual tiene una función que sea dos veces diferenciable y el gradiente con función diferenciable

Algoritmo de Newton:

Paso 0: Elegir $\varepsilon > 0$ para tener una precisión mínima. Para comenzar se elige como punto inicial x^1.

Paso 1:

$$\text{Calcular } \nabla f(x^1) \; y \; Hf(x^1), Hallar \; x^2 = x^1 - Hf(x^1)^{-1}\nabla f(x^1), \text{Ir al Paso 2.}$$

Paso 2:

Si $\|x^2 - x^1\| < \varepsilon$ ó $\|\nabla f(x^1)\| < \varepsilon$,

entonces *parar*, decimos que x^1 es una solución para el problema. Pero si no se cumple debemos tomar $x^2 = x^1$ y luego ir al Paso 1.

Condiciones:

1. Solo se puede aplicas si las funciones son 2 veces diferenciables.

2. Se aplicará en una función con aproximación cuadrática.

3. Veremos que puntos de la función da cero en la gradiente.

4. Verificar si hay inversa en cada punto del hessiano.

Calcular el vector d^k como solución del sistema de ecuaciones:

$$D^2 f(x) d^k = -\nabla f(x) \qquad (2.9)$$

Pero este método es más eficiente por su velocidad de convergencia en cada iteración, este necesita hallar segundas derivadas parciales y resolver el sistema de ecuaciones. Adicionalmente sabemos que, d^k es una dirección de descenso está garantizada cuando $D^2 f(x^k)$ es definida positiva.

Ejemplo 2.6

Problema con algoritmo de Newton.

$Max \;\; x^2y + 2xy - y^2 - 3y$

$$\nabla f(x,y) = \begin{pmatrix} \frac{\partial f(x,y)}{\partial x} \\ \frac{\partial f(x,y)}{\partial y} \end{pmatrix} = \begin{pmatrix} 2xy + 2y \\ x^2 + 2x - 2y - 3 \end{pmatrix}$$

$$Hf(x,y) = \begin{bmatrix} \frac{\partial^2 f(x,y)}{\partial x^2} & \frac{\partial^2 f(x,y)}{\partial x \partial y} \\ \frac{\partial^2 f(x,y)}{\partial y \partial x} & \frac{\partial^2 f(x,y)}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 2y & 2x+2 \\ 2x+2 & -2 \end{bmatrix}$$

$$x^2 = x^1 - Hf(x^1)^{-1} \nabla f(x^1)$$

Iteración 1

$$x^1 = \begin{pmatrix} -1 \\ -1 \end{pmatrix} \;\rightarrow\; \nabla f(x^1) = \begin{pmatrix} 0 \\ -2 \end{pmatrix}$$

$$Hf(x^1)^{-1} = \begin{pmatrix} -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}$$

Iteración 2

$$x^2 = \begin{pmatrix} -1 \\ -1 \end{pmatrix} - \begin{pmatrix} -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 \\ -2 \end{pmatrix} = \begin{pmatrix} -1 \\ -2 \end{pmatrix}$$

$x^2 = \begin{pmatrix} -1 \\ -2 \end{pmatrix} \rightarrow \nabla f = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ parar

Al aplicar el método en el ejemplo 2.6, se tiene:

Cuando un problema tiene funciones diferenciables, podemos aplicar el Algoritmo de Newton en programación no lineal y hallar algoritmos que utilizan derivadas en la solución. Usando conceptos básicos como el Hessiano el cual tiene una función que sea dos veces diferenciable y el gradiente con función diferenciable

Denotamos $f : R^n \to R$, como el *gradiente* de f , (∇f)

$$\nabla f(x_1, x_2, \dots x_n) = \left(\frac{\partial f(x_1, .., x_n)}{\partial x_1}, \frac{\partial f(x_1, .., x_n)}{\partial x_2}, \dots, \frac{\partial f(x_1, .., x_n)}{\partial x_n}\right)^t$$

$$f(\boldsymbol{x_k}) = (x_1 - 2)^4 + (x_1 - 2x_2)^2$$

$$\nabla f(\boldsymbol{x_k}) = \begin{bmatrix} 4(x_1 - 2)^3 + 2(x_1 - 2x_2) \\ -4(x_1 - 2x_2) \end{bmatrix}$$

El gradiente de una función nos indica la dirección del máximo crecimiento. la función en un punto del hiperplano es tangente al vector normal de la función en el punto inicial $(0,3)$.

$$\nabla f(x_1, x_2) = \begin{bmatrix} 4(x_1 - 2)^3 + 2(x_1 - 2x_2) \\ -4(x_1 - 2x_2) \end{bmatrix}$$

$$D^2 f(x) d^k = -\nabla f(x)$$

$$\mathbf{f}(x_1, x_2) = (x_1 - 2)^4 + (x_1 - 2x_2)^2 = \quad (0 - 2)^4 + (0 - 2(3))^2 = 52$$

$$D^2 f(x) d^k = -\nabla f(x_1, x_2) = \begin{bmatrix} 4(x_1 - 2)^3 + 2(x_1 - 2x_2) \\ -4(x_1 - 2x_2) \end{bmatrix}$$

$$= \begin{bmatrix} 4(0 - 2)^3 + 2(0 - 2(3)) \\ -4(0 - 2(3)) \end{bmatrix} = \begin{bmatrix} -44 \\ 24 \end{bmatrix}$$

Tabla 2

Algoritmo de Newton

Iteración k	x_k	$f(x_k)$	$\nabla f(x_k)$
1	(0.00,3.00)	52.00	(-44.00,24.00)
2	(0.67,0.33)	3.13	(-9.39,-0.04)
3	(1.11,0.56)	0.63	(-2.84,-0.04)
4	(1.41,0.70)	0.12	(-0.80,-0.04)
5	(1.61,0.80)	0.02	(-0.22,-0.04)
6	(1.74,0.87)	0.05	(-0.07, 0.00)
7	(1.83,0.91)	0.0009	(-0.0003,-0.04)

Autor: J.Moré and S.Wright (1993)

2.3.3 Restricciones de Igualdad con condiciones de optimalidad

En el problema:

$$\text{P)} \quad Min\ f(x) \qquad (2.10)$$

$$\text{s.a.} \quad g_1(x) = b_1$$

$$g_2(x) = b_2$$

$$g_m(x) = b_n \qquad m \leq n$$

Definición 2.3.6 Decimos que un punto regular $x \in \mathbb{R}^n$ de las restricciones P) si y solo si:

$g_i(x) = b_i \qquad i = 1,2,\dots,m$

$\nabla g_1(x), \nabla g_2(x), \dots, \nabla g_m(x), son\ vectores\ linealmente\ independientes$

Dada la definición 2.2.4, la que usaremos para mostrar los mínimos locales varios casos teóricos, que nos ayuda para verificar el cumplimiento de algunas condiciones de regularidad de P).

Definición 2.3.7 Función Lagrangiana asociada a P):

$$L(x,\lambda) = f(x) + \sum_{i=1}^{m} \lambda_i (g_i(x) - b_i)$$

Considerando $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_m)^T$ como el vector de los Multiplicadores de Lagrange, se debe a Joseph Louis Lagrange, cuando queremos maximizar o minimizar funciones con varias variables, y está restringida ciertos recursos

escasos. Este método nos ayuda a tener un problema restringido en *n* variables en otro irrestricto de *n* + 1 variables y así de esta manera tomar las ecuaciones para su resolución. Tomamos (λ) como una variable escalar, llamado multiplicador de Lagrange, donde las restricciones estan determinados por una combinación lineal, las cuales incluyen a los multiplicadores (λ) como coeficientes. Rockafellar, R. T. (1976 ii)

Veremos los resultados teóricos consideran algunas propiedades que cumple un mínimo local, las que consideran, de manera particular en una función lagrangiana veremos que se trata un punto estacionario ver definición 2.3.27.

Teorema 2.3.9 (Condiciones necesarias de primer orden):

Sean las funciones continuamente diferenciales $f(x)\ y\ g_1(x), g_2(x), \dots, g_m(x)$, donde x^ un punto regular en las restricciones P), además mínimo local diremos que ∃ un vector $\lambda^{\wedge} \in \mathbb{R}^m$, de multiplicadores de Lagrange:

$$\nabla_x\ L(x^{\wedge}, \lambda^{\wedge}) = \nabla f(x^{\wedge}) + \sum_{i=1}^{m} \lambda_i^{\wedge}\ \nabla g_i(x^{\wedge}) = 0$$

Teorema 2.3.10 (Condiciones necesarias y suficientes de 2do orden).

Dadas las funciones $f, g_1, g_2, \dots, g_m$ son dos veces continuamente diferenciables con un punto regular $x^{\wedge} \in \mathbb{R}^n$ en las restricciones de P) con $\lambda^{\wedge} \in \mathbb{R}^n$, cumplen:

$\nabla_x L(x^{\wedge}, \lambda^{\wedge}) = \nabla f(x^{\wedge}) + \sum_{i=1}^{m} \lambda_i^{\wedge} \nabla g_i(x^{\wedge}) = 0$

y que $D^2 f(x^{\wedge}) + \sum_{i=1}^{m} \lambda_i^{\wedge}\ D^2 g_i(x^{\wedge})$

es una matriz definida positiva donde x^ es un mínimo local de P).

Ejemplo 2.7

$$\begin{aligned} Min \quad & x_1^2 + x_2^2 \\ sa: \quad & x_1 + x_2 = 4 \\ & \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2 \end{aligned}$$

Hallaremos la solución óptima utilizando las condiciones de optimalidad:

$f(x) = x_1^2 + x_2^2 \quad ; m = 1 \qquad\qquad g(x) = x_1 + x_2 - 4 = 0$

$$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i (g_i(x) - b_i)\, L(X_1, X_2, \lambda) = X_1^2 + X_2^2 + \lambda (X_1 + X_2 - 4)$$

$$\nabla f(x) = (2x_1,\ 2x_2)^T \qquad\qquad \nabla g(x) = (1,\ 1)^T$$

Luego, las condiciones de primer orden son, ver Teorema 2.3.9:

$$\nabla_x\ L(x^{\wedge}, \lambda^{\wedge}) = \nabla f(x^{\wedge}) + \sum_{i=1}^{m} \lambda_i^{\wedge}\ \nabla g_i(x^{\wedge}) = 0$$

$\frac{\partial \mathrm{L}}{\partial X_1} = 2X_1 + \lambda = 0$

$\frac{\partial \mathrm{L}}{\partial X_2} = 2X_2 + \lambda = 0 \quad \ldots . Factibilidad$

Solucionando el sistema de ecuaciones $x_1 = x_2 = 2$, $\lambda = -4$: de un problema (P) en el sistema:

$$x_1 = 2, x_2 = 2$$

Luego verificamos las condiciones de segundo orden ver Teorema 2.3.10, se cumplen pues:

$$\nabla_x\ L(x^{\wedge}, \lambda^{\wedge}) = \nabla f(x^{\wedge}) + \sum_{i=1}^{m} \lambda_i^{\wedge}\ \nabla g_i(x^{\wedge}) = 0$$

$D^2 f(x^{\wedge}) + \sum_{i=1}^{m} \lambda_i^{\wedge}\ D^2 g_i(x^{\wedge})$

$D^2 f(x) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}; \qquad D^2 g(x) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

$$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} + \lambda \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

es definida positiva.

Notar que en x* se tiene:

$$-\nabla f(x^*) = \sum_{i=1}^{m} \lambda_i^*\ \nabla g(x^*)$$

$\nabla f(x) = (2x_1,\ 2x_2)^T \qquad\qquad \nabla g(x) = (1,\ 1)^T, donde\ x_1 = 2, x_2 = 2, \lambda = -4$

$$\nabla f(x) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix} \qquad \nabla g(x^*) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$-\nabla f(x^*) = \sum_{i=1}^{m} \lambda_i^* \ \nabla g(x^*) \ \Rightarrow -[4 \quad 4]^T = -4[1 \quad 1]^T$$

Ejemplo 2.8

Como invertir un recurso, donde el riesgo asociado (en miles de soles) se da en la siguiente tabla, de tal forma que se minimice la función $800y^2 + 400x^2 + 200xy + 400yz. + 1600z^2$

Alternativa	Porcentaje % del riesgo
X	10
Y	10
Z	15
Riesgo por retorno esperado	12

Solución

X_i : Porcentajes de los recursos a invertir en miles de soles en la alternativa tipo i (X, Y, Z)

$Min\ Z = 400{X_1}^2 + 800{X_2}^2 + 200X_1X_2 + 1600{X_3}^2 + 400X_2X_3$

$s.a$

$X_1 + X_2 + X_3 = 1 \qquad \ldots\ldots\ldots\ldots \ (\propto)$

$10X_1 + 10\,X_2 + 15\,X_3 = 12 \ \rightarrow \ 1{,}0X_1 + 1{,}0\,X_2 + 1{,}5\,X_3 = 1{,}2 \ldots. \ (\beta)$

$X_1, X_2, X_3 \geq 0$

Veremos la factibilidad de las restricciones, igualando $(\propto)\ y\ (\beta)$ se obtiene:

$X_3 = \frac{0{,}2}{0{,}5} = 0{,}40 \ldots\ldots (\varphi),$

Reemplazando en el Multiplicador de Lagrange:

$$L(x,\lambda) = f(x) + \sum_{i=1}^{m} \lambda_i (g_i(x) - b_i)$$

$$L(X_1, X_2, X_3, \lambda_1, \lambda_2)$$
$$= 400{X_1}^2 + 800{X_2}^2 + 200X_1X_2 + 1600{X_3}^2 + 400X_2X_3$$
$$+ \lambda_1 (X_1 + X_2 + X_3 - 1) + \lambda_2 (X_1 + X_2 + 1{,}5X_3 - 1{,}2)$$

y, luego las condiciones de primer orden de KKT se escriben:

$$\frac{\partial \mathrm{L}}{\partial X_1} = 800X_1 + 200X_2 + \lambda_1 + \lambda_2 = 0$$

$$\frac{\partial \mathrm{L}}{\partial X_2} = 1600X_2 + 400X_3 + \lambda_1 + \lambda_2 = 0$$

$$\frac{\partial \mathrm{L}}{\partial X_3} = 3200X_3 + 400X_2 + \lambda_1 + 1{,}5\,\lambda_2 = 0 \ldots Factibilidad$$

Resolviendo el sistema con las condiciones de primer orden de KKT se obtiene la solución óptima:

Por ejemplo

$$\frac{\partial \mathrm{L}}{\partial X_1} = 800X_1 + 200X_2 + \lambda_1 + \lambda_2$$
$$= 800\,(0.50) + 200(0.10) - 1380 + 1800 = 0$$

$$X_1 = 0{,}50$$

$$X_2 = 0{,}10$$

$$X_3 = 0{,}40$$

$$\lambda_1 = -1380$$

$$\lambda_2 = 1800$$

Verificando si se dan las condiciones de segundo orden:

$$\nabla_x L(\hat{x}, \hat{\lambda}) = \nabla f(\hat{x}) + \sum_{i=1}^{m} \hat{\lambda_i} \nabla g_i(\hat{x}) = 0$$

Luego, como

$$\nabla g_1 = (1,1,1)^T \; y \; \nabla g_2 = (1,1,1.5)^T \; son \; linealmente \; independientes,$$

Entonces, el problema siempre se cumple, por esto los valores reemplazados en la función objetivo es 390 000 soles.

Por el teorema de Weierstrass, ver **Teorema 2.3.20,** obtenemos una solución, si consideramos las variables $(X_1 \geq 0 \; , X_2 \geq 0, X_3 \geq 0$ en reemplazo de $X_1 > 0 \; , X_2 > 0, X_3 > 0$ respectivamente, es decir relajamos la positividad). Además, como se cumple $por \; (\varphi), \ldots X_3 = 0{,}40$

Por lo tanto, vemos que $X_1 = 0{,}50, X_2 = 0{,}10, X_3 = 0{,}40$ es la solución del problema. Con las condiciones de KKT podemos comprobar que el problema tiene una función objetivo y restricciones de la forma convexa los que nos indican que es óptimo.

2.3.4 Restricciones con Desigualdad y condiciones de optimalidad

Dado $P)$ un problema con restricciones de desigualdad

$$(P) \; min\{f(x) : \; g(x) \; \leq \; 0\}$$

donde $f : \mathbb{R}^n \to \mathbb{R}$ es una función con valor real, $g : \mathbb{R}^n \to \mathbb{R}^p$ es una función tal que $g(x) = (g_1(x), g_2(x), \ldots, g_p(x))$, con $g_i : \mathbb{R}^n \to \mathbb{R}$ y la notación $g(x) \leq 0$ significa que $g_i(x) \leq 0$, para todo $i = 1,2, \ldots, p$.

Definición 2.3.8 Para cada $\bar{x} \in X = \{x \in \mathbb{R}^n : g(x) \leq 0\}$, g_i es llamada restricción activa en $\bar{x}$ si $g_i(\bar{x}) = 0$. Análogamente llamamos a g_i restricción inactiva en $\bar{x}$ si $g_i(\bar{x}) < 0$.

Dado $\bar{x}$ denotaremos $I(\bar{x}) = \{i \in \{1,2, \ldots, p\} : \; g_i(\bar{x}) = 0\}$.

Definición 2.3.9 Un punto llamado regular de $(\mathrm{P}), \mathrm{donde} \; \bar{x} \in X = \{x \in \mathbb{R}^n :$ $g(x) \leq 0\}$ y $\{\nabla g_i(\bar{x})\} i \in I(\bar{x}) \; \exists$ y son linealmente independientes.

Teorema 2.3.11 (Condición de Karush-Kuhn-Tucker). Dado un mínimo local regular $x*$ del problema (P), $con\, f$ y $g_i, i \in I\,(x*)$, son continuas y diferenciables en $x*$ y $g_i, i \notin I\,(x*)$, decimos que existen $\lambda_i* \in \mathbb{R}$ únicos , para todo $i \in I\,(x*)$, dado

$$\begin{cases} \nabla f(x^*) + \sum_{i \in I(x^*)} \lambda_i^* \nabla g_i(x^*) = 0 \\ \lambda_i^* \geq 0, \;\; para\, todo \;\; i \in I(x^*). \end{cases}$$

Corolario 2.3.3 Con las hipótesis del teorema 2.3.9 y la función g_i diferenciables $\forall i \notin I(x^*)$, por lo que λ_i^*≥0 son únicos para todo i = 1,2,...,m :

$$\begin{cases} \nabla f(x^*) + \sum_{i=1}^{p} \lambda_i^* \nabla g_i(x^*) = 0 \\ \lambda_i^* g_i(x^*) = 0, \;\; i = 1,2 \ldots, p \\ \lambda_i^* \geq 0, \;\; i = 1,2 \ldots, p. \end{cases}$$

Observación 2.3.6 El resultado dado en el corolario 2.3.2, nos ofrece un método para obtener los candidatos a solución mediante la solución del sistema que denominaremos condiciones de optimalidad de (Karush-Kuhn-Tucker) KKT.

$$\begin{cases} \nabla f(x) + \sum_{i=1}^{p} \lambda_i \nabla g_i(x) = 0 \\ g_i(x) \leq 0, \;\; i = 1,2 \ldots, p \\ \lambda_i g_i(x) = 0, \;\; i = 1,2 \ldots, p \\ \lambda_i \geq 0, \;\; i = 1,2 \ldots, p. \end{cases}$$

Teorema 2.3.12 Si $f, g_i\,, i \in I(x^*)$, son dos veces diferenciables en x^*, g_i, $i \notin I(x^*)$siendo continuas en los puntos mínimo local regular x^* y x^* de (P), por lo tanto existen puntos únicos $\lambda_i^* \geq 0, i = 1, \ldots, m$, de manera que 1.

$$1.\, \nabla f(x^*) + \sum_{i \in I(x^*)} \lambda_i^* \nabla g_i(x^*) = 0.$$

$$2.\, v^T \nabla_{xx}^2 L(x^*, \lambda^*) v \geq 0, para\, todo\, v \in \mathbb{R}^n : \nabla g_i(x^*)^T v = 0\; y\; todo\; i \in I(x^*),$$

$$donde\; L(x, \lambda) = f(x) + \sum_{i=1}^{p} \lambda_i g_i(x).$$

Teorema 2.3.13 Sean f, g_i, $i \in I(x^*)$, dos veces diferenciables en x^*. Si $x^* \in X = \{x \in \mathbb{R}^n : g(x) \leq 0\}$ satisface: Entonces x^* es un mínimo local estricto de f en X.

$i. \nabla f(x^*) + \sum_{i \in I(x^*)} \lambda_i^* \nabla g_i(x^*) = 0, para\ algunos\ \lambda_i^* \geq 0, i \in I(x^*)$

$ii. v^T \left(\nabla^2 f(x^*) + \sum_{i \in I(x^*)} \lambda_i^* \nabla^2 g_i(x^*) \right) v > 0, para\ todo\ v \in \mathbb{R}^n, v \neq 0: \nabla g_i(x^*)^T v$

$= 0, y\ todo\ i \in I(x^*).$

Entonces x^* es un minimo local estricto de f en X.

Método de multiplicadores de Lagrange (λ)

Usaremos las derivadas parciales ó podemos usar diferenciales totales y la regla de la cadena. Podemos obtener una función que sea igual a cero usando alguna función implícita en las que se dan condiciones de Kuhn-Tucker, ver Bazaara MS, Sherali HD, Shetty CM (1993).

Las condiciones de Kuhn-Tucker, consideran:

Dado (x, y) un punto máximo local $f(x, y) \leq 0$, hay un valor λ no negativo donde qué λ y (x, y) cumplen con las condiciones de optimalidad de kuhn-tucker, ver Teorema 2.3.11

$$\frac{df(x,y)}{dx} - \lambda \frac{dg(x,y)}{dx} = 0$$

$$\frac{df(x,y)}{dy} - \lambda \frac{dg(x,y)}{dy} = 0$$

$$\lambda g(x, y) = 0$$

$$g(x, y) \leq 0$$

Esta condición es suficiente, si $f(x, y)\ y\ g(x, y)$ son cóncavas hacia abajo por esto se trata de un punto máximo $f(x, y)$ análogamente diremos que este es punto mínimo de $-f(x, y)$.

Ejemplo 2.9

Minimizar el costo, $f(x, y) = 6x^2 + 3y^2$,

s.a.

si $x + y \geq 24$

Aplicando Kuhn-Tucker tenemos

$g(x,y) = x + y - 24 \geq 0$

$$\frac{df(x,y)}{dx} - \lambda \frac{dg(x,y)}{dx} = 0 \quad \rightarrow \quad \frac{d(6x^2 + 3y^2)}{dx} - \lambda \frac{d(x + y - 24)}{dx} = 0$$

$$\frac{d(6x^2+3y^2)}{dx} - \lambda \frac{d(x+y-24)}{dx} = 0 \rightarrow 12x - \lambda(1) = 0 - - - - \alpha$$

$$\frac{df(x,y)}{dy} - \lambda \frac{dg(x,y)}{dy} = 0 \quad \rightarrow \quad \frac{d(6x^2 + 3y^2)}{dy} - \lambda \frac{d(x + y - 24)}{dy} = 0$$

$$\frac{d(6x^2 + 3y^2)}{dy} - \lambda \frac{d(x + y - 24)}{dy} = 0 \rightarrow 6y - \lambda(1) = 0 - - - - \beta$$

$Verificamos\ las\ condiciones\ de\ KKT$

$\lambda g(x,y) = 0 \quad \rightarrow \quad \lambda (x + y - 24) = 0$

$g(x,y) \leq 0 \quad \rightarrow \quad (x + y - 24) \leq 0,$

$\lambda (x + y - 24) = 0$

$(x + y - 24) \leq 0$

$De \quad \alpha \rightarrow 12x - \lambda = 0 \rightarrow \lambda = 12x$

$De \ \beta \rightarrow 6y - \lambda(1) = 0 \rightarrow \lambda = 6y$, Igualando los términos $\rightarrow \ 12x = 6y \rightarrow y = 2x$

$Para\ que\ se\ cumpla\ \lambda(x + y - 24) = 0 \rightarrow$ Decimos que $\lambda = 0$

$(x + y - 24) = 0$

Si $\lambda = 0$ entonces $x = y = 0$, vemos que el punto (0, 0) no cumple la condición de KKT, ver Teorema 2.3.9.

Remplazando en el punto (0, 0)

Reemplazando vemos que $(0 + 0 - 24) \leq 0$, por lo tanto, el punto P (0,0) no es óptimo .

Pero si $\lambda \neq 0$, $(x + y - 24) = 0 \rightarrow (x + 2x - 24) = 0 \rightarrow x = 8 \ y \ y = 16$

Donde los puntos P (8, 16) satisfacen la condición de KUHN-TUCKER

$(8 + 16 - 24) \geq 0$, entonces el punto (8, 16) es óptimo

Vemos que la función de costo $f(x, y) = (6x^2 + 3y^2)$ evaluada en $(8, 16)$, es un mínimo ya que la funcion de costo es cóncava hacia arriba sujeta a la restricción $x + y - 24 \geq 0$

2.3.5 Elementos de convexidad

Definición 2.3.10 Un conjunto $C \subseteq \mathbb{R}^n$ es llamado convexo si $C = \emptyset$ o para todo $x, y \in C$ y todo $\alpha \in [0,1]$ tenemos $\alpha x + (1 - \alpha)y \in C$. Donde $A \subset \mathbb{R}^n$ es un conjunto llamado *afín*, $\forall\, x, y \in A$ y $\forall\, t \in \mathbb{R}, tx + (1 - t)y \in A$. Teniendo un subconjunto $K \in \mathbb{R}^n$ será un *cono* si, $\forall\, d \in K\ y\ t \geq 0, td \in K$.

Definición 2.3.11 Sea un conjunto convexo C no vacío y $f: C \subset \mathbb{R}^n \to \mathbb{R}$ con función f llamada convexa si cumple:

$f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2)$, para todo $x_1, x_2 \in C$ y todo $\alpha \in [0,1]$. f es *estrictamente convexa* en C si $f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2)$, es estricta $\forall\, x_1, x_2 \in C, x_1 \neq x_2\ y\ \alpha \in (0,1)$.

Definición 2.3.12: Dado un conjunto convexo C no vacío y $f: C \subset \mathbb{R}^n \to \mathbb{R}$ una función f llamada *fuertemente convexa* con módulo $\gamma > 0$ si

$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y) - \gamma\alpha(1 - \alpha)\|x - y\|^2$, para todo $x, y \in C$ y $\alpha \in [0,1]$.

Observación 2.3.7 Observemos que de las definiciones 2.3.9, 2.3.10 decimos que si f es fuertemente convexa → f es estrictamente convexa implica que si f es convexa, pero los recíprocos no siempre son verdaderos veremos en el ejemplo 2.10.

Ejemplo 2.10 Mostraremos varias funciones convexas que son fuertemente convexas y estrictamente convexas.

1. $f(x) = x^2$ es fuertemente convexa en $\mathbb{R}$.
2. $f(x) = e^x$ no es fuertemente convexa en $\mathbb{R}$ (pero si estrictamente convexa).
3. $f(x) = x$ no es estrictamente convexa en $\mathbb{R}$ (pero es convexa).

Definición 2.3.13 Sea un conjunto convexo C no vacío y $f: C \subset \mathbb{R}^n \to \mathbb{R}$ una función f es llamada cóncava (estrictamente cóncava) si $-f$ es convexa (estrictamente convexa).

Definición 2.3.14 Sea un conjunto convexo C no vacío y $f: C \subset \mathbb{R}^n \to \mathbb{R}$ una función f es llamada *fuertemente cóncava* si $-f$ es fuertemente convexa en C.

Definición 2.3.15 Una función $f : \mathbb{R}^n \to \mathbb{R}$ es *afín* si para todo $x, y \in \mathbb{R}$ y $\alpha \in \mathbb{R}$ se tiene que

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

Teorema 2.3.14 Dado un conjunto convexo $C \neq \emptyset$. $f: C \to \mathbb{R}$, es una función convexa ssi el conjunto $Epi(f) := \{(x, r) \in C \times \mathbb{R} : f(x) \leq r\}$ es convexo en $\mathbb{R}^n \times \mathbb{R}$. T

Teorema 2.3.15 Dado un conjunto convexo $C \neq \emptyset$, con $f: C \to \mathbb{R}$, una función convexa. Vemos que el conjunto de nivel $L_f(\alpha) := \{x \in C: f(x) \leq \alpha\}$ es convexo en $\mathbb{R}^n$ para todo $\alpha \in \mathbb{R}$.

Observación 2.3.8 El recíproco del Teorema 2.3.13 no es verdadero, considere por ejemplo 2.11, la función $f(x) = lnx$, no es convexa, pero podemos observar que el conjunto de nivel $L_f(\alpha)$ es convexo,$\forall \alpha \in \mathbb{R}$.

Proposición 2.3.1 Sea $f : \mathbb{R}^n \to \mathbb{R}$. Entonces f es convexa si, y solamente si, para todo $x \in \mathbb{R}^n$, $d \neq 0, g(t) = f(x + td)$ es convexa en $\mathbb{R}$.

Definición 2.3.16 Sea un conjunto convexo no vacío,$D \subseteq \mathbb{R}^n$, es si y solo si:

$$\lambda x + (1-\lambda)y \in D, para\ todo\ x \in D, y \in D\ , con\ \lambda \in [0,1]\ .$$

Definición 2.3.17 Decimos que un operador $T: H \rightrightarrows H$ es de Lipschitz con constante $L > 0$, si para todo $x, x' \in D$

$$\|y - y'\| \leq L\|x - x'\|, para\ todo\ y \in T(x), y\ para\ todo\ y' \in T(x'),$$

Teorema 2.3.16 Sea $f: C \to \mathbb{R}$ una función convexa y $x^0 \in int(C)$, entonces f es *localmente Lipschitziana* en x^0.

Corolario 2.3.4 Dado C conjunto convexo y $f: C \to \mathbb{R}$ una función convexa, si $\bar{x} \in int(C)$, entonces f es continua en $\bar{x}$.

Corolario 2.3.5 Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función convexa, entonces f es continua en $\mathbb{R}^n$.

Teorema 2.3.17 Sean C un conjunto convexo con $int(C) \neq \emptyset$ y $f: C \to \mathbb{R}$ una función convexa, entonces para $x \in int(C)$ y $d \in \mathbb{R}^n, d \neq 0$, donde:

1. Donde φ es una función no decreciente

2. Si $f'(x,d)$ la derivada direccional ∃.

Lema 2.3.1 Sean C un conjunto convexo, $f: C \to \mathbb{R}$ una función convexa y $x \in int(C)$, entonces $f'(x,.) : \mathbb{R}^n \to \mathbb{R}$ es *homogénea de grado 1*, convexa y para todo $y \in C$: $f(y) \geq f(x) + f'(x, y-x)$.

Definición 2.3.18 Sea C un conjunto convexo de $\mathbb{R}^n$ y $f: C \to \mathbb{R}$ una función. El punto $s \in \mathbb{R}^n$ es un *subgradiente* de f en $\bar{x} \in C$ si

$f(y) \geq f(\bar{x}) + \langle s, y - \bar{x} \rangle$, para todo $y \in C$.

El conjunto de subgradientes de f en $\bar{x}$ denominado *subdiferencial* de f en $\bar{x}$ y es denominado $\partial f(\bar{x})$, donde, $\partial f(\bar{x}) = \{s \in \mathbb{R}^n : f(y) \geq f(\bar{x}) + \langle s, y - \bar{x} \rangle$ para todo $y \in C\}$.

Definición 2.3.19

Decimos que $(\hat{\partial}\,)$ subdiferencial de Fréchet es:

$$\hat{\partial} f(x) = \left\{ s \epsilon \mathbb{R}^n : \lim\ inf_{y \to x} \frac{f(y) - f(x) - \langle s, y - x \rangle}{\|y - x\|} \geq 0 \right\}$$

Definición 2.3.20 Un operador $T: H \rightrightarrows H$ es denominado monótono, si para todo

$x, x' \in D(T)$. $\langle y - y', x - x' \rangle \geq 0$, para todo $y \in T(x)$ y para todo $y' \in T(x')$.

El operador T es denominado estrictamente monótono cuando la desigualdad anterior es estricta para $x \neq x'$.

Definición 2.3.21 Un operador $T: H \rightrightarrows H$ es denominado monótono maximal (m.m) si T es monótono y el grafo de T no está contenido propiamente en el grafo de cualquier otro operador monótono. En otras palabras, si T es monótono y para todo T' monótono tal que $T(x) \subset T'(x)$ para todo x, se cumple que $T = T'$.

Ejemplo 2.12

Sea $f: \mathbb{R} \to \mathbb{R}$ definida por $f(x) = |x|$, entonces,

$$\partial f(x) = \begin{cases} +1, & \text{si } x > 0 \\ +1, & \text{si } x > 0 \end{cases}$$

Teorema 2.3.18 Sea C un conjunto convexo y $f: C \to \mathbb{R}$ una función convexa y $\bar{x} \in int(C)$ entonces $\partial f(\bar{x})$ es no vacio y acotado.

Corolario 2.3.6. Dado un conjunto convexo C y $f: C \to \mathbb{R}$ una función convexa, inferior semicontinua, $\bar{x} \in int(C) \to \partial f(\bar{x})$ es convexo no vacio y compacto.

Proposición 2.3.2 Sea C un conjunto convexo, $f: C \to \mathbb{R}$ una función convexa y $\bar{x} \in int(C)$. donde f es diferenciable en $\bar{x}$, $\to \partial f(\bar{x}) = \{\nabla f(\bar{x})\}$.

Teorema 2.3.19 Sea C un conjunto convexo y abierto de $\mathbb{R}^n$ y $f: C \to \mathbb{R}$ una función diferenciable en C. Entonces, f es convexa ssi:

$f(x) \geq f(y) + \langle \nabla f(y), x - y \rangle$, para todo $x, y \in C$.

En la solución del problema de programación no lineal nos asegura la unicidad y existencia.

Teorema 2.3.20 (Weierstrass). Dado una función continua f es y conjunto cerrado no vacío D, acotado de $\mathbb{R}^n$, → un problema P) tiene óptima solución.

Teorema 2.3.21 Dada una función continua f y D es un conjunto cerrado no vacío y además f cumple que: $\lim_{|x| \to +\infty} f(x) = +\infty$, entonces un problema P) tiene solución óptima.

La unicidad en la solución óptima nos confirma bajo ciertas condiciones muy especiales.

También es posible asegurar que, si el problema tiene un mínimo local, es mínimo global.

Dado el problema P) debemos ver si el problema es convexo, vemos que la función objetivo y la región factible también son un conjunto convexo.

Definición 2.3.22. Decimos que $f: \mathbb{R}^n \to \mathbb{R}$ es una función convexa ssi:

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda) f(y)$$

para todo $x, y \in D(x \neq y)\ con\ \lambda \in [0,1]$

Si $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$ se verifica en forma estricta, entonces f es estrictamente convexa.

Definición 2.3.23

Hallaremos la longitud de paso que de un "descenso suficiente" de f aunque esta búsqueda es inexacta en relación con el valor $f(x_k)$.Ver Dennis J. and R. B. Schnabel (1996).

Consideramos f diferenciable, x_k, d_k,dados y fijos. Definimos $\phi(\mathrm{t}) = f(x_k + td_k)$ .Para un punto fijo $\alpha \in (0,1)$, tomamos los t, de manera que:

$$\phi(\mathrm{t}) \leq \phi(0) + \alpha\, \mathrm{t}\, \phi(0).$$

Las condiciones de Armijo se dan en términos de f.

$f(x_k + td_k) \leq f(x_k) + \alpha\, \mathrm{t}\nabla f(x_k)^T d_k$

Esta condición de Armijo es equivalente al valor de la derivada direccional y la longitud del paso t_k, .

Donde el algorítmo, se da de la forma:

1. Iniciar con t = 1.

2. Si $f(x_k + td_k) \leq f(x_k) + \alpha\, \mathrm{t}\nabla f(x_k)^T d_k$, entonces definir $t_k = t$;

3. Caso contrario, debemos reducir t con tomando diferentes criterios hasta que se de la condición.

Posible manera de cumplir con t:

calcular t (puede ser el punto medio del intervalo)

Teorema 2.3.22 Dada una función dos veces diferenciable y continua, las condiciones son equivalentes:

1. Una función f es convexa
2. $f(x) \geq f(y) + \nabla f^T(y)(x-y)$ para dos puntos x e y.

3. Las segundas derivadas parciales de f, se trabajan en la hessiana denotada por $D^2f(x)$, $\forall$ x es positiva semi definida .

Podemos tomar un conjunto cualquiera convexo o no, cumpliendo la definición siguiente:

Teorema 2.3.23 Dado un conjunto convexo C abierto de $\mathbb{R}^n$ y una $función$ $f: C \to \mathbb{R}$,*dos veces continuamente diferenciable* en C. Donde f es convexa sii $\nabla^2 f(x) \geq 0, \forall\ x \in C$.

2.3.6 Optimización Convexa

El problema de optimización convexa en el espacio euclidiano $\mathbb{R}^n$ consiste en resolver el siguiente problema de optimización:

$$opt\{f(x) : x \in X\}$$

donde f es una función convexa $f: X \subset \mathbb{R}^n \to \mathbb{R}$ y X un conjunto convexo en $\mathbb{R}^n$.

El problema de optimización lineal

Decimos que $min\{\langle c, xi\ \rangle: Ax = b, x \geq 0\}$,es un problema de optimización convexa, sujeta $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m, c \in \mathbb{R}^n$.

Ejemplo 2.13 El problema de optimización

$$Min\left\{\frac{1}{2}\langle Qx, x\rangle - \langle c, x\rangle: Ax\ =\ b, x\ \geq\ 0\right\},$$

Donde una matriz simétrica Q es positiva semidefinida, del problema de optimización convexa.

Una mportante propiedad de minimización convexa es que un mínimo local es mínimo global el que mostramos el resultado siguiente.

Teorema 2.3.24 Si una función f convexa $f: X \subset \mathbb{R}^n \to \mathbb{R}$ definida en un subconjunto convexo $X \to$ todo mínimo local es global.

Teorema 2.3.25 Considere el problema de minimización convexa

$min\{f(x) : x \in X \subset \mathbb{R}^n\}$.

Diremos que el conjunto de soluciones óptimas, $X*$,de un probelma es convexo.

Teorema 2.3.26 Dada una función estrictamente convexa $f: X \subset \mathbb{R}^n \to \mathbb{R}$ está definida en el subconjunto convexo X donde $\bar{x}$ es un mínimo de $f \to \bar{x}$ es el punto mínimo global único.

Teorema 2.3.27 Considere el problema de minimización

$$min\{f(x) : x \in X\}$$

donde una función convexa $f: X \subset \mathbb{R}^n \to \mathbb{R}$ es un conjunto convexo $X \neq \emptyset$ y f es diferenciable en $\mathbb{R}^n \to \bar{x}$ es un mínimo global de f en X si y solo si $\langle \nabla f(\bar{x}), x - \bar{x} \rangle \geq 0$, para todo $x \in X$.

Corolario 2.3.7 Considere el problema de minimización sin restricciones

$$min\{f(x) : x \in \mathbb{R}^n\},$$

donde f es una función diferenciable en $\mathbb{R}^n$. Entonces, $\bar{x}$ es un mínimo global si, y solo si, $\nabla f(\bar{x}) = 0$.

Definición 2.3.24 Una función diferenciable pseudoconvexa, $f: \mathbb{R}^n \to \mathbb{R}$ es, $\forall$ par de puntos distintos x, y tenemos $\langle \nabla f(x), y - x \rangle \geq 0$, entonces $f(y) \geq f(x)$.

Definición 2.3.25 Dado un conjunto convexo C y $f: C \to \mathbb{R}$ una función real. f llamada función cuasi convexa en C, $\forall\ x, y \in C$, $t \in [0,1]$, se cumple que

$$f(tx + (1-t)y) \leq max\{f(x), f(y)\}.$$

Teorema 2.3.28 Sea C un conjunto convexo y $f: C \to \mathbb{R}$ una función. f es cuasiconvexa en C si, y solamente si, para cada $c \in \mathbb{R}$ el conjunto de nivel inferior de f, $Lf(c) = \{x \in C: f(x) \leq c\}$, es convexo.

Definición 2.3.26

Si una función f: $R^n \to R$, se denomina Hessiano de f, (Hf)

$$Hf(x_1, x_2.., x_n) = \begin{bmatrix} \frac{\partial^2 f(x_1,..,x_n)}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 f(x_1,..,x_n)}{\partial x_1 \partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial^2 f(x_1,..,x_n)}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(x_1,..,x_n)}{\partial x_n \partial x_n} \end{bmatrix}$$

La función Hessiana nos ayuda para dar la condición suficiente de un punto estacionario el cual pude ser un máximo o mínimo (relativo)

Definición 2.3.27

Si una función en un punto tiene gradiente cero, entonces decimos que es una condición necesaria donde el punto es un máximo o mínimo (local), Guler,O [1991], Kaplan y Tichatsche(1998) probaron que la sucesion finaliza en un número finito de iteraciones en un punto estacionario o de acumulación de $\{x_k\}$ es estacionario.

2.3.7 Optimización Cuasi-convexa

El problema de optimización cuasi-convexa ver definición 2.3.25, en el espacio euclidiano $\mathbb{R}^n$, Arrow KJ, Enthoven AC (1961), consiste en resolver el siguiente problema de optimización:

$$opt\{f(x) : x \in X\}$$

Decimos que una función $f: X \subset \mathbb{R}^n \to \mathbb{R}$ cuasi convexa y X un conjunto convexo en $\mathbb{R}^n$.

Ejemplo 2.14

Una función cuasi convexa en economía, es la función de Coob-Douglas $f(x_1, x_2) = -x_1^{\alpha_1} x_2^{\alpha_2}$ cuando $\alpha_1 + \alpha_2 > 1$

Dado una función convexa h(x) y el conjunto convexo $D = \{x \in \mathbb{R}^n / h(x) \leq d\}$ para un escalar real d.

Se puede probar que la intersección de conjuntos convexos es también un conjunto convexo.

(P) $\min f(x)$

$$s.a. \quad h_r(x) \leq d_r \quad (r = 1,2, \dots, l)$$

Tenemos funciones convexas $f(x)\ y\ h_r(x)$ para d, $r = 1,2, \dots, l$ es un problema convexo, donde la intersección de los conjuntos convexos es el dominio de factibilidad.

$$D_r = \{x \in \mathbb{R}^n / h_r(x) \leq d_r \ \ para\ r = 1, \dots, l\}$$

Teorema 2.3.29 Si un problema convexo, donde x* un mínimo local de un problema → x* es una solución óptima o mínimo global de un problema, si x* es una solución óptima única cuando la función es estrictamente convexa .

Los problemas de programación no lineal tienen como principal dificultad la inclusión de restricciones no lineales de la forma $g(x) = b$, dado una sucesión de puntos $\{x \in \mathbb{R}^n : g(x) = b\}$ cuando g(x) es una función no lineal casi siempre no es convexo. Vemos que algunos de los problemas de programación no lineal son no convexos por eso no podemos garantizar que la solución de un problema sea óptima.

Ejemplo 2.15 Como puede verse en este ejemplo, resolviendo gráficamente, esta geometría de los problemas también cambia respecto a lo visto en programación lineal.

Veremos en el problema siguiente:

$$Min \quad (x_1 - 3)^2 + (x_2 - 4)^2$$

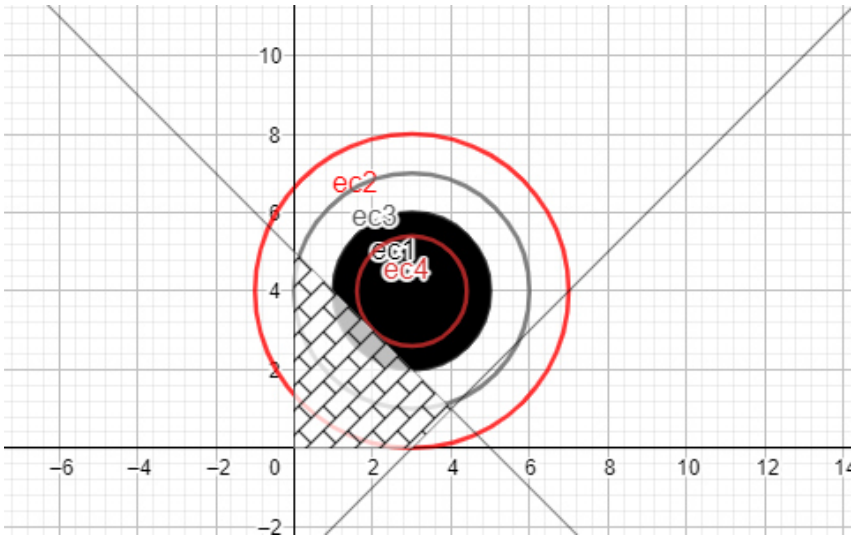
$$\text{s.a.} \quad (x_1 + x_2) \leq 5$$

$$(x_1 - x_2) \leq 3$$

$$x_1 \geq 0, x_2 \geq$$

Figura 2.3

Gráfica de un modelo no lineal, dominio de factibilidad del problema



Nota: En la figura 2.3., tenemos x* como solución óptima $x_1^* = 2, x_2^* = 3$, con un único punto en la curva de nivel con un valor menor cruza con puntos factibles de la región factible.

Verificamos que los valores no coinciden con los vértices del dominio de factibilidad, aunque esta solución se alcanza en el limite del conjunto.

Además, vemos que a diferencia de programación lineal, no siempre se da.

Ejemplo 2.16 Modificando el problema del ejemplo 2.15:

$$\text{Min} \quad (x_1 - 2)^2 + (x_2 - 2)^2$$

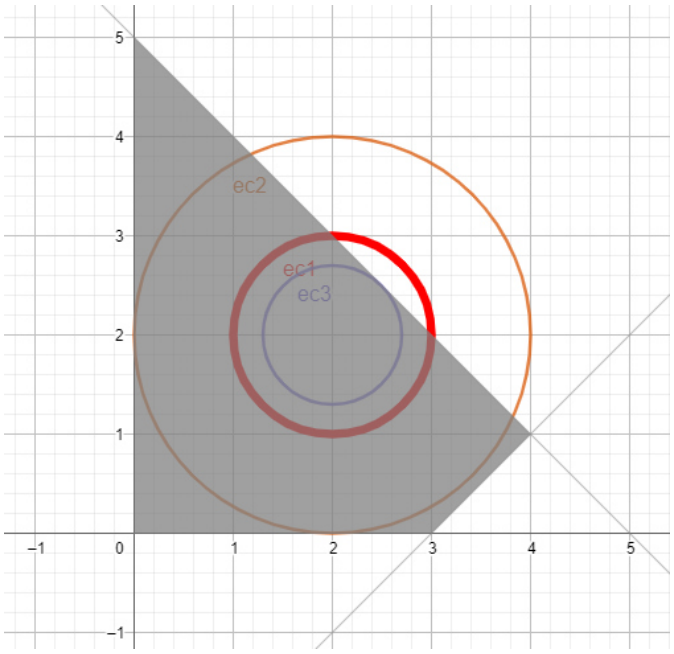
$$\text{s.a} \quad (x_1 + x_2) \leq 5$$

$$(x_1 - x_2) \leq 3$$

$$x_1 \geq 0, x_2 \geq 0$$

Figura 2.4

Gráfica de un modelo no lineal con dominio de factibilidad del problema



Nota: Vemos que la solución cambia a lo hallado en la *figura 2.3*, cuando la solución óptima era $x_1^* = 2, x_2^* = 2$, ahora vemos que la nueva solución cuyo dominio de factibilidad pertenece al interior de la región.

Observamos gráficamente en la *figura 2.3*, que la solución óptima no esta en el interior del dominio de factibilidad.

2.3.8 Métodos de optimización con restricciones

a) **Método de activación de restricciones**

Se usa este método en problemas que sólo con restricciones de desigualdad.

Un problema sin restricción tiene una solución óptima que no necesariamente cumple con algunas de las restricciones, se toma k restricciones como si fueran ecuaciones resolviendo el problema con restricciones hasta encontrar los valores partiendo de restricciones activas cumpliendo además con las restricciones no utilizadas.

1: Resolver el problema sin restricciones. Parar si el punto óptimo cumple con todas las restricciones, si no se cumple, tomar k=1 ir al paso siguiente.

2: Tomar una de las k restricciones y encontrar la solución factible que cumpla con condiciones de optimalidad KKT entonces parar. Si no activar otro conjunto de k restricciones y repetir. Cuando se probaron con todos los conjuntos de k restricciones y no se encuentra una solución factible entonces ir a 3.

3: Tomemos k = N (donde N es el número total de restricciones) no hay una solución factible. Si no, tomar k= k+1 luego ir a 2:

Ejemplo 2.17. Dado el problema:

$$Min \quad (2x_1 - 5)^2 + (2x_2 - 1)^2$$

s.a. $\quad x_1 + 2x_2 \leq 2$

$$x_1, x_2 \geq 0$$

Tomando el problema irrestricto con solución.

$$x_1^* = 5/2 \quad y \quad x_2^* = 1/2$$

Cuando se usa: $\nabla f(x) = 0$

Vemos que el punto no cumple la restricción:

$$h_1(x_1, x_2) = x_1 + 2x_2 \leq 2$$

Reemplazando esto en la restricción lineal activa:

$$\nabla f(x_1, x_2) + \mu_1 \nabla h_1(x_1, x_2) = 0$$

$$h_1(x_1, x_2) = 2$$

Hallamos la solución óptima:

$$\hat{x_1} = 2.2$$

$$\hat{x_2} = -0.10$$

$\mu_1^{\wedge} = -2.4$ pero no cumple que $x_2 \geq 0$

Pero si tomamos a $x_1 = 0$ vemos que podemos hallar un mínimo local:

$x_1 = 0, x_2 = 1 / 2$

$\mu_1 = 0, \mu_2 = 0, \mu_3 = 0$

Se puede hallar otro mínimo local tomando una restricción activa

$x_1 + 2x_2 \leq 2 \;\; y \;\; x_2 = 0$

Hallando el punto $x_1 = 2, x_2 = 0$

b) ***Método de Frank – Wolfe.***

Usamos este método cuando la función objetivo es convexa no lineal con todas las restricciones lineales. Se sustituye la función objetivo aproximando en una secuencia de funciones lineales, que dan como resultado una secuencia de problemas de programación lineal.

Ejemplo 2.18

$Min\ Z = -10{X_1}^2 - {X_2}^2 + 10X_1 + 25X_2 - 4X_1 X_2$

s.a.

$$X_1 + 2X_2 \leq 10$$

$$X_1 + X_2 \leq 9$$

$X_1, X_2 \geq 0$

Aplicando multiplicadores de Lagrange

$$L(X_1, X_2, \lambda_1, \lambda_2) = -10{X_1}^2 - {X_2}^2 + 10X_1 + 25X_2 - 4X_1 X_2 + \lambda_1 (X_1 + 2X_2 - 10) + \lambda_2 (X_1 + X_2 - 9) - \mu_1(X_1) - \mu_2(X_2)$$

Las primeras derivadas parciales son:

$$\frac{\partial \mathrm{L}}{\partial X_1} = 10 - 20X_1 - 4X_2 + \lambda_1 + \lambda_2 - \mu_1 = 0$$

$$\frac{\partial \mathrm{L}}{\partial X_2} = 25 - 2X_2 - 4X_1 + 2\lambda_1 + \lambda_2 - \mu_2 = 0$$

El problema de programación lineal equivalente al original de acuerdo con el método de Wolfe ʎ, μ

$$Min\ W = A_1 + A_2$$

S.a.

$$20X_1 + 4X_2 - \lambda_1 - \lambda_2 + \mu_1 + A_1 = 10$$

$$4X_1 + 2X_2 - 2\lambda_1 - \lambda_2 + \mu_2 + A_2 = 25$$

$$X_1 + 2X_2 + S_3 = 10$$

$$X_1 + X_2 + S_4 = 9$$

En el software POM, usaremos esta nomenclatura,

$$\lambda_1 = L1; \lambda_2 = L2; \mu_1 = M1; \mu_2 = M2$$

	X1	X2	L1	L2	M1	M2	A1	A2	S3	S4		RHS	Equation form
Minimize	0	0	0	0	0	0	1	1	0	0			Min A1 + A2
Constraint 1	20	4	-1	-1	1	0	1	0	0	0	=	10	20X1 + 4X2 - L1 - L2 + M1 + A1
Constraint 2	4	2	-2	-1	0	1	0	1	0	0	=	25	4X1 + 2X2 --2L1 - L2 + M2 + A2
Constraint 3	1	2	0	0	0	0	0	0	1	0	=	10	X1 + 2X2 + S3 = 10
Constraint 4	1	1	0	0	0	0	0	0	0	1	=	9	X1 + X2 + S4 = 9

Con las siguientes restricciones de holgura complementaria:

$$\lambda_1 x_1 = 0$$

$$\lambda_2 x_2 = 0$$

Utilizando el método simplex se tiene que la solución básica inicial es:

Cj	Basic Variables	0 X1	0 X2	0 L1	0 L2	0 M1	0 M2	1 A1	1 A2	0 artfcl 1	0 artfcl 2	0 slack 3	0 slack 4	Quantity
Iteration 1														
0	artfcl 1	20	4	-1	-1	1	0	1	0	1	0	0	0	10
0	artfcl 2	4	2	-2	-1	0	1	0	1	0	1	0	0	25
0	slack 3	1	2	0	0	0	0	0	0	0	0	1	0	10
0	slack 4	1	1	0	0	0	0	0	0	0	0	0	1	9
	zj	-24	-6	3	2	-1	-1	0	0	0	0	0	0	35
	cj-zj	24	6	-3	-2	1	1	1	1	0	0	0	0	

$$W = -35;\ A_1 = 10;\ A_2 = 25; S_3 = 10; S_4 = 9$$

En la iteración 2, entra X_1. El punto extremo luego se recalcula

Cj	Basic Variables	0 X1	0 X2	0 L1	0 L2	0 M1	0 M2	1 A1	1 A2	0 artfcl 1	0 artfcl 2	0 slack 3	0 slack 4	Quantity
Iteration 2														
0	X1	1	0.2	-0.05	-0.05	0.05	0	0.05	0	0.05	0	0	0	0.5
0	artfcl 2	0	1.2	-1.8	-0.8	-0.2	1	-0.2	1	-0.2	1	0	0	23
0	slack 3	0	1.8	0.05	0.05	-0.05	0	-0.05	0	-0.05	0	1	0	9.5
0	slack 4	0	0.8	0.05	0.05	-0.05	0	-0.05	0	-0.05	0	0	1	8.5
	zj	0	-1.2	1.8	.8	.2	-1	1.2	0	1.2	0	0	0	23
	cj-zj	0	1.2	-1.8	-0.8	-0.2	1	-0.2	1	-1.2	0	0	0	

$W = -23;\ X_1 = 0.5; A_2 = 23; S_3 = 9.5\ ; S_4 = 8.5$

En la iteración 3, entra X_2 y sale X_1. El punto extremo luego se recalcula

Cj	Basic Variables	0 X1	0 X2	0 L1	0 L2	0 M1	0 M2	1 A1	1 A2	0 artfcl 1	0 artfcl 2	0 slack 3	0 slack 4	Quantity
Iteration 3														
0	X2	5	1	-0.25	-0.25	0.25	0	0.25	0	0.25	0	0	0	2.5
0	artfcl 2	-6.0	0	-1.5	-0.5	-0.5	1	-0.5	1	-0.5	1	0	0	20.0
0	slack 3	-9.0	0	0.5	0.5	-0.5	0	-0.5	0	-0.5	0	1	0	5.0
0	slack 4	-4.0	0	0.25	0.25	-0.25	0	-0.25	0	-0.25	0	0	1	6.5
	zj	6	0	1.5	.5	.5	-1	1.5	0	1.5	0	0	0	20.0
	cj-zj	-6.0	0	-1.5	-0.5	-0.5	1	-0.5	1	-1.5	0	0	0	

$W = -20;\ X_2 = 2.5;\ A_2 = 20; S_3 = 5; S_4 = 6.5$

La última iteración (A_1=0 y A_2=0), Luego el punto extremo es:

Cj	Basic Variables	0 X1	0 X2	0 L1	0 L2	0 M1	0 M2	1 A1	1 A2	0 artfcl 1	0 artfcl 2	0 slack 3	0 slack 4	Quantity
Iteration 5														
0	X2	5	1	-0.25	-0.25	0.25	0	0.25	0	0.25	0	0	0	2.5
0	M2	-6.0	0	-1.5	-0.5	-0.5	1	-0.5	1	-0.5	1	0	0	20.0
0	slack 3	-9.0	0	0.5	0.5	-0.5	0	-0.5	0	-0.5	0	1	0	5.0
0	slack 4	-4.0	0	0.25	0.25	-0.25	0	-0.25	0	-0.25	0	0	1	6.5
	zj	0	0	0	0	0	0	2	2	0	0	0	0	0
	cj-zj	0	0	0	0	0	0	-1	-1	0	0	0	0	

$W = 0; X_2 = 2.5;\ \mu_2 = 20; S_3 = 5\ ; S_4 = 6.5$

$Min\ Z = -10{X_1}^2 - {X_2}^2 + 10X_1 + 25X_2 - 4X_1 X_2 = -(6.25) + 25(2.5) = 56.25$

La solución óptima es: $X_1 = 0;\ X_2 = 5; Z = 56.25$

c) Problemas con restricciones de desigualdad e igualdad.

Tomamos el problema P):

P) Min $f(x)$

s.a. $g_i(x) = b_i \quad i = 1,2,\dots,m$

$h_r(x) = d_r \quad r = 1,2,\dots,l$

En este caso decimos que x^ es un punto regular de las restricciones del problema ssi:

$g_i(\hat{x}) = b_i, \quad i = 1,2,\dots,m$

$h_r(x^{\wedge}) \neq d_r \ , \quad r = 1,2, \dots, l$

$\nabla g_1(x^{\wedge}), \nabla g_2(x^{\wedge}), \dots, \nabla g_m(x^{\wedge}), \nabla h_i(x^{\wedge})$ *vectores linealmente independientes*

$Con\ j \ \in \{r \ / \ h_r\,(x^{\wedge}) \leq d_r \ \ para \ r = 1, \dots, l\,\}$

Teorema 2.3.30 (Condiciones necesarias de 1er orden de KKT).
Suponer que las funciones son continuas $f, g_1, g_2, \dots, g_m, h_1, h_2, \dots, h_l,$ y diferenciables. Donde x^ un mínimo local y punto regular de P), tomamos los multiplicadores de Lagrange: $\lambda_1, \lambda_2, \dots, \lambda_m, \mu_1, \mu_2, \dots, \mu_l \geq 0$:

$$\nabla f(x^{\wedge}) + \sum_{i=1}^{m} \lambda_i \ \nabla g_i(x^{\wedge}) + \sum_{r=1}^{l} \mu_r \nabla h_r(x^{\wedge}) = 0$$

$$\mu_r(h_r(x^{\wedge}) - d_r) = 0 \quad ; \ r = 1, \ 2, \ \dots, \ l$$

2.4. El Método de los Multiplicadores como caso particular del APP

Una amplia variedad de problemas puede ser considerados como casos especiales basado en el articulo de Souza SS, Oliveira PR, da Cruz Neto JX, Soubeyran A (2010), algunos ejemplos son los problemas convexos de programación no lineal, problemas de complementariedad monótono, problemas de flujos en redes, desigualdades variacionales y sistemas de ecuaciones e inecuaciones lineales.

La aplicación más importante se encuentra en el método de los multiplicadores de Hestenes y Powell (1969). Este algoritmo fue propuesto para resolver problemas de programación convexa Hestenes, (1969). A continuación, mostraremos que el método de los multiplicadores de Hestenes y Powell es un caso particular del algoritmo de punto proximal.

Consideremos el problema

$$\min f(x) \tag{2.11}$$

$$s.a.\ g_i(x) \leq 0 \quad (1 \leq i \leq m) \tag{2.12}$$

donde $f, g_i: \mathbb{R}^n \to \mathbb{R}$ son funciones convexas y diferenciables.

El lagrangiano estandar para este problema es $L: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ definido por

$$L(x,y) = \begin{cases} f(x) + \sum_{i=1}^{m} y_i g_i(x) & , \text{si } y \geq 0 \\ +\infty & , \text{caso contrario} \end{cases} \tag{2.13}$$

Para cada $x \in \mathbb{R}^n$, sea $x^+ \in \mathbb{R}^n$ definido por $x_j^+ = \max\{x_j, 0\}$. y sea $\alpha > 0$ y definiremos el Lagrangiano aumentado $L_\alpha: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ como

$$L_\alpha(x,y) = f(x) + \frac{1}{4\alpha} \sum_{i=1}^{m} \{[(y_i + 2\alpha g_i(x))^+]^2 - y_i^2\} \cdot$$

Es claro que, como f y g_i son diferenciables entonces L_α es diferenciable y

$$\nabla_x L_\alpha(x,y) = \nabla f(x) + \sum_{i=1}^{m} \left(y_i + 2\alpha g_i(x)\right)^+ \nabla g_i(x). \tag{2.14}$$

El método de los multiplicadores genera sucesiones $\{x^k\} \subset \mathbb{R}^n, \{y^k\} \subset \mathbb{R}^m$ mediante: Dados $y^0 \in \mathbb{R}_+^m$ y $y^k \in \mathbb{R}^m$

$$x^k = \arg\min_{x \in \mathbb{R}^n} L_\alpha(x, y^k) \tag{2.15}$$

$$y_i^{k+1} = \left(y_i^k + 2\alpha g_i(x^k)\right)^+ \cdot \tag{2.16}$$

Consideremos ahora la función objetivo dual $\varphi: \mathbb{R}^m \to \mathbb{R}$ definido por

$$\varphi(y) = \min_{x \in \mathbb{R}^n} L(x,y) \tag{2.17}$$

con L definido en (2.13). Es claro que φ es una función cóncava.

A continuación, probaremos que la sucesión $\{y^k\}$ generada por (2.15), (2.16) es la misma que la sucesión generada por el "algoritmo punto proximal" aplicado $a - \varphi$.

Teorema 2.4.1 *Sea* $\{\bar{y}^k\} \subset \mathbb{R}^m$ *la sucesión generada por el algoritmo de punto proximal para* $min_{y\in\mathbb{R}^m}\{-\varphi(y)\}$, *con* $\lambda_k = \frac{1}{4\alpha}$, $y\{y^k\}$ *la sucesión generada.*

Si $\bar{y}^0 = y^0$ *entonces* $\bar{y}^k = y^k$ *para todo* k.

Demostración. La prueba se hará por inducción, para $k = 0$, por hipótesis $y^0 = \bar{y}^0$. Asumamos$y^k = \bar{y}^k$, así

$$\bar{y}^{k+1} = \arg\min_{y\in\mathbb{R}^m}\left\{-\varphi(y) + \frac{1}{4\alpha}\|y - y^k\|^2\right\}$$

luego como $-\varphi$ y $\|\cdot - y^k\|^2$ son funciones propias y convexas, por la Proposición 2.3.1 se tiene que

$$0 \in \partial\left(-\varphi(\cdot) + \frac{1}{4\alpha}\|\cdot - y^k\|^2\right)(\bar{y}^{k+1})$$

además, como $r.i\big(dom(-\varphi)\big) \cap r.i\Big(dom(\|\cdot - y^k\|^2)\Big) \neq \emptyset$,

Luego obtenemos

$$\partial\left(-\varphi(\bar{y}^{k+1}) + \frac{1}{4\alpha}\|\bar{y}^{k+1} - y^k\|^2\right) = \partial\Big(-\varphi(\bar{y}^{k+1})\Big) + \partial\left(\frac{1}{4\alpha}\|\bar{y}^{k+1} - y^k\|^2\right),$$

así,

$$0 \in \partial\Big(-\varphi(\bar{y}^{k+1})\Big) + \partial\left(\frac{1}{4\alpha}\|\bar{y}^{k+1} - y^k\|^2\right),$$

pero, como $\|\cdot - y^k\|^2$ es diferenciable, entonces

$$0 \in \partial\Big(-\varphi(\bar{y}^{k+1})\Big) + \frac{1}{2\alpha}(\bar{y}^{k+1} - y^k)$$

es decir,

$$\frac{1}{2\alpha}(\bar{y}^{k+1} - y^k) \in \partial_\varphi(\bar{y}^{k+1}). \qquad (2.18)$$

De la definición de ∂_φ, (2.48) es equivalente a

$\varphi(\bar{y}^{k+1}) \geq \varphi(u) + \frac{1}{2\alpha}\langle \bar{y}^{k+1} - y^k, \bar{y}^{k+1} - u\rangle$ para todo $u \geq 0$.

es decir,

$$\varphi(u) \leq \varphi(\bar{y}^{k+1}) - \frac{1}{2\alpha}\langle \bar{y}^{k+1} - y^k, \bar{y}^{k+1} - u\rangle \text{ para todo } u \geq 0. \qquad (2.19)$$

Debido a que (2.19) determina a $\bar{y}^{k+1}$ de manera única, es suficiente mostrar que sustituyendo $\bar{y}^{k+1}$ por y^{k+1} también se cumpla (2.19).

De (2.44), (2.45) y (2.46), se tiene que

$$\begin{aligned} 0 &= \nabla_x L_\alpha(x^k, y^k) \\ &= \nabla f(x^k) + \sum_{i=1}^{m} \left(y_i^k + 2\alpha g_i(x^k)\right)^+ \nabla g_i(x^k) \\ &= \nabla f(x^k) + \sum_{i=1}^{m} y_i^{k+1} \nabla g_i(x^k) \end{aligned}$$

Luego, de (2.13)

$$0 = \nabla f(x^k) + \sum_{i=1}^{m} y_i^{k+1} \nabla g_i(x^k) \in \partial_x L(x^k, y^{k+1})$$

así, se tiene que x^k es un minimizador de $L(\cdot, y^{k+1})$ y por (2.17)

$$\varphi(y^{k+1}) = f(x^k) + \sum_{i=1}^{m} y_i^{k+1} g_i(x^k). \qquad (2.20)$$

Aquí, tenemos una expresión para el primer término de (2.19) con y^{k+1} en vez de $\bar{y}^{k+1}$. Ahora analizaremos el segundo término, es decir, $\frac{1}{2\alpha}\langle y^{k+1} - y^k, y^{k+1} - u\rangle$.

De (2.46), se tiene que

$$y_i^{k+1} = \left(y_i^k + 2\alpha g_i(x^k)\right)^+ = \max\{0, y_i^k + 2\alpha g_i(x^k)\},$$

Así

$$y_i^{k+1} - y_i^k = -y_i^k + \max\{0, y_i^k + 2\alpha g_i(x^k)\}$$

$$= \max\{-y_i^k + 0, -y_i^k + y_i^k + 2\alpha g_i(x^k)\}$$

$$= \max\{-y_i^k, 2\alpha g_i(x^k)\}$$

$$\geq 2\alpha g_i(x^k) \tag{2.21}$$

esto implica que

$$y_i^{k+1}(y_i^{k+1} - y_i^k) = 2\alpha y_i^{k+1} g_i(x^k). \tag{2.22}$$

Así, para cualquier $u \geq 0$, de (2.21) y (2.22), se tiene que

$$\frac{1}{2\alpha}(y_i^{k+1} - y_i^k)(y_i^{k+1} - u_i) = \frac{1}{2\alpha}(y_i^{k+1} - y_i^k)(y_i^{k+1}) - \frac{1}{2\alpha}(y_i^{k+1} - y_i^k)(u_i)$$

$$= g_i(x^k)y_i^{k+1} - \tfrac{1}{2\alpha}(y_i^{k+1} - y_i^k)(u_i) \quad \leq y_i^{k+1} g_i(x^k) - u_i g_i(x^k) \tag{2.23}$$

De (2.20) y (2.23) se tiene

$$\varphi(y^{k+1}) - \frac{1}{2\alpha}\langle y^{k+1} - y^k, y^{k+1} - u\rangle$$

$$\geq f(x^k) + \sum_{i=1}^{m} y_i^{k+1} g_i(x^k) - \sum_{i=1}^{m} y_i^{k+1} g_i(x^k). + \sum_{i=1}^{m} u_i g_i(x^k)$$

$$= f(x^k) + \sum_{i=1}^{m} u_i g_i(x^k)$$

$$= L(x^k, u)$$

$$\geq \min_{x \in \mathbb{R}^n} L(x, u)$$

$$= \varphi(u)$$

así,

$\varphi(u) \leq \varphi(y^{k+1}) - \frac{1}{2\alpha}\langle y^{k+1} - y^k, y^{k+1} - u\rangle$ para todo $u \geq 0$

esto es equivalente a

$$y^{k+1} = \arg\min_{y\in\mathbb{R}^m}\left\{-\varphi(y) + \frac{1}{4\alpha}\|y - y^k\|^2\right\}$$

$$x^{k+1} = P_k\left(x^k\right) \qquad (2.24)$$

por lo tanto, $\bar{y}^{k+1} = y^{k+1}$ y el procedimiento de inducción está completo.

Observación 2.4.1 *Pero si bien teoricamente es posible ejecutar la iteración proximal (2.24), en la práctica se puede realizar sólo en unos pocos casos. Dos situaciones en los que el paso proximal puede calcularse fácilmente es cuando:*

(i) $T(x) = \mathcal{N}_C(x)$ *es el operador normal asociado a algún conjunto convexo cerrado* $C \subset H$.

(ii) $T(x) = Ax + b$ *es una función afín.*

En (i), el cono normal a C en x es definido por

$$\mathcal{N}_C(x) = \begin{cases}\{z \in H: \langle z, x' - x\rangle \leq 0 \text{ para todo } x' \in C\} & \text{, si } x \in C \\ \emptyset & \text{, caso contrario}\end{cases}$$

y la función proximal asociado $P = (I + \mathcal{N}_C)^{-1}$ es la proyección ortogonal sobre C. Veamos esto.

Sea $P: H \to H$ un operador, probaremos que $P(x) = \arg\min\{\|y - x\|: y \in C\}$.

Debido a que

$$\arg\min\{\|y - x\|: y \in C\} = \arg\min\{\|y - x\| + \delta_C(y): y \in H\}$$

$$= \arg\min\{(1/2)\|y - x\|^2 + \delta_C(y): y \in H\}$$

y como $\frac{1}{2}\|\cdot - x\|^2 + \delta_C(\cdot)$ es coerciva, continua y estrictamente convexa entonces existe $\bar{x}$ Tal que

$$\bar{x}\arg\min\{\|y - x\|: y \in C\} = \arg\min\{(1/2)\|y - x\|^2 + \delta_C(y): y \in H\}$$

luego

$$0 \in \partial\big((1/2)\|\cdot - x\|^2 + \delta_C(\cdot)\big)(\bar{x}),$$

como $\frac{1}{2}\|\cdot - x\|^2 + \delta_C$ son funciones propias, convexas y $r.i\left(dom\left(\frac{1}{2}\|\cdot - x\|^2\right)\right) \cap r.i\big(dom(\delta_C)\big) \neq \emptyset$, entonces se tiene que

$$0 \in \partial\big((1/2)\|\cdot - x\|^2\big)(\bar{x}) + \partial\delta_C(\bar{x})$$

luego como $\frac{1}{2}\|\cdot - x\|^2$ Es diferenciable y de la definición 2.3.21, se tiene que $\partial\delta_C(\bar{x}) = \mathcal{N}_C(\bar{x})$, así

$$0 \in \bar{x} - x + \mathcal{N}_C(\bar{x})$$

$$x \in \bar{x} + \mathcal{N}_C(\bar{x})$$

de esta manera

$$x \in (I + \mathcal{N}_C)(\bar{x})$$

así,

$$\bar{x} \in (I + \mathcal{N}_C)^{-1}(x)$$

pero, como $(I + \mathcal{N}_C)^{-1}$ es un operador unívoco, entonces $\bar{x} = (I + \mathcal{N}_C)^{-1}(x) = P(x)$

Por lo tanto, $P = (I + \mathcal{N}_C)^{-1}$ es la proyección ortogonal sobre C.

En (ii), la iteración proximal en (*2.54*) presenta la misma dificultad que resolver un sistema de ecuaciones lineales.

Veamos esto, dado $x^k \in H$, de la iteración proximal (*2.54*) se tiene

$$x^{k+1} = (I + T)^{-1}(x^k)$$

$$x^k = (I + T)(x^{k+1})$$

$$x^k = x^{k+1} + T(x^{k+1})$$

$$x^k = x^{k+1} + Ax^{k+1} + b$$

$$x^k = (I + A)x^{k+1} + b$$

$$x^k - b = (I + A)x^{k+1}$$

Denotando $\tilde{A} = I + A$ y $\tilde{b} = x^k - b$, entonces debemos resolver el sistema lineal $\tilde{A}x^{k+1} = \tilde{b}$ para obtener la siguiente iteración proximal.

2.5 Distancia Generalizada

Presentaremos la definición de la distancia proximal y de la distancia proximal inducida, que fueron introducidas por Auslender A, Teboulle M (1999) el que adaptaremos a nuestra situación para el conjunto R_+^n.

Definición 2.5.1 *Una función* $d: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+ \cup \{+\infty\}$ *conocida como distancia proximal en* $\mathbb{R}_{++}^n$, $si\ \forall y \in \mathbb{R}_{++}^n$ *se cumplen las propiedades:*

1. $d(\cdot, y)$ *es una distancia estrictamente convexa, semicontinua inferior, propia, y continuamente diferenciable en* $\mathbb{R}_{++}^n$*;*

2. $dom\ d(\cdot, y) \subset\ \mathbb{R}_+^n\ y\ dom\ \partial_1 d(\cdot, y) = \mathbb{R}_{++}^n$, *donde* $\partial_1 d(\cdot, y)$ *denota la función subdiferencial clásica de la función* $d(\cdot, y)$*con respecto a la primera variable;*

3. $d(\cdot, y)$ *es coerciva en* $\mathbb{R}^n (i.e., \lim_{\|u\| \to \infty} d(u, y) = +\infty$ *;*

4. $d(y, y) = 0$.

Donde $\mathcal{D}(\mathbb{R}_{++}^n)$ *es una familia de funciones que cumplen con los puntos dados.*

Definición 2.5.2 *Dado una distancia proximal* $d \in \mathcal{D}(\mathbb{R}_{++}^n)$, *donde la función* $H: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+ \cup \{+\infty\}$ *es inducida por* d, *se dice que* H *es una función de valores finitos en* $\mathbb{R}_{++}^n \times \mathbb{R}_{++}^n$ *y para cada* $a, b \in \mathbb{R}_{++}^n$ *se satisface*

(i) $H(a, a) = 0$.

(ii) $\langle c - b, \nabla_1 d(b, a) \rangle \leq H(c, a) - H(c, b) - H(a, b), \forall c \in \mathbb{R}_{++}^n,$

El gradiente $\nabla_1 d(\cdot, y)$ *de la función* $d(\cdot, y)$ *se refiere a la primera variable.*

Denotemos por $(d, H) \in \mathcal{F}(\mathbb{R}_{++}^n)$ *a la distancia proximal e inducida cumple con las condiciones ver* definición 2.5.1, *también denotemos por* $(d, H) \in \mathcal{F}(\mathbb{R}_{++}^n)$ *si existe un* H *tal que:*

(Iiii) H *es de valores finitos en* $\mathbb{R}^n_{++} \times \mathbb{R}^n_{++}$ *satisfaciendo (Ii) y (Iii), para cada* $c \in \mathbb{R}^n_+$.

(iv) Para cada $c \in \mathbb{R}^n_+, H(c,\cdot)$ *tiene conjuntos de nivel acotados en* $\mathbb{R}^n_{++}$.

Luego tenemos $(d,H) \in \mathcal{F}_+(\mathbb{R}^n_+)$, *cuando*

(v) $(d,H) \in \mathcal{F}(\mathbb{R}^n_+)$.

(vi) $\forall y \in \mathbb{R}^n_+ \; y \; \forall\{y^k\} \subset \mathbb{R}^n_{++}$ *acotada con* $\lim_{k\to+\infty} H(y,y^k) = 0$, *entonces*

$$\lim_{k\to+\infty} y^k = y.$$

(vii) $\forall y \in \mathbb{R}^n_+$, $donde \; \forall\{y^k\} \subset \mathbb{R}^n_{++}$ *tal que* $\lim_{k\to+\infty} y^k = y$, *entonces*

$$\lim_{k\to+\infty} H(y,y^k) = 0.$$

Algunos casos de distancias proximales que satisfacen las condiciones mostradas, como las distancias de Bregman, distancias proximales basadas en $\phi-$ divergencias, distancias auto-proximales, y distancias basadas en distancias proximales logaritmo cuadraticos, ver Auslender A, Teboulle M (2006)

Distancias Bregman: *Sea* $h: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *una función propia y diferenciable en int(domh). Definimos la función* $D_h(\cdot,\cdot): \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *por*

$$D_h(x,y) \coloneqq h(x) - h(y) - \langle \nabla h(y), x - y \rangle.$$

La función h *es denominada una función Bregman con zona* S *si:*

(1) h *es continua y estrictamente convexa en* $\bar{S}$.

(2) h *es diferenciable y continua en* S;

(3) Dado cualquier $x \in \bar{S} \; y \; \delta \in \mathbb{R}$, *el conjunto de nivel inferior*

$L_{D_h}(x,\delta) = \{y \in S: D_h(x,y) \leq \delta\}$, *es acotada;*

(4) Si $\{y^k\} \subset S$ *converge a* y *entonces* $D_h(y,y^k)$ *converge a* 0.

Cuando h *es una función Bregman entonces* D_h *es denominada distancia Bregman asociado a* h*, ver* Eckstein J (1998).

Además, se ha probado (Bauschke y Borwein (1997)) que las condiciones $(1)-(4)$ *implican lo siguiente:*

(5) El conjunto de nivel inferior $L_{D_h}(\delta, y) = \{x \in \bar{S}: D_h(x, y) \leq \delta\}$ *es acotado para todo* $y \in S$*.*

(6) Si $\{z^k\} \subset \bar{S}$ *es acotada,* $\{y^k\} \subset S$ *converge a* $y, y \lim_{k\to+\infty} D_h(z^k, y^k) = 0$*, entonces* $\lim_{k\to+\infty} z^k = y$*.*

Sea $\theta: \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ *una función propia convexa y semicontinua inferior* ver definición 2.3.3, *con* $(0, +\infty) \subset dom\theta \subset [0, +\infty)$ *tal que* $\theta \in C^2\big((0, +\infty)\big), \theta''^{(t)} > 0, \forall t > 0\ y \lim_{t\to 0^+} \theta'(t) = -\infty$ *. Denotemos por* Θ_0 *a la clase de funciones que satisfacen las propiedades antes mencionadas y* $\theta(0) < +\infty$ *y definir* $\bar{h} = \sum_{i=1}^{n} \theta(x_i)$*, así*

$$D_{\bar{h}}(x, y) = \sum_{i=1}^{n} \big(\theta(x_i) - \theta(x_i) - \theta'(y_i)(x_i - y_i)\big).$$

esto es, D_h *es separable. Algunos ejemplos de funciones que pertenecen a* θ_0 *son los siguientes:*

(a) $\theta_1(t) = t \log t$ *para el cual*

$$D_{\bar{h}_1}(x, y) = \sum_{i=1}^{n} \left(x_i \log \frac{x_i}{y_i} + y_i - x_i\right).$$

Esta función es denominada distancia divergente Kullback-Leibler.

(b) $\theta_2(t) = t^p - t^q$*, si* $t \geq 0\ y\ \theta_2(t) = +\infty$*, si* $t < 0$*; con* $p > 1\ y\ 0 < q < 1$*.*

Para $p = 2, q = \frac{1}{2}$*, tenemos*

$$D_h(x, y) = \|x - y\|^2 + \sum_{i=1}^{n} \frac{1}{\sqrt[2]{y_i}} \left(\sqrt{x_i} - \sqrt{y_i}\right)^2.$$

(c) $\theta_3(t) = \frac{1}{1-p}(pt - t^p)$*, si* $t \geq 0\ y\ \theta_3(t) = +\infty$*, si* $t < 0$*; con* $p \in (0,1)$ *Así*

$$D_h(x,y) = \sum_{i=1}^{n} \left(y_i^p + \frac{1}{(1-p)} \left(p y_i^{p-1} x_i - x_i^p \right) \right).$$

Es fácil probar que la distancia Bregman separable $D_{\bar{h}}$ *satisface las condiciones de la* ver teorema 4.4.2*, para cada* $\theta \in \Theta_0$*, por lo tanto* $D_{\bar{h}} \in \mathcal{D}(\mathbb{R}^n_{++})$.

Ahora definimos $H(x,y) \coloneqq D_{\bar{h}}(x,y)$*, así* $H(a,a) = 0, \forall a \in \mathbb{R}^n_{++}$ *y de la identidad de los tres puntos,* Solodov MV, Svaiter BF (2000)*, obtenemos*

$$\langle c - b, \nabla_1 D_{\bar{h}}(b,a) \rangle \leq H(c,a) - H(c,b) - H(a,b).$$

Además, como $D_{\bar{h}}(c,\cdot)$ *tiene conjuntos de nivel acotado en* $\mathbb{R}^n_{++}$*, entonces* $(D_{\bar{h}}, H) \in \mathcal{F}(\mathbb{R}^n_{++})$*. Finalmente, de* $(B6)$ *y* $(B4)$ *tenemos, respectivamente*

(i) $\forall\, y \in \mathbb{R}^n_+, \forall \{y^k\} \subset \mathbb{R}^n_{++}$ *acotada con* $\lim_{k\to+\infty} H(y, y^k) = D_{\bar{h}}(y, y^k) = 0$*, entonces* $\lim_{k\to+\infty} y^k = y.$

(ii) $\forall\, y \in \mathbb{R}^n_+\, y\, \forall \{y^k\} \subset \mathbb{R}^n_{++}$ *tal que* $\lim_{k\to+\infty} y^k = y,$ *entonces* $\lim_{k\to+\infty} H(y, y^k) = D_{\bar{h}}(y, y^k) = 0.$

Esto es,$(D_{\bar{h}}, H) \in \mathcal{F}_+(\mathbb{R}^n_+)$.

2.6 Variables utilizadas en el modelo de Cobb Douglas

Una función de producción muy conocida usado en microeconomía para ver procesos de producción, este se uso por primera vez en un estudio empírico en Estados Unidos donde se veía la relación en la productividad con el capital y la fuerza de trabajo. Tomando la función de producción original homogénea se consideró de rango 2 1 en ambos factores, con rendimientos constantes a escala.

La formula de producción se puede expresar como:

$$Q = A * L^{\alpha} K^{1-\alpha}$$

Teniendo un valor positivo $\alpha < 1$ que indica la elasticidad del factor trabajo (L).

La elasticidad del factor capital (K) seria el complemento del factor elasticidad del trabajo $1-\alpha = \beta$, corresponde a la existencia de constantes a escala de rendimientos. La constante A seria un parámetro.

La homogeneidad de grado 1 es una de las condiciones que posee la función Cobb- Douglas, además de tener facilidad de estimación y los rendimientos marginales decrecientes en cada factor de la producción.

Podemos escribir la función Cobb-Douglas de la forma:

$$Q = F(L, K) = A * L^{\alpha} K^{\beta} \qquad (2.25)$$

de manera que puede tomar distintos rendimientos a escala:

$$\lambda^n Q = F(\lambda L, \lambda K) = \lambda^{\alpha+\beta}\left(A * L^{\alpha} K^{\beta}\right)$$

Entonces diremos que los rendimientos son crecientes si $\alpha + \beta > 1$, pero tendrá rendimiento decreciente si $\alpha + \beta < 1$, pero si los rendimientos constantes a escala;será cuando si $\alpha + \beta = 1$.

Cuando queremos trabajar con los datos podemos hacer uso de una transformación logarítmica:

$$\text{Log Q} = \log \text{A} + \alpha \text{ Log L} + \beta \text{Log K} \qquad (2.26)$$

Caracterización de las variables utilizadas en la especificación

Como queremos hallar una estimación en el caso de Perú de la función producción para el, se analizó cuales son las fuentes que nos ayudaría a extraer los datos con el objeto de obtener los datos de fuerza de trabajo asi como el capital. Por eso definimos las características de estas variables.

Factor de trabajo

Denotamos L como el factor trabajo, por eso se buscó la información estimada por el INEI sobre la PEA ocupada, obteniendo la información anualizada, debido a que la no disponibilidad de la información del INEI se usaron los datos correspondientes al periodo 2001-2019.

Factor de capital

Como no se contaba con el dato inicial correspondiente al capital tuvo que estimarla. Por esto, se estimó nivel inicial de stock de capital de 13 032 (Miles de soles) para el año 1994. Donde el acervo de capital para los siguientes años se

calculó con el método de inventarios perpetuos considerando la inversión bruta fija, Ana Azofeifa V. & Marlene Villanueva S. (1996) y una tasa de depreciación. Tomamos una tasa de depreciación anual de 5 %, Shirley Miller L. (2002).

Producción

Se uso el Producto Interno Bruto (PBI) en miles de soles como indicador a precios constantes.

Tomamos la inversión bruta fija no residencial como variable relevante en el cálculo del stock de capital. Como no se tiene esta variable se consideró la inversión bruta interna Zivot, E. and D. W. Andrews (1992)., en los que se considera los inventarios y la inversión residencial, los cuales no ayudan a incrementar la capacidad productiva.

Los inventarios perpetuos se utilizan para poder estimar el stock de capital mediante la forma recursiva: $K_t = (1-\delta)K_{t-1} + I_{t-1}$, tal que denotamos a la inversión bruta fija como I,ver Ana Azofeifa V. & Marlene Villanueva S. (1996) y δ es la tasa de depreciación, Shirley Miller L. (2002).

Productividad

Como no se puede medir directamente el nivel de productividad debemos obtenerlos como resto de la ecuación, ver Hamilton, J. D. (1994). Ver (2.25), cuando se hallan los parámetros de elasticidad $\alpha \ y \ \beta$ (en cuanto a la fuerza laboral y el capital), Perron, P. (1994). Ver figura 4.1 (PBI-Millones de Soles-2007)

Capítulo 3

3. METODOLOGíA

Se tomaron las siguientes partes en el trabajo:

3.1 El Universo

El universo en espacios Euclidianos en nuestra investigación de optimización no convexa y su aplicación a una clase de modelos económicos de decisión como problemas de la cartera, problemas de producción. En Investigación de Operaciones uno de los casos que se estudia son problemas de optimización no convexa y asi hallar los valores máximos o mínimos de funciones no convexas restringidas en un conjunto convexo. Por otro lado, los modelos económicos de decisión representan situaciones donde el agente económico trata de obtener la mejor decisión dentro de todas las posibilidades.

3.2 Técnicas de recopilación de datos

En la investigación a trabajar fue necesaria una búsqueda en revistas publicadas de información científica, uso de hemerotecas y bibliotecas especializadas, asistencia a eventos científicos y reuniones con algunos especialistas en el tema.

Los datos de microeconomía (Modelo Cobb Douglas), se obtuvieron de la información en las tablas extraídas de estadísticas publicadas por INEI y BCR**.**

3.3 Técnicas descriptivas para la contrastación o demostración de las hipótesis

Luego se tomo como metodología un enfoque de tipo inductivo-deductivo tomando en cuenta investigaciones recientes, con el fin de poder aplicarlo a nuestros requerimientos y asi lograr los resultados que detallaremos en las siguientes secciones.

Encontramos resultados de algoritmos utilizados para hallar resultados en problemas de optimización en funciones cuasi-convexas, pero no se encuentra en el caso genérico no convexa de minimización.

Estudiamos los artículos en mención y observamos que podíamos extender más aún el método considerando algunas hipótesis apropiadas.

Observamos también que no se habían publicado experimentos computacionales aplicados a modelos económicos salvo algunas experiencias de funciones no convexas, por esto se mostró una aplicación de Cobb Douglas.

Capítulo 4

4. RESULTADOS Y DISCUSIÓN

4.1 Análisis, interpretación y discusión de resultados

Los resultados de esta tesis son los siguientes:

- Convergencia débil de la extensión del método proximal propuesto a un punto de KKT generalizado.

- Aplicación del método para la solución de problemas económicos de decisión.

- Presentación de algunos resultados computacionales del método propuesto en la solución de algunos problemas académicos para constatar los resultados teóricos.

Los resultados anteriores podemos interpretarlos como la extensión del método proximal en el caso de problemas de optimización no convexa en el ortante no negativo.

En casos prácticos es importante resolver problemas no convexa de optimización en el ortante no negativo pues estas restricciones reflejan la

cantidad elegída en la decisión de un agente económico en teoría de decisión, cantidad producida en producción, etc.

En el capítulo 4, Cobb, C. W. y Douglas, P. H. (1928): se muestra un caso con este método ilustrando su utilidad en las aplicaciones.

Con respecto a la discusión de los resultados podemos mencionar lo siguiente:

- La aplicación del método a problemas económicos presentados en esta tesis se puede ampliar desde compañías individuales hasta cuestiones económicas globales.

4.2 Pruebas de hipótesis

Usando un concepto de subdiferencial de Fréchet y sus propiedades geométricas fue posible extender el "Método del Punto Proximal ", para solucionar problemas no convexos de optimización y en particular cuando la función es convexa el método se reduce al clásico método de optimización propuesta por Martinet (1970).

Usando la teoría de sucesiones y análisis matemático fue posible obtener la convergencia generada en la sucesión por la extensión del "Método del Punto Proximal ", Articulo PESQUIMAT De la Cruz, L, Papa, E. (2013).

Mostramos en las aplicaciones que los modelos de decisión en economía están íntimamente relacionados a problemas de optimización no convexa, en particular usando Cobb Douglas.

4.3 Presentación de resultados de Cobb Douglas

Presentamos los resultados obtenidos en el modelo de Economía (Cobb Douglas) en el periodo 1994 -2019.

Ver anexos 1-8 (Se muestran las tablas extraídas de estadísticas del INEI y BCR)

Figura 4.1

PBI-Millones de Soles-2007

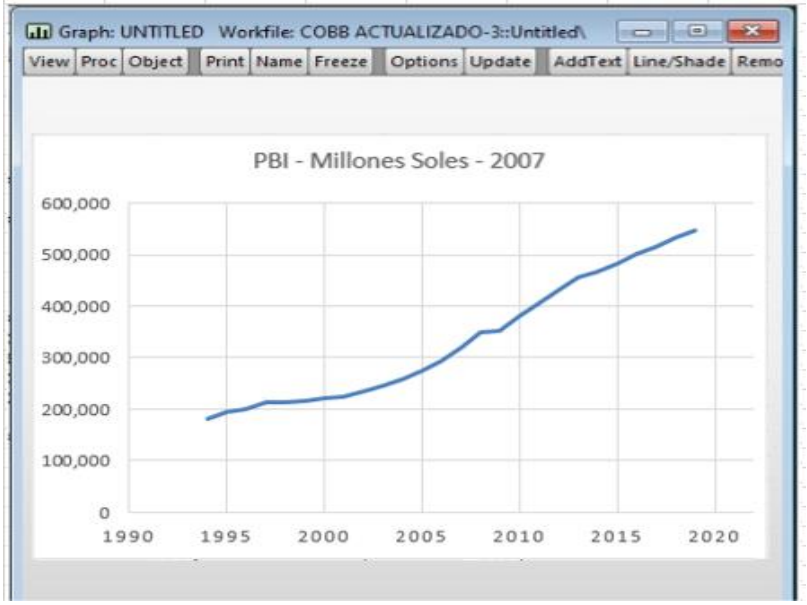
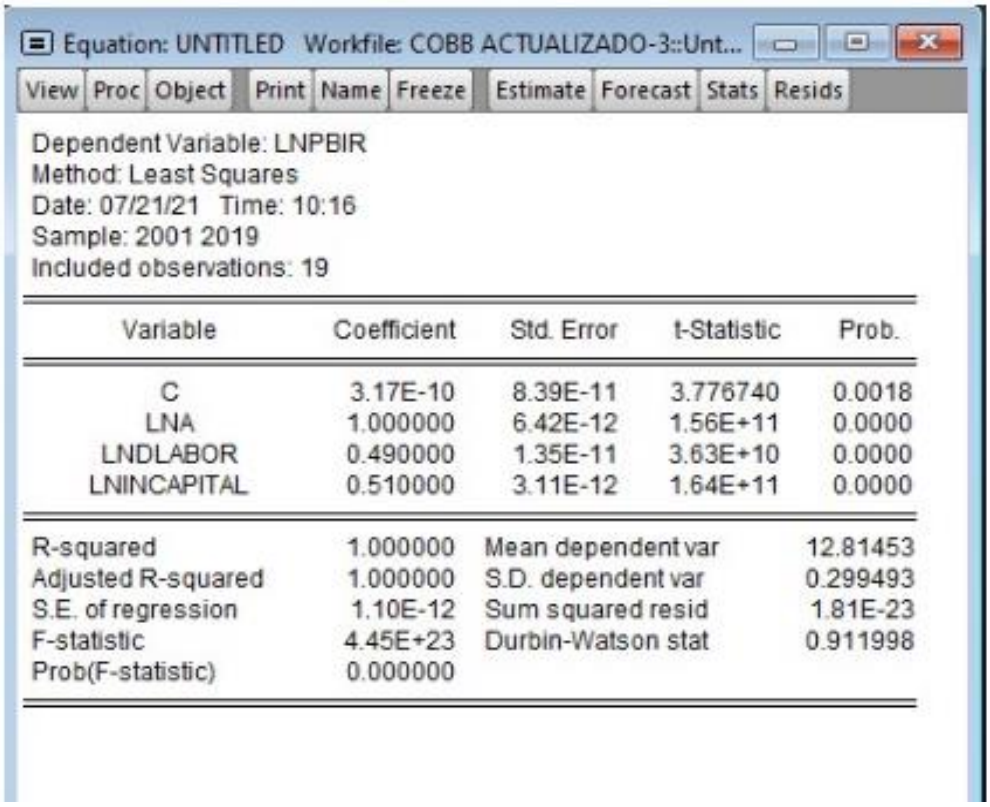


Figura 4.2

Los estadísticos Cusum y Cusum cuadrado



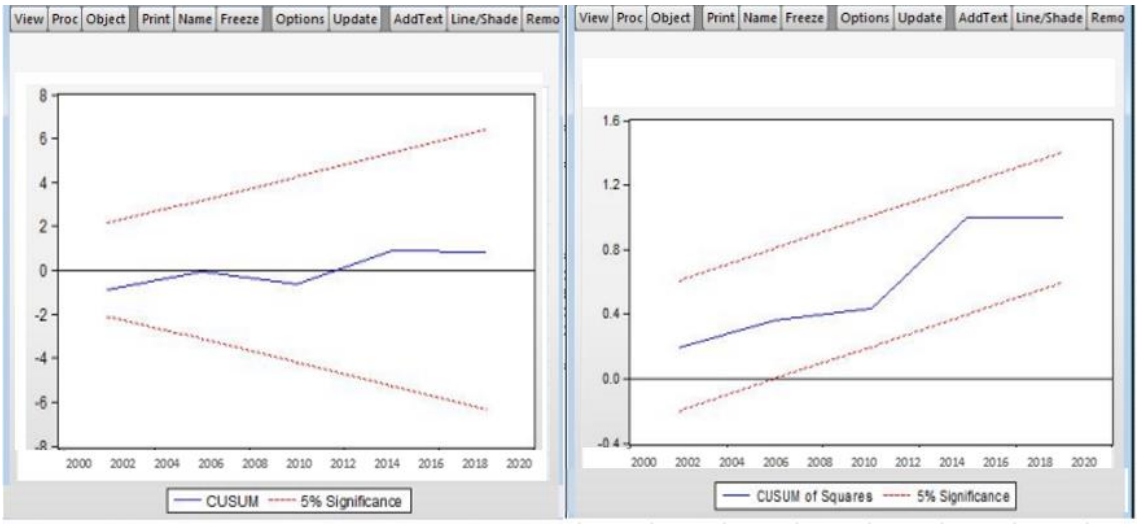
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.17E-10	8.39E-11	3.776740	0.0018
LNA	1.000000	6.42E-12	1.56E+11	0.0000
LNDLABOR	0.490000	1.35E-11	3.63E+10	0.0000
LNINCAPITAL	0.510000	3.11E-12	1.64E+11	0.0000

R-squared	1.000000	Mean dependent var	12.81453
Adjusted R-squared	1.000000	S.D. dependent var	0.299493
S.E. of regression	1.10E-12	Sum squared resid	1.81E-23
F-statistic	4.45E+23	Durbin-Watson stat	0.911998
Prob(F-statistic)	0.000000		

Nota: Tomando estadísticos Cusum y Cusum cuadrado los cuales nos indican que los coeficientes son estables en la regresión del ajuste, podemos observar que el coeficiente de determinación (R^2ajustado) es bastante alto, lo que no da la idea que la mayor parte de la producción puede darse por el aporte de los factores de producción.

Figura 4.3

Gráfica de los estadísticos Cusum y Cusum cuadrado



4.4 Algoritmo Proximal para Minimización

Estamos interesados en resolver el problema

$$\min\{f(x): x \geq 0\} \qquad (4.1)$$

donde $f: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ es una función propia ver **definición 2.3.5,** tal que $dom\, f \cap \mathbb{R}^n_+ \neq \emptyset\; y\; x \geq 0$ nos indica que x, x_i, es no negativo.

En esta subsección presentamos una extensión del "Método del Punto Proximal ". con una distancia proximal y asi hallar los resultados del problema planteado basado en el articulo de Papa Quiroz y Oliveira (2012i).

4.4.1 El Algoritmo

Tenemos un punto inicial y una sucesión de parámetros positivos $\{\lambda_k\}$ y

$$x^0 \in \mathbb{R}^n_{++}.$$

Para cada $k = 1,2,...$, si $0 \in \hat{\partial} f(x^{k-1})$, hallamos el valor. Si no hallar un $x^k \in \mathbb{R}^n$ de manera

$$0 \in \hat{\partial}\Big(f(\cdot) + \lambda_k d(\cdot, x^{k-1})\Big)(x^k),$$

donde d es una distancia proximal tal que $(d, H) \in F_+(\mathbb{R}^n_+)$.

Observación 4.4.1 *Observamos que el algoritmo es una extensión (para funciones no convexas)* "Método del Punto Proximal "*. donde, la función f es convexa decimos que las iteraciones se dan de la forma.*

$$x^k = \arg\min_{x \geq 0}\{f(x) + \lambda_k d(x, x^{k-1})\},$$

Observación 4.4.2 *Como estamos interesados en resolver (4.1) donde f es no convexa es necesario ver que en el algoritmo sólo se necesita hallar un punto estacionario, en cada iteración, ver* definición 2.3.27 (*no un mínimo global) de la función regularizadora $f(\cdot) + \lambda_k d(\cdot, x^{k-1})$ vemos asi que los algoritmos locales se pueden hallar en cada iteración de manera satisfactoria*.

Observación 4.4.3 *Dado (4.3), con $(\lambda_k/2)d(\cdot, x^{k-1})$, es continuamente diferenciable, obtenemos que*

$$0 \in \hat{\partial} f(x^k) + \lambda_k \nabla_1 d(x^k, x^{k-1}).$$

Así, existe $g^k \in \hat{\partial} f(x^k)$ tal que

$$g^k = -\lambda_k \nabla_1 d(x^k, x^{k-1}).$$

4.4.2 Resultados de Convergencia

Teorema 4.4.1 *Sea f una función semicontinua inferior y acotada inferiormente donde la sucesión $\{x^k\}$, generada por* "Método del Punto Proximal ", *en el ortante $x^k \in \mathbb{R}^n_{++}$.esta bien definida.*

Demostración. Como f es acotada inferiormente y semicontinua inferior y la distancia $d(\cdot, y)$ es coerciva, ver Definición 2.3.4, $f(x) + \lambda_k d(x, x^{k-1})$ semicontinua inferior, ver **definición 2.3.3** y coerciva se confirma la existencia de $\{x^k\}$,la positividad de x^k es inmediata de la definición de d.

Hipótesis A. $f: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ es una función propia ver **definición 2.3.5,** acotada inferiormente y semicontinua inferior ver **definición 2.3.3**.

Para hallar el punto límite tomamos una clase especial de las distancias proximales a otra de las distancias proximales separables,

$$d(x, y) = \sum_{i=1}^{n} d_i(x_i, y_i),$$

donde $d_i: \mathbb{R} \times \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$.

Definamos también:

$$J(\bar{x}) = \{i \in \{1,2, \dots, n\}: \bar{x}_i > 0\}.$$

Teorema 4.4.2 *Digamos que la hipótesis* $\boldsymbol{A}$ *es satisfecha,* d *es separable. Si* $0 < \lambda_k < \bar{\lambda}$*, para todo* k *y algún* $\bar{\lambda} > 0$*,* → *la sucesión* $\{x^k\}$ *converge debilmente a un punto* $\bar{x}$ *tal que:*

i) $\bar{x}_i \geq 0, \quad \forall i = 1,2, \dots, n;$

ii) $\liminf_{k \to +\infty} g_i^k \geq 0, \quad \forall i \in J(\bar{x})$

iii) $\liminf_{k \to +\infty} x_i^k g_i^k \geq 0, \quad \forall i \in J(\bar{x}).$

4.5 Aplicaciones a Economía

Al aplicar modelos con el "Método del Punto Proximal “, se dan en diversas líneas de investigación en optimización fraccionaria, teoría de localización, teoría de decisiones y en casos de economía. En esta tesis damos un ejemplo de la teoría del consumidor, en un modelo de microeconomía.

En la mayoría de los problemas se hacen hipótesis de tal forma que los modelos sean convexos, y luego se usa cualquier metodo de optimización de problemas convexos, restringiendo su aplicación a casos más generales.

Según nuestra investigación no existe aún un software que trate problemas específicamente no convexos, es por eso por lo que esta tesis es un aporte para este fin.

4.5.1 Aplicación del Método a la Teoría Económica de decisión.

En la teoría económica de la decisión se dan modelos matemáticos que se da en la elección óptima de un agente económico (un Estado, empresa, etc.) con una variedad de alternativas a escoger. Tomando el agente económico en un problema de decisión:

• Teniendo alternativas el que nos ayuda a elegir el universo de ellos.

• En el agente económico evalúa las diferentes alternativas tomando el criterio de valorización de ellos. Este criterio se usa en una relación binaria reflexiva y transitiva ($\precsim$), definida económicamente como “es al menos tan preferible” el que nos indica sus preferencias sobre el conjunto de elección. En economía llamamos preferencia a esta relación.

• Los recursos escasos delimitan al conjunto de oportunidades en la que el agente económico efectivamente se puede elegir.

Donde el conjunto de elección X, se dice que $\mu : X \to \mathbb{R}$ llamada función de utilidad ($\precsim$) si $\forall x, y \in X$

$$y \precsim x \text{ , sii } \mu(y) \leq \mu(x).$$

Las preferencias ($\precsim$), sujetas a condiciones sobre X pueden ser dados por una función de utilidad μ . de manera que se pueda elegir la mejor alternativa para maximizar la función μ . podemos decir que resolver el problema de decisión del agente económico es equivalente a resolver el problema de maximización:

$$max\{\mu(x) : x \in X\}.$$

Un de los casos de una función de utilidad es la función cuasi-cóncava, donde , μ cumple con:

$\mu(\lambda x + (1-\lambda)y) \geq min\{\mu(x), \mu(y)\}$, para todo $x, y \in X$; y todo $\lambda \in [0,1]$.

Podemos probar que la función está relacionada íntimamente a la hipótesis de convexidad de la preferencia, osea $\precsim$, es convexa si para $x, y, z \in X$ con $x \precsim y$ y $x \precsim z$ en X y $0 \leq \alpha \leq 1$ entonces

$$x \precsim \alpha y + (1-\alpha)z.$$

La definición $\precsim$ en el caso económico, se puede decir como la tendencia natural del agente económico a diversificar las mejores entre todas posibles elecciones, donde, si $z\ e\ y$ son dos alternativas al $\precsim x$, entonces, podemos pensar que la combinación $\alpha y + (1-\alpha)z,\ \forall \alpha \in [0,1]$, es continúa siendo al menos tan preferible como x.

La condición de convexidad de la preferencia $\precsim$ y ciertas condiciones apropiadas sobre X, permiten ver la existencia de una función de utilidad μ cuasi cóncava y por lo que tener un problema de optimización cuasi-cóncavo y las restricciones son cuasi-cóncavas.

Ejemplo 4.5.1

Una empresa puede producir n productos en cantidades $x_1, x_2, \dots, x_n$ siendo sus costos de producción (por unidad) respectivamente igual a $p_1, p_2, \dots, p_n$, donde $p_n > 0$, para todo $i = 1,2,\dots,n$. Con un capital de b soles en la producción y una función continua f nos indica la utilidad de producción de los n productos, el problema de conocer una utilidad máxima este dado:

$$max\{f(x) : p^T x \leq b, x_i \geq 0\}.$$

En economía una de las funciones más conocidas que expresan la función de producción de una empresa es la función de Cobb-Douglas, Zivot, E. and D. W. Andrews (1992),

$f(x) = k.{x_1}^{\alpha_1}{x_2}^{\alpha_2} \dots {x_n}^{\alpha_n}$,

donde $k > 0$, $\alpha_1, \alpha_2, \dots, \alpha_n > 0$ y $\alpha_1 + \alpha_2 + \cdots + \alpha_n = 1$.

Cobb, C. W. y Douglas, P. H. (1928): publicaron un estudio en el cual modelaban el crecimiento de la economía estadounidense durante el perÍodo 1899-1922. Ellos tomaron un punto de vista simple en la economía con la producción determinada por el número de mano de obra relacionada x_1 y el numero de capital invertido x_2. Hay varios factores que pueden afectar el rendimiento económico, este caso resultó notablemente exacto. La función utilizada en el modelo de producción de manera que

$$f(x) = k.{x_1}^{\alpha_1}{x_2}^{\alpha_2},$$

Donde la producción total $f(x_1, x_2)$ (el valor monetario de todos los bienes producidos anualmente), la mano de obra x_1 es la cantidad total de HH trabajadas al año) y el capital invertido x_2 es el valor monetario de toda la maquinaria, equipo, edificios, etc). Usando derivadas parciales para demostrar cómo la forma particular del modelo se infiere de ciertas suposiciones que plantearon con respecto a la economía.

Si la función de producción $f(x_1, x_2)$, con derivada parcial de $\frac{\partial f}{\partial x_1}$ es la razón a la cual cambia la producción con respecto a la cantidad de mano de obra. Los economistas conocen como producción marginal respecto a la mano de obra o productividad marginal de la mano de obra, ver De Lamo, A. R. y Dolado, J. J. (1993); análogamente tenemos la derivada parcial de $\frac{\partial f}{\partial x_2}$ es la razón a la cual cambia la producción en referencia a la cantidad de capital invertido l que se denomina productividad marginal del capital. En estos términos las suposiciones que plantearon Cobb y Douglas se pueden plantear como sigue:

1) Si la mano de obra o el capital disminuye, entonces pasa lo mismo con la producción.

2) La producción marginal respecto a la mano de obra es equivalente a la cantidad de producción por unidad de mano de obra.

3) La productividad marginal del capital es proporcional a la cantidad de producción por unidad de capital.

Debido a que la producción por unidad de mano de obra es $\frac{f}{x_1}$ la hipótesis 2) plantea que

$$\frac{\partial f}{\partial x_1} = \alpha \frac{f}{x_1},$$

Para algún $\alpha > 0$. Si tenemos x_2 constante diremos que la ecuación diferencial parcial se convierte en una ecuación diferencial ordinaria

$$\frac{df}{dx_1} = \alpha \frac{f}{x_1},$$

Si resolvemos esta ecuación diferencial separable obtenemos:

$$g(x_1) = f(x_1, x_2) = c(x_2).{x_1}^{\alpha}$$

De forma análoga con respecto a la otra variable da

$$h(x_2) = f(x_1, x_2) = c(x_1).{x_2}^{\beta}$$

Para algún $\beta > 0$. Al comparar las dos ecuaciones anteriores obtenemos:

$$f(x) = k.{x_1}^{\alpha}{x_2}^{\beta},\ldots\ldots\ldots\ldots(\theta)$$

donde k es una constante que es independiente tanto de x_1 e x_2.

En la ecuación anterior, la mano de obra $x_1\ y$ el capital x_2 aumenta un factor *m* es independiente tanto de $x_1,\ x_2$, entonces en (θ) tenemos:

$$f(mx_1, mx_2,) = k(mx_1)^{\alpha}(mx_2)^{\beta} = km^{\alpha+\beta}\,{x_1}^{\alpha}{x_2}^{\beta} = m^{\alpha+\beta}\,f(\,x_1, x_2).$$

Si, además, $\alpha + \beta = 1$, entonces:

$$f(mx_1, mx_2,) = mf(\,x_1, x_2),$$

condición que es llamada en economía función homogénea de grado 1 y que se interpreta como la producción se incrementa en un valor *m*. por eso Cobb y Douglas tomaron $\alpha + \beta = 1$, por lo que, se tiene:

$$f(x) = k.{x_1}^{\alpha}{x_2}^{1-\alpha},$$

Tomamos la función de dos variables en Cobb y Douglas en producción, ver Felipe, Jesús y Adams, F. Gerard (2005).

Cobb y Douglas tomaron los datos que publicó el gobierno en el año 1899 como una línea de referencia f, x_1 y x_2 en el año 1899 se le dio un valor de 100 para la Tabla 3.

Los datos de otros años se tomaron como porcentajes de los valores de 1899.

Cobb, C. W. y Douglas, P. H. (1928): aplicaron el método de los mínimos cuadrados para ajustar los datos, ver la Tabla Nro: 3 a la función:

$$f(x) = 1.01.{x_1}^{0.75}{x_2}^{0.25}.$$

En efecto, de la función $f(x) = k.{x_1}^{\alpha}{x_2}^{1-\alpha}$, dividiendo por x_2

$$\frac{f}{x_2} = k.{x_1}^{\alpha}.\left(\frac{{x_2}^{1-\alpha}}{x_2}\right) = k.{x_1}^{\alpha}.\left(\frac{1}{{x_2}^{\alpha}}\right)$$

tomando logaritmos a ambos lados tenemos:

$$ln\left(\frac{f}{x_2}\right) = \ln(k) + \alpha ln\left(\frac{x_1}{x_2}\right).$$

Si hacemos $x = ln\left(\frac{x_1}{x_2}\right)$

$y = ln\left(\frac{f}{x_2}\right)$, tenemos que $y = \alpha x + \ln(k)$.

Tabla 3.

Crecimiento de la Economía Estadounidense durante el período 1899 – 1922

Año	$f(x_1, x_2)$	x_1	x_2
1899	100	100	100
1900	101	105	107
1901	112	110	114
1902	122	117	122
1903	124	122	131
1904	122	121	138
1905	143	125	149
1906	152	134	163
1907	151	140	176
1908	126	123	185
1909	155	143	198
1910	159	147	208

1911	153	148	216
1912	177	155	226
1913	184	156	236
1914	169	152	244
1915	189	156	266
1916	225	183	298
1917	227	198	335
1918	223	201	366
1919	218	196	387
1920	231	194	407
1921	179	146	417
1922	240	161	431

Autor: Charles Cobb y Paul Douglas (1928,)

Usando la Tabla 3 y elaborando una tabla de valores de $x = ln\left(\frac{x_1}{x_2}\right)$ y $y = ln\left(\frac{f}{x_2}\right)$ y luego usando una regresión cuadrática para calcular la recta que pasa por los puntos $x = ln\left(\frac{x_1}{x_2}\right)$ y $y = ln\left(\frac{f}{x_2}\right)$, se obtienen los valores de $\alpha = 0.75$ y $k = 1.01$.

Consideremos el siguiente problema de optimización:

La producción total f de un determinado producto tiene como variable dependiente la cantidad x_1 de mano de obra usada y de la cantidad x_2 de inversión de capital. Considerando los supuestos de Cobb y Douglas dados anteriormente:

La función de Cobb Douglas esta dada, ver Felipe, Jesús y Adams, F. Gerard (2005).

$$f(x) = -k.{x_1}^{\alpha}{x_2}^{1-\alpha},$$

donde $k > 0$ y $0 < \alpha < 1$. Donde m *es el* costo unitario de mano de obra y *n* el costo unitario de capital, la empresa puede gastar su presupuesto pero sólo hasta *p* dólares del total, tomando *P* como función de producción que se maximiza bajo determinadas condiciones:

$$mx_1 + nx_2 \leq p$$

y por lo tanto, el problema de optimización consiste en:

$$\min\{f(x) = -k.{x_1}^{\alpha}{x_2}^{1-\alpha}: mx_1 + nx_2 \leq p, x_1 \geq 0 x_2 \geq 0\} \quad (4.2)$$

Usaremos el "Método del Punto Proximal", presentado en este artículo De la Cruz, L, Papa E. (2013), para encontrar la solución del problema.

4.5.2. Pruebas econométricas para obtener el valor de α y β

Se definen α y β, ver **2.6,** debemos estimar estos valores de las funciones de producción utilizando análisis de regresión lineal clásico, también se puede usar el método de regresión no lineal. En los dos casos se estiman los parámetros de la función mediante varios algoritmos; uno de ellos es con Mínimos Cuadrados Ordinarios (MCO), los cuales se obtienen de minimizar una de las desviaciones cuadráticas de los datos observados y los datos estimados de la variable dependiente. Otro caso se puede usar la Máxima Verosimilitud, (MV), el cual tiene algunas propiedades estadísticas más fuertes que MCO esto resulta de hallar los parámetros de manera que la probabilidad de observar un valor determinado de la variable dependiente sea máxima. En el caso de estimar el modelo de una función de produccción en el caso de Perú utilizando el software econométrico Eviews 8.

Como el tamaño muestral utilizado es pequeño, recomendamos tener cuidado al evaluar los resultados anteriores. Se sabe que la literatura dice que existen varias limitaciones con muestras pequeñas, porque pueden surgir sesgos al estimar por Mínimos Cuadrados Ordinarios (MCO).

4.5.3. Resultados con función Cobb-Douglas

Para hallar la relación entre las funciones del tipo Cobb-Douglas, se uso la transformación logarítmica en un modelo lineal:

$$Log\ Q_t\ = Log\ A_t\ + \propto Log\ L_t\ + \beta Log\ K_t \quad (4.3)$$

Para la especificación teórica 3.1 se presenta el siguiente modelo econométrico:

$$Y_t = C_0 + C_1 X_1 + C_2 X_2 + C_3 X_3 + \xi_t \qquad (4.4)$$

Donde $x_1 = Log\ A_t, x_2 = Log\ L_t,\ x_3 = Log\ K_t,$

Además de ello se estimó la serie en primeras diferencias para corregir el problema de la no estacionariedad de las variables.

En la función de producción se hallo que la participación del factor trabajo en el producto es aproximadamente a 0,50 (Ver Figura 4.2)

Quedando la formulación de esta manera $minf(x) = -{x_1}^{0.5}{x_2}^{0.5}$ (4.5)

Los valores de α y β, son las elasticidades de la producción en relación a los factores, Ver (Figura 4.2). muestran que los valores del estadístico "F", que corresponden a cada parámetro, vemos que ambos factores son significativos. La variable ficticia con el signo esperado, resultó significativa y, esto nos dice que, en efecto, la crisis del período 2008 recibió el peso del ajuste, el cual se manifestó fuertemente en una reducción de la producción. El uso de esta variable se determina al momento de observar la serie del PBI, se observa un cambio en el periodo 2008, la razón fundamental de éste quiebre es el impacto que tuvo de la crisis financiera internacional en la economía nacional. (Ver Figura 4.1) Los estadísticos Cusum y Cusum cuadrados indica que los coeficientes de regresión del ajuste son estables. (Ver Figura 4.3). El valor (R^2ajustado) es bastante alto, lo cual indica que la mayor parte de la producción podría explicarse por el aporte de los factores productivos. (Ver Figura 4.2), se utilizó el paquete econométrico Eviews 8.

Para la implementación usamos una computadora Core i5, 2.30 GHz, 3Gb de Ram, Windows 7 como sistema operativo con 64 bits usamos el Software MATLAB 7.10 .

En los ejemplos que estamos usando, tomaremos $\lambda_k = 0.125$, el punto inicial (1, 2) y el error de tolerancia 0.0001.

Problema 1: Consideremos el problema aplicando al resultado de Cobb-Douglas (Ver 4.2 y 4.5):

$$\min\{f(x) = -{x_1}^{0.5}{x_2}^{0.5} : 2x_1 + x_2 \leq 6, x_1 \geq 0, x_2 \geq 0\}$$

Notemos que la función $f: \mathbb{R}^2 \to \mathbb{R} \cup \{+\infty\}$ definido por $f(x_1, x_2) = {x_1}^{0.5}{x_2}^{0.5}$, f es propia, pues

$$dom\, f = \{(x_1, x_2) \in \mathbb{R}^2 / {x_1}^{0.5}{x_2}^{0.5} < +\infty\} \neq \emptyset$$

$$\forall\, (x_1, x_2) \in dom\, f: -\infty < {x_1}^{0.5}{x_2}^{0.5}.$$

Además, la función f es continua (así es semicontinua inferior) y el punto óptimo, $x^* = (3/2,3)$, es el punto de mínimo global de f.

Tomamos para este problema la distancia proximal:

$$D(x, z) = d\big(y(x), y(z)\big),$$

Donde:

D es la distancia proximal, donde d será la distancia dada con Kullback-Leibler, ver definición 2.2.1.:

$$d(u, v) = \sum_{i=1}^{2} u_i \log\left(\frac{u_i}{v_i}\right) + v_i - u_i,$$

Que reemplazada en la función D, da:

$$D(x, z) = (6 - 2x_1 - x_2) log\left(\frac{6 - 2x_1 - x_2}{6 - 2z_1 - z_2}\right) + (2z_1 + z_2 - 2x_1 - x_2) +$$

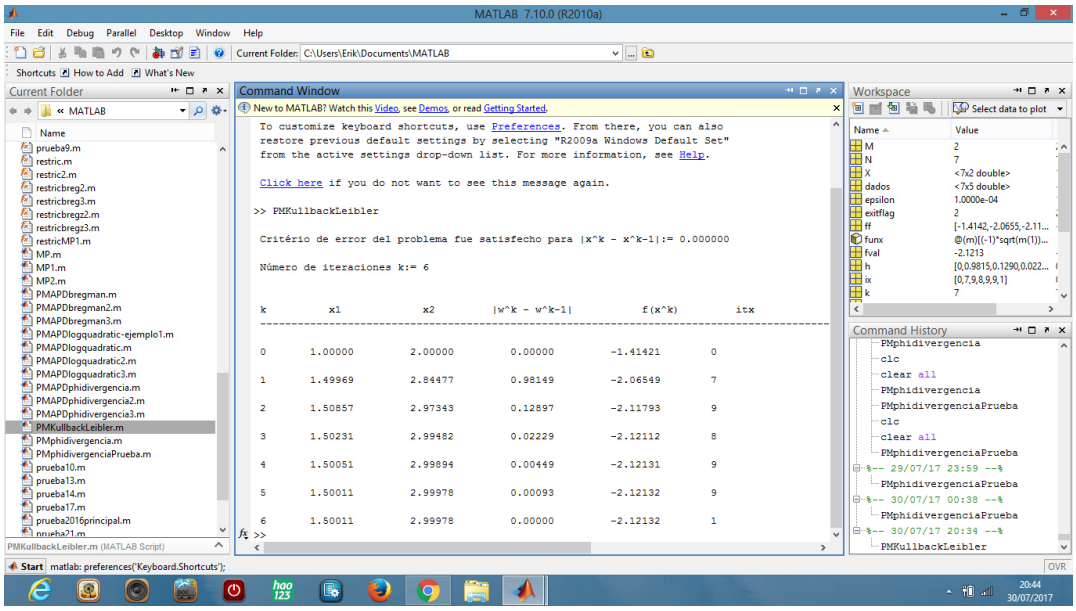
$$\left(x_1 log\left(\frac{x_1}{z_1}\right) + z_1 - x_1\right) + \left(x_2 log\left(\frac{x_2}{z_2}\right) + z_2 - x_2\right)\text{h}$$

La función regularizada es:

$f_k = -{x_1}^{0.5}{x_2}^{0.5} + \lambda_k D(x, x^{k-1})$.

Los resultados de las iteraciones se dan en la figura 4.4.

Figura 4.4 *Distancia proximal de Kullback-Leibler:*



Distancia proximal de Kullback-Leibler

N^0	x_s	FObj	Criterio de error
0	(1.00000 2.00000)	-1.41421	0.77587
1	(1.49969 2.84477)	-2.06549	0.98149
2	(1.50857 2.97343)	-2.11793	0.12897
3	(1.50231 2.99482)	-2.12112	0.02229
4	(1.50051 2.99894)	-2.12131	0.00449
5	(1.50011 2.99978)	-2.12132	0.00093
6	(1.50011 2.99978)	-2.12132	0.00000

Problema 2: Consideremos el problema aplicando al resultado de Cobb-Douglas (Ver 4.2 y 4.5) donde hallaremos las variables y la función objetivo de:

$$\min\{f(x) = -{x_1}^{0.5}{x_2}^{0.5}: 2x_1 + x_2 \leq 6, x_1 \geq 0 x_2 \geq 0\}$$

Tomamos para este problema la distancia proximal:

$$D(x,z) = d\big(y(x), y(z)\big),$$

Pero ahora la distancia d estará dada por la distancia φ -divergencia, ver definición 2.2.2.

$$d(u,v) = \sum_{i=1}^{2} \left(v_i log\left(\frac{v_i}{u_i}\right) + u_i - v_i \right)$$

Que reemplazada en la función D, da:

$$D(x,z) = (6 - 2z_1 - z_2) log\left(\frac{6 - 2z_1 - z_2}{6 - 2x_1 - x_2}\right) + (2x_1 + x_2 - 2z_1 - z_2) +$$

$$\left(z_1 log\left(\frac{z_1}{x_1}\right) + x_1 - z_1\right) + \left(z_2 log\left(\frac{z_2}{x_2}\right) + x_2 - z_2\right).$$

La función regularizada es:

$$f_k = -{x_1}^{0.5}{x_2}^{0.5} + \lambda_k D(x, x^{k-1}).$$

Las iteraciones las obtenemos en la figura 4.5:

Figura 4.5

Distancia φ – divergencia

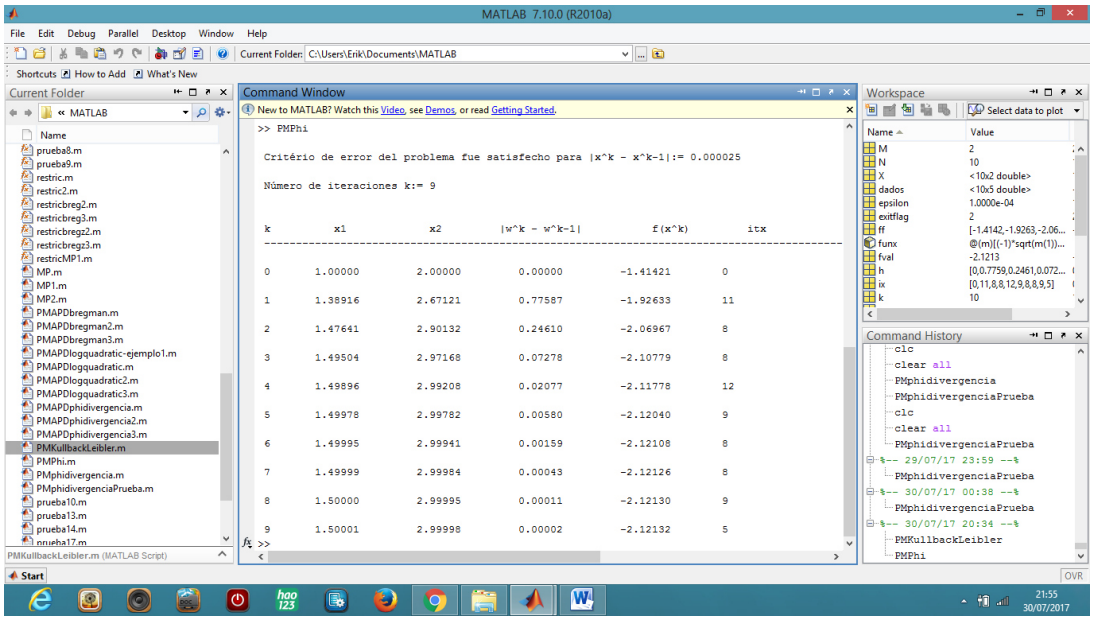


Tabla 5

Distancia φ − divergencia

N^0	x_s		FObj	Criterio de error
0	(1.00000	2.00000)	-1.41421	0.77587
1	(1.38916	2.67121)	-1.92633	0.77587
2	(1.47641	2.90132)	-2.06967	0.24610
3	(1.49504	2.97168)	-2.10779	0.07278
4	(1.49896	2.99208)	-2.11778	0.02077
5	(1.49978	2.99782)	-2.12040	0.00580
6	(1.49995	2.99941)	-2.12108	0.00159
7	(1.49999	2.99984)	-2.12126	0.00043
8	(1.50000	2.99995)	-2.12130	0.00011
9	(1.50001	2.99998)	-2.12132	0.00002

Problema 3: Consideremos el problema:

$$\min\{f(x) = -{x_1}^{0.5}{x_2}^{0.5}: 2x_1 + x_2 \leq 6, x_1 \geq 0 x_2 \geq 0\}$$

Ya sabemos que la solución óptima es:

$x_1 = 3/2, x_2 = 3$

Tomamos para este problema la distancia proximal:

$$D(x,z) = d\big(y(x), y(z)\big),$$

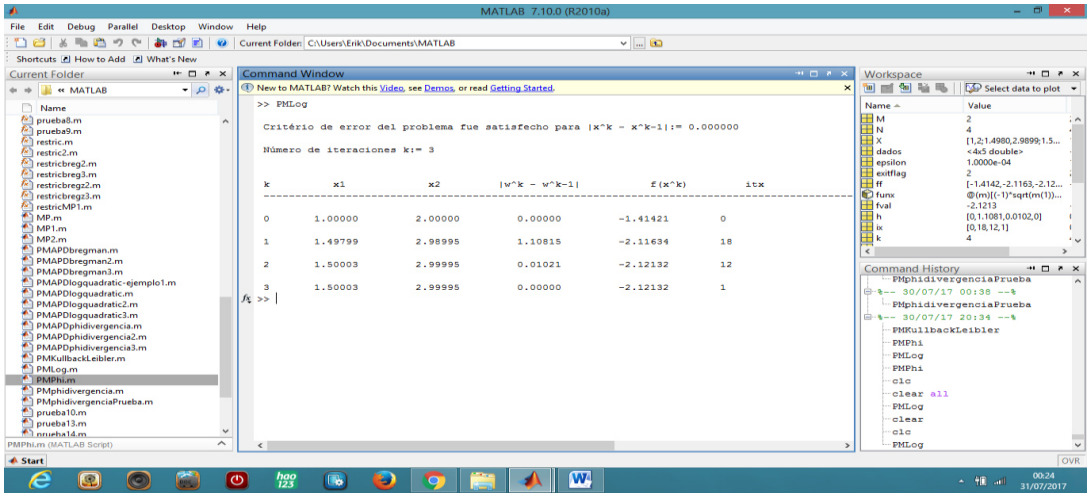
Pero ahora la distancia d, estará dada por la distancia Logaritmo cuadrático, ver Xu Y, Bingsheng H, Xiaoming Y (2006), ver definición 2.2.3.

.

$$d(u,v) = \sum_{i=1}^{2}\left({u_i}^2 - u_i v_i - {v_i}^2 \log\left(\frac{v_i}{u_i}\right)\right),$$

Figura 4.6

Distancia Logaritmo cuadrático



N^0	x_s		FObj	Criterio de error
0	(1.00000	2.00000)	-1.41421	0.00000
1	(1.49799	2.98995)	-2.11634	1.10815
2	(1.50003	2.99995)	-2.12132	0.01021
3	(1.50003	2.99995)	-2.12132	0.00000

4.6 Experimentos Computacionales

En cada problema académico se aplicará el metodo proximal usando las distintas distancias proximales como son las distancias ϕ –divergencias, distancias homogéneas de segundo orden llamado *distancia Logaritmo cuadrático y Distancia proximal de Kullback-Leibler con* $\lambda_k = 1/k, \forall k = 1,2,\ldots$, asi como los programas del método usado.

Ejemplo 4.5.2 Consideremos el problema

$$\begin{cases} \min\ (\ln x_1)^2 + (\ln x_2)^2 \\ \text{s. a} \\ (x_1, x_2) \in \mathbb{R}^2_{++} \end{cases} \qquad \ldots(4.6)$$

Notemos que la función $f: \mathbb{R}^2 \to \mathbb{R} \cup \{+\infty\}$ definido por $f(x_1, x_2) = (\ln x_1)^2 + (\ln x_2)^2$, f es propia, pues

$$dom\ f = \{(x_1, x_2) \in \mathbb{R}^2 / (\ln x_1)^2 + (\ln x_2)^2 < +\infty\} \neq \emptyset$$

$$\forall\,(x_1, x_2) \in dom\, f\colon\ -\infty < \ (\ln x_1)^2\ + (\ln x_2)^2\ .$$

Además, la función f es continua (así es semicontinua inferior) y el punto $x^* = (1,1)$, es el punto de mínimo global de f.

Para aplicar el algoritmo consideremos como punto inicial $x^0 = (4,9)$ y los escalares $\lambda_k = \ 1/\ k$ para cada $k\ =\ 1,2,\ldots$, debemos obtener un x^k tal que

$$0\ \in \hat{\partial}\ (f(\cdot)\ +\ \lambda_k d(\cdot, x^{k-1}))(x^k)$$

resolveremos estos subproblemas usando una distancia homogénea de segundo orden**,** Pan S, Chen JS (2007).

$$d(x, y)\ = \sum_{i=1}^{2}(x_i^2 -\ x_i y_i -\ y_i^2 \log\,(\tfrac{x_i}{y_i}))$$

la cual es llamada distancia Logaritmo-cuadrático.

Primera Iteración:

Como $\nabla f(x^0)\ =\ (0.6931,0.4883)\ \neq (0,0)$ tomamos un punto $x^0\ =\ (x_1, x_2)$, tal que

$$0\ \in \hat{\partial}\ (f(\cdot)\ +\ 1.d(\cdot,(4,9)))(x)$$

Ver **definición 2.2.3**

Ejemplo 4.5.3

Resolver el sistema del ejemplo 4.5.2

$$\nabla f(x_1, x_2)\ +\ 1.\nabla_1 d((x_1, x_2),(4,9))\ =\ (0,0)$$

es decir, debemos resolver el sistema

$$\tfrac{2}{x_1}\, lnx_1\ +\ 1.(2x_1\ - 4 - \tfrac{16}{x_1})\ =\ 0$$

$$\tfrac{2}{x_2}\, lnx_2\ +\ 1.(2x_2\ - 9 - \tfrac{81}{x_2})\ =\ 0$$

cuya solución x^1 , es

$$(x_1, x_2) = (3.7701, 8.8367)$$

con $x^1 = (3.7701, 8.8367)$, pasamos a la segunda iteración.

Segunda Iteración:

Como $\nabla f(x^1) = (0.7040, 0.4932) \neq (0,0)$ debemos obtener un punto $x^1 = (x_1, x_2)$, tal que

$$0 \in \hat{\partial} f(\cdot) + 1/2.d(\cdot, (3.7701,8.8367)))(x)$$

notamos que debemos resolver el sistema

$$\nabla f(x_1, x_2) + \frac{1}{2}.\nabla_1 d((x_1, x_2), (3.7701,8.8367)) = (0,0)$$

es decir, debemos resolver el sistema

$$\frac{2}{x_1} lnx_1 + \frac{1}{2}.(2x_1 - 3.7701 - \frac{14.2137}{x_1}) = 0$$

$$\frac{2}{x_2} lnx_2 + \frac{1}{2}.(2x_2 - 8.8367 - \frac{78.0873}{x_2}) = 0$$

cuya solución es

$(x_1, x_2) = (3.3092, 8.5053)$

así $x^2 = (3.3092,8.5053)$ y pasamos a la tercera iteración.

Tercera Iteración:

Como $\nabla f(x^2) = (0.7233, 0.5034) \neq (0,0)$ debemos obtener un punto $x^2 = (x_1, x_2)$, tal que

$$0 \in \hat{\partial} f(\cdot) + 1/3.d(\cdot, (3.3092,8.5053)))(x)$$

notamos que debemos resolver el sistema

$$\nabla f(x_1, x_2) + \frac{1}{3}.\nabla_1 d((x_1, x_2), (3.3092,8.5053)) = (0,0)$$

es decir, debemos resolver el sistema

$$\frac{2}{x_1} lnx_1 + \frac{1}{3}.(2x_1 - 3.3092 - \frac{10.9508}{x_1}) = 0$$

$$\frac{2}{x_2} lnx_2 + \frac{1}{3}.(2x_2 - 8.5053 - \frac{72.3401}{x_2}) = 0$$

cuya solución es

$$(x_1, x_2) = (2.6319, 7.9962)$$

así $x^3 = (2.6319, 7.9962)$ y pasamos a la siguiente iteración.

De esta manera procede el algoritmo hasta que $\|\nabla f(x^k)\| < \xi$, donde ξ es un margen de error permitido por el programador. En el programa siguiente tomaremos $\xi = 10^{-6}$.

A continuación, mostraremos el programa del algoritmo que fue implementado en el software MATLAB para resolver este ejemplo y seguidamente mostraremos los resultados numéricos de la implementación.

Programa del método

$x = [4; 9];$ % punto inicial

$s = 1;\ sigma = 0.1;\ beta = 0.5;$ % parámetros dados, ver definición 2.3.21.

$\xi = 0.000001;$ % error permitido para resolver el problema

$tol = 0.000001;$ % error permitido para resolver los subproblemas

$Nmax = 100;$ % número máximo de iteraciones para resolver cada subproblema

$Mmax = 100;$ % número máximo de iteraciones para resolver el problema

$B = [1\ 0\ ;\ 0\ 1\ ];$ % matriz inicial para el método Cuasi-Newton

$M = 1;$ %contador del número de iteraciones

fprintf('\n Punto Inicial: %f', x)

$a = [4; 9];$

while M<=Mmax

fprintf('\n Iteración N: %d',M)

$G = [2 * (1/x(1)) * log(x(1)); 2 * (1/x(2)) * log(x(2))]$

Normadf= $sqrt(G(1,:)^2 + G(2,:)^2)$

if Normadf >= $epsilon$ % compara con el máximo error permitido "epsilon"

% empieza el algoritmo Cuasi-Newton

$k = 1;$

while k<=Nmax,

$G = [2 * (1/x(1)) * log(x(1)) + (1/M) * (2 * x(1) - a(1) - (a(1)^(2)/x(1)));$

$2 * (1/x(2)) * log(x(2)) + (1/M) * (2 * x(2) - a(2) - (a(2)^(2)/$
$x(2)))];$%gradienteenx

$P = -inv(B) * G$;

V=B; % salvando el valor de B

Normadf= $sqrt(G(1,:)^2 + G(2,:)^2);$

if Normadf >= tol

m=0; % iniciamos la búsqueda unidimensional

$factual = log(x(1))^2 + log(x(2))^2 + (1/M) * (x(1)^2 - x(1) * a(1)$
$- (a(1)^2$

$* log(x(1)/a(1))) + x(2)^2 - x(2) * a(2) - (a(2)^2 * log(x(2)/a(2))));$

fnuevo $= log(x(1) + P(1))^2 + log(x(2) + P(2))^2 + (1/M) * ((x(1) + P(1))^2$

$(x(1) + P(1)) * a(1) - (a(1)^2 * log((x(1) + P(1))/a(1))) + (x(2) + P(2))^2$

$-(x(2)+P(2))*a(2)-(a(2)\text{^}2*log((x(2)+P(2))/a(2))));$

dif=fnuevo-factual;

lambda=1;

test=sigma*lambda*G'*P ;

while dif > test,

m=m+1;

lambda=beta^m*s ;

x1=x(1)+lambda*P(1);

x2=x(2)+lambda*P(2);

dif=log(x1)^2+log(x2)^2+(1/M)*((x1)^2-(x1)*a(1)-(a(1)^2*log(x1/a(1)))

+(x2)^2-(x2)*a(2)-(a(2)^2*log((x2)/a(2))))-1*(log(x(1))^2+log(x(2))^2

+(1/M)*(x(1)^2-x(1)*a(1)-(a(1)^2*log(x(1)/a(1)))+ x(2)^2-x(2)*a(2)

(a(2)^2*log(x(2)/a(2)))));

test=sigma*lambda*G'*P;

end %terminamos la búsqueda unidimensional

w=x; % salvando los valores de x

x=w+lambda*P;

fprintf('\n nuevo punto x1=: %f',x(1));

fprintf('\n nuevo punto x2=: %f',x(2));

%gradiente en el antiguo x

$T=[2*(1/w(1))*log(w(1))+(1/M)*(2*w(1)-a(1)-(a(1)\text{^}2/w(1)));$

$2 * (1/w(2)) * log(w(2)) + (1/M) * (2 * w(2) - a(2) - (a(2)\text{^}2/w(2)))];$

%gradiente en el nuevo x

$H = [2 * (1/x(1)) * log(x(1)) + (1/M) * (2 * x(1) - a(1) - (a(1)\text{^}2/x(1)));$

$2 * (1/x(2)) * log(x(2)) + (1/M) * (2 * x(2) - a(2) - (a(2)\text{^}2/x(2)))];$

S=x-w;

y=H-T;

$B = V - ((V * S) * (V * S)')/(S' * V * S) + (y * y')/(y' * S);$

$k = k + 1;$

else

fprintf('\n criterio de error del subproblema satisfecho=:%f',Normadf)

fprintf('\n el punto x1 aproximado es: %f',x(1))

fprintf('\n el punto x2 aproximado es: %f',x(2))

fprintf('\n número de iteraciones del subproblema es:%d',k-1)

k=1000000;

end

end

% termina el algoritmo Cuasi-Newton

a=x;

M=M+1;

else

fprintf('\n criterio de error satisfecho Normagradiente=:%f',Normadf)

```
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 número de iteraciones es: %d',M-1) %

M =1000000;

end

end

if M > 999999

fprintf('\n resolvimos el problema aproximadamente:%f')

else

fprintf('\n : número de iteraciones excedido a%d',M-1)

end
```

Los resultados numéricos son presentados en la tabla 6.

Tabla 6

Distancia Logaritmo-cuadrático.

N^0	x_s	FObj	NorGrad
1	(3.770004, 8.836622)	6.5088	0.8596
2	(3.309218, 8.505339)	6.0147	0.8812

3	(2.631944, 7.996157)	5.2586	0.9006
4	(1.853270, 7.292258)	4.3281	0.8604
5	(1.285544, 6.367920)	3.4903	0.7006
6	(1.064830, 5.188019)	2.7144	0.6455
7	(1.011773, 3.730841)	1.7337	0.7062
8	(1.001868, 2.186358)	0.6119	0.7156
9	(1.000267, 1.271684)	0.0578	0.3780
10	(1.000035, 1.039644)	0.0015	0.0748
11	(1.000004, 1.004863)	2.3276e-005	0.0096
12	(1.000001, 1.000538)	2.8955e-007	0.0011
13	(1.000000, 1.000056)	3.1013e-009	1.1137e-004
14	(1.000000, 1.000005)	2.9057e-011	1.0781e-005
15	(1.000000, 1.000001)	3.2871e-013	1.1467e-006
16	(1.000000, 1.000000)	2.7250e-015	1.0440e-007

En los resultados de la tabla 6, podemos observar de los ejemplos 4.5.2 y 4.5.3 que eligiendo a la distancia homogénea de segundo orden (Logaritmo-cuadrático) en la definición del algoritmo, resolvimos el problema en una mayor cantidad de iteraciones en comparación de la distancia de Bregman (Kullback-Leibler) ver definición 2.2.1.

Ejemplo 4.5.4 Para el mismo problema del ejemplo 4.5.3, si consideramos en cada subproblema los parámetros $\lambda_k = 1$ para cada $k = 1,2,\ldots$, cambiando en el programa anterior 1 /M por 1, obtenemos los resultados numéricos que se presentan en la tabla 7.

Tabla 7

Distancia de Bregman (Kullback-Leibler)

N^0	x_s	FObj	NorGrad
1	(3.770004, 8.836622)	6.5088	0.8596
2	(3.537013, 8.671606)	6.2617	0.8709
3	(3.301437, 8.504898)	6.0088	0.8814
4	(3.063943, 8.336445)	5.7508	0.8905
5	(2.825573, 8.166190)	5.4889	0.8973
6	(2.587902, 7.994072)	5.2251	0.9003
7	(2.353257, 7.820028)	4.9624	0.8976
8	(2.124957, 7.643992)	4.7050	0.8868
9	(1.907515, 7.465895)	4.4585	0.8652
10	(1.706601, 7.285662)	4.2295	0.8304
11	(1.528470, 7.103219)	4.0238	0.7829
12	(1.378629, 6.918483)	3.8442	0.7277
13	(1.260051, 6.731372)	3.6892	0.6750
14	(1.171984, 6.541797)	3.5529	0.6349
15	(1.110301, 6.349667)	3.4275	0.6120
16	(1.069141, 6.154886)	3.3069	0.6036
.	.	.	.
.	.	.	.
.	.	.	.
64	(1.000000, 1.000001)	1.1461e-012	2.1412e-006
65	(1.000000, 1.000001)	4.1261e-013	1.2847e-006
66	(1.000000, 1.000000)	1.4854e-013	7.7082e-007

En el ejemplo 4.5.4 notamos la importancia que toma la elección de los parámetros λ_k para definir la función objetivo de cada subproblema, ya que en la tabla 7, podemos observar que la solución del ejemplo 4.5.4 se ha obtenido con más iteraciones, cuando tomamos $\lambda_k = 1$ que cuando tomamos $\lambda_k = 1/k$ para cada $k = 1,2,\ldots$.

Así podemos concluir que con una mejor elección de los parámetros λ_k, en este caso $\lambda_k = 1/k$, podemos obtener mejores resultados computacionales.

Ejemplo 4.5.5 Consideremos el problema

$$\begin{cases} \min\ (\ln x_1)^2\ +\ (\ln x_2)^2 \\ \text{s. a} \\ (x_1, x_2)\ \in\ IR^2_{++} \end{cases}$$ Para aplicar el algoritmo 2.6. escojamos como punto inicial $x^0 = (4,9)$ y los escalares $\lambda_k = 1/k$ para cada $k = 1,2,\ldots$, debemos obtener un x^k tal que

$$0 \in \hat{\partial}\{\, f(\cdot) + \lambda_k\, d(.\,, x^{k-1})\,\}(x^k)$$

resolveremos estos subproblemas usando una distancia φ -divergencia definida por:

$$d(x,y)\ = \sum_{i=1}^{2} \left(y_i \log\left(\frac{y_i}{x_i}\right) + x_i - y_i \right)$$

Teóricamente para la

Primera Iteración:

Como $\nabla f(x^0) = (0.6931, 0.4883) \neq (0,0)$ debemos obtener un punto $x^0 = (x_1, x_2)$, tal que

$$0\ \in \hat{\partial}\ f(\cdot)\ +\ 1.d(\cdot,(4,9)))(x)$$

notamos que por la definición 2.2.2, debemos resolver el sistema

$$\nabla f(x_1, x_2)\ +\ 1.\nabla_1 d((x_1, x_2),(4,9))\ =\ (0,0)$$

es decir, debemos resolver el sistema

$$\frac{2}{x_1}\, lnx_1\ +\ 1.(1\ -\ \frac{4}{x_1})\ =\ 0$$

$$\frac{2}{x_2} lnx_2\ +\ 1.(1\ -\frac{9}{x_2})\ =\ 0$$

cuya solución es

$$(x_1, x_2)\ =\ (2.3184, 5.5665)$$

así $x^1\ =\ (2.3184, 5.5665)$ y pasamos a la segunda iteración.

Segunda Iteración:

Como $\nabla f(x^1) = (0.7254,0.6168) \neq (0,0)$ tomamos el punto $x^1 = (x_1, x_2)$, tal que

$$0 \in \hat{\partial} f(\cdot) + \frac{1}{2} . d(\cdot, (2.3184,5.5665)))(x)$$

notamos que debemos resolver el sistema

$$\nabla f(x_1, x_2) + \frac{1}{2} . \nabla_1 d((x_1, x_2), (2.3184,5.5665)) = (0,0)$$

es decir, debemos resolver el sistema

$$\frac{2}{x_1} lnx_1 + \frac{1}{2} . (1 - \frac{2.3184}{x_1}) = 0$$

$$\frac{2}{x_2} lnx_2 + \frac{1}{2} . (1 - \frac{5.5665}{x_2}) = 0$$

cuya solución es

$$(x_1, x_2) = (1.2924, 2.2763)$$

así $x^2 = (1.2924,2.2763)$ y pasamos a la tercera iteración.

Tercera Iteración:

Como $\nabla f(x^2) = (0.3969,0.7227) \neq (0,0)$ debemos obtener un punto $x^2 = (x_1, x_2)$, tal que

$$0 \in \hat{\partial} f(\cdot) + \frac{1}{3} . d(\cdot, (1.2924,2.2763)))(x)$$

notamos que debemos resolver el sistema

$$\nabla f(x_1, x_2) + \frac{1}{3} . \nabla_1 d((x_1, x_2), (1.2924,2.2763)) = (0,0)$$

es decir, debemos resolver el sistema

$$\frac{2}{x_1} lnx_1 + \frac{1}{3} . (1 - \frac{1.2924}{x_1}) = 0$$

$$\frac{2}{x_2} lnx_2 + \frac{1}{3} . (1 - \frac{2.2763}{x_2}) = 0$$

cuya solución es

$(x_1, x_2) = (1.0426,1.1971)$

así $x^3 = (1.0426,1.1971)$ y pasamos a la cuarta iteración.

Cuarta Iteración:

Como $\nabla f(x^3) = (0.0094,0.0428) \neq (0,0)$ tomamos el punto $x^3 = (x_1, x_2)$, tal que

$0 \in \hat{\partial} f(\cdot) + + \frac{1}{4}.d(\cdot,(1.0426,1.1971)))(x)$

notamos que debemos resolver el sistema

$\nabla f(x_1, x_2) + \frac{1}{4}.\nabla_1 d((x_1, x_2),(1.0426,1.1971)) = (0,0)$

es decir, debemos resolver el sistema

$\frac{2}{x_1} lnx_1 + \frac{1}{4}.(1 - \frac{1.0426}{x_1}) = 0$

$\frac{2}{x_2} lnx_2 + \frac{1}{4}.(1 - \frac{1.1971}{x_2}) = 0$

cuya solución es

$(x_1, x_2) = (1.0047,1.0221)$

Así $x^4 = (1.0047,1.0221)$ y pasamos a la siguiente iteración.

De esta manera procede el algoritmo hasta que $\|\nabla f(x^k)\| < \xi$, donde ξ es un margen de error permitido por el programador. En el programa siguiente tomaremos $\xi = 10^{-6}$. A continuación, mostraremos el programa del ver definición 2.2.2, que fue implementado en el software MATLAB para resolver este ejemplo y seguidamente mostraremos los resultados numéricos de la implementación.

Programa del metodo restricto

x=[4;9]; % punto inicial

s=1; sigma=0.1; beta=0.5; % parámetros dados, ver definición 2.3.21

ξ=0.000001; % error permitido para resolver el problema

tol=0.000001; % error permitido para resolver los subproblemas

Nmax=100; % número máximo de iteraciones para resolver cada subproblema

Mmax=100; % número máximo de iteraciones para resolver el problema

B=[1 0 ; 0 1]; % matriz inicial para el método Cuasi-Newton

M=1; %contador del número de iteraciones

fprintf('\n Punto Inicial: %f',x)

a=[4;9];

while M<=Mmax

fprintf('\n Iteración N: %d',M)

G=[2*(1/x(1))*log(x(1));2*(1/x(2))*log(x(2))]

Normadf=sqrt(G(1,:)^2+G(2,:)^2)

if Normadf >=epsilon % compara con el máximo error permitido "epsilon"

% empieza el algoritmo Cuasi-Newton

k=1;

while k<=Nmax,

$G = [2 * (1/x(1)) * log(x(1)) + (1/M) * (-(a(1)/x(1)) + 1);$

$2 * (1/x(2)) * log(x(2)) + (1/M) * (-(a(2)/x(2)) + 1)];\ \%\ gradiente\ en\ x$

P=-inv(B)*G;

V=B; %salvando el valor de B

Normadf=sqrt(G(1,:)^2+G(2,:)^2);

if Normadf >= tol

m=0; % iniciamos la búsqueda unidimensional

factual=log(x(1))^2+log(x(2))^2+(1/M)*(a(1)*log(a(1)/x(1))

-a(1)+x(1)+ a(2)*log(a(2)/x(2))-a(2)+x(2));

fnuevo=log(x(1)+P(1))^2+log(x(2)+P(2))^2+(1/M)*(a(1)*log(a(1)/(x(1)+P(1))

$-a(1) + (x(1) + P(1)) + \ a(2) * log(a(2)/(x(2) + P(2))) - a(2) + x(2) + P(2));$

dif=fnuevo-factual;

lambda=1;

test=sigma*lambda*G'*P;

while dif > test,

m=m+1;

lambda=beta^m*s;

x1=x(1)+lambda*P(1);

x2=x(2)+lambda*P(2);

$dif = log(x1)^2 + log(x2)^2 + (1/M) * (a(1) * log(a(1)/(x1)) - a(1) + (x1)$
$+ a(2) *$

$log(a(2)/(x2)) - a(2) + x2) - 1 * (log(x(1))^2 \ + \ log(x(2))^2 + (1/M) *$

$(a(1) * log(a(1)/x(1)) - a(1) + x(1) + \ a(2) * log(a(2)/x(2)) - a(2) + x(2)));$

test=sigma*lambda*G'*P ;

end % terminamos la búsqueda unidimensional

```
w=x; % salvando los valores de x

x=w+lambda*P;

fprintf('\n nuevo punto x1=: %f',x(1));

fprintf('\n nuevo punto x2=: %f',x(2));

T=[2*(1/w(1))*log(w(1))+(1/M)*(-(a(1)/w(1))+1);

2*(1/w(2))*log(w(2))+(1/M)*(-(a(2)/w(2))+1)];%gradienteenelantiguox

H=[2*(1/x(1))*log(x(1))+(1/M)*(-(a(1)/x(1))+1);

2*(1/x(2))*log(x(2))+(1/M)*(-(a(2)/x(2))+1)]; %gradiente en el nuevo x

S=x-w;

y=H-T;

B=V-((V*S)*(V*S)')/(S'*V*S)+(y*y')/(y'*S);

k=k+1;

else

fprintf('\n criterio de error del subproblema satisfecho=:%f',Normadf)

fprintf('\n el punto x1 aproximado es: %f',x(1))

fprintf('\n el punto x2 aproximado es: %f',x(2))

fprintf('\n número de iteraciones del subproblema es:%d',k-1)

k=1000000;

end

end

% termina el algoritmo Cuasi-Newton
```

```
a=x;

M=M+1;

else

fprintf('\n criterio de error satisfecho Normagradiente=:%f',Normadf)

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 número de iteraciones es: %d',M-1) %

M =1000000;

end

if M > 999999

fprintf('\n resolvimos el problema aproximadamente:%f')

else

fprintf('\n : número de iteraciones excedido a%d',M-1)

end
```

Los resultados numéricos son presentados en la Tabla 8.

Tabla 8.

Distancia φ -divergencia con $\lambda_k = 1$

N^0	N_s	x_s	FObj	NorGrad

1	10	(2.318317, 5.566476)	3.6543	0.9522
2	9	(1.292379, 2.276289)	0.7424	0.8245
3	7	(1.042522, 1.197063)	0.0341	0.3110
4	5	(1.004735, 1.022110)	5.0057e-004	0.0438
5	3	(1.000430, 1.002012)	4.2247e-006	0.0041
6	2	(1.000033, 1.000155)	2.5036e-008	3.1641e-004
7	2	(1.000002, 1.000010)	1.1128e-010	2.1098e-005
8	1	(1.000000, 1.000001)	5.1155e-013	1.4305e-006
9	1	(1.000000, 1.000000)	1.7746e-015	8.4252e-008

Ejemplo 4.5.7 Para el mismo problema del ejemplo 4.5.6 si consideramos en cada subproblema los parámetros $\lambda_k = 1$ para cada k = 1,2,..., cambiando en el programa anterior 1/ M por 1, obtenemos los resultados numéricos presentados en la tabla 9.

Tabla 9

Distancia φ -divergencia con $\lambda_k = 1/M$

N^0	N_s	x_s	FObj	NorGrad
1	10	(2.318317, 5.566476)	3.6543	0.9522
2	6	(1.503168, 3.224759)	1.5370	0.9063
3	5	(1.177085, 1.920053)	0.4521	0.7338
4	7	(1.060190, 1.337880)	0.0881	0.4489
5	5	(1.020197, 1.116852)	0.0126	0.2017
6	4	(1.006748, 1.039456)	0.0015	0.0756

7	3	(1.002251, 1.013210)	1.7728e-004	0.0263
8	3	(1.000751, 1.004410)	1.9923e-005	0.0089
9	3	(1.000250, 1.001471)	2.2221e-006	0.0030
.				
15	1	(1.000000, 1.000002)	4.1749e-012	4.0865e-006
16	1	(1.000002, 1.000001)	4.6387e-013	1.3622e-006
17	1	(1.000000, 1.000000)	5.1541e-014	4.5405e-007

Capítulo 5

5. IMPACTOS

• Vemos que el trabajo nos ayuda en problemas de minimización con funciones objetivos cuasi-convexas en el progreso y la obtención de una solución eficiente usando una extensión del método de punto proximal. Anteriores trabajos resultan ser casos particulares de nuestro enfoque general**.**

• De los ejemplos numéricos podemos ver que con una mejor elección de los parámetros λ_k, podemos obtener mejores resultados computacionales, en todos los casos la mejor elección es $\lambda_k = 1/k$

Esta tesis tiene dos contribuciones:

El “Método del Punto Proximal” funciona bien en problemas grandes no diferenciables, convexos

Capítulo 6

CONCLUSIONES

Se presentó el Método del Punto Proximal, para minimizar funciones cuasi-convexas que aparecen en la microeconomía, presentado es globalmente convergente, que significa que, dado cualquier punto inicial, la sucesión generada por el algoritmo converge en el caso cuasi-convexo y converge débilmente para el caso no cuasi-convexo, en el ortante no negativo, damos una condición suficiente en la distancia proximal para obtener la convergencia del problema y la solución óptima en modelos económicos para la toma de decisiones, con funciones no convexas.

Uno de los resultados más importantes de esta investigación es que se ha podido extender el método proximal clásico a un método proximal para resolver problemas de maximización y minimización con funciones objetivos no convexas, en cada iteración para resolver los subproblemas solo pedimos hallar en la función regularizadora un punto estacionario, y que no siempre sea un mínimo global.

Se ha realizado experimentos computacionales para verificar que método de punto proximal propuesto a modelos de decisión en microeconomía en un modelo de Cobb Douglas es implementable y aplicable a problemas académicos tanto para el caso restricto e irrestricto.

RECOMENDACIONES

La obtención de la convergencia a un punto generalizado de KKT es aceptable pero no óptima, en futuros trabajos en este tema se puede considerar modificar un poco el método para obtener convergencia a un punto óptimo.

La desventaja es que la convergencia es solamente lineal, por esto se recomienda en futuros trabajos usar otras variantes para mejorar la velocidad de la convergencia para funciones no diferenciables y no convexas.

Capítulo 7

7. REFERENCIAS BIBLIOGRÁFICAS

Alvarez F, Bolte J, Brahic O (2004) Hessian riemannian gradient flows in convex programming. SIAM J. Optim., 43, 2: 477-501

Arrow KJ, Enthoven AC (1961) Quasi-concave programming. Econometrica, 29, 4: 779-800

Attouch H, Bolte J (2009) On the convergence of the proximal algorithm for nonsmooth functions involving analytic features. Mathematical Programming, 116, 1-2, Ser. B: 5-16.

Attouch H, Bolte J, Svaiter B (2010) Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods. Mathematical Programming, Ser A, online first doi:10.1007/s10107-011-0484-9.

Attouch H, Teboulle M (2004) Regularized Lotka-Volterra dynamical system as continuous proximal-like method in optimization. Journal Optimization Theory and Applications, 121: 541-580

Auslender A, Teboulle M (2006) Interior gradient and proximal methods for convex and conic optimization. SIAM J. Optim., 16, 3: 697-725, 2006

Auslender A, Teboulle M, Ben-Tiba S (1999) Interior proximal and multiplier methods based on second order homogeneous functionals. Math Oper. Research 24, 3: 645-668

Azofeifa V Ana. & Marlene Villanueva S. (19 Referencias [1] Ana Azofeifa V. & Marlene Villanueva S. (1996). Estimación de una función de producción: Caso de Costa Rica. Banco Central de Costa Rica

Barbolla R., Cerdá E., Sanz P., Optimización: cuestiones ejercicios y aplicaciones a la economía, Editorial Prentice Hall, 2000.

Bauschke, HH, Borwein JM (1997) Legendre functions and the method of random Bregman projections. Journal of Convex Analysis, 4: 27-67

Bazaara MS, Sherali HD, Shetty CM (1993) Nonlinear programming: theory and algorithms, 2nd Edition, John Wiley and Sons, New York, NY

Cobb, C. W. y Douglas, P. H. (1928): “A Theory of Production”; American Economic Review, vol. 18, pp. 139 – 165. Comisión Europea (2007); European Economy, nº 8.

Cunha FGM, da Cruz Neto JX, Oliveira PR (2010) A proximal point algorithm with $\phi-$divergence to quasiconvex programming, 59: 777-792

Chen JS, Pan S (2008) A proximal-like algorithm for a class of nonconvex programming. Pacific Journal of Optimization, 4: 319-333

De la Cruz,L, Papa E. (2013) Métodos del punto proximal y su aplicación a modelos económicos, PESQUIMAT, Revista de la F.C.M. de la Universidad Nacional Mayor de San Marcos Vol. XV N°l, pp. 68-78, Lima - Perú, Marzo 2013

De Lamo, A. R. y Dolado, J. J. (1993); “Un Modelo del Mercado de Trabajo y la Restricción de Oferta en la Economía Peruana”; Investigaciones Económicas, Vol. XVII (1), pp. 87 – 118.

Dennis J. and R. B. Schnabel. Numerical methods for unconstrained optimization and nonlinear equations. SIAM, 1996.

Douglas, Paul (1976): The Cobb – Douglas Production Function Once Again: Its History, Its Testing, and Some Empirical Values. Journal of Political Economy, vol. 84, pp. 903 – 915.

Eckstein J (1998) Approximate iterations in Bregman-function-based proximal algorithms. Mathematical Programming 83: 113-123

" Encyclopedia of optimization ", Springer Science and BUSINESS Media LLC,2009

Fletcher R., Luenberger, David G.ISBN: 978-0-471-49463-8 Julio de 2000 Métodos prácticos de optimización, segunda edición Pp.(299,268)

Guler, O (1991) On the convergence of the proximal point algorithm for convex minimization. SIAM J. Control Optim., 29(2):403–419, 1991.

Hamilton, J. D. (1994). Time Series Analysis. Princeton University Press.

Hestenes y Powell (1969) Revista de teoría y aplicaciones de la optimización volumen 12 , paginas555–562 (1973)

Hestenes, Mr , métodos de multiplicador y gradiente , Journal of Optimization Theory and Applications, vol. 4, págs. 303-320, 1969.

Humes Jr C, Silva PJS (2005) Inexact proximal algorithms and descent methods in optimization. Optimization and Engineering, 6: 257-271.

Kaplan A, Tichatschke R (1998) Proximal point methods and nonconvex optimization. Journal of Global Optimization, 13: 389-406.

Kaplan A, Tichatschke R (2004) On inexact generalized proximal methods with a weakened error tolerance criterion, Optimization 53: 3-17.

Kiwiel KC (1997) Proximal minimization methods with generalized Bregman functions. SIAM Journal of Control and Optimization, 35: 1142-1168.

Langenberg N, Tichatschke R (2012) Interior proximal methods for quasiconvex optimization. J Glob Optim 52: 641-661.

Luenberger D.. Adisson Wesley Iberoamericana 1989.Introducción a la Programación Lineal y No Lineal,

Luenberger and Ye (2008), Linear and Nonlinear Programming :111-143,

Martinet B. (1970). Régularization, d'inéquations variationelles par approximations successives, Revue Francaise q'Informatique et de Recherche Operationelle: 154-159.

Minoux M.. John Wiley & Sons (1986), Theory and Algorithms, Inc., New York, Mathematical Programming

Moré J.and S.Wright (1993),Optimization Software Guide, SIAM Frontiers in Applied Mathematics 14, SIAM Publications, Philadelphia.

Miller L Shirley. (2002).Métodos alternativos para la estimación del PBI potencial: Una aplicación para el caso de Perú. Banco Central de Reserva del Perú [2] Greene, W. (2007). Econometric Analysis. New York: McMillan, Sixth Edition

Pan S, Chen JS (2007) Entropy-like proximal algorithms based on a second-order homogeneous distances function for quasiconvex programming. Journal of Global Optimization, 39: 555-575.

Papa Quiroz EA, Oliveira PR (2012i) An extension of proximal methods for quasiconvex minimization on the nonnegative orthant. European Journal of Operational Research 216: 26-32.

Papa Quiroz, E.A. and Oliveira, P.R (2012ii) Full convergence of the proximal point method for quasiconvex function on Hadamard manifolds, ESAIM: Control, Optimisation and Calculus of Variations, 18, 483-500, 2012.

Papa Quiroz, E.A. and Oliveira, P.R (2012iii), Proximal point methods for minimizing quasiconvex locally lipschitz functions on Hadamard manifolds, Nonlinnear Analysis. 75, 5924-5932, 2012.

Perron, P. (1994). «Trend, Unit Root and Structural Change in Macroeconomic Time Series», in Cointegration for the Applied Economist. B. B. Rao (Editor), Macmillan Press, 113-146.

Rockafellar, R.T. (1976 i) Monotone operations and the proximal point method. SIAM J. Control Optimization, 14: 877-898.

Rockafellar, R. T. (1976 ii), Augmented Lagrangians and Application of Proximal Point Algorithm in Convex Programming, Mathematics of Operation Research, Vol.1, pp. 97-116, 1976.

Solodov MV, Svaiter BF (2000) An inexact hybrid generalized proximal point algorithm and some new results on the theory of Bregman functions. Mathematics of Operations Research 25: 214-230

Souza SS, Oliveira PR, da Cruz Neto JX, Soubeyran A (2010) A proximal method with separable Bregman distance for quasiconvex minimization on the nonnegative orthant. European Journal of Operational Research, 201: 365-376

Xu Y, Bingsheng H, Xiaoming Y (2006) A hybrid inexact logarithmic-quadratic proximal method for nonlinear complementarity problems. J. Math. Anal. 322: 276-287.

Zivot, E. and D. W. Andrews (1992). «Furhter Evidence on the Great Crash, the Oil Price Shock and the Unit Root Hypothesis». Journal of Business and Economic Statistics 10, 251-270.

ANEXOS

Anexo 1

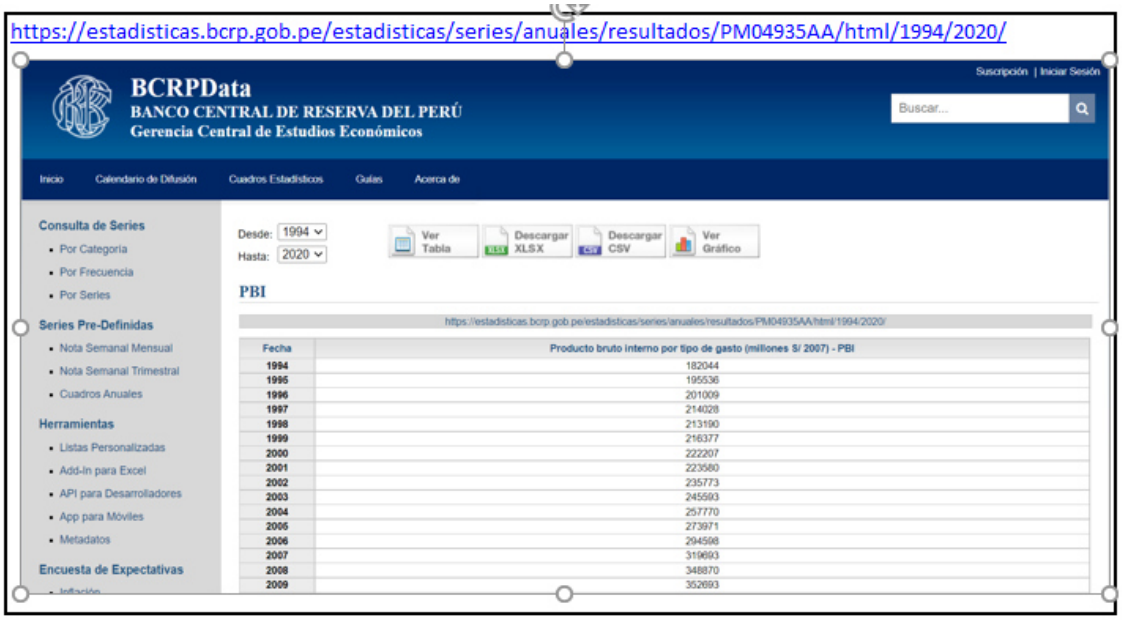
Demanda interna -inversión bruta interna



Fecha	Producto bruto interno por tipo de gasto (millones S/ 2007) - Demanda Interna - Inversión Bruta Interna
1994	32818
1995	39205
1996	37427
1997	43561
1998	43278
1999	36797
2000	36035
2001	33580
2002	35514
2003	38435
2004	38026
2005	38878
2006	54437
2007	71188
2008	91521
2009	69388

Anexo 2

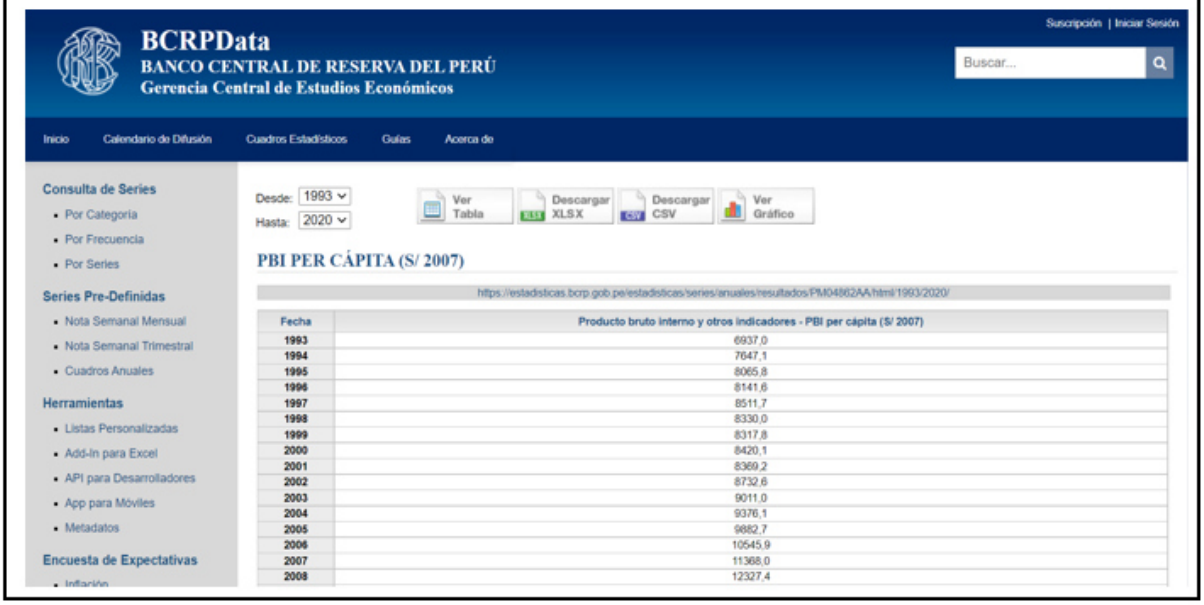
PBI (S/. 2007)



Anexo 3

PBI per cápita (S/. 2007)

https://estadisticas.bcrp.gob.pe/estadisticas/series/anuales/resultados/PM04862AA/html/1993/2020/



BCRPData
BANCO CENTRAL DE RESERVA DEL PERÚ
Gerencia Central de Estudios Económicos

Suscripción | Iniciar Sesión

Buscar...

Inicio | Calendario de Difusión | Cuadros Estadísticos | Guías | Acerca de

Consulta de Series
- Por Categoría
- Por Frecuencia
- Por Series

Series Pre-Definidas
- Nota Semanal Mensual
- Nota Semanal Trimestral
- Cuadros Anuales

Herramientas
- Listas Personalizadas
- Add-In para Excel
- API para Desarrolladores
- App para Móviles
- Metadatos

Encuesta de Expectativas
- Inflación

Desde: 1993 Hasta: 2020

Ver Tabla | Descargar XLSX | Descargar CSV | Ver Gráfico

PBI PER CÁPITA (S/ 2007)

https://estadisticas.bcrp.gob.pe/estadisticas/series/anuales/resultados/PM04862AA/html/1993/2020/

Fecha	Producto bruto interno y otros indicadores - PBI per cápita (S/ 2007)
1993	6937,0
1994	7647,1
1995	8065,8
1996	8141,6
1997	8511,7
1998	8330,0
1999	8317,8
2000	8420,1
2001	8369,2
2002	8732,6
2003	9011,0
2004	9376,1
2005	9882,7
2006	10545,9
2007	11368,0
2008	12327,4

Anexo 4

PERÚ: Producto bruto interno total y por habitante, 1994-2019

https://www.inei.gob.pe/estadisticas/indice-tematico/economia/

PERÚ: PRODUCTO BRUTO INTERNO TOTAL Y POR HABITANTE, 1994-2019

Año	Producto Bruto Interno		Población [1/]	Producto Bruto Interno por habitante	
	Millones de Soles	Tasas Anuales de		Soles	Tasas Anuales de
1994	182,044	12.3	23,501,974	7,746	10.3
1995	195,536	7.4	24,242,600	8,066	4.1
1996	201,009	2.8	24,689,213	8,142	0.9
1997	214,028	6.5	25,145,317	8,512	4.5
1998	213,190	-0.4	25,592,876	8,330	-2.1
1999	216,377	1.5	26,013,829	8,318	-0.1
2000	222,207	2.7	26,390,142	8,420	1.2
2001	223,580	0.6	26,714,547	8,369	-0.6
2002	235,773	5.5	26,999,085	8,733	4.3
2003	245,593	4.2	27,254,632	9,011	3.2
2004	257,770	5.0	27,492,091	9,376	4.1
2005	273,971	6.3	27,722,342	9,883	5.4
2006	294,598	7.5	27,934,784	10,546	6.7
2007	319,693	8.5	28,122,158	11,368	7.8
2008	348,870	9.1	28,300,372	12,327	8.4
2009	352,693	1.1	28,485,319	12,382	0.4
2010	382,081	8.3	28,692,915	13,316	7.5
2011	406,256	6.3	28,905,725	14,055	5.5
2012	431,199	6.1	29,113,162	14,811	5.4
2013	456,435	5.9	29,341,346	15,556	5.0
2014	467,308	2.4	29,616,414	15,779	1.4
2015	482,506	3.3	29,964,499	16,103	2.1
2016	501,581	4.0	30,422,831	16,487	2.4
2017	514,215	2.5	30,973,992	16,602	0.7
2018	534,625	4.0	31,562,130	16,939	2.0
2019	546,408	2.2	32,131,400	17,005	0.4

[1/] Hasta el año 1994 se tomo del documento Perú: Estimaciones y Proyecciones

Fuente: Instituto Nacional de Estadística e Informática.

Anexo 5

PERÚ: Formación bruta de capital fijo, pública y privada, 1994-2019

PERÚ: FORMACIÓN BRUTA DE CAPITAL FIJO PÚBLICA Y PRIVADA, 1994

Valores a precios constantes de 2007
(Millones de soles)

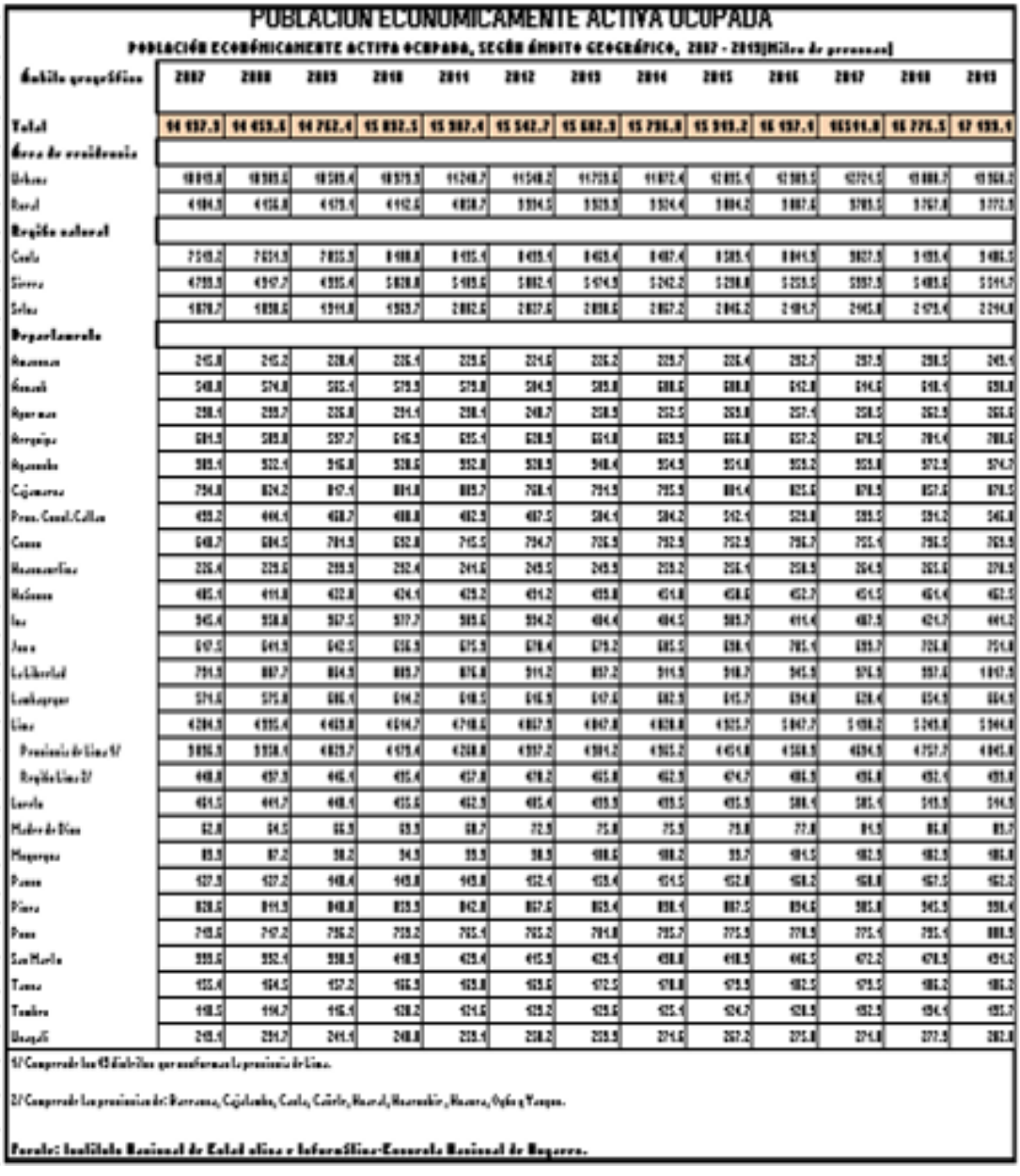
Año	Total	Pública	Privada
1994	33,234	9,046	24,188
1995	40,335	9,408	30,927
1996	39,189	8,853	30,336
1997	45,167	9,880	35,287
1998	44,635	10,196	34,439
1999	39,700	10,859	28,841
2000	37,654	9,176	28,478
2001	34,602	7,038	27,564
2002	34,772	6,756	28,016
2003	36,725	7,134	29,591
2004	39,430	7,334	32,096
2005	43,816	8,236	35,580
2006	52,084	9,496	42,588
2007	63,892	11,433	52,459
2008	81,666	15,289	66,377
2009	78,886	19,826	59,060
2010	96,141	23,067	73,074
2011	105,687	20,465	85,222
2012	121,423	24,526	96,897
2013	127,954	27,540	100,414
2014	125,380	27,841	97,539
2015	116,404	26,261	90,143
2016	111,207	26,277	84,930
2017	112,639	25,261	87,378
2018	117,898	26,922	90,976
2019	121,232	26,769	94,463

Fuente: Instituto Nacional de Estadística e Informática.

Anexo 6

Población económicamente activa ocupada, según ámbito geográfico, 2007 - 2019(miles de personas)

https://www.inei.gob.pe/estadisticas/indice-tematico/employed-economically-active-population/



POBLACION ECONOMICAMENTE ACTIVA OCUPADA

POBLACIÓN ECONÓMICAMENTE ACTIVA OCUPADA, SEGÚN ÁMBITO GEOGRÁFICO, 2007 - 2019(Miles de personas)

Ámbito geográfico	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Total	14 937.3	14 459.6	14 762.4	15 032.5	15 387.4	15 542.7	15 682.3	15 796.8	15 949.2	16 137.1	16591.8	16 776.5	17 199.1
Área de residencia													
Urbana	10 015.8	10 303.6	10 589.4	10 879.3	11 240.7	11 540.2	11 719.6	11 872.4	12 035.1	12 303.5	12724.5	13 001.7	13 364.2
Rural	4 934.5	4 156.0	4 173.1	4 152.6	4 036.7	3 974.5	3 962.8	3 924.4	3 914.2	3 887.6	3 789.3	3 767.8	3 772.9
Región natural													
Costa	7 519.2	7 634.5	7 815.5	8 108.8	8 435.1	8 435.1	8 453.4	8 487.4	8 583.1	8 844.5	9027.5	9 153.4	9 486.5
Sierra	4 788.5	4 547.7	4 555.4	4 828.8	5 015.6	5 082.1	5 174.5	5 242.2	5 298.8	5 253.5	5357.5	5 483.6	5 541.7
Selva	1 878.7	1 858.5	1 911.8	1 953.7	2 082.6	2 027.6	2 038.6	2 067.2	2 046.2	2 084.7	2 065.8	2 073.4	2 214.8
Departamento													
Amazonas	245.8	245.2	238.4	235.1	233.6	234.6	235.2	233.7	235.4	232.7	237.5	238.2	243.1
Áncash	548.8	574.8	582.1	579.5	579.8	584.5	585.8	588.6	588.8	592.8	614.6	618.1	638.8
Apurímac	258.1	259.7	256.8	254.1	258.1	248.7	258.5	252.5	253.8	257.1	258.2	262.5	266.6
Arequipa	684.5	688.8	697.7	696.5	695.1	698.5	694.8	693.5	666.8	697.2	678.5	704.4	708.6
Ayacucho	588.1	582.1	596.8	598.6	592.8	598.5	568.4	564.5	564.8	553.2	553.8	572.5	574.7
Cajamarca	794.8	824.2	847.1	864.8	885.7	768.1	794.5	795.5	804.4	825.6	878.5	857.6	878.5
Prov. Const. Callao	455.2	466.1	458.7	488.8	482.5	487.5	584.1	584.2	542.1	529.8	588.5	594.2	545.8
Cusco	648.7	684.5	704.5	652.8	745.5	754.7	735.5	752.5	752.5	756.7	722.1	756.5	763.5
Huancavelica	226.4	225.6	255.5	252.4	244.6	245.5	245.5	255.2	256.1	258.5	264.5	263.6	278.5
Huánuco	485.1	498.8	482.8	474.1	475.2	474.2	455.8	454.8	458.6	452.7	454.5	454.4	462.5
Ica	545.4	558.8	567.5	577.7	585.6	594.2	584.4	584.5	585.7	615.4	587.5	624.7	644.2
Junín	647.5	645.5	642.5	656.5	675.5	678.4	675.2	685.5	658.1	705.1	695.7	726.8	754.8
La Libertad	794.5	807.7	814.5	805.7	876.8	911.2	857.2	915.5	918.7	945.5	974.5	957.4	1 047.5
Lambayeque	574.6	575.8	586.1	614.2	648.5	646.5	647.6	682.5	645.7	654.8	628.4	654.5	654.5
Lima	4 284.5	4 255.4	4 425.8	4 514.7	4 748.6	4 857.5	4 847.8	4 828.8	4 825.7	5 047.7	5 458.2	5 245.8	5 344.8
Provincia de Lima 1/	3 836.5	3 858.1	4 025.7	4 075.4	4 268.8	4 357.2	4 384.2	4 365.2	4 454.8	4 568.5	4834.5	4 757.7	4 845.8
Región Lima 2/	448.8	457.5	445.1	435.4	457.8	478.2	455.8	452.5	474.7	486.5	456.8	452.1	498.8
Loreto	454.5	444.7	448.1	435.6	452.5	435.4	455.5	455.5	455.5	388.1	385.1	345.5	344.5
Madre de Dios	62.8	64.5	66.5	69.5	68.7	72.5	75.8	75.5	75.8	77.8	84.5	86.8	85.7
Moquegua	85.5	87.2	88.2	84.5	85.5	88.5	488.6	488.2	85.7	484.5	482.5	482.5	486.8
Pasco	427.5	427.2	448.4	445.8	445.8	452.1	455.4	454.5	452.8	458.2	458.8	457.5	452.2
Piura	828.5	844.5	848.8	855.5	842.8	857.5	865.4	858.1	887.5	854.5	585.8	845.5	858.4
Puno	748.5	747.2	756.2	758.2	765.1	765.2	784.8	785.7	775.5	778.5	775.1	785.1	888.5
San Martín	555.5	552.1	558.5	448.5	425.4	445.5	425.1	458.8	448.5	445.5	472.2	478.5	454.2
Tacna	455.4	454.5	457.2	455.5	455.8	455.5	472.5	478.8	475.5	482.5	475.5	485.2	485.2
Tumbes	448.5	448.7	445.1	438.2	424.5	425.2	425.5	425.1	424.7	428.5	452.5	454.1	455.7
Ucayali	245.1	254.7	244.1	248.8	255.1	258.2	255.5	274.5	267.2	275.8	274.8	277.5	282.8

1/ Comprende los 43 distritos que conforman la provincia de Lima.

2/ Comprende las provincias de: Barranca, Cajatambo, Canta, Cañete, Huaral, Huarochirí, Huaura, Oyón y Yauyos.

Fuente: Instituto Nacional de Estadística e Informática-Encuesta Nacional de Hogares.

Anexo 7

Data de Cobb Douglas

AÑO	Y	K	L	A
	PBI - Millones Soles - 2007	Formación Bruta de Capital - Millones Soles - 2007	PEA Ocupada - Miles de personas	A(t)=Y(t)/[L(t)^alfa][K(t)^beta] alfa= 0,31 a 0,65 [0,49: 1950-2000] beta= 1- alfa
2001	223,580	34,602	11,862	10.92
2002	235,773	34,772	12,034	11.40
2003	245,593	36,725	12,837	11.19
2004	257,770	39,430	13,060	11.23
2005	273,971	43,816	13,120	11.29
2006	294,598	52,084	13,683	10.89
2007	319,693	63,892	14,198	10.46
2008	348,870	81,666	14,460	9.98
2009	352,693	78,886	14,762	10.16
2010	382,081	96,141	15,093	9.85
2011	406,256	105,687	15,307	9.91
2012	431,199	121,423	15,543	9.72
2013	456,435	127,954	15,683	9.98
2014	467,308	125,380	15,797	10.29
2015	482,506	116,404	15,919	10.99
2016	501,581	111,207	16,197	11.59
2017	514,215	112,639	16,511	11.70
2018	534,625	117,898	16,776	11.79
2019	546,408	121,232	17,133	11.76

INEI | DATOS INEI | Data INEI | BCRP | Data BCRP | Hoja1