

# Apuntes de Optimización

Prof. Juan Carlos Ferrer, Departamento de Ingeniería Industrial y de Sistemas, P.U.C.

Prof. Juan Carlos Muñoz, Departamento de Ingeniería de Transporte, P.U.C.

Marzo 2006



# Contents

<b>1</b>	<b>Introducción</b>	<b>1</b>
1.1	Introducción al Modelamiento . . . . .	2
1.2	¿Qué es un modelo? . . . . .	2
1.3	¿Por qué se construyen modelos? . . . . .	3
1.4	Modelos de Programación Matemática . . . . .	4
1.5	Caracterización de puntos extremos de una función . . . . .	5
1.6	Existencia de solución óptima . . . . .	6
1.7	Problemas Equivalentes . . . . .	9
1.7.1	Equivalencia I . . . . .	9
1.7.2	Equivalencia II . . . . .	10
1.7.3	Equivalencia III . . . . .	12
1.7.4	Equivalencia IV . . . . .	19
1.7.5	Equivalencia V . . . . .	21
1.7.6	Equivalencia VI . . . . .	22
1.7.7	Equivalencia VII . . . . .	22
1.8	Nociones Básicas de Convexidad . . . . .	23
1.8.1	Conjuntos Convexos . . . . .	23
1.8.2	Funciones Convexas . . . . .	24
1.8.3	Criterios prácticos de convexidad de funciones . . . . .	31
1.9	Problemas resueltos . . . . .	34
<b>2</b>	<b>Programación No lineal</b>	<b>45</b>
2.1	Optimización de una función sin restricciones . . . . .	45
2.1.1	Condiciones Necesarias y suficientes para extremos . . . . .	45
2.1.2	Métodos de búsqueda de soluciones . . . . .	49
2.2	Optimización de una función con restricciones . . . . .	54
2.2.1	Caso 1: Problema Unidimensional . . . . .	54
2.2.2	Caso 2: Restricciones de igualdad . . . . .	58
2.2.3	Caso 3: Restricciones de desigualdad . . . . .	70
2.3	Problemas resueltos . . . . .	79
2.3.1	Optimización de una función sin restricciones . . . . .	79

2.3.2	Optimización de una función con restricciones . . . . .	86
<b>3</b>	<b>Programación Lineal</b>	<b>105</b>
3.1	Introducción . . . . .	105
3.2	El Método Simplex . . . . .	112
3.2.1	Pasos en el método Simplex . . . . .	115
3.2.2	Solución inicial factible básica . . . . .	121
3.2.3	Análisis Matricial del Método Simplex . . . . .	126
3.2.4	Casos especiales en el desarrollo de Simplex . . . . .	131
3.3	Análisis de Sensibilidad de los Resultados . . . . .	134
3.3.1	Rango de variación de los costos . . . . .	134
3.3.2	Rango de variación del nivel de recursos . . . . .	136
3.3.3	Ejemplo . . . . .	137
3.4	Dualidad . . . . .	138
3.4.1	Relaciones Primal-Dual . . . . .	141
3.4.2	Análisis Matricial del problema dual . . . . .	142
3.5	Problemas resueltos . . . . .	144
3.5.1	Geometría de PL . . . . .	144
3.5.2	Método Simplex . . . . .	150
3.5.3	Dualidad . . . . .	166

# Chapter 1

## Introducción

El objetivo principal de estos apuntes es que permitan aprender y entender la esencia de los métodos de optimización en forma clara y didáctica. El objetivo no es sólo que los alumnos sepan modelar y resolver problemas, sino que también puedan comprender su resolución.

Los métodos de optimización son un conjunto de herramientas cuantitativas muy poderosas para enfrentar, modelar y resolver importantes problemas reales en los más diversos ámbitos (no se limita sólo a ingeniería). El objetivo de este texto es apoyar al curso de Optimización de la Escuela de Ingeniería de la Pontificia Universidad Católica de Chile para que los alumnos puedan llegar un poco más lejos que las clases, o para que lo puedan utilizar como apoyo en el futuro para enfrentar problemas de optimización en otros cursos de la carrera o en sus trabajos.

Creemos muy necesario que la recopilación y creación de estos apuntes de métodos de optimización debe ser muy clara y didáctica para explicar los conceptos. El lector podrá entender el mecanismo detrás de los diferentes métodos y así estar preparado para poder discernir entre soluciones correctas e incorrectas.

Durante nuestros años de doctorado en M.I.T. y Berkeley (1997-2002) tuvimos la oportunidad de tomar varios cursos de investigación de operaciones donde juntamos bastante material de primer nivel, que se ha incorporado en estos apuntes.

El contenido de estos apuntes a nivel general es:

- Introducción: Se dan algunas nociones de qué es un modelo y de cómo se construyen. Se presentan algunos modelos equivalentes, y se revisan algunas nociones de convexidad.
- Programación No Lineal: Condiciones necesarias y suficientes para un mínimo local o global; Métodos de búsqueda de soluciones óptimas (Gradiente, Newton, etc.)
- Programación Lineal: Formulación y forma estándar de problemas lineales; Geometría de problemas lineales; Método Simplex; Análisis de sensibilidad; Teoría de dualidad.
- Programación Entera: Formulación de modelos de programación entera; Algoritmo Branch & Bound; Algoritmo de planos cortantes.
- Flujo en Redes: Modelos clásicos como ruta más corta y máximo flujo.

- Programación Dinámica: Método de resolución cuando hay relaciones intertemporales entre las variables del problema.

A pesar de que estos apuntes no pretenden ser una guía de ejercicios, incorporan ejercicios resueltos en su gran mayoría a modo de ilustrar mejor cada uno de los conceptos tratados. Además se incorporará breves instrucciones de cómo operar el módulo de optimización de Microsoft Excel, y de lenguajes de modelación algebraica como AMPL y SML.

Este texto se encuentra en formato PDF en la página web del curso de Optimización, de manera que los alumnos (y ex-alumnos) puedan tener acceso a la más reciente versión.

## 1.1 Introducción al Modelamiento

Cotidianamente nos enfrentamos a la necesidad de **decidir**, evaluando opciones o cursos de acción a seguir entre los cuales escoger la mejor alternativa. Si la opción es única, la decisión ya está tomada, pero si son muchas o infinitas, entonces identificarlas y escoger la preferida puede ser muy complejo. Incluso, cuando la opción debe ser tomada por un grupo de personas, unos pueden tener diferentes visiones que otros. Para poder comparar las distintas alternativas hay que establecer explícitamente los objetivos que se persiguen con la decisión, y así lograr ordenar las opciones disponibles de la mejor forma. Esta tarea (identificar opciones y clasificar objetivos) suele ser no trivial.

Cuando el problema es complejo y resulta difícil identificar las opciones disponibles y/o escoger la mejor, muchas veces se recurre a modelos como un recurso de apoyo. Los modelos permiten tomar una decisión más eficientemente; esto es más rápidamente, económicamente, informadamente, etc.

## 1.2 ¿Qué es un modelo?

El término **modelo** se usa para hablar de una estructura que se ha construido y que exhibe características de un sistema. Philippi (1981) define un modelo como “una simplificación de alguna realidad concreta orientada a nuestros propósitos”. En otras palabras, es una caricatura de la realidad, que captura los factores dominantes que determinan el comportamiento del sistema en estudio.

Un modelo debe seleccionar las características más relevantes de la realidad, pues normalmente no pueden considerarse todos sus aspectos. Las características a considerar dependerán del uso que se le dará al modelo. Esto produce un **trade-off** entre la complejidad del modelo y las características consideradas. Así, la simplicidad debe estar en la mente de todo modelador.

Algunos modelos son tangibles como por ejemplo los aeromodelos, mapas y maquetas arquitectónicas. Por ejemplo los aeromodelos reúnen las características de vuelo de los aviones reales, ya que ese es el uso que se les dará, es decir, la atención se centra en su diseño, resistencia, aerodinámica, por lo que, en este caso, los asientos no se ponen ya que son irrelevantes para el objetivo.

Normalmente en Investigación de Operaciones (IO<sup>1</sup>) se usan modelos abstractos. Estos por lo general, serán matemáticos, donde los símbolos algebraicos reflejarán las relaciones internas del sistema modelado. Así, una característica esencial de un modelo matemático en IO, es que contiene las relaciones matemáticas relevantes (ecuaciones, inecuaciones, dependencias lógicas, etc) que se extraen al observar el mundo real.

### 1.3 ¿Por qué se construyen modelos?

i) **Clarifica** las relaciones existentes, no siempre claras para el observador. Se comprende mejor el sistema. Ayuda a identificar las alternativas. Es como una metodología de aprendizaje. Por ejemplo en un mapa se clarifican las rutas posibles para llegar de un lugar a otro. En muchas disciplinas (la física por ejemplo) los modelos no se usan para decidir sino que para comprender mejor.

ii) Permite un **análisis metódico** cuyo fin sea sugerir líneas de acción, no evidentes de otro modo. Por ejemplo una maqueta de una casa ayuda a decidir mejor como orientar los espacios, la iluminación, etc.

iii) Un modelo permite **experimentar en él**, cosa que no siempre es posible en el sistema real (Ej: avión, planta manufacturera, economía de un país, etc).

Estas tres razones son también válidas para construir modelos matemáticos que permiten experimentar con sistemas complejos de gran tamaño, considerar muchas alternativas simultáneamente (sin necesidad de enumerarlas a priori) e identificar un mejor (u óptimo) curso de acción.

Es importante distinguir entre modelo y datos. El modelo queda definido por las relaciones que tiene incorporado. La idea es que los modelos tengan validez independiente de los datos que se incorporarán. Sin embargo, hay veces que los modelos tienen validez sólo en un rango de datos permitido. (Ej: Fluidez de un sólido sujeto a tensiones (Resistencia de materiales)).

Los modelos de Programación Matemática son los más comúnmente usados. Otros ejemplos de modelos de uso frecuente son: simulación (modelos muy complejos de resolver), planificación en red, econométricos (predecir una variable en función de otra), series de tiempo (modelos que indican qué hacer en base a datos pasados).

Los problemas pueden ser modelados en más de una forma. Es interesante observar qué tipo de modelación es más eficiente (tiempo, resultado, memoria).

Cuando se plantea modelar un sistema existen muchos conceptos errados al respecto. Hay quienes se niegan a su utilidad argumentando que hay muchas hipótesis cuestionables como que hay factores difíciles de cuantificar, o bien, que siempre los datos carecen de precisión. Es clave ver si el resultado es sensible o no respecto a dicha hipótesis.

En otro extremo hay quienes profieren una fe metafísica a los modelos matemáticos para la toma de decisiones (y aún más si provienen de un computador). Obviamente la calidad de la solución

---

<sup>1</sup>IO: Ciencia de la toma de decisiones, o, la aplicación de métodos científicos a la administración y gestión de organizaciones.

dependerá de la precisión de la estructura y de los datos del modelo.

Hay que considerar que la fe sin crítica a un modelo es obviamente peligrosa. No se recomienda en absoluto aceptar la primera respuesta que un modelo matemático produce sin un análisis y cuestionamiento posterior profundo. Normalmente un modelo será una herramienta dentro de un conjunto, a la hora de tomar una decisión.

La respuesta de un modelo deberá enfrentarse a un cuidadoso examen, y si ésta es inaceptable, deberá comprenderse el por qué y probablemente incorporarse a un modelo modificado. Si la respuesta es aceptable, puede ser prudente mantenerlo como una opción. Por medio de un sucesivo cuestionamiento de las respuestas y alternando el modelo, es posible clarificar las opciones factibles y comprender mejor lo que es posible.

## 1.4 Modelos de Programación Matemática

En muchas ocasiones uno enfrenta problemas en que es necesario determinar la mejor opción dentro de un conjunto de alternativas disponibles. A esta acción se le denomina **Optimizar**. Dentro del proceso de optimizar existen algunos métodos que son de mucha utilidad, los métodos de **Programación Matemática**, que se basan en la creación de un modelo que represente el sistema de interés para luego trabajar en torno a esta reducción o caricatura de la realidad, y así facilitar la toma de decisión.

A los estudiantes de ingeniería la palabra “programar” seguramente les evoca largas horas frente al computador diseñando un software. Sin embargo en Programación Matemática (PM) la palabra se usa con el significado de planificar. De hecho, PM en su esencia no tiene nada que ver con computadores. Lo que sucede posteriormente, es que los problemas de PM son tan grandes que requieren apoyo computacional para su resolución.

La característica común de todos los problemas de PM es que pretenden optimizar algo (minimizar o maximizar). Lo que se pretende optimizar se le denomina **función objetivo**. Cuando uno observa que se requiere optimizar considerando más de un objetivo, tenemos una **función multi-objetivo**, o con objetivos difusos. En ese caso existen al menos dos alternativas para abordar el problema: priorizar o ponderar. Priorizar se refiere a resolver múltiples problemas en forma secuencial de acuerdo a un orden (priorización) de los objetivos. Ponderar apunta a la resolución conjunta de los objetivos, llevándolos a una unidad común.

La Figura 1.1 muestra una clasificación general de los modelos matemáticos. En este texto trataremos principalmente modelos determinísticos de programación no-lineal y de programación lineal, tal como se destacan en dicha figura.

El problema general de Programación Matemática consiste en determinar un vector  $x^T = (x_1, \dots, x_n)$  que perteneciendo a un conjunto  $D$ , subconjunto de un espacio vectorial de orden  $n$ ,



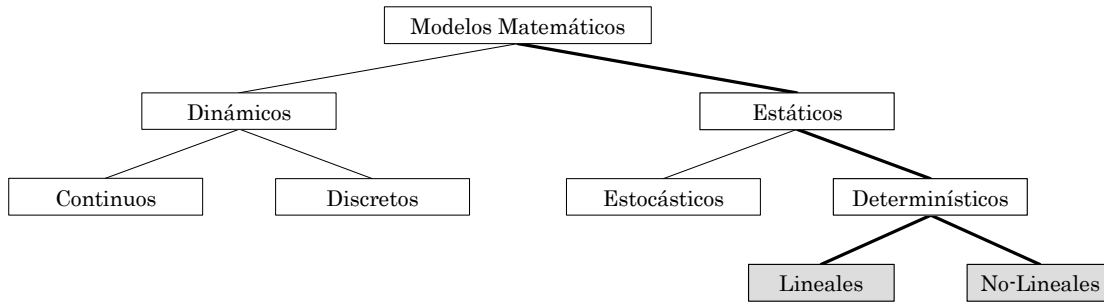


Figure 1.1: Clasificación de Modelos Matemáticos

maximiza (o minimiza) una función objetivo  $f(x_1, \dots, x_n)$ . Esto es,

$$\begin{aligned} P) \quad & \max f(x^T) \\ \text{s.a.} \quad & x^T \in D, \end{aligned} \tag{1.1}$$

donde el **dominio** o **set de oportunidades**  $D$  estará formado por los siguientes tipos de restricciones: (i) restricciones funcionales

$$\begin{aligned} h_i(x^T) &= 0 \quad \text{con } i = 1, \dots, m \\ g_j(x^T) &\geq 0 \quad \text{con } j = 1, \dots, s \end{aligned}$$

y (ii) restricciones de conjuntos

$$x \in \Omega \quad : \quad \Omega \subset E^n.$$

A la función objetivo  $f(x^T)$  también se le denomina función económica o de utilidad (costos). Cabe mencionar que maximizar una función equivale a minimizar dicha función, pero con signo contrario. Por otra parte, en Programación Lineal,  $f(x)$  y las restricciones  $h(x)$  y  $g(x)$ , son lineales.

En esta sección se entregan conceptos matemáticos básicos para la solución de problemas de programación matemática. En particular, se caracterizan las soluciones óptimas de un problema y se discute en qué casos es posible garantizar la existencia de una solución óptima. Más adelante se introduce el concepto de modelos equivalentes, esto es, modelos cuya solución óptima es necesariamente equivalente. Por último, se introduce la convexidad en conjuntos y funciones, pues ésta será de gran ayuda para decidir si un óptimo local es solución al problema.

## 1.5 Caracterización de puntos extremos de una función

**Definición 1** *Un punto extremo de una función definida sobre un dominio  $D$  puede ser local o global, estricto o no estricto, según se define a continuación (la Figura 1.2 muestra gráficamente cada una de estos puntos):*

$x^*$  es máximo global si  $x^*$  es punto factible y si  $f(x^*) \geq f(x) \forall x \in D$ .

$x^*$  es *máximo global estricto* si  $x^*$  es punto factible y si  $f(x^*) > f(x) \forall x \in D \setminus \{x^*\}$ .

$x^*$  es *máximo local* si  $x^*$  es punto factible y si  $f(x^*) \geq f(x) \forall x \in \text{Vecindad de } x^*$  y  $x \in D$ .

$x^*$  es *máximo local estricto* si  $x^*$  es punto factible y si  $f(x^*) > f(x) \forall x \in \text{Vecindad de } x^*$  y  $x \in D \setminus \{x^*\}$ .

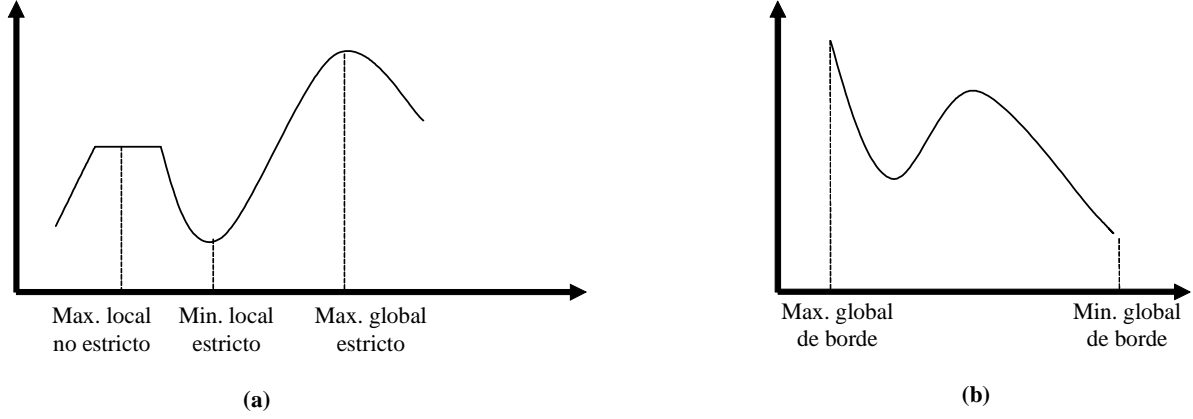


Figure 1.2: Soluciones locales vs. globales

## 1.6 Existencia de solución óptima

**Definición 2** Todo punto  $x \in D$  define una solución factible o realizable de  $P$ .

**Definición 3** Un punto factible  $\hat{x}$  define una solución óptima del problema de optimización  $\min_{x \in D} f(x)$  si  $f(\hat{x}) \leq f(x) \forall x \in D$  y  $\hat{x} \in D$ .

Nota: En el caso particular en que  $D = \mathbb{R}^n$ , se dice que el problema es no-restringido o irrestricto.

**Definición 4** En un problema de maximización,  $\hat{v}$  define el valor óptimo de  $P$  si:  $\hat{v} = \sup(f(x))$ ,  $x \in D$ . (*sup*: supremo, cota superior de la función sobre el dominio). En particular, si  $\hat{x}$  es solución óptima de  $P \implies \hat{v} = f(\hat{x})$ . Análogamente para un problema de minimización,  $\hat{v} = \inf(f(x))$ ,  $x \in D$ . (*inf*: ínfimo, cota inferior de la función sobre el dominio).

**Nota:** No todos los problemas tienen solución óptima. Podemos pensar en cuatro familias de problemas que podrían carecer (o carecen) de ella:

- Problemas cuyo conjunto de restricciones define un dominio de puntos factibles vacío.
- Problemas que contemplan alguna restricción de desigualdad estricta, es decir que define una región que no incluye a su borde. Si en este caso la función objetivo mejora en cuanto más se acerque al borde, se estará en una situación en que la solución óptima no se podrá identificar.
- Problemas que contemplan un dominio no acotado, es decir que incluye puntos cuyas variables toman valores tan grandes (en valor absoluto) como sea necesario. Si en este caso la función objetivo

mejora en cuanto más se acerque a estos puntos de valor absoluto ilimitado, entonces la solución óptima no se podrá identificar.

d) Problemas con una función objetivo discontinua sobre el dominio factible. Si en este caso la función objetivo evaluada en el punto de discontinuidad no es óptima, pero ésta mejora en cuanto más se acerque al punto de discontinuidad desde alguna dirección, entonces la solución óptima no se podrá identificar.

Evidentemente, no todos los problemas con dominio no acotado carecen de solución. Tampoco todos los problemas con restricciones de desigualdad estricta o todos los con función objetivo discontinua. Sin embargo, en presencia de estos eventos, la solución óptima no está garantizada.

Estas observaciones nos conducen naturalmente al siguiente teorema de existencia de solución óptima.

**Teorema 1** (*Existencia de soluciones óptimas*) Consideremos el siguiente problema de optimización

$$\begin{aligned} P) \quad & \text{Min } f(x) \\ & x \in D \end{aligned}$$

Si  $f(x)$  es continua sobre  $D$ , con  $D$  cerrado y no vacío sobre  $\mathbb{R}^n$ , entonces bajo la hipótesis:

$$H) f(x) \rightarrow +\infty \text{ si } \|x\| \rightarrow +\infty, x \in D \quad (1.2)$$

$P)$  admite al menos una solución óptima.

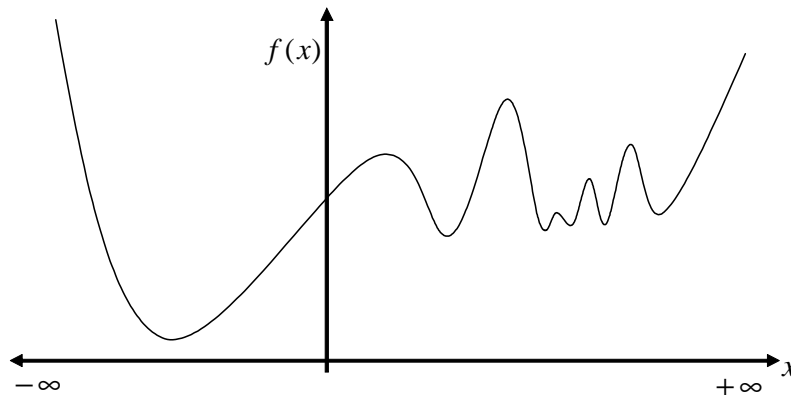


Figure 1.3: Ejemplo de Teorema de Existencia

**Nota:** Un problema  $P)$  puede tener valor óptimo finito y, sin embargo, no admitir solución óptima. Para ilustrar esto veamos el siguiente ejemplo:

$$P) \min_{x \geq 0} e^{-x}$$

donde  $\hat{v} = v(P) = 0$ . ¿ $\exists \hat{x} \geq 0 / e^{-x} = 0$ ? No! Por lo tanto  $P$  no tiene solución, aunque  $\hat{v} \in \mathbb{R}$  (ver

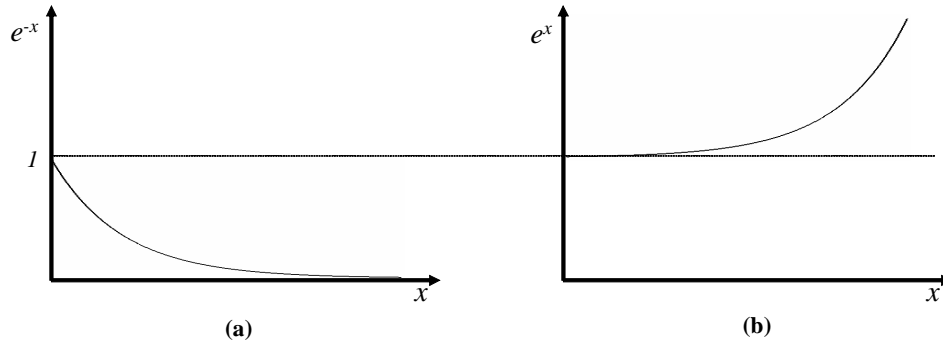


Figure 1.4: Ejemplos de valores y soluciones óptimas

Figura 1.4a).

Se podría pensar que no existe solución óptima porque  $D$  es no acotado. Sin embargo, muchos problemas no acotados admiten solución óptima. Basta considerar,

$$P) \min_{x \geq 0} e^x,$$

en que  $v(P) = 1$  y  $\hat{x} = 0$  es la solución óptima (ver Figura 1.4b).

**Nota:** Un conjunto es cerrado si contiene a todos los puntos frontera. El caso de  $D = \mathbb{R}$  es cerrado, ya que el conjunto de los puntos frontera es vacío, y todo conjunto contiene al conjunto vacío. Como ejemplo de un conjunto no cerrado considere  $x > 0$  que no contiene a su frontera ( $x = 0$ ).

**Corolario 2** (*Bolzano-Weierstrass*) Si  $f$  es continua, sobre un dominio no vacío, cerrado y acotado, entonces el problema necesariamente tendrá solución óptima, ya que la imagen de una función continua sobre un set compacto, es también compacta (cerrada y acotada).

**Demostración.**  $D$  es acotado  $\Rightarrow \nexists$  ninguna sucesión de puntos  $\{x^k \in D\}_{k \geq 0}$  tal que:

$$\begin{aligned} \|x^k\| &\rightarrow +\infty \\ x^k &\in D \\ k &\rightarrow \infty \end{aligned}$$

Luego, la hipótesis  $H$  en (1.2) se cumple por vacuidad en este caso. ■

**Nota:** Es importante notar que estos teoremas identifican condiciones suficientes para existencia de optimalidad, pero no condiciones necesarias. Así, muchos problemas que no satisfacen las condiciones identificadas en el Teorema 1 o en el Corolario 2 tienen solución óptima.

## 1.7 Problemas Equivalentes

Un problema se puede modelar de varias formas diferentes. En programación matemática se denominan problemas equivalentes a problemas que garantizan el mismo conjunto de soluciones óptimas. Evidentemente, para un mismo problema, se busca aquel modelo que facilita su resolución.

Considere el problema de localizar una estación de bomberos en una ciudad con el siguiente criterio: que el tiempo máximo de respuesta de la bomba a eventos en cinco edificios de asistencia masiva sea lo más breve posible. En este caso para cada punto posible de instalar la bomba se evalúa el tiempo de respuesta a cada uno de los cinco edificios; el mayor de estos cuatro valores representa la función objetivo evaluada en el punto. Es interesante notar que en este caso la función objetivo es minimizar el valor maximo entre otras cinco funciones, es decir si denominamos la posición de la bomba  $x$ , entonces la función objetivo es:

$$\text{Min } (\max \{d_1(x), d_2(x), d_3(x), d_4(x), d_5(x)\}).$$

Una función objetivo como esta presenta serios problemas pues es difícil de tratar algebraicamente y presenta discontinuidades en su derivada. Así, sería deseable poder modelar este problema mediante un modelo equivalente que no presente estas dificultades.

A continuación se presentan siete equivalencias que pueden facilitar la modelación de muchos problemas, entre ellos el problema de la localización de la estación de bomberos.

### 1.7.1 Equivalencia I

$$\begin{array}{ccc} P) & \text{Max } f(x) & \sim \quad \tilde{P}) \text{Min } (-f(x)) \\ & x \in D. & x \in D. \end{array}$$

$\hat{x}$  es solución óptima de  $P$ ), con valor óptimo  $\hat{v}$ , si y sólo si  $\hat{x}$  es solución óptima de  $\tilde{P}$ ), con  $v(\tilde{P}) = -\hat{v}$ .

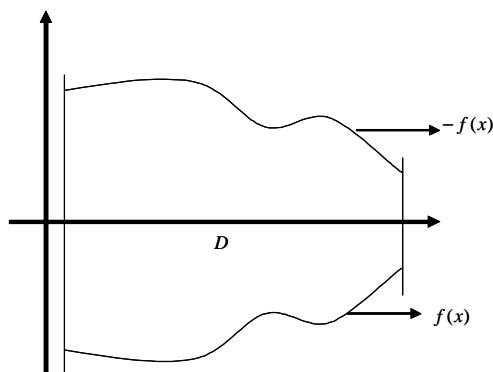


Figure 1.5: Equivalencia I

**Ejemplo 1** *Algunos ejemplos de esta equivalencia son:*

*Maximizar utilidad = Minimizar pérdida.*

*Maximizar probabilidad de sano = Minimizar probabilidad de enfermo.*

Por lo tanto, dado un problema de maximización, procedemos a cambiar el signo de la función de costo, para luego resolver el problema de minimización equivalente. Conocida la solución óptima  $x^*$  y el valor óptimo  $v(\tilde{P})$ , la solución óptima de  $P$  es  $x^*$  y el valor óptimo  $v(P) = -v(\tilde{P})$ .

### 1.7.2 Equivalencia II

$$\begin{aligned} P) \quad & \text{Min } f(x) \\ & x \in D \subset \mathbb{R}^n \end{aligned}$$

Equivalente a,

$$\begin{aligned} \tilde{P}) \quad & \text{Min } \mu \\ & f(x) \leq \mu \\ & x \in D \subset \mathbb{R}^n \\ & \mu \in \mathbb{R} \end{aligned}$$

$$\begin{aligned} \tilde{\tilde{P}}) \quad & \text{Min } \mu \\ & f(x) = \mu \\ & x \in D \subset \mathbb{R}^n \\ & \mu \in \mathbb{R} \end{aligned}$$

Considerando el problema  $\tilde{P}$  o  $\tilde{\tilde{P}}$ , el nuevo dominio será:  $\tilde{D} = D \times \mathbb{R} \subset \mathbb{R}^{n+1}$ , donde  $\hat{x}$  es solución óptima de  $P$  con valor óptimo  $\hat{v} = v(P)$ , si y sólo si  $(\hat{x}, \mu)$  es solución óptima de  $\tilde{P}$  con valor óptimo  $\hat{v}$ .

Comprobaremos esta equivalencia por contradicción. Supongamos que el problema  $P$  tiene solución óptima  $x_0$  y definimos  $f(x_0) = \mu_0$ . Por una parte el problema  $\tilde{P}$  o  $\tilde{\tilde{P}}$  no puede tener valor óptimo  $\mu_1 > \mu_0$  pues el punto  $(x_0, \mu_0)$  es factible para el problema y evaluado en la función objetivo es mejor que  $\mu_1$ . Por otra parte, el problema  $\tilde{P}$  o  $\tilde{\tilde{P}}$  no puede tener valor óptimo  $\mu_1 < \mu_0$  pues entonces existiría un punto  $x_1$  en  $D$  tal que  $f(x_1) \leq \mu_1$  en cuyo caso la solución óptima a  $P$  sería  $x_1$  y no  $x_0$ .

Es importante destacar que en el nuevo problema el objetivo es encontrar un  $\mu$  lo más pequeño posible tal que exista un  $x$  que satisfaga  $f(x) \leq \mu$ . Así, en la Figura 1.6 se observa que  $\mu^*$  cumple con ser el óptimo al problema pues existe un  $x^* \in D$  tal que  $f(x^*) \leq \mu^*$  y para todo valor de  $\mu$  menor a  $\mu^*$  no existe un  $x$  en  $D$  tal que  $f(x) \leq \mu$ .

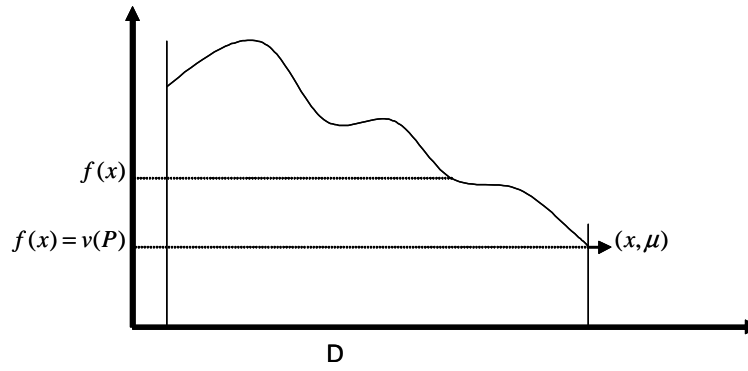


Figure 1.6: Equivalencia II

**Ejemplo 2** En un callejón de 100 metros de largo hay un prófugo que necesita ubicarse en el lugar que tenga menos luz. En el callejón hay 4 focos con diferentes características (altura, potencia, posición). Suponga que la intensidad de luz que llega al prófugo es solamente la del foco que alumbra más en dicho punto y que para estos efectos los postes (si los hubiera) se pueden considerar transparentes (ver Figura 1.7). Suponga que la intensidad de luz de cada foco se puede considerar inversamente proporcional a la distancia entre el foco y el prófugo y directamente proporcional a la potencia del foco. De esta forma, al prófugo le interesará pararse en el lugar que haya menos luz,

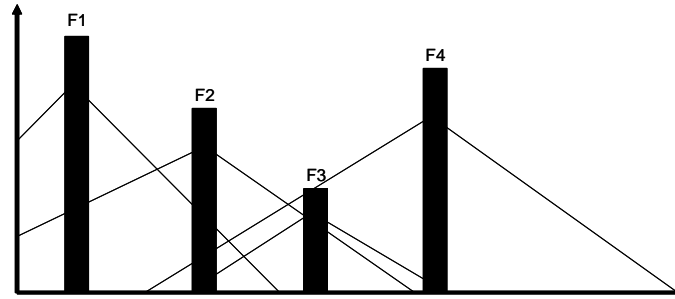


Figure 1.7: Ejemplo de Equivalencia II

es decir, su función objetivo será:

$$\text{Min } (\max \{f_1, f_2, f_3, f_4\})$$

donde

$$f_i = \frac{k_i}{\sqrt{h_i^2 + (x_i - x)^2}}$$

$$0 \leq x \leq 100$$

en que  $k_i$ ,  $x_i$  y  $h_i$  corresponden a la potencia, ubicación y altura respectivamente de cada foco. La

única variable del problema es la ubicación del prófugo,  $x$ .

Al igual que el problema de la estación de bomberos, este modelo es complicado de abordar pues presenta una función objetivo cuya derivada no es continua. Así, resulta deseable identificar un modelo equivalente para este problema que no presente dicha complicación. La siguiente sección presenta dicha equivalencia.

### 1.7.3 Equivalencia III

Consideremos un problema como el anterior donde el objetivo consiste en encontrar un punto tal que el máximo entre  $n$  funciones evaluadas en el punto sea mínimo. Algebraicamente, esto puede expresarse del siguiente modo:

$$P) \quad \text{Min}_{x \in D} \left( \max_{i=1, \dots, n} f_i(x) \right)$$

Gráficamente, el problema se ilustra en la Figura 1.8, en que las funciones  $f_i(x)$  se dibujan con líneas tenues mientras la función objetivo se ilustra con una línea gruesa. Es importante destacar que esta última función no gozará necesariamente de una derivada continua (e.g., en los puntos A y B de la figura 1.8). Así, estamos minimizando una función no diferenciable.

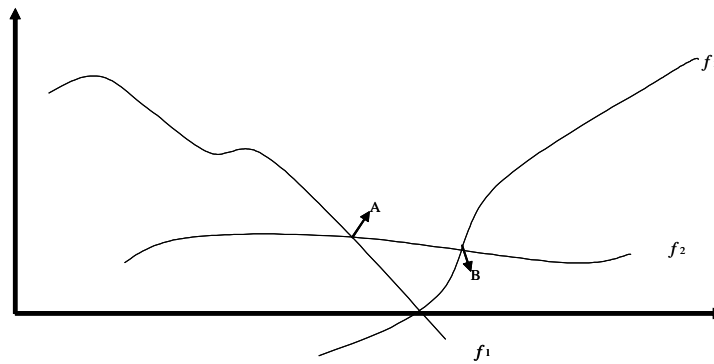


Figure 1.8: Equivalencia III

Sin embargo, de acuerdo a la equivalencia II, el problema  $P)$  puede convertirse en el siguiente problema equivalente:

$$\begin{aligned} \tilde{P}) \quad & \text{Min } \mu \\ & \max_{i=1, \dots, n} f_i(x) \leq \mu \\ & x \in D \\ & \mu \in \mathbb{R} \end{aligned}$$



Naturalmente la nueva desigualdad puede expresarse como  $n$  desigualdades individuales:

$$\begin{aligned} & \tilde{\tilde{P}}) \text{ Min } \mu \\ & f_i(x) \leq \mu \quad \forall i = 1, \dots, n \\ & x \in D \\ & \mu \in \mathbb{R} \end{aligned}$$

Este último problema contempla sólo funciones diferenciables.

Consideremos el siguiente ejemplo cuya función objetivo es no lineal y cuya derivada no es continua.

$$\begin{aligned} P) \text{ Max } & \left( \min \left\{ \frac{7x_1 + 6x_2 + 5x_3}{4}, \frac{5x_1 + 9x_2 + 4x_3}{3} \right\} \right) \\ & \text{s.a.} \quad \begin{aligned} 8x_1 + 5x_2 + 3x_3 &\leq 100 \\ 6x_1 + 9x_2 + 8x_3 &\leq 200 \\ x_1, x_2, x_3 &\geq 0 \end{aligned} \end{aligned}$$

A continuación identificaremos un problema equivalente cuyas funciones tengan derivadas continuas. De la regla I),  $P)$  es equivalente a:

$$\tilde{P}) \min_{x \in D} (-\min \{f_1(x), f_2(x)\})$$

Es fácil notar que:  $-\min \{f_1(x), f_2(x)\} = \max \{-f_1(x), -f_2(x)\}$ , por lo tanto  $\tilde{P})$  es equivalente a

$$\tilde{\tilde{P}}) \min_{x \in D} (\max \{-f_1(x), -f_2(x)\})$$

de la regla III  $\tilde{\tilde{P}})$  es equivalente a

$$\begin{aligned} & \tilde{\tilde{\tilde{P}}}) \min \mu \\ & \text{s.a.} \quad \begin{aligned} -f_1(x) &\leq \mu \\ -f_2(x) &\leq \mu \\ x &\in D \\ \mu &\in \mathbb{R}. \end{aligned} \end{aligned}$$

o equivalentemente:

$$\begin{aligned}
 & \tilde{\tilde{P}}) \max \mu \\
 s.a. \quad & \mu \leq f_1(x) \\
 & \mu \leq f_2(x) \\
 & x \in D \\
 & \mu \in \mathbb{R}.
 \end{aligned}$$

Y este problema es derivable en el espacio  $D \times \mathbb{R}$ . Es decir,

$$\begin{aligned}
 & \tilde{P}) \max(\min_{i=1,\dots,n} \{f_i(x)\}) \\
 s.a. \quad & x \in D
 \end{aligned}$$

es equivalente a

$$\begin{aligned}
 & \tilde{\tilde{P}}) \max \mu \\
 s.a. \quad & \mu \leq f_i(x) \quad \forall i = 1, \dots, n \\
 & x \in D \\
 & \mu \in \mathbb{R}.
 \end{aligned}$$

Como se observa, este problema lineal es equivalente al problema original. Estas equivalencias suelen resultar muy útiles en presencia de módulos de funciones en la función objetivo. Por ejemplo:

$$\begin{aligned}
 & P) \min |3x - 2| \\
 s.a. \quad & -1 \leq x \leq 2 \\
 & x \in \mathbb{R}
 \end{aligned}$$

En este caso vemos que la función objetivo no es derivable en  $x = 2/3$  (ver Figura 1.9).

Ocupando la regla II vemos que  $P)$  es equivalente a:

$$\begin{aligned}
 & \tilde{P}) \min \mu \\
 s.a. \quad & |3x - 2| \leq \mu \quad \text{(Restricción no-lineal)} \\
 & -1 \leq x \leq 2 \\
 & x \in \mathbb{R} \\
 & \mu \in \mathbb{R}
 \end{aligned}$$

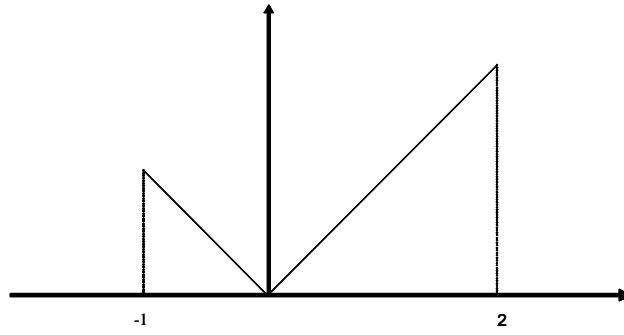


Figure 1.9: Ejemplo de minimización del módulo

Pero también  $\tilde{\tilde{P}}$  es equivalente a:

$$\begin{aligned}
 & \tilde{\tilde{P}}) \quad \min \mu \\
 & s.a. \quad 3x - 2 \leq \mu \\
 & \quad \quad -(3x - 2) \leq \mu \\
 & \quad \quad -1 \leq x \leq 2 \\
 & \quad \quad x \in \mathbb{R} \\
 & \quad \quad \mu \in \mathbb{R}
 \end{aligned}$$

Ahora todas las funciones que intervienen son derivables (y en este caso lineales). Es importante notar que en esta formulación se utilizó la equivalencia  $|3x - 2| \leq \mu \Leftrightarrow 3x - 2 \leq \mu$  y  $-(3x - 2) \leq \mu$ . Esta equivalencia es posible ya que la expresión  $|3x - 2| \leq \mu$  es satisfecha por un conjunto convexo de puntos  $(x, \mu)$ . Este conjunto queda bien recogido por las dos restricciones equivalentes. Sin embargo una expresión como  $|f(x)| \geq \mu$  ( $\mu \geq 0$ ) no tiene un conjunto de restricciones diferenciables equivalente ya que el conjunto de puntos  $(x, \mu)$  que la satisface no es convexo. Dicha expresión será equivalente al conjunto de puntos  $(x, \mu)$  tal que  $f(x) \geq \mu$  o bien  $-f(x) \geq \mu$ .

En el ejemplo anterior se puede ver que en el óptimo,  $|3x - 2| = \mu \Leftrightarrow 3x - 2 = \mu$  o bien  $3x - 2 = -\mu$ ; esta relación puede representarse mediante  $(3x - 2 - \mu)(3x - 2 + \mu) = 0$ .

Por lo tanto,  $\tilde{\tilde{\tilde{P}}}$  también es equivalente a:

$$\begin{aligned}
 & \tilde{\tilde{\tilde{P}}}) \quad \min \mu \\
 & s.a. \quad (3x - 2 - \mu)(3x - 2 + \mu) = 0 \\
 & \quad \quad -1 \leq x \leq 2 \\
 & \quad \quad x \in \mathbb{R} \\
 & \quad \quad \mu \in \mathbb{R}
 \end{aligned}$$

En este caso, como el problema  $\tilde{\tilde{\tilde{P}}}$  es lineal, lo preferimos a este último problema.

Es importante notar que aunque  $-1 \leq x \leq 2$  está demás en este caso, muchas veces no es recomendable eliminar (o relajar) restricciones reales de un problema aún cuando se sepa que la solución óptima del problema permanecería inalterable (a menos que la relajación simplifique considerablemente la resolución del problema). Esto se debe a que a veces falla la intuición del modelador y porque el modelo puede ser usado más adelante con otros parámetros que activen la restricción relajada.

En términos generales, los siguientes modelos resultan equivalentes:

$$P) \quad \min_{x \in D} |f(x)|$$

$$\begin{aligned} \tilde{P}) \quad & \min \mu \\ & f(x) \leq \mu \\ & -f(x) \leq \mu \\ & x \in \mathbb{R} \\ & \mu \in \mathbb{R} \end{aligned}$$

$$\begin{aligned} \tilde{\tilde{P}}) \quad & \min \mu \\ (f(x) - \mu)(f(x) + \mu) &= 0 \\ & x \in \mathbb{R} \\ & \mu \in \mathbb{R} \end{aligned}$$

Al realizar estas equivalencias, el lector debe cuidarse de no incurrir en equivalencias erróneas. Por ejemplo:

$$\begin{aligned} P) \quad & \max(\max_{i=1, \dots, n} \{f_i(x)\}) \\ s.a. \quad & x \in D \end{aligned}$$

no es equivalente a

$$\begin{aligned} \tilde{P}) \quad & \max \mu \\ s.a. \quad & \mu \leq f_i(x) \quad \forall i = 1, \dots, n \\ & x \in D \\ & \mu \in \mathbb{R}. \end{aligned}$$

pues todo lo que se puede decir es que  $P$ ) es equivalente a

$$\begin{aligned} & \widetilde{P}) \max \mu \\ \text{s.a. } & \mu \leq \max_{i=1, \dots, n} \{f_i(x)\} \\ & x \in D \\ & \mu \in \mathbb{R}. \end{aligned}$$

Sin embargo, que  $\mu$  sea inferior al máximo de un conjunto de funciones no es equivalente a que  $\mu$  sea inferior a cada una de ellas (equivaldría a ser inferior al mínimo de ellas). Lo que se requeriría es exigir que  $\mu$  sea inferior a al menos una de las funciones, es decir  $\mu \leq f_1(x)$  ó  $\mu \leq f_2(x)$  ó ....  $\mu \leq f_n(x)$ . En cambio al exigir  $\mu \leq f_i(x) \forall i = 1, \dots, n$  se está exigiendo  $\mu \leq f_1(x)$  y  $\mu \leq f_2(x)$  y .... y  $\mu \leq f_n(x)$ . Incorporar las relaciones lógicas asociadas al operador “ó” típicamente requieren de variables binarias asociadas a cada relación. Estas variables binarias no son continuas por lo que tampoco son diferenciables. Así, no existe un modelo equivalente diferenciable a  $P$ ) en este caso.

Por ejemplo, consideremos el siguiente problema

$$\begin{aligned} & P) \max |3x - 2| \\ \text{s.a. } & -1 \leq x \leq 2 \\ & x \in \mathbb{R} \end{aligned}$$

La solución óptima a este problema será  $x = -1$  con un valor óptimo  $\hat{v} = 5$ . Podemos ver que el siguiente modelo:

$$\begin{aligned} & P) \max \mu \\ \text{s.a. } & \mu \leq 3x - 2 \\ & \mu \leq -(3x - 2) \\ & -1 \leq x \leq 2 \\ & x \in \mathbb{R} \end{aligned}$$

no es equivalente pues la solución  $x = -1, \mu = 5$  no es factible, ya que no satisface la primera restricción.

Considere el siguiente problema:

$$\begin{aligned} & P) \min z = (\max \{|x_1|, |x_2|\}) \\ \text{s.a. } & x_1 + 2x_2 \geq 1 \\ & x_1, x_2 \geq 0 \end{aligned}$$

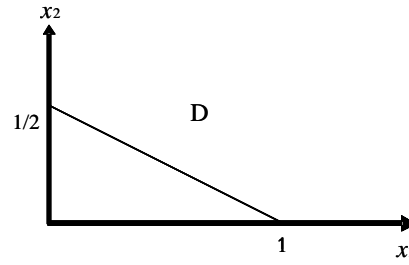


Figure 1.10: Ejemplo equivalencia III

Aplicando la regla III,  $P$ ) es equivalente a:

$$\begin{aligned}
 \tilde{P}) \quad & \min \mu \\
 & x_1 \leq \mu \\
 & -x_1 \leq \mu \\
 & x_2 \leq \mu \\
 & -x_2 \leq \mu \\
 & x_1 + 2x_2 \geq 1 \\
 & x_1, x_2 \geq 0 \\
 & (x_1, x_2) \in \mathbb{R}^2 \\
 & \mu \in \mathbb{R}
 \end{aligned}$$

Una ventaja de los problemas bidimensionales como el anterior es que pueden resolverse mediante curvas de nivel. En este procedimiento se dibuja en el plano  $\mathbb{R}^2$  un conjunto de curvas tal que cada una de ellas represente el lugar geométrico de todos los puntos de  $\mathbb{R}^2$  que reemplazados en la función objetivo den el mismo valor o cota. A continuación se debe graficar en el mismo plano los puntos que satisfacen las restricciones del dominio. Así, es posible identificar visualmente la solución óptima al problema. Por ejemplo, en la Figura 1.10 podemos ver el dominio del problema  $P$ ), y la Figura 1.11 ilustra el dominio al agregar las curvas de nivel.

En este caso, las curvas de nivel permiten observar que la función objetivo corresponde a una pirámide invertida de base cuadrada en que las aristas de la base son paralelas a los ejes cartesianos. El problema  $P$ ) busca el punto del dominio que pertenezca a la curva de nivel de mínima cota (ver Figura 1.11).

La figura permite identificar que en el óptimo  $x_1^* = x_2^*$ , y como  $x_1^* + 2x_2^* = 1$ , entonces  $x_1^* = \frac{1}{3}$  y  $x_2^* = \frac{1}{3}$ . Por lo tanto  $\hat{v}(P) = \frac{1}{3}$ .

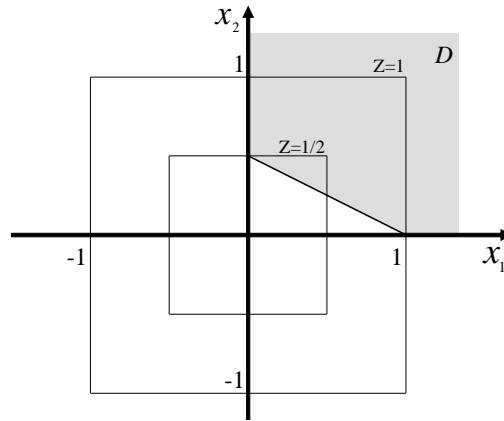


Figure 1.11: Curvas de nivel

#### 1.7.4 Equivalencia IV

Considere el siguiente problema:

$$P) \quad \min \sum_{i=1}^r f_i(\vec{x})$$

$$\vec{x} \in D \subset \mathbb{R}^n$$

Este problema puede expresarse en forma equivalente como:

$$\tilde{P}) \quad \min \sum_{i=1}^r \mu_i$$

$$f_i(\vec{x}) \leq \mu_i; \quad \forall i = 1, \dots, r$$

$$\vec{x} \in D$$

$$\vec{\mu} \in \mathbb{R}^r$$

Esta equivalencia puede demostrarse por contradicción al igual que la equivalencia II. Se deja al lector como ejercicio.

Esta equivalencia es de especial interés en casos en que algunas funciones  $f_i(x)$  son no diferenciables y al pasarlas al dominio permiten encontrar una equivalencia que sí lo sea. Por ejemplo:

$$P) \quad \min(|x_1| + |x_2|)$$

$$s.a. \quad x_1 + 2x_2 \geq 1$$

$$\begin{aligned}
\tilde{P}) \quad & \min \mu_1 + \mu_2 \\
s.a. \quad & |x_1| \leq \mu_1 \\
& |x_2| \leq \mu_2 \\
& x_1 + 2x_2 \geq 1
\end{aligned}$$

Lo que también equivale a: (Problema Lineal Equivalente)

$$\begin{aligned}
\tilde{\tilde{P}}) \quad & \min \mu_1 + \mu_2 \\
s.a. \quad & x_1 \leq \mu_1 \\
& -x_1 \leq \mu_1 \\
& x_2 \leq \mu_2 \\
& -x_2 \leq \mu_2 \\
& x_1 + 2x_2 \geq 1 \\
& \mu_1, \mu_2 \in \mathbb{R}
\end{aligned}$$

En términos generales se puede decir que:

$$\begin{aligned}
P) \quad & \min \sum_{i=1}^r |f_i(x)| \\
s.a. \quad & x \in D \subset \mathbb{R}^n
\end{aligned}$$

$$\begin{aligned}
\tilde{P}) \quad & \min \sum_{i=1}^r \mu_i \\
& f_i(x) \leq \mu_i; \quad \forall i = 1, \dots, r \\
& -f_i(x) \leq \mu_i; \quad \forall i = 1, \dots, r \\
& x \in D \\
& \mu \in \mathbb{R}^r
\end{aligned}$$

Ejemplo: ¿son los siguientes problemas equivalentes?:

$$\begin{aligned}
P) \quad & \min x - 2 \\
s.a. \quad & x \geq 0
\end{aligned}$$

$$\begin{aligned}
\tilde{P}) \quad & \min (x - 2)^2 \\
s.a. \quad & x \geq 0
\end{aligned}$$

Claramente no lo es porque la función  $g(x) = x^2$  es estrictamente creciente sobre los números



no negativos. Sin embargo, aquí se aplica sobre una función que puede arrojar números negativos (e.g. si  $x = 0$ , entonces  $f(x) = -2$ ). De hecho la solución óptima al primer problema es  $x = 0$ , mientras el del segundo es  $x = 2$ .

### 1.7.5 Equivalencia V

El problema de minimización  $\text{Min}_{x \in D} f(x)$  es equivalente al problema  $\text{Min}_{x \in D} g(f(x))$ , donde la función  $g : f(D) \subset \mathbb{R} \rightarrow \mathbb{R}$  es estrictamente creciente sobre  $f(D) = \{y \in \mathbb{R} / \exists x \in D : y = f(x)\}$ , por lo que la solución óptima será la misma en ambos casos, pero no así los valores óptimos. Para comprobar esta equivalencia, considere a  $x^*$  como el óptimo del problema original. Esto significa que  $f(x^*) \leq f(x) \forall x \in D$ . Como  $g(x)$  es estrictamente creciente, si  $x_1 \leq x_2$  entonces  $g(x_1) \leq g(x_2)$ . Así, necesariamente  $g(f(x^*)) \leq g(f(x)) \forall x \in D$  y por lo tanto ambos problemas arrojarán la misma solución óptima. Por ejemplo, al cambiar la escala de unidades de la función objetivo de dólares a pesos ( $\text{US\$} \rightarrow \$$ ), se está aplicando una función  $g(x)$  lineal y creciente sobre todo el dominio por lo que la solución óptima al problema es la misma independiente de las unidades.

**Ejemplo 3** Determinar el punto más cercano al origen  $(0,0)$  que satisfaga  $2x_1 + x_2 \geq 2$ .

En este caso la función de costos sería  $\sqrt{x_1^2 + x_2^2}$ , y el dominio  $D = \{(x_1, x_2) / 2x_1 + x_2 \geq 2\}$ . La función raíz cuadrada suele presentar complicaciones pues no es diferenciable en cero. Por lo tanto, resulta conveniente modificar la función objetivo con el fin de eliminar la raíz.

Tomemos  $g(y) = y^2$ . Esta función es estrictamente creciente sobre el recorrido de la función objetivo original, esto es, los números no negativos  $\mathbb{R}^+$ , ya que  $f(x) \geq 0 \forall x \in D$ .

Por lo tanto, los siguientes problemas son equivalentes:

$$\begin{array}{ll} P) & \min \sqrt{x_1^2 + x_2^2} \\ \text{s.a} & 2x_1 + x_2 \geq 2 \end{array}$$

$$\begin{array}{ll} \tilde{P}) & \min x_1^2 + x_2^2 \\ \text{s.a} & 2x_1 + x_2 \geq 2 \end{array}$$

y este último problema presenta la ventaja que su función objetivo:  $\tilde{f}(x_1, x_2) = x_1^2 + x_2^2$  es diferenciable sobre todo  $\mathbb{R}^2$ .

**Nota:** Si el problema  $P$ ) está bien formulado (i.e. el argumento de la raíz sea no negativo para todo  $x \in D$ ) se puede siempre prescindir de la raíz.

$$\begin{array}{ll} P) & \min_{x \in D} \sqrt{r(x)} \\ \text{con} & r(x) \geq 0 \quad \forall x \in D \end{array}$$

$$\tilde{P}) \quad \min_{x \in D} r(x)$$

En este caso es importante recordar que si  $\hat{x}$  es solución óptima de  $\tilde{P}$  con  $v(\tilde{P}) = r(\hat{x})$ , entonces  $\hat{x}$  es solución óptima de  $P$  con  $v(P) = \sqrt{r(\hat{x})}$ .

### 1.7.6 Equivalencia VI

Basándose en las equivalencias anteriores podemos decir que el problema

$$\begin{aligned} P) \quad & \min_{x \in D} \frac{1}{f(x)} \\ \text{con} \quad & f(x) \neq 0, \forall x \in D \\ \text{y} \quad & f(x) > 0 \end{aligned}$$

es equivalente a

$$\tilde{P}) \quad \max_{x \in D} f(x).$$

Para probar lo anterior basta con suponer el caso particular en que  $g(y) = \ln y$ , estrictamente creciente  $\forall y > 0$ . Aplicando la regla V  $P$ ) es equivalente a:

$$\tilde{P}) \quad \min_{x \in D} \ln \left( \frac{1}{f(x)} \right) = \min_{x \in D} -\ln f(x)$$

Finalmente, de la regla I,  $\tilde{\tilde{P}}$ ) es equivalente a:

$$\tilde{\tilde{P}}) \quad \max_{x \in D} \ln f(x)$$

De la regla V, con  $g(y) = e^y$ ,  $\tilde{\tilde{P}}$ ) es equivalente a:

$$\tilde{\tilde{\tilde{P}}}) \quad \max_{x \in D} f(x).$$

Es importante destacar que esta equivalencia ocurre sólo si  $f(x)$  es estrictamente positivo para todo  $x$ . Si la función puede tomar valores negativos para algunos puntos del dominio, entonces aplicar esta equivalencia puede conducir a error.

### 1.7.7 Equivalencia VII

Aunando las equivalencias anteriores, podemos decir que el problema

$$P) \quad \min_{x \in D} (g(x) + \max_{i=1, \dots, r} \{f_i(x)\})$$

Geoméricamente

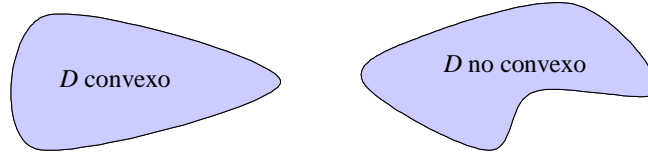


Figure 1.12: Conjuntos convexos y no convexos

es equivalente a:

$$\begin{aligned}
 \tilde{P}) \quad & \min_{x \in D} \mu_1 + \mu_2 \\
 & g(x) \leq \mu_1 \\
 & f_i(x) \leq \mu_2 \quad \forall i = 1, \dots, r \\
 & \mu_1, \mu_2 \in \mathbb{R}.
 \end{aligned}$$

## 1.8 Nociones Básicas de Convexidad

Los problemas de optimización pueden ser sumamente complejos y contener múltiples soluciones óptimas locales lo que puede dificultar enormemente su resolución. Estas dificultades pueden reducirse si se identifican adecuadamente algunas características tanto de la función objetivo como del dominio del problema. En esta sección se revisan las características de los conjuntos convexos, las funciones convexas y los problemas convexas.

### 1.8.1 Conjuntos Convexos

Intuitivamente un conjunto se dice **convexo** si para cualquier par de puntos en el conjunto, todos los puntos en la línea recta que los une también pertenecen al conjunto. Basta que haya un par de puntos en el dominio que no satisfaga esta condición para que el conjunto se diga **no convexo**. En la Figura 1.12 se observa un ejemplo de un conjunto convexo y uno no convexo. Formalmente,

**Definición 5** Un conjunto  $D \subset \mathbb{R}^n$  se dice convexo si para todo par de puntos  $x_1 \in D$  y  $x_2 \in D$  y cada número real  $\lambda$  con  $0 \leq \lambda \leq 1$ ,  $x = \lambda x_1 + (1 - \lambda)x_2 \in D$ .

**Definición 6** Segmento  $[x_1, x_2] = \{z \in \mathbb{R}^n / z = (1 - \lambda)x_1 + \lambda x_2, \lambda \in [0, 1]\}$ . Dado lo anterior, un dominio  $D$  será convexo si:  $\forall x_1$  y  $x_2 \in D \Rightarrow [x_1, x_2] \subseteq D$ .

**Claim 1** Todos los poliedros definidos por desigualdades lineales son convexas. Esta demostración se deja al lector.

**Proposición 3** El conjunto formado por la intersección de conjuntos convexas es convexo (ver Figura 1.13).

**Demostración.** Sean  $A$  y  $B$  conjuntos convexos, y su intersección denominémosla  $C = A \cap B$ . Se debe demostrar que si  $x_1$  y  $x_2 \in C$ , entonces  $[x_1, x_2] \in C$ . Si  $x_1$  y  $x_2 \in C$ , entonces  $x_1$  y  $x_2 \in A$  y  $x_1$  y  $x_2 \in B$ . Como  $A$  y  $B$  son convexos,  $[x_1, x_2] \subset A$  y  $[x_1, x_2] \subset B$ , por lo tanto,  $[x_1, x_2] \subset C$ .  $\Rightarrow C$  es convexo. Ver Figura 1.13. ■

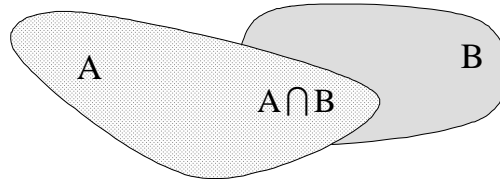


Figure 1.13: Intersección de Conjuntos

### 1.8.2 Funciones Convexas

En optimización, estamos especialmente interesados en identificar puntos extremos de la función objetivo. Para algunas funciones esto resulta particularmente sencillo. Desde esta perspectiva, en la Figura 1.14 se entrega una clasificación del universo de funciones. En esta figura se destacan las funciones convexas, estrictamente convexas y cuasi-convexas que a continuación se describen.

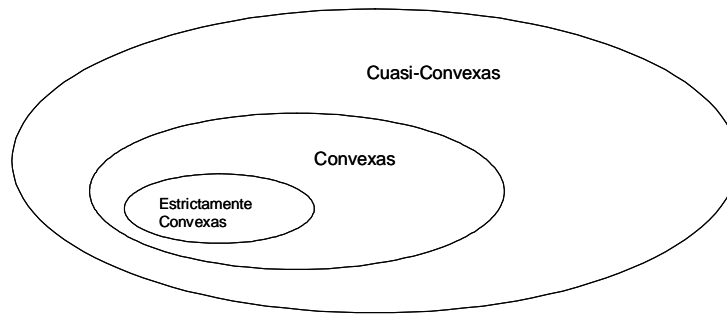


Figure 1.14: Universo de funciones

**Definición 7 (*Función convexa*)** Sea  $f(x) : D \rightarrow \mathbb{R}$ , con  $D$  convexo. Entonces,  $f(x)$  es convexa sobre  $D$  si:

$$f((1 - \lambda)x_1 + \lambda x_2) \leq (1 - \lambda)f(x_1) + \lambda f(x_2) \quad \forall x_1, x_2 \in D, \lambda \in [0, 1].$$

En otras palabras, lo que dice la Definición 7, es que si en cualquier intervalo  $[x_1, x_2]$  todo punto del grafo de  $f(x)$  está siempre bajo la cuerda que une los puntos  $(x_1, f(x_1))$  y  $(x_2, f(x_2))$ , entonces  $f(x)$  es una función convexa. La Figura 1.15 lo muestra gráficamente. Es importante notar que una función definida sobre un dominio no convexo no puede ser convexa.

¿Qué ocurre con la función de la Figura 1.16? El grafo de  $f(x)$  sobre la región  $[x_1, x_2]$  está por encima de la cuerda, por lo tanto  $f(x)$  no es convexa sobre  $D$ .

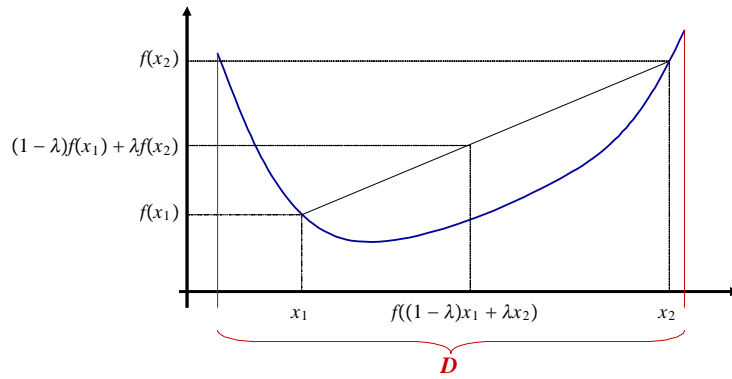


Figure 1.15: Función convexa

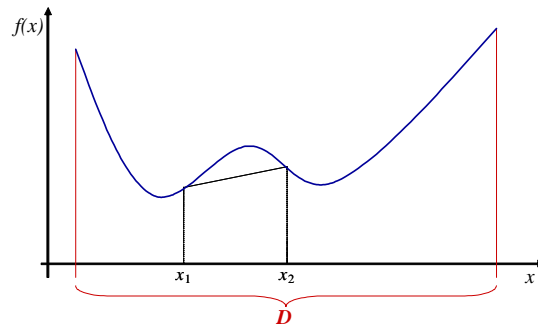


Figure 1.16: Función no convexa

Al observar los gráficos de funciones convexas y no convexas, se deduce la siguiente propiedad:

**Proposición 4** Si la función  $f(x)$  es diferenciable y convexa sobre  $D$ , entonces la tangente en cualquier punto de  $f(x)$  no podrá exceder a  $f(x)$  para cualquier  $x$  en el dominio (ver Figura 1.17). Es decir, para que  $f(x)$  sea convexa,  $f(x) \geq f(\bar{x}) + \frac{df(\bar{x})}{dx} (x - \bar{x}), \forall x, \bar{x} \in D$ .

**Proposición 5** La suma de funciones convexas es también convexa.

**Demostración.** Sean  $f_1$  y  $f_2$  las funciones. Sea  $x_1, x_2 \in D$ , y  $0 \leq \lambda \leq 1$ , tal que

$$\begin{aligned} f_1((1-\lambda)x_1 + \lambda x_2) &\leq (1-\lambda)f_1(x_1) + \lambda f_1(x_2) \\ f_2((1-\lambda)x_1 + \lambda x_2) &\leq (1-\lambda)f_2(x_1) + \lambda f_2(x_2). \end{aligned}$$

Sumando obtenemos

$$f_1((1-\lambda)x_1 + \lambda x_2) + f_2((1-\lambda)x_1 + \lambda x_2) \leq (1-\lambda)f_1(x_1) + \lambda f_1(x_2) + (1-\lambda)f_2(x_1) + \lambda f_2(x_2).$$

Si  $f_1(x) + f_2(x) = g(x)$  entonces  $g((1-\lambda)x_1 + \lambda x_2) \leq (1-\lambda)g(x_1) + \lambda g(x_2)$ . Por lo tanto  $g(x)$ , es decir  $f_1(x) + f_2(x)$ , es convexa. ■

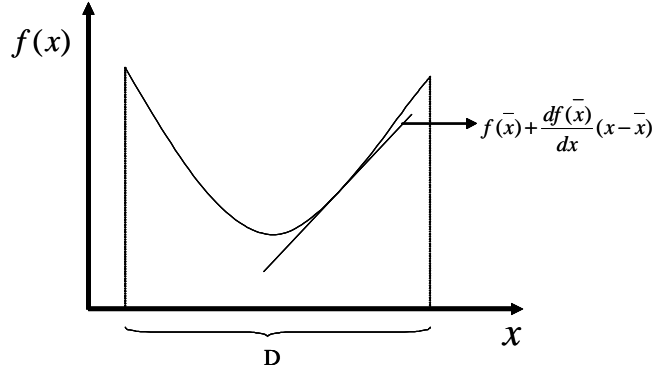
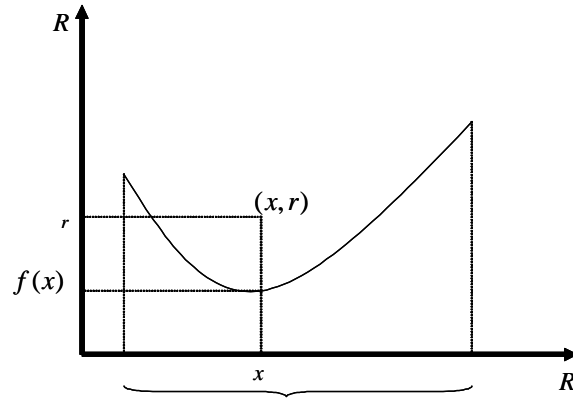


Figure 1.17: Tangente de función convexa

Figure 1.18: Epígrafo de la función  $f(x)$ 

Análogamente, es fácil demostrar que una función convexa multiplicada por cualquier parámetro positivo sigue siendo convexa.

**Definición 8 (Función cóncava)** Sea  $f(x) : D \rightarrow \mathbb{R}$ , con  $D$  convexo. Entonces,  $f(x)$  es cóncava sobre  $D$  si:

$$f((1 - \lambda)x_1 + \lambda x_2) \geq (1 - \lambda)f(x_1) + \lambda f(x_2) \quad \forall x_1, x_2 \in D, \lambda \in [0, 1].$$

Así, resulta claro que si  $f(x)$  es una función convexa, entonces  $-f(x)$  será una función cóncava.

**Proposición 6** Si  $D$  es convexo,  $f : D \rightarrow \mathbb{R}$  es convexa sobre  $D$  si y solo si el epígrafo de  $f$  sobre  $D$ ,  $E_D(f) = \{(x, r) : x \in D, r \geq f(x)\}$ , es una parte convexa de  $\mathbb{R}^n \times \mathbb{R}$ . Ver Figura 1.18.

**Proposición 7** Sean  $D \in \mathbb{R}^n$  convexo, y  $f_i(x) \forall i = 1, \dots, p$  convexas sobre  $D$ , entonces  $f(x) = \max \{f_i(x)\}$  define una función convexa sobre  $D$ .

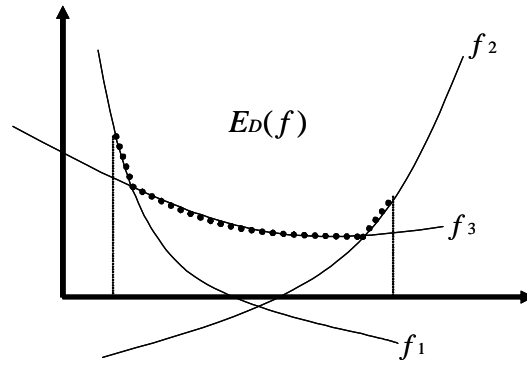


Figure 1.19: Intersección de epígrafos

**Demostración.** Dado que la intersección (cualquiera) de conjuntos convexos es un conjunto convexo, podemos decir:

$$\begin{aligned} f(x) &= \max_{i=1,\dots,p} \{f_i(x)\} \\ E_D(f) &= \bigcap_{i=1}^p E_D(f_i). \end{aligned}$$

Como cada uno de los epígrafos de las funciones  $f_i(x)$  es convexo, su intersección también lo es. Así,  $E_D(f)$  es convexo, por lo que  $f(x)$  es convexa sobre  $D$  (ver Figura 1.19). ■

Las funciones convexas presentan favorables propiedades al optimizar. Por ejemplo, un punto mínimo local es siempre mínimo global. Sin embargo, no necesariamente este óptimo será único. Por ejemplo, la función  $f(x)$  en la Figura 1.20 tiene varias soluciones óptimas.

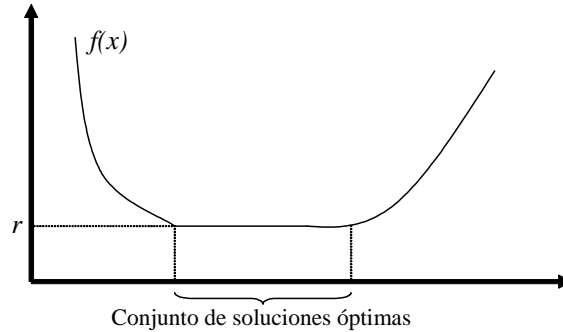


Figure 1.20: Ejemplo de múltiples soluciones óptimas

Así, surge la necesidad de definir las funciones estrictamente convexas.

**Definición 9 (Función estrictamente convexa)** Sea  $f(x) : D \rightarrow \mathbb{R}$ , con  $D$  convexo. Entonces,  $f(x)$  es estrictamente convexa sobre  $D$  si:

$$f((1 - \lambda)x_1 + \lambda x_2) < (1 - \lambda)f(x_1) + \lambda f(x_2) \quad \forall x_1, x_2 \in D, \lambda \in [0, 1].$$

Es decir toda función estrictamente convexa es también convexa. En forma análoga se definen las funciones estrictamente cóncavas.

La Figura 1.21a presenta una nueva función convexa, pero no estricta. Cabe mencionar que toda función lineal  $f(x) = ax + b$  es cóncava y convexa a la vez, pero ni estrictamente cóncava ni estrictamente convexa. Por otra parte, la Figura 1.21b muestra una función que no es ni cóncava ni convexa, pero que localmente cerca de  $x_1$  es cóncava y localmente cerca de  $x_2$  es convexa.

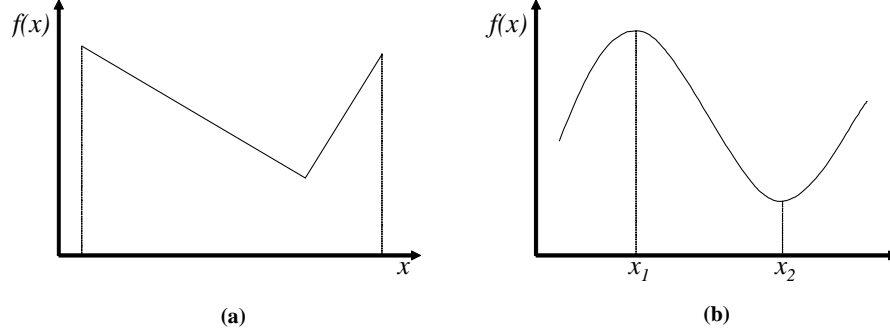


Figure 1.21: Casos especiales de funciones

A continuación se demostrará que si  $f(x)$  es convexa, entonces el lugar geométrico de los puntos que satisfacen  $f(x) \leq \alpha$  debe constituir un conjunto convexo para cualquier  $\alpha$ .

**Proposición 8** Si  $\theta : D \rightarrow \mathbb{R}, D \subset \mathbb{R}^n, D$  convexo,  $\theta$  convexa, entonces los conjuntos de nivel  $C_\alpha(\theta) = \{x \in D / \theta(x) \leq \alpha\}$  son convexos  $\forall \alpha \in \mathbb{R}$ .

**Demostración.** Sean  $x_1, x_2 \in C_\alpha(\theta)$ , con  $\alpha \in \mathbb{R}$  fijo, y  $\lambda \in [0, 1]$ . Mostremos que  $(1 - \lambda)x_1 + \lambda x_2 \in C_\alpha(\theta)$ . Si  $x_1$  y  $x_2 \in C_\alpha(\theta) \Rightarrow x_1, x_2 \in D$ , y

$$\theta(x_1) \leq \alpha \quad y \quad \theta(x_2) \leq \alpha. \quad (1.3)$$

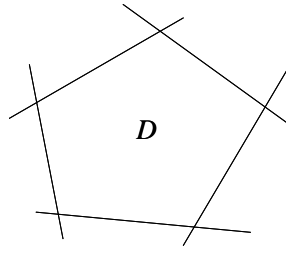
Como  $D$  es convexo  $\Rightarrow (1 - \lambda)x_1 + \lambda x_2 \in D$ ,  $\lambda \in [0, 1]$  y como  $\theta(\cdot)$  es convexa sobre  $D$ , entonces  $\theta((1 - \lambda)x_1 + \lambda x_2) \leq (1 - \lambda)\theta(x_1) + \lambda\theta(x_2)$ . Por lo tanto de (1.3) tenemos que  $\theta((1 - \lambda)x_1 + \lambda x_2) \leq (1 - \lambda)\alpha + \lambda\alpha = \alpha$ . Así,  $(1 - \lambda)x_1 + \lambda x_2 \in C_\alpha(\theta)$ . ■

**Corolario 9** Si el dominio de restricción está definido como:  $D = \{x \in \mathbb{R}^n : g_i(x) \leq 0, i = 1, \dots, m\}$  con  $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$  funciones convexas, entonces  $D$  es una parte convexa de  $\mathbb{R}^n$ .

**Demostración.** Dado que las funciones  $g_i(x)$  son convexas, las regiones definidas por las restricciones  $g_i(x) \leq 0$  son también convexas (conjuntos de nivel). El dominio  $D$  queda definido como la intersección de las regiones de puntos definidos por cada una de las  $m$  restricciones, formalmente  $D = \cap_{i=1}^m C_0(g_i)$ . Como cada uno de las regiones  $C_0(g_i)$  es convexa y la intersección de conjuntos convexas es convexa, entonces  $D$  es convexo. ■

Es importante destacar que restricciones del tipo  $h_i(x) = 0$ , sólo definen regiones convexas si la función es lineal (no basta que  $h_i(x)$  sea convexa).



Figure 1.22: Dominio  $D$  formado por semi-espacios

**Caso particular:** Si tenemos restricciones lineales del tipo  $g_i(x) = a_i x - b_i$  (funciones lineales son cóncavas y convexas a la vez), entonces  $D$  es un poliedro convexo (intersección finita de semi-espacios, ver Figura 1.22). Las funciones convexas no son las únicas en satisfacer que el lugar geométrico de los puntos que cumplen con  $f(x) \leq \alpha$  son conjuntos convexas, para cualquier  $\alpha$ . Existen muchas funciones que satisfacen la propiedad sin ser convexas. Al conjunto de funciones que satisface esta propiedad se le denomina funciones cuasi-convexas. Alternativamente, estas funciones pueden describirse del siguiente modo:

**Definición 10 (Función cuasi-convexa)** Si  $D$  es una parte convexa de  $\mathbb{R}^n$ , una función  $f : D \rightarrow \mathbb{R}$  se dice cuasi-convexa sobre  $D$  si:

$$f(z) \leq \max\{f(x), f(y)\} \quad \forall x, y \in D, \quad x \neq y, \quad \forall z \in ]x, y[.$$

Es decir, tomando cualquier par de puntos en el dominio, todos los puntos en la recta que los une tendrán una imagen no superior a las imágenes de los dos puntos. Cuando  $f(x)$  es una función cuasi-convexa, ésta posee a lo más un valle, como es el caso de la función cuasi-convexa de la Figura 1.23. Es importante notar que esta función no es convexa.

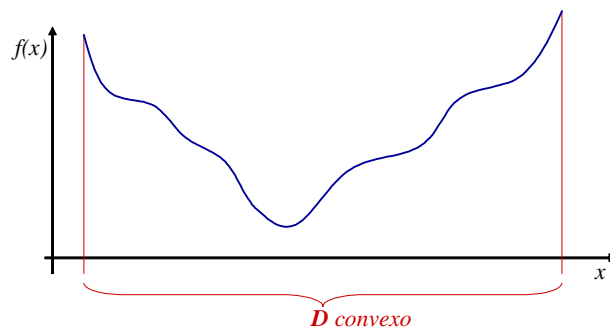


Figure 1.23: Función cuasi-convexa

Una característica importante de las funciones cuasi-convexas radica en que un mínimo local de estas funciones es siempre también su mínimo global. Como se mencionó anteriormente, otra característica importante es que para toda función  $f(x)$  cuasi-convexa se cumple que el lugar geométrico

de los puntos tal que satisfacen  $f(x) \leq a$  conforman necesariamente un conjunto convexo. Así, un conjunto definido por una serie de restricciones del tipo  $f_i(x) \leq a_i$  en que todas las funciones  $f_i(x)$  son cuasi-convexas es un conjunto convexo.

**Definición 11** (*problema convexo*) El problema de minimización

$$\begin{aligned} P) \quad & \text{Min } f(x) \\ & x \in D. \end{aligned}$$

se dice convexo, si  $D$  es una parte convexa de  $\mathbb{R}^n$  y  $f(x)$  es convexo sobre  $D$ .

Es importante notar que basta una variable discreta que incida en el conjunto  $D$  para que éste no sea convexo. Por ende los problemas con variables discretas típicamente no son convexos.

**Proposición 10** Si  $P)$  es convexo, todo punto mínimo local del problema es también su óptimo global.

**Demostración.** Sea  $\bar{x}$  punto mínimo local de  $P)$ . Entonces definamos un conjunto vecindad de  $\bar{x}$  que denominaremos  $B(\bar{x}, R)$  en que todos los puntos del conjunto están a una distancia menor a  $R$  de  $\bar{x}$  y tal que  $f(\bar{x}) \leq f(x) \forall x \in D \cap B(\bar{x}, R)$ ; de otro modo  $\bar{x}$  no sería mínimo local. Supongamos que  $\bar{x}$  no es mínimo global. En ese caso existiría un punto  $\bar{z} \in D$  ( $\bar{z} \neq \bar{x}$ ) tal que  $f(\bar{z}) < f(\bar{x})$ . Definimos ahora el conjunto de puntos en la recta entre  $\bar{x}$  y  $\bar{z}$  como:  $y(\lambda) = (1 - \lambda)\bar{x} + \lambda\bar{z} \in D \quad \forall \lambda \in [0, 1]$ . Todos estos puntos pertenecen a  $D$  pues  $\bar{x} \in D$ ,  $\bar{z} \in D$  y  $D$  es convexo.

Entonces

i) debe existir un  $\lambda$  suficientemente pequeño tal que  $y(\lambda) \in B(\bar{x}, R)$ , y por lo tanto,  $f(y(\lambda)) > f(\bar{x})$

Sin embargo,

ii)  $f(x)$  es convexa  $\Rightarrow f(y(\lambda)) \leq (1 - \lambda)f(\bar{x}) + \lambda f(\bar{z})$ , y como  $f(\bar{z}) < f(\bar{x})$ , esto implicaría que  $f(y(\lambda)) \leq f(\bar{x})$  lo que contradice el punto i) anterior.

Así, se demuestra que  $\bar{x}$  debe ser un punto mínimo global de  $P)$ . ■

Esta propiedad de los problemas convexos es altamente deseable, pues permite focalizar el objetivo en encontrar un punto localmente óptimo. Esto se traduce en que dado un punto candidato a óptimo global del problema, no importa cuan grande sea el dominio, nos basta compararlo con los puntos de su vecindad inmediata para determinar su optimalidad. Por último,

**Proposición 11 Corolario 12** Si  $f(x)$  es estrictamente convexa sobre  $D$  (convexo), entonces todo punto mínimo local de  $f(x)$  es también su único mínimo global.

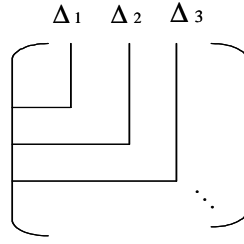


Figure 1.24: Determinantes menores de una matriz

### 1.8.3 Criterios prácticos de convexidad de funciones

Sea  $f(x) = f(x_1, x_2, \dots, x_n)$  continua y dos veces diferenciable. Su matriz Hessiana será:

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & & \\ \vdots & & \ddots & \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & & & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

la cual es una matriz simétrica. La Figura 1.24 muestra las matrices menores de esta matriz, cuyos determinantes se denominarán  $\Delta_i$ .

**Definición 12** *Caracterización de la matriz Hessiana*

1. Si todos los determinantes de las matrices menores son estrictamente positivos ( $\Delta_i > 0$ )  $\Rightarrow H$  se dirá definida positiva. Esto significa que para cualquier vector  $d$  en  $\mathbb{R}^n$  se cumple que  $d^T H d > 0$ .
2. Si todos los determinantes de las matrices menores son no negativos ( $\Delta_i \geq 0$ )  $\Rightarrow H$  se dirá semidefinida positiva. Esto significa que para cualquier vector  $d$  en  $\mathbb{R}^n$  se cumple que  $d^T H d \geq 0$ .
3. Si los determinantes de las matrices menores impares son estrictamente negativos y los de las matrices menores pares estrictamente positivos ( $\Delta_1 < 0, \Delta_2 > 0, \Delta_3 < 0, \dots$ )  $\Rightarrow H$  es definida negativa. Esto significa que para cualquier vector  $d$  en  $\mathbb{R}^n$  se cumple que  $d^T H d < 0$ .
4. Si los determinantes de las matrices menores impares son no positivos y los de las matrices menores pares no negativos ( $\Delta_1 \leq 0, \Delta_2 \geq 0, \Delta_3 \leq 0, \dots$ )  $\Rightarrow H$  es semidefinida negativa. Esto significa que para cualquier vector  $d$  en  $\mathbb{R}^n$  se cumple que  $d^T H d \leq 0$ .

Sea  $f()$  dos veces diferenciable sobre  $D$ , con  $D \in \mathbb{R}^n$  y convexo. Entonces  $f()$  es convexa sobre  $D$  si y solo si  $D^2 f(x) = H$  es semidefinida positiva  $\forall x \in D$ . Es decir,  $d^T D^2 f(x) d \geq 0, \forall d \in \mathbb{R}^n, \forall x \in D$ .

**Ejemplo:** Sea  $f(x_1, x_2) = 2x_1 - 3x_2 + x_1^4 + x_1x_2 + x_2^2$ . Las derivadas parciales son:

$$\begin{aligned}\frac{\partial f}{\partial x_1} &= 2 + 4x_1^3 + x_2 \\ \frac{\partial f}{\partial x_2} &= -3 + x_1 + 2x_2\end{aligned}$$

y las segundas derivadas serían:

$$\frac{\partial^2 f}{\partial x_1^2} = 12x_1^2 \quad \frac{\partial^2 f}{\partial x_1 \partial x_2} = 1 \quad \frac{\partial^2 f}{\partial x_2 \partial x_1} = 1 \quad \frac{\partial^2 f}{\partial x_2^2} = 2.$$

De esta forma, la matriz Hessiana es:

$$H = \begin{bmatrix} 12x_1^2 & 1 \\ 1 & 2 \end{bmatrix}.$$

Los determinantes menores son  $\Delta_1 = 12x_1^2$ , no negativo para todo punto del dominio, y  $\Delta_2 = 24x_1^2 - 1$ . Pero, ¿ $\Delta_2 \geq 0$ ? Sólo si  $x_1^2 \geq \frac{1}{24}$ , es decir, si  $|x_1| \geq \sqrt{\frac{1}{24}}$ .

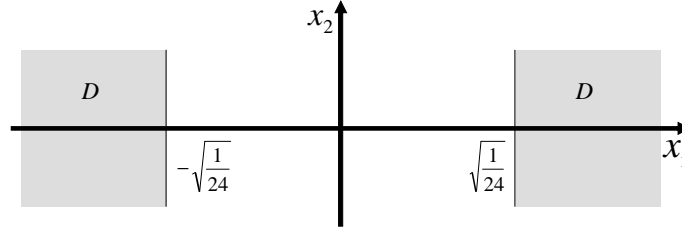


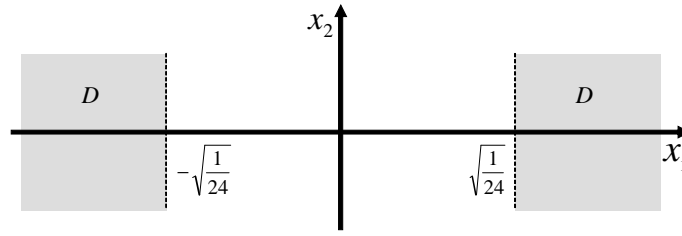
Figure 1.25: Dominio de  $f(x_1, x_2)$

Es importante destacar que  $f(x)$  no es convexa sobre toda la región achurada en la Figura 1.25, ya que esta región no es convexa. Asimismo, basta que la función se defina sobre un dominio convexo que incluya alguna región no achurada (algún punto en  $[-\sqrt{\frac{1}{24}}, \sqrt{\frac{1}{24}}]$  para que la función no sea convexa pues el Hessiano no sería semidefinido positivo).

**Proposición 13** Si  $D$  es una parte convexa de  $\mathbb{R}^n$  y  $D^2f(x)$  es definida positiva, es decir,  $d^T D^2f(x)d > 0$ ,  $\forall d \in \mathbb{R}^n, d \neq 0$ , entonces  $f(x)$  es estrictamente convexa sobre  $D$ .

En el ejemplo anterior,  $f()$  es estrictamente convexa sobre toda región convexa contenida en la región  $x_1 < \frac{-1}{\sqrt{24}}$  o en la región  $x_1 > \frac{1}{\sqrt{24}}$ ;  $x_2 \in \mathbb{R}$ . Estas regiones excluyen explícitamente a los puntos en que  $x_1 = \frac{-1}{\sqrt{24}}$  ó  $x_1 = \frac{1}{\sqrt{24}}$  (ver Figura 1.26).

Es importante destacar la causalidad de esta proposición: si el Hessiano es definido positivo, entonces la función es estrictamente convexa, pero no al revés. Es decir, si  $f(x)$  es estrictamente convexa y dos veces diferenciable en una región, no implica que  $D^2f(x)$  sea definida positiva en la región. Por ejemplo, si  $f(x) = x^4$  que es estrictamente convexa,  $D^2f(x) = 12x^2$  no es definida positiva en  $x = 0$ .

Figure 1.26: Dominio no convexo de  $f(x_1, x_2)$ 

**Ejemplo:** Considere el siguiente problema de optimización:

$$\begin{aligned} P) \quad & \min x^2 + y^2 + z^2 - 6000(x + y + z) \\ s.a \quad & x + 2y + 4z \leq 4000 \\ & x, y, z \geq 0 \end{aligned}$$

Tenemos un problema con función objetivo cuadrática, estrictamente convexa. Veamos su matriz Hessiana para comprobar esto.

$$D^2 f(x, y, z) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = H$$

Por lo tanto,  $H$  es definida positiva (constante), y como  $D$  es un dominio acotado, cerrado y no vacío, entonces por el Teorema de Existencia tenemos que  $P$  admite solución óptima. Como observamos que  $H$  es definida positiva en cualquier punto, decimos que  $f()$  es estrictamente convexa. Como además el dominio es convexo, decimos que el problema admite solución óptima única.

**Ejemplo:** En una fábrica de chocolates necesitan determinar el precio de venta óptimo para su producto de modo de maximizar la utilidad obtenida. El precio no puede superar el precio de 60¢ de un producto de la competencia. El costo de producción de la barra es de 8,5¢, y la demanda se representa mediante la función  $f(P) = \frac{1000}{P^2}$ , donde  $P$  es el precio de venta expresado en centavos. De esta forma, la formulación del modelo sería:

$$\begin{aligned} P) \quad & \text{Max}(P - 8,5) \left( \frac{1.000}{P^2} \right) \\ s.a \quad & 8,5 \leq P \leq 60 \end{aligned}$$

¿Existe un precio óptimo? Sí, pues  $f(P)$  es continua sobre el intervalo  $[8,5; 60]$ , cerrado, acotado y no vacío (notar que el dominio es convexo). La condición de primer orden es:

$$\frac{df(P)}{dP} = \frac{1}{P^3} (17.000 - 1.000P) = 0$$

de la cual se obtiene  $P^* = 17\phi$  como un punto de derivada nula y por lo tanto posible precio óptimo. A continuación es necesario analizar la convexidad de  $f(P)$  para ver si  $P^*$  corresponde efectivamente a un máximo. Para esto se analiza la segunda derivada de  $f(P)$ ,

$$f''(P) = 2.000P^{-3} - 51.000P^{-4}. \quad (1.4)$$

¿Es  $\frac{d^2f(P)}{dP^2} \leq 0 \forall P \in [8, 5; 60]$ ? No, si se iguala (1.4) a cero se obtiene que en  $P = 25,5$  hay un cambio de signo en la segunda derivada de  $f(P)$ , donde  $f''(P) < 0$  si  $P < 25,5$  y  $f''(P) > 0$  si  $P > 25,5$ . Por lo tanto,  $P$  es no convexo globalmente (ver Figura 1.27). Por lo anterior se deduce que en  $P^* = 17\phi$  la función es cóncava localmente, por lo cual este precio corresponde a una solución óptima local. Para garantizar que la solución es óptima global se puede argumentar que la función es estrictamente decreciente para valores superiores a  $P = 17\phi$  por lo que este precio maximiza la utilidad ( $\hat{v} = 29,4\phi$ ).

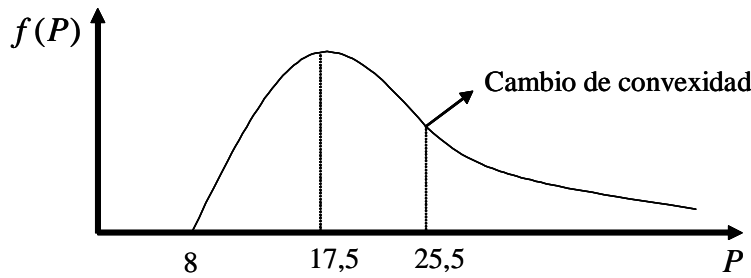


Figure 1.27: Convexidad de la función de utilidad  $f(P)$

## 1.9 Problemas resueltos

**Problema 14 (Equivalencia)** Considere el siguiente problema y escriba uno equivalente que sea diferenciable en el dominio.

$$\begin{aligned} P) \quad & \text{Max} \quad \ln(|x_1 + x_2 + 1|)^{-1} \\ & \text{s.a.} \quad x_1 + 2x_2 \leq 5 \\ & \quad \quad x_1, x_2 \geq 0 \end{aligned}$$

**Solución:**

Aplicando la regla I de Problemas equivalentes (y reglas básicas de logaritmos) transformamos  $P$  en un problema de minimización.

$$\begin{aligned} \hat{P}) \quad & \text{Min} \quad \ln(|x_1 + x_2 + 1|) \\ & x \in D \end{aligned}$$

Utilizando la equivalencia V aplicamos una función  $g(y) = e^y$  que es creciente  $\forall x$ , para eliminar el  $\ln$ . Luego:

$$\begin{aligned} & \hat{P}) \text{ Min } |x_1 + x_2 + 1| \\ & x \in D \end{aligned}$$

Y aplicando la regla II llegamos a

$$\begin{aligned} & \hat{P}) \text{ Min } \mu \\ & \text{s.a.} \\ & x_1 + x_2 + 1 \leq \mu \\ & -x_1 - x_2 - 1 \leq \mu \\ & x_1 + 2x_2 \leq 5 \\ & x_1, x_2 \geq 0 \end{aligned}$$

**Problema 15** Se quiere instalar una antena de telecomunicaciones, cuya cobertura incluya a cinco ciudades aledañas de ubicaciones  $(x_i, y_i, z_i)$ . Para lograrlo se debe buscar una posición en el que la intensidad de la señal sea lo mayor posible para las ciudades, sin salir de los ciertos límites dados  $((x, y, z) \in D)$ . Además se sabe que la intensidad en cierta ciudad es proporcional a la potencia de la antena ( $k$ ), e inversamente proporcional a la distancia desde la antena hasta la ciudad. ¿En qué ubicación debe instalarse la antena? Modele un problema de optimización que sea diferenciable en el dominio.

### Solución:

La intensidad de la antena en la ciudad de posición  $\vec{x}_i$  será

$$I_i(\vec{x}) \propto \frac{k}{|\vec{x} - \vec{x}_i|}, \quad \text{es decir,} \quad I_i(\vec{x}) \propto \frac{k}{\sqrt{(z - z_i)^2 + (x - x_i)^2 + (y - y_i)^2}}$$

Así la intensidad de la señal que recibe la ciudad con menor cobertura será

$$I(\vec{x}) = \min_{i=1 \dots 5} I_i(\vec{x})$$

Así el modelo de optimización será

$$\text{Max}_{\vec{x} \in D} I(\vec{x}) = \text{Max}_{\vec{x} \in D} \left[ \min_{i=1 \dots 5} \frac{k}{\sqrt{(z - z_i)^2 + (x - x_i)^2 + (y - y_i)^2}} \right]$$

que se puede reescribir como

$$\text{Max}_{\vec{x} \in D} I(\vec{x}) = \text{Min}_{\vec{x} \in D} \left[ \max_{i=1 \dots 5} \frac{-k}{\sqrt{(z - z_i)^2 + (x - x_i)^2 + (y - y_i)^2}} \right]$$

Esto es equivalente al problema

$$\begin{aligned} & \text{Min} \quad \mu \\ \text{s.a.} \quad & \frac{-k}{\sqrt{(z - z_i)^2 + (x - x_i)^2 + (y - y_i)^2}} \leq \mu, \quad \forall i = 1 \dots 5 \\ & \vec{x} \in D, \quad \mu \in \mathbb{R} \end{aligned}$$

**Problema 16** Considere los siguientes problemas de optimización, y escriba uno equivalente que sea diferenciable en el dominio.

$$\begin{aligned} (a) \quad & \text{Max} \quad -[7x^3 + |x^2 + y| - \min\{x^2, -|x + y|\}] \\ \text{s.a.} \quad & x^3 + 8y^2 \leq 250, \quad x \geq 0 \end{aligned}$$

$$(b) \quad \text{Max}_{x \geq 0} \min\{-|x^2 - 7x - 10|, \min\{x^3 - 4x, x^4 - 7\}\}$$

**Solución:**

(a)

$$\begin{aligned} & \text{Max} \quad -[7x^3 + |x^2 + y| - \min\{x^2, -|x + y|\}] \tag{1.5} \\ \Leftrightarrow & \text{Min} \quad 7x^3 + |x^2 + y| - \min\{x^2, -|x + y|\} \\ \text{s.a.} \quad & (x, y) \in D, \text{ con } D = \{(x, y) / x^3 + 8y^2 \leq 250, x \geq 0\} \\ \Leftrightarrow & \text{Min} \quad 7x^3 + |x^2 + y| + \max\{-x^2, |x + y|\} \\ \text{s.a.} \quad & (x, y) \in D \\ \Leftrightarrow & \text{Min} \quad \mu_1 + \mu_2 + \mu_3 \\ \text{s.a.} \quad & 7x^3 \leq \mu_1, \quad |x^2 + y| \leq \mu_2, \\ & \max\{-x^2, |x + y|\} \leq \mu_3, \quad (x, y) \in D, \quad \mu_i \in \mathbb{R} \quad \forall i \\ \Leftrightarrow & \text{Min} \quad \mu_1 + \mu_2 + \mu_3 \\ \text{s.a.} \quad & 7x^3 \leq \mu_1, \quad x^2 + y \leq \mu_2, \quad -x^2 - y \leq \mu_2, \\ & -x^2 \leq \mu_3, \quad |x + y| \leq \mu_3, \quad (x, y) \in D, \quad \mu_i \in \mathbb{R} \quad \forall i \\ \Leftrightarrow & \text{Min} \quad \mu_1 + \mu_2 + \mu_3 \\ \text{s.a.} \quad & 7x^3 \leq \mu_1, \quad x^2 + y \leq \mu_2, \quad -x^2 - y \leq \mu_2, \\ & -x^2 \leq \mu_3, \quad x + y \leq \mu_3, \quad -(x + y) \leq \mu_3, \\ & (x, y) \in D, \quad \mu_i \in \mathbb{R} \quad \forall i \end{aligned}$$



(b)

$$\begin{aligned}
& \text{Max}_{x \geq 0} \min \{ -|x^2 - 7x - 10|, \min \{ x^3 - 4x, x^4 - 7 \} \} \\
\Leftrightarrow & \text{Min}_{x \geq 0} - \min \{ -|x^2 - 7x - 10|, \min \{ x^3 - 4x, x^4 - 7 \} \} \\
\Leftrightarrow & \text{Min}_{x \geq 0} \max \{ |x^2 - 7x - 10|, -\min \{ x^3 - 4x, x^4 - 7 \} \} \\
\Leftrightarrow & \text{Min}_{x \geq 0} \max \{ |x^2 - 7x - 10|, \max \{ -x^3 + 4x, -x^4 + 7 \} \} \\
\Leftrightarrow & \text{Min } \mu \\
& \text{s.a.} \quad |x^2 - 7x - 10| \leq \mu, \quad -x^3 + 4x \leq \mu, \\
& \quad \quad -x^4 + 7 \leq \mu, \quad x \geq 0, \quad \mu \in \mathbb{R} \\
\Leftrightarrow & \text{Min } \mu \\
& \text{s.a.} \quad x^2 - 7x - 10 \leq \mu, \quad -(x^2 - 7x - 10) \leq \mu, \\
& \quad \quad -x^3 + 4x \leq \mu, \quad -x^4 + 7 \leq \mu, \quad x \geq 0, \quad \mu \in \mathbb{R}
\end{aligned}$$

**Problema 17** Considere el siguiente problema de optimización, y escriba uno equivalente que sea diferenciable en el dominio.

$$\begin{aligned}
& \text{Max} \quad -\frac{1}{2} \ln [7x^3 + |x^2 + y| - \min \{ x^2, -|x + y| \}] \\
& \text{s.a.} \quad x^3 + 8y^2 \leq 250, x \geq 0
\end{aligned}$$

**Solución:**

$$\begin{aligned}
\Leftrightarrow & \text{Min } \ln [7x^3 + |x^2 + y| - \min \{ x^2, -|x + y| \}]^{1/2} \\
& \text{s.a.} \quad (x, y) \in D, \text{ con } D = \{ (x, y) / x^3 + 8y^2 \leq 250, x \geq 0 \}
\end{aligned}$$

Utilizando la función estrictamente creciente  $e^x$  se obtiene

$$\begin{aligned}
\Leftrightarrow & \text{Min } [7x^3 + |x^2 + y| + \max \{ -x^2, |x + y| \}]^{1/2} \\
& \text{s.a.} \quad (x, y) \in D.
\end{aligned}$$

Luego utilizando la función estrictamente creciente en  $x \geq 0$ ,  $x^2$  se obtiene

$$\begin{aligned}
\Leftrightarrow & \text{Min } [7x^3 + |x^2 + y| + \max \{ -x^2, |x + y| \}] \\
& \text{s.a.} \quad (x, y) \in D.
\end{aligned}$$

Este problema es el mismo que se describe en la Ecuación 1.5, y puede ser resuelto del modo ahí expuesto. Otras funciones estrictamente crecientes de interés:  $\ln x, -1/x, -e^{-x}$ , etc.

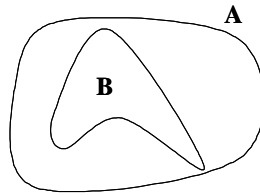


Figure 1.28:

**Problema 18 (Convexidad (I21°03))** *Responda:*

- (a) ¿Qué condiciones debe cumplir un problema para ser denominado convexo?
- (b) ¿Por qué interesa verificar si un problema es convexo?
- (c) ¿Si una función definida sobre un conjunto  $A$  convexo, es convexa, entonces definida sobre un conjunto  $B \subset A$  también lo es?

**Solución:**

- (a) Debe tener una función objetivo convexa sobre un dominio convexo.
- (b) Porque esto garantiza que un óptimo local sea global.
- (c) No, ya que  $B$  debe ser convexa también. Por ejemplo, como vemos en la figura 1.28,  $B \subset A$ , pero el conjunto  $A$  es convexo y  $B$  no lo es.

**Problema 19 (Convexidad)** *Comente acerca de la convexidad de la función  $f(x, y) = 2x^2 + 3y^2 - y^3$ .*

**Solución:** Para poder entender la convexidad de esta función lo primero que hacemos es identificar su Hessiano:

$$\begin{aligned}\frac{\partial f}{\partial x} &= 4x \\ \frac{\partial f}{\partial y} &= 6y - 3y^2\end{aligned}$$

entonces

$$H = \begin{bmatrix} 4 & 0 \\ 0 & 6 - 6y \end{bmatrix}$$

con lo que  $\Delta_1 = 4$ ,  $\Delta_2 = 24 - 24y$ , por lo que el Hessiano es semi definido positivo  $\forall x \in \mathbb{R}, y \leq 1$  y por lo tanto la función es convexa en esta región. Al excluir el borde  $y = 1$  ( $\forall x \in \mathbb{R}, y < 1$ ) el Hessiano se vuelve definido positivo y la función, estrictamente convexa.

**Problema 20 (Caracterización de soluciones)** *Encuentre y caracterice la(s) solución(es) óp-*

tima(s) del siguiente problema:

$$\begin{aligned} P) \text{ Min } f(x, y, z) &= x^2 + y^2 + z^2 + xz + yz \\ (x, y, z) &\in \mathbb{R}^3 \end{aligned}$$

**Solución:**

$$\begin{aligned} \frac{\partial f}{\partial x} &= 2x + z = 0 \\ \frac{\partial f}{\partial y} &= 2y + z = 0 \\ \frac{\partial f}{\partial z} &= 2z + x + y = 0 \end{aligned}$$

Por lo tanto hay sólo un punto extremo:  $\mathbf{x}^* = (0, 0, 0)$

Veamos el Hessiano:

$$H = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

$\Delta_1 = 2 > 0$ ,  $\Delta_2 = 4 > 0$ ,  $\Delta_3 = 4 > 0 \implies H$  es positivo definido. Por lo tanto,  $f(x, y, z)$  es estrictamente convexa en todo el dominio, por lo tanto  $\mathbf{x}^* = (0, 0, 0)$  es un mínimo global estricto.

**Problema 21** Considere la función

$$f(x, y) = 2x^3 + y + 20.$$

- a) ¿Existe alguna región en que  $f(x, y)$  sea definida positiva?
- b) ¿Existe alguna región en que  $f(x, y)$  sea semidefinida positiva?
- c) ¿Existe alguna región en que  $f(x, y)$  sea semidefinida negativa?

**Solución:**

En primer lugar se ve que

$$\frac{\partial f}{\partial x} = 6x^2 \quad \frac{\partial f}{\partial y} = 1 \quad ; \quad H = \begin{bmatrix} 12x & 0 \\ 0 & 0 \end{bmatrix}$$

a) Se necesita  $\Delta_i > 0$ . En este caso  $\Delta_1 = 12x > 0 \implies x > 0$ , pero  $\Delta_2 = 0$  por lo que nunca podrá ser definida positiva.

b) Se necesita  $\Delta_i \geq 0$ . Esto se cumple  $\forall x \geq 0$ .

c) Se necesita  $\Delta_{2i-1} \leq 0$  y  $\Delta_{2i} \geq 0$ . Esto se cumple  $\forall x \leq 0$ .

**Problema 22** ¿Qué condiciones deben cumplir los parámetros  $a$  y  $b$  para que un punto de la función

$$f(x, y) = x^2 + axy + by^2 + x + y$$

cumpla con las condiciones necesarias, pero no suficientes para ser mínimo?

**Solución:**

i) Condición necesaria de 1er orden:

$$\begin{aligned}\frac{\partial f}{\partial x} &= 2x + ay + 1 \equiv 0 \\ \frac{\partial f}{\partial y} &= ax + 2by + 1 \equiv 0\end{aligned}$$

ii) Condición necesaria de 2do orden:

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 2 & a \\ a & 2b \end{bmatrix}$$

Para que se cumpla la segunda condición necesaria, se necesita que  $H$  sea semidefinida positiva ( $\Delta_i \geq 0$ ). En este caso  $\Delta_1 = 2 > 0$  y se necesita  $\Delta_2 = 4b - a^2 \geq 0$ .

iii) Condición suficiente:  $H$  definida positiva ( $\Delta_i > 0$ ). En este caso  $\Delta_1 = 2 > 0$  y se necesitaría  $\Delta_2 = 4b - a^2 > 0$ .

$\Rightarrow$  Para que se cumpla ii) pero no iii) se necesita:  $4b - a^2 = 0 \Rightarrow 4b = a^2$

**Problema 23** Responda las siguientes preguntas:

a) Sean  $f(x)$  y  $g(x)$  funciones convexas. ¿Es el siguiente problema de optimización convexo?

$$\begin{aligned} \text{Min} \quad & f(x) \\ \text{s.a.} \quad & g(x) \geq 0 \end{aligned}$$

b) Demuestre que si el dominio de restricciones está definido como

$$D = \{x \in \mathbb{R}^n : g_i \leq 0, \forall i = 1 \dots m\}$$

donde  $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$  son funciones convexas, entonces  $D$  es una parte convexa de  $\mathbb{R}^n$ .

c) Comente acerca de la convexidad de la función

$$f(x, y) = 2x^2 + 3y^2 - y^3.$$

d) ¿Qué condiciones debe cumplir un problema para ser denominado convexo?

e) Si una función definida sobre un conjunto  $A$  convexo, es convexa, entonces definida sobre un conjunto  $B \subseteq A$  también lo es?

f) Encuentre y caracterice las soluciones óptimas del siguiente problema

$$\begin{aligned} \text{Min} \quad & f(x, y, z) = x^2 + y^2 + z^2 + xz + yz \\ \text{s.a.} \quad & (x, y, z) \in \mathbb{R}^3 \end{aligned}$$

**Solución:**

a) Para que el problema sea convexo,  $f(x)$  debe ser convexa y las restricciones deben formar un dominio convexo. En este caso, pese a  $g(x)$  ser convexa, el dominio no necesariamente lo es. Ver Figura 1.29. Así el problema de optimización no necesariamente es convexo.

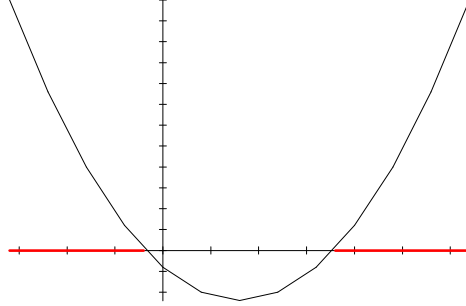


Figure 1.29: Ejemplo de dominio no convexo

b) Si  $g_i(x)$  es una función convexa, entonces el conjunto  $g_i(x) \leq 0$  es convexo. Además se sabe que la intersección de conjuntos convexos, es un conjunto convexo  $\implies D$  es convexo.

c) En primer lugar se ve que

$$\frac{\partial f}{\partial x} = 4x \quad \frac{\partial f}{\partial y} = 6y - 3y^2 \quad ; \quad H = \begin{bmatrix} 4 & 0 \\ 0 & 6 - 6y \end{bmatrix}$$

En  $y = 1$ , el hessiano es semidefinido positivo, por lo que la función es convexa en la región  $y \leq 1$ . Podemos agregar que en la región  $y < 1$ , la función es estrictamente convexa.

d) Para que el problema sea convexo,  $f(x)$  debe ser convexa y estar definida sobre un dominio  $D$  convexo. Para que la función objetivo sea convexa se requiere que

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2) \quad \forall x_1, x_2 \in D; \lambda \in [0, 1]$$

e) El conjunto  $B \subseteq A$  no necesariamente es convexo, por lo que la función definida en el dominio  $B$  no necesariamente sería convexa.

f) En primer lugar buscamos los puntos candidatos a solución óptima utilizando las primeras derivadas, donde se obtiene

$$\frac{\partial f}{\partial x} = 2x + z \equiv 0, \quad \frac{\partial f}{\partial y} = 2y + z \equiv 0, \quad \frac{\partial f}{\partial z} = 2z + x + y \equiv 0$$

Despejando este sistema, se encuentra el punto candidato  $(x, y, z) = (0, 0, 0)$ . Luego se estudia el hessiano

$$H = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \Rightarrow \Delta_1 = 2 > 0, \Delta_2 = 4 > 0, \Delta_3 = 4 > 0 \quad \forall x, y, z$$

Así  $H$  es definida positiva en todo el dominio, por lo que la función objetivo es estrictamente convexa. Esto implica que el punto  $(0,0,0)$  es un mínimo global.

**Problema 24** Encuentre y caracterice las soluciones óptimas del siguiente problema:

$$f(x, y) = 2x^3 + 3x^2 - 12x + 2y^2 - 8y$$

**Solución:**

Las derivadas parciales son

$$\frac{\partial f}{\partial x} = 6x^2 + 6x - 12 \equiv 0, \quad \frac{\partial f}{\partial y} = 4y - 8 \equiv 0$$

Resolviendo este sistema, se obtienen los puntos candidatos  $(1, 2)$  y  $(-2, 2)$ . Para caracterizar estos puntos se estudia el hessiano, dado por

$$H = \begin{bmatrix} 12x + 6 & 0 \\ 0 & 4 \end{bmatrix}$$

Así el primer punto será un mínimo local estricto, y el segundo no es caracterizable. Al observar el hessiano se puede apreciar que la función es convexa en la región  $x \geq -1/2$ .

**Problema 25** Encuentre y caracterice las soluciones óptimas de la función

$$f(x, y) = \sin x + \sin y + \sin(x + y)$$

**Solución:**

En primer lugar se buscan los puntos críticos por medio de las primeras derivadas, que son

$$\frac{\partial f}{\partial x} = \cos x + \cos(x + y) \equiv 0, \quad \frac{\partial f}{\partial y} = \cos y + \cos(x + y) \equiv 0$$

Se ve que los puntos críticos son tales que  $\cos x = \cos y$ , es decir donde  $x = y + 2n\pi$ ,  $n=0,1,2,\dots$ . Despejando para  $x$ , se obtiene  $2\cos^2 x + \cos x - 1 = 0$ , considerando que  $\cos 2x = 2\cos^2 x - 1$  y que  $\cos(x - 2n\pi) = \cos x$ . Así los puntos críticos serán las soluciones de esta ecuación cuadrática, y serán  $\cos x = -1$  y  $\cos x = 1/2$ . Es decir,  $x = \pi, \pi/3, 5\pi/3$  (Y sus respectivos períodos  $2m\pi$ ,  $m=0,1,2,\dots$ ). Así los puntos críticos serán

$$(\pi + 2m\pi, \pi + 2n\pi), (\pi/3 + 2m\pi, \pi/3 + 2n\pi), (5\pi/3 + 2m\pi, 5\pi/3 + 2n\pi)$$

Para determinar la naturaleza de estos puntos, se calcula el hessiano resultando

$$H = \begin{bmatrix} -\sin x - \sin(x + y) & -\sin(x + y) \\ -\sin(x + y) & -\sin y - \sin(x + y) \end{bmatrix}$$

Luego éste se evalúa en los puntos críticos:

i)

$$H = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Se podría decir que el hessiano es semidefinido positivo y/o negativo. Así, no se puede caracterizar este punto.

ii)

$$H = \begin{bmatrix} -\sqrt{3} & \frac{-\sqrt{3}}{2} \\ \frac{-\sqrt{3}}{2} & -\sqrt{3} \end{bmatrix} \Rightarrow \Delta_1 = -\sqrt{3} < 0, \Delta_2 = 3 - 3/4 = 9/4 > 0$$

Así en el segundo punto crítico, el hessiano es definido negativo, por lo que el punto corresponde a un máximo local estricto,  $\forall n, m$ .

iii)

$$H = \begin{bmatrix} \sqrt{3} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \sqrt{3} \end{bmatrix} \Rightarrow \Delta_1 = \sqrt{3} > 0, \Delta_2 = 3 - 3/4 = 9/4 > 0$$

Así en el tercer punto crítico, el hessiano es definido positivo, por lo que el punto corresponde a un mínimo local estricto,  $\forall n, m$ .





## Chapter 2

# Programación No lineal

En este capítulo se presentan técnicas para identificar soluciones óptimas a problemas de variables continuas con función objetivo o restricciones no lineales. Inicialmente se abordan problemas sin restricciones, luego se analizan problemas unidimensionales con restricciones, luego problemas multidimensionales con restricciones de igualdad para finalmente analizar problemas con restricciones generales. Las condiciones de optimalidad encontradas durante este capítulo serán luego aplicadas a problemas lineales en el capítulo siguiente.

### 2.1 Optimización de una función sin restricciones

Considere el problema de minimización

$$P) \quad \min_{x \in \mathbb{R}^n} f(x)$$

donde  $f(x)$  es una función diferenciable.

Un punto  $\bar{x} = (x_1, x_2, \dots, x_n)$  es un mínimo local si  $f(\bar{x}) \leq f(\bar{x} + h) \forall h = (h_1, h_2, \dots, h_n)$  tal que  $|h_j|$  es suficientemente pequeño para todo  $j$ . Es decir,  $\bar{x}$  es un mínimo local si el valor de  $f(x)$  en cada punto del entorno o vecindad de  $\bar{x}$  no es menor a  $f(\bar{x})$ .

#### 2.1.1 Condiciones Necesarias y suficientes para extremos

En esta sección se presentan las condiciones necesarias y suficientes que debe cumplir un punto para ser considerado punto extremo del problema. Es importante recordar lo que se entiende por condiciones necesarias y suficientes. Condiciones necesarias son aquéllas que todo punto extremo debe cumplir. Es decir una vez definidas estas condiciones, no existe un punto extremo al problema que no cumpla con estas condiciones. Condiciones suficientes son aquéllas que si un punto las cumple entonces el punto constituye sin lugar a dudas un punto extremo.

Adicionalmente distinguiremos entre las condiciones de primer y segundo orden. Las condiciones de primer orden son aquéllas que se pueden desprender a partir de un análisis de primeras derivadas

de las funciones involucradas en el problema. Las condiciones de segundo orden son aquéllas que se pueden desprender a partir de un análisis de primeras y segundas derivadas de las funciones involucradas en el problema.

**Teorema 26** (Condición necesaria de 1<sup>er</sup> orden). Si  $\bar{x} \in \mathbb{R}^n$  es un punto mínimo local de  $f(x)$ , entonces debe cumplirse que  $\nabla f(\bar{x}) = \vec{0}$ , es decir,

$$\frac{\partial f(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)}{\partial x_i} = 0 \quad \forall i = 1, \dots, n.$$

Esta condición es necesaria, pero no suficiente para un mínimo: también la satisfacen otros puntos extremos como máximos y puntos de inflexión. Así, para que  $\bar{x}$  sea un punto mínimo, debe cumplirse una segunda condición necesaria.

**Teorema 27** (Condición necesaria de 2<sup>do</sup> orden) Si  $\bar{x} \in \mathbb{R}^n$  es un punto mínimo local de  $f(x)$ , y  $f(x)$  es dos veces diferenciable, entonces debe cumplirse que  $\nabla f(\bar{x}) = \vec{0}$  (1<sup>er</sup> orden) y que la matriz Hessiana  $D^2 f(\bar{x})$  sea semidefinida positiva (2<sup>do</sup> orden).

**Teorema 28** (Condición suficiente de 2<sup>do</sup> orden) Si  $\bar{x}$  verifica:  $\nabla f(\bar{x}) = \vec{0}$  y  $D^2 f(\bar{x})$  es definida positiva, entonces  $\bar{x}$  es un punto mínimo local estricto de  $f()$ .

Estas condiciones se pueden desprender de la expansión de Taylor de la función objetivo en torno al punto  $\bar{x}$  de la siguiente forma:

$$f(x) = f(\bar{x}) + (x - \bar{x})\nabla f(\bar{x}) + \frac{(x - \bar{x})D^2 f(\bar{x})(x - \bar{x})^T}{2} + R_2(x)$$

Nos interesa determinar las condiciones que debe satisfacer  $\bar{x}$  para ser considerado punto mínimo local. Sabemos que si analizamos puntos en la vecindad de  $\bar{x}$ , entonces  $R_2(x)$  es un orden de magnitud más pequeño que el término de segundo orden de esta expansión de Taylor por lo que podemos ignorarlo para un análisis de optimalidad local. Adicionalmente, para que un punto sea considerado mínimo, el gradiente de la función en el punto debe ser nulo. Así, para que  $f(\bar{x})$  sea menor a la función objetivo evaluada en cualquier  $x$  en la vecindad de  $\bar{x}$ , debe suceder que

$$f(\bar{x}) \leq f(x) = f(\bar{x}) + \frac{(x - \bar{x})D^2 f(\bar{x})(x - \bar{x})^T}{2} + R_2(x).$$

Es decir, nos interesará que

$$\frac{(x - \bar{x})D^2 f(\bar{x})(x - \bar{x})^T}{2} = \frac{(\Delta x)D^2 f(\bar{x})(\Delta x)^T}{2} \geq 0$$

lo cual equivale a que la matriz  $D^2 f(\bar{x})$  sea semidefinida positiva (i.e.,  $\Delta x D^2 f(\bar{x}) \Delta x^T \geq 0, \forall \Delta x$ ) para que esto se cumpla y basta con que sea definida positiva (i.e.,  $\Delta x D^2 f(\bar{x}) \Delta x^T > 0, \forall \Delta x$ ) para que se cumpla estrictamente.

**Comentario:** Si la segunda derivada de  $f(x)$  es continua en  $\bar{x}$ , entonces si  $\bar{x}$  satisface la condición del teorema 28, entonces  $f(x)$  es estrictamente convexa en la vecindad de  $\bar{x}$ .

Es importante destacar que estas condiciones sólo sirven para describir un punto mínimo local. Para determinar la optimalidad global del problema es necesario comparar la función objetivo evaluada en cada uno de los puntos óptimos locales y adicionalmente verificar que se cumplan las condiciones de existencia de solución óptima para el problema.

**Ejemplo 4** Considere el problema

$$\begin{aligned} P) \quad & \min -2x_1x_2 - 2x_2 + x_1^2 + 2x_2^2 \\ & \text{s.a. } (x_1, x_2) \in \mathbb{R}^2 \end{aligned}$$

Las condiciones de primer orden serán

$$\begin{aligned} \frac{\partial f}{\partial x_1} &= -2x_2 + 2x_1 = 0 \\ \frac{\partial f}{\partial x_2} &= -2x_1 - 2 + 4x_2 = 0 \end{aligned}$$

lo que implica que  $\nabla f(x) = 0 \Leftrightarrow x_1 = 1, x_2 = 1$ .

Veamos el Hessiano de la función objetivo en  $(1, 1)$  para comprobar que este punto sea mínimo local:

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix}.$$

Como podemos ver,  $H$  es definida positiva en todo el dominio (no depende de  $x_1$  ni de  $x_2$ ), y en particular en  $(1, 1)$ , por lo tanto este punto corresponde a un punto mínimo local estricto de  $P$ ), pero como además  $f(x)$  es estrictamente convexa (el hessiano es definido positivo en todo el dominio), podemos asegurar que este punto es **mínimo global único** de  $P$ ).

**Ejemplo 5** Considere la función  $f(x, y) = x^3 + y$  (ver Figura 2.1). La matriz Hessiana está dada por

$$D^2 f(x, y) = \begin{bmatrix} 6x & 0 \\ 0 & 0 \end{bmatrix}$$

Interesa identificar los determinantes de las matrices menores para tipificar esta matriz Hessiana. En este caso observamos que  $\Delta_1 \geq 0$  y que  $\Delta_2 = 0$ .

¿Existe alguna región en que  $f(x, y)$  sea positiva definida?

No.

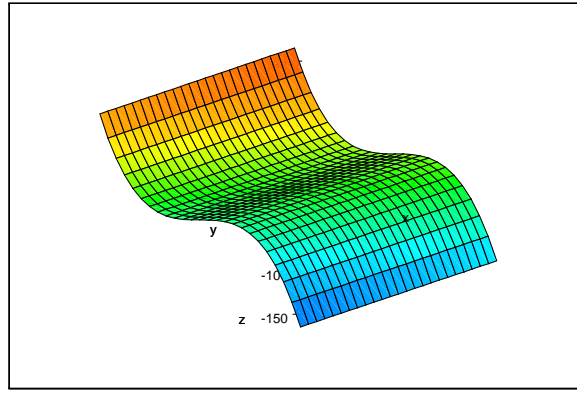


Figure 2.1: Ejemplo de convexidad

¿Existe alguna región en que  $f(x, y)$  sea semidefinida positiva?

$$\text{Sí, } \forall x \geq 0.$$

¿Existe alguna región en que  $f(x, y)$  sea semidefinida negativa?

$$\text{Sí, } \forall x \leq 0.$$

Es decir con esta información no se puede garantizar la existencia de un punto mínimo o máximo local. De hecho se aprecia en la figura que esta función no posee mínimos ni máximos locales.

**Ejemplo 6** Considere el problema

$$P) \quad \min x^4$$

En este caso,

$$\begin{aligned} f'(x) &= 4x^3 \\ f''(x) &= 12x^2 \end{aligned}$$

En  $\bar{x} = 0$ ,  $f'(\bar{x}) = 0$  y  $f''(\bar{x}) = 0$ . Por lo tanto, cumple con la condición necesaria de 2º orden. De esta información no se podría inferir nada más, pero como:  $f''(x) = 12x^2 \geq 0 \quad \forall x \Rightarrow f(x)$  es convexa, y como además es diferenciable, entonces  $\bar{x} = 0$  es un punto mínimo local de  $P$ ). En efecto, si  $f(x)$  es convexa diferenciable,  $f(x) \geq f(\bar{x}) + f'(\bar{x})(x - \bar{x}) \quad \forall x \in \mathbb{R}$ , con  $\bar{x} = 0 : f(x) \geq f(0) + f'(0)(x - 0)$ . Por lo tanto,  $f(0) \leq f(x) \quad \forall x \in \mathbb{R}$ . Así,  $\bar{x} = 0$  es un punto mínimo global de  $P$ ), de hecho es único, pues  $\forall x \neq 0 \Rightarrow x^4 > 0$ .

Algunas funciones son difíciles de optimizar por la vía de encontrar puntos de gradiente nulo (incluso cuando éstas son convexas). En estos casos, es necesario buscar métodos alternativos para identificar sus puntos extremos como los que se describen en la próxima sección.

### 2.1.2 Métodos de búsqueda de soluciones

Los métodos que se presentan aquí son métodos iterativos que comienzan de un punto factible y a través de pasos sucesivos se aproximan a un punto que satisfaga condiciones locales de optimalidad. Estos métodos no garantizan optimalidad e incluso en algunos casos tampoco garantizan convergencia. Muchas veces el desenlace del método es altamente sensible al punto de partida escogido.

Los dos métodos que se presentan en esta sección sirven para resolver problemas de optimización no restringidos. Sin embargo, pueden generalizarse para incorporar restricciones. Estos métodos se usan para minimizar funciones cuyas raíces (aquellos puntos de derivada nula) son difíciles de detectar.

#### Método de Newton

Este método opera para funciones dos veces diferenciables. A pesar que puede usarse para funciones de múltiples variables, en esta sección se presenta el método simple para funciones con sólo una variable. Así, el problema es  $\min(x)$  y por lo tanto, se asume  $f'(x)$  y  $f''(x)$  conocidas.

Sea  $x_k$  un punto factible (todos los puntos son factibles en este caso). Se puede aproximar  $f(x)$  entorno a  $x_k$ , a través de una expansión de Taylor de 2<sup>do</sup> grado. Así, la función puede aproximarse mediante la siguiente función cuadrática (ver Figura 2.2):

$$q(x) = f(x_k) + f'(x_k)(x - x_k) + \frac{1}{2}f''(x_k)(x - x_k)^2 \quad (2.1)$$

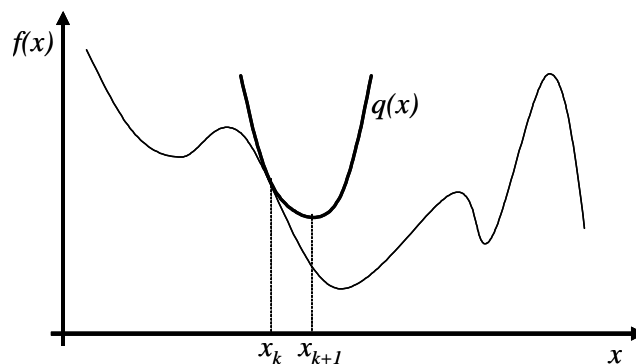


Figure 2.2: Método de Newton

La función  $q(x)$  es una buena aproximación de segundo grado para  $f(x)$  ya que:

1.  $q(x_k) = f(x_k)$
2.  $q'(x_k) = f'(x_k)$
3.  $q''(x_k) = f''(x_k)$

Si  $f''(x_k) > 0 \Rightarrow q(x)$  es convexa, y si  $f''(x_k) < 0 \Rightarrow q(x)$  es cóncava. Ahora, en vez de  $\min f(x)$ , se procede a resolver el problema  $\min q(x)$  que resulta mucho más sencillo ( $q(x)$  es una parábola). La condición de primer orden será:

$$\frac{dq}{dx} = f'(x_k) + f''(x_k)(x - x_k) = 0$$

de donde despejando nos queda la siguiente relación:

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}. \quad (2.2)$$

Este método iterativo típicamente se aproxima sucesivamente a un óptimo local. Sin embargo, es posible que tarde infinitas iteraciones en alcanzarlo. Por esto, es importante definir algún criterio para saber cuándo parar de iterar. Existen diferentes criterios de parada, los más comunes son los siguientes: parar si  $(x_k - x_{k+1}) \rightarrow 0$ , o si  $(f(x_k) - f(x_{k+1})) \rightarrow 0$ . Asimismo, se puede contemplar una combinación de estas exigencias.

El método de Newton también se puede interpretar como un método para identificar raíces de una función (en este caso, las raíces de la derivada de la función a minimizar, es decir, de  $f'(x)$ ) sólo conociendo las condiciones locales en cada punto; esto es  $f'(x_k)$  y  $f''(x_k)$ . Gráficamente, el método consiste en que en el espacio de la derivada de  $f(x)$  se trace una recta que pase por el punto  $(x_k, f'(x_k))$  y que tenga pendiente  $f''(x_k)$ , es decir

$$y - f'(x_k) = (x - x_k)f''(x_k).$$

Luego, el punto de intersección de esta recta con el eje  $x$  determinará  $x_{k+1}$  (ver Figura 2.3), es decir, igualando  $y = 0$  obtenemos el resultado en (2.2).

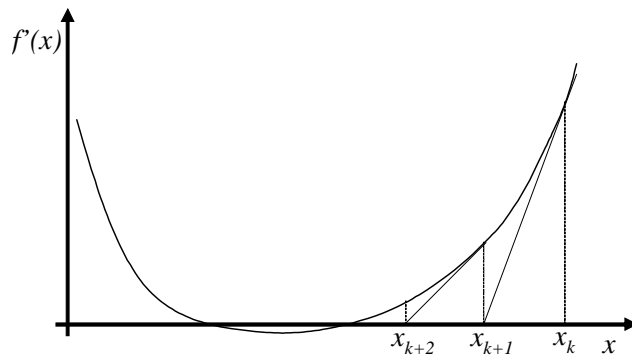


Figure 2.3: Método de Newton en el espacio de la derivada

**Observación:** Es importante notar que el método busca puntos extremos sean estos mínimos o máximos. Para distinguir si el proceso está buscando un mínimo o un máximo basta con revisar el signo de la segunda derivada de la función en cada punto. Este deberá ser positivo al buscar

mínimos y negativo al buscar máximos.

**Observación:** Este método no garantiza convergencia. Para algunas funciones diverge, mientras que otras no converge nunca. La Figura 2.4 presenta ejemplos gráficos de divergencia y de no convergencia.

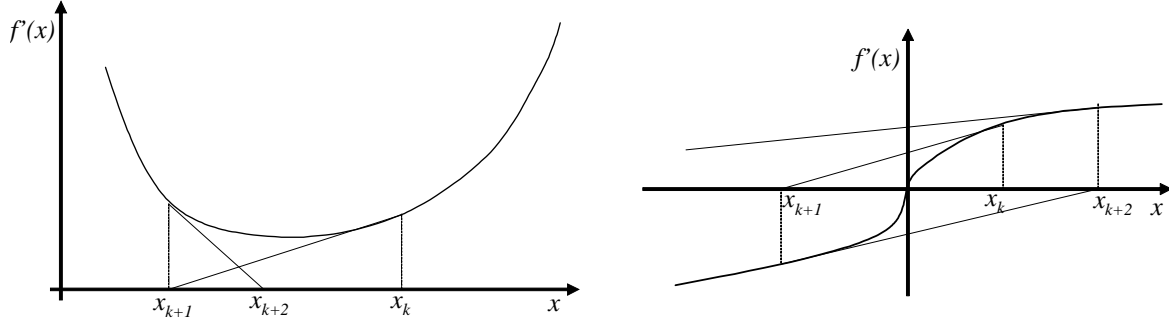


Figure 2.4: Ejemplos de divergencia y de no convergencia

**Ejemplo 7** Considere el problema  $\min 2x^3 - 21x^2 + 60x - 7$ .

Rápidamente podemos ver que la primera y segunda derivada corresponden a

$$\begin{aligned} f'(x) &= 6x^2 - 42x + 60 \\ f''(x) &= 12x - 42 \end{aligned}$$

respectivamente. Por lo tanto, si  $x > 3,5 \Rightarrow f''(x) > 0$ , luego,  $f(x)$  es estrictamente convexa. Y si  $x < 3,5 \Rightarrow f''(x) < 0$ , por lo que  $f(x)$  es estrictamente cóncava. Por lo tanto,  $x = 3,5$  corresponde a un punto de inflexión. Aplicando el método de Newton, con  $x_0 = 4$  (cercano al punto crítico  $x = 3,5$ ), sucede lo siguiente:

$$\begin{aligned} x_{k+1} &= x_k - \frac{6x_k^2 - 42x_k + 60}{12x_k - 42} \\ &= \frac{12x_k^2 - 42x_k - 6x_k^2 + 42x_k - 60}{6(2x_k - 7)} \\ &= \frac{6x_k^2 - 60}{6(2x_k - 7)} \end{aligned}$$

Por lo tanto,

$$x_{k+1} = \frac{x_k^2 - 10}{(2x_k - 7)}$$

Iterando vemos que el método converge a  $x_k = 5$ ,

$k$	0	1	2	3	4	5
$x_k$	4	6	7,2	5,65	5,099	$\sim 5$

que corresponde a un mínimo pues  $f''(x)$  en el punto es positiva. Ahora, cambiando el punto inicial a  $x_0 = 3$  y volvemos a aplicar el método, vemos que éste converge a  $x_k = 2$ ,

$k$	0	1	2	3	4	5
$x_k$	3	1	1,8	1,988	1,999	$\sim 2$

que corresponde a un máximo pues  $f''(x)$  en el punto es negativa.

**Ejemplo 8** Considere el problema  $\min \frac{x^3}{4} - \frac{7x^2}{8} + \ln x$  sujeto a  $x > 0$ .

Derivando encontramos

$$\begin{aligned} f'(x) &= \frac{3}{4}x^2 - \frac{7}{4}x + \frac{1}{x} \\ f''(x) &= \frac{3}{2}x - \frac{7}{4} - \frac{1}{x^2} \end{aligned}$$

$\Rightarrow x_{k+1} = x_k - \frac{(\frac{3}{4}x^2 - \frac{7}{4}x + \frac{1}{x})}{(\frac{3}{2}x - \frac{7}{4} - \frac{1}{x^2})}$ . Es fácil ver que  $\frac{3}{4}x^2 - \frac{7}{4}x + \frac{1}{x} = 0$  puede escribirse como  $(x - 2)(x - 1)(3x + 2) = 0, x \neq 0$  expresión que muestra claramente las raíces de la función objetivo. Iterando con el método desde dos puntos iniciales diferentes obtenemos los siguientes resultados:

$$\begin{array}{llll} x_0 = 0,5 & \rightarrow 0,7625 & \rightarrow 0,9401 & \rightarrow 0,9954 \\ x_0 = 3 & \rightarrow 2,3052 & \rightarrow 2,0512 & \rightarrow 2,0021. \end{array}$$

El primero de éstos corresponde a un máximo local mientras el segundo a un mínimo local.

### Método del Gradiente o de Cauchy

El método del gradiente también sirve para resolver problemas de optimización multidimensionales sin restricciones. Su algoritmo base es:

- i) Escoger un punto inicial de  $\mathbb{R}^n$
- ii) Determinar una dirección de movimiento
- iii) Moverse en esa dirección de acuerdo a algún criterio determinando un nuevo punto.
- iv) Verificar criterio de parada. Si no se satisface, volver a i).

Sea  $x_0$  nuestro punto de partida. Habrá que escoger alguna dirección  $\vec{d}$  tal que:  $f(\vec{x}_0 + \alpha\vec{d}) < f(\vec{x}_0)$ . Pero, ¿cómo escoger  $\vec{d}$  y  $\alpha$  de modo que  $f(\vec{x}_0 + \alpha\vec{d})$  sea lo más pequeño posible? Sabemos que  $\nabla f(\vec{x}_0)$  corresponde a la dirección de máximo crecimiento de la función en el punto  $\vec{x}_0$ . Por lo tanto, la dirección a escoger en un problema de minimización es la opuesta,  $\vec{d} = -\nabla f(\vec{x}_0)$  corresponde a la dirección de máximo decrecimiento.

Luego, nuestros pasos a seguir son:

- i) Tomar un punto conocido  $\vec{x}_0$ .
- ii) Escoger como dirección de movimiento al gradiente de la función en dicho punto, con signo contrario.



iii) Determinar el mínimo relativo de  $f(\vec{x})$  en esta dirección.

De esta forma  $\vec{x}_{k+1} = \vec{x}_k - \alpha \nabla f(\vec{x}_k)$ . El siguiente paso consiste en encontrar el  $\alpha$  que minimiza  $f(\vec{x}_{k+1})$ , es decir,

$$\min_{\alpha \geq 0} f(\vec{x}_k - \alpha \nabla f(\vec{x}_k)),$$

lo que se reduce a un problema unidimensional (la única variable en este problema es  $\alpha$ ) que puede ser resuelto algebraicamente o mediante el método de Newton.

La velocidad de convergencia de este método podría ser extremadamente lenta. Por ejemplo, es muy lenta ante una función objetivo con curvas de nivel elípticas. En cambio, en curvas de nivel circulares es inmediata (sólo una iteración). Este método se asemeja a una bolita que cae por un cerro y va tomando la mayor pendiente para descender. La diferencia es que la bolita cambia su dirección continuamente, no así este método que lo hace discretamente.

**Ejemplo 9** Considere el problema

$$P) \quad \min_{(x_1, x_2) \in \mathbb{R}^2} -2x_1x_2 - 2x_2 + x_1^2 + 2x_2^2$$

Es fácil ver que  $P$ ) es estrictamente convexo, ya que

$$H = \begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix}$$

es definida positiva, con lo que sabemos que cualquier punto de partida debiera llevarnos al óptimo global. Apliquemos el método de Cauchy a partir del punto inicial  $x^0 = (0, 0)$ . El gradiente de  $f(x)$  es

$$\nabla f(x) = [-2x_2 + 2x_1, -2x_1 - 2 + 4x_2]$$

con lo que

$$x^1 = (0, 0) - \alpha(0, -2) = (0, 2\alpha).$$

Así, el problema unidimensional a resolver para determinar cuánto avanzar es:

$$\min f(0, 2\alpha) = -4\alpha + 8\alpha^2$$

Derivando con respecto a  $\alpha$ , y luego igualando a cero obtenemos  $\alpha = \frac{1}{4}$ . Por lo tanto,  $x^1 = (0, \frac{1}{2})$ . Para la siguiente iteración tenemos que  $x^2 = (0, \frac{1}{2}) - \alpha(-1, 0) = (\alpha, \frac{1}{2})$ . Luego, resolviendo

$$\min f(\alpha, \frac{1}{2}) = -\alpha - 1 + \alpha^2 + \frac{1}{2}$$

obtenemos  $\alpha = \frac{1}{2}$ , y así  $x^2 = (\frac{1}{2}, \frac{1}{2})$ . Si seguimos iterando obtendremos

$$x^3 = (\frac{1}{2}, \frac{3}{4}), \quad x^4 = (\frac{3}{4}, \frac{3}{4}), \quad x^5 = (\frac{3}{4}, \frac{7}{8}),$$

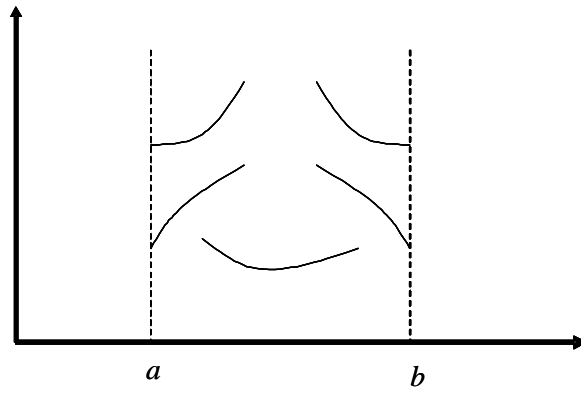


Figure 2.5:

lo cual lo podemos generalizar de la siguiente forma. Si  $k$  es par,

$$x^k = \left( \frac{2^{\frac{k}{2}} - 1}{2^{\frac{k}{2}}}, \frac{2^{\frac{k}{2}} - 1}{2^{\frac{k}{2}}} \right), \quad (2.3)$$

y si  $k$  es impar,

$$x^k = \left( \frac{2^{\frac{k-1}{2}} - 1}{2^{\frac{k-1}{2}}}, \frac{2^{\frac{k+1}{2}} - 1}{2^{\frac{k+1}{2}}} \right), \quad (2.4)$$

y al tomar límite a cualquiera de estas dos expresiones podemos ver que  $\lim_{k \rightarrow \infty} x^k = (1, 1)$ , y comprobar que  $\nabla f(1, 1) = 0$ .

## 2.2 Optimización de una función con restricciones

### 2.2.1 Caso 1: Problema Unidimensional

Considere el problema de optimización

$$\begin{aligned} P) \quad & \text{Min } f(x) \\ & a \leq x \leq b \\ & x \in \mathbb{R} \end{aligned}$$

con  $f(x)$  continua y derivable. En este caso, el óptimo del problema no necesariamente debe presentar un gradiente nulo, pues puede encontrarse en alguno de los bordes del dominio. Este análisis será de mucha utilidad para determinar las condiciones de optimalidad en problemas más complejos. Así, las condiciones de optimalidad deben modificarse para incorporar esta posibilidad.

**Teorema 29** (*Condición Necesaria de 1<sup>er</sup> orden*) Si  $\bar{x} \in [a, b]$  es punto mínimo local de  $P$ ) entonces:

$$i) \text{ Si } \bar{x} = a \Rightarrow f'(\bar{x}) \geq 0$$

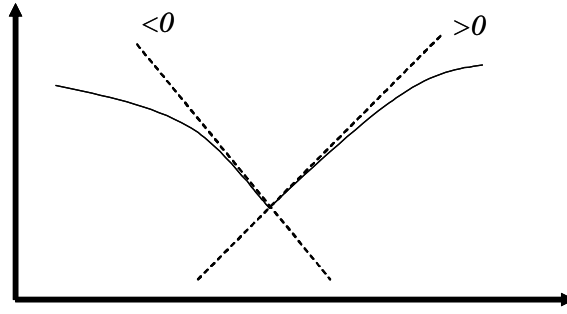


Figure 2.6:

- ii) Si  $\bar{x} = b \Rightarrow f'(\bar{x}) \leq 0$
- iii) Si  $a < \bar{x} < b \Rightarrow f'(\bar{x}) = 0$

Así, basta con una derivada no negativa en  $x = a$  o no positiva en  $x = b$  para que estos puntos sean considerados candidatos a mínimo del problema. Por otra part, las condiciones para puntos interiores del dominio permanecen intactas.

A diferencia del problema irrestricto, en los bordes es posible identificar condiciones suficientes de optimalidad local con sólo observar la primera derivada en esos puntos:

**Teorema 30** (Condiciones suficientes de 1<sup>er</sup> orden)

- i) Si  $\bar{x} = a$  y  $f'(a) > 0$  entonces  $\bar{x} = a$  es punto mínimo local estricto de  $P$
- ii) Si  $\bar{x} = b$  y  $f'(b) < 0$  entonces  $\bar{x} = b$  es punto mínimo local estricto de  $P$

**Observación:** ¿Qué pasa si la primera derivada no es continua en el dominio?

Para que un punto sea mínimo en esas condiciones, habrá que agregar a las condiciones necesarias de primer orden anteriores que:

Si  $f'_-(x) \neq f'_+(x) \Rightarrow f'_-(x) \leq 0$  y  $f'_+(x) \geq 0$  y para las condiciones suficientes se tendría que agregar que si  $f'_-(x) \neq f'_+(x) \Rightarrow f'_-(x) < 0$  y  $f'_+(x) > 0$ .

**Teorema 31** (Condición Necesaria de 2<sup>o</sup> orden). Sea  $f(x)$  dos veces diferenciable, y  $\bar{x}$  punto mínimo local de  $P$ , entonces:

- i) Si  $\bar{x} = a \Rightarrow (f'(a) \geq 0)$  y  $(f''(a) \geq 0)$  si  $f'(a) = 0$
- ii) Si  $\bar{x} = b \Rightarrow (f'(b) \leq 0)$  y  $(f''(b) \geq 0)$  si  $f'(b) = 0$
- iii) Si  $a < \bar{x} < b \Rightarrow f'(\bar{x}) = 0$  y  $f''(\bar{x}) \geq 0$

**Teorema 32** (Condición Suficiente de 2<sup>o</sup> orden). Sea  $\bar{x} \in [a, b]$ , entonces:

- i) Si  $\bar{x} = a$  y  $(f'(a) > 0)$  ó  $(f'(a) = 0$  y  $f''(a) > 0)$ , entonces  $\bar{x} = a$  es punto mínimo local estricto de  $P$ .
- ii) Si  $\bar{x} = b$  y  $(f'(b) < 0)$  ó  $(f'(b) = 0$  y  $f''(b) > 0)$ , entonces  $\bar{x} = b$  es punto mínimo local estricto de  $P$ .
- iii) Si  $a < \bar{x} < b$  y  $(f'(\bar{x}) = 0$  y  $f''(\bar{x}) > 0)$ , entonces  $\bar{x}$  es punto mínimo local estricto de  $P$ .

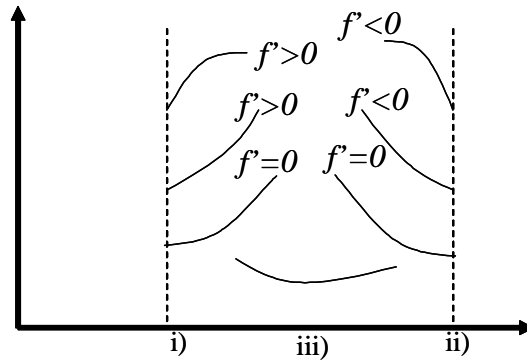


Figure 2.7:

**Ejemplo 10**

$$\begin{aligned} P) \quad & \text{Min } x^2 \\ \text{s.a.} \quad & x \geq -1 \end{aligned}$$

De acuerdo a lo anterior, se debe proceder a analizar todos los puntos candidatos a óptimo. Estos son aquéllos de derivada cero y los bordes del intervalo. En este caso  $f'(x) = 2x \Rightarrow \bar{x} = 0$ . Esta satisface la condición suficiente de 2° orden, ya que  $f'(0) = 0$  y  $f''(0) = 2 > 0$ .

Por lo tanto,  $\bar{x} = 0$  es punto mínimo local estricto por teorema de condición suficiente de 2° orden. En este caso, se puede argumentar que  $f''(x) > 0 \forall x \in D \Rightarrow f(x)$  es estrictamente convexa y por ende  $\bar{x} = 0$  es un punto mínimo global único de  $P$ ).

**Ejemplo 11**

$$\begin{aligned} P) \quad & \text{Max } -2x^3 + 3x^2 + 12x + 6 \\ \text{s.a.} \quad & 0 \leq x \leq 8 \end{aligned}$$

El problema se puede reescribir como:

$$\begin{aligned} \tilde{P}) \quad & \text{Min } 2x^3 - 3x^2 - 12x - 6 \\ \text{s.a.} \quad & 0 \leq x \leq 8 \end{aligned}$$

Inicialmente se debe determinar si este problema admite solución óptima. El problema satisface las condiciones de Bolzano-Weierstrass por lo que podemos seguir adelante. A continuación se buscan los puntos de derivada nula:  $\frac{\partial f}{\partial x} = 6x^2 - 6x - 12 = 0 \Rightarrow (x+1)(x-2) = 0$  y se obtiene:  $\bar{x}_1 = -1$  (punto fuera de  $D$ ), y  $\bar{x}_2 = 2$ .

La segunda derivada indica:  $\frac{\partial^2 f}{\partial x^2} = 12x - 6 \Rightarrow f''(2) = 18 > 0$ , por lo que se desprende que  $x_2$  es mínimo local del problema. ¿Qué pasará en los bordes?

$f'(0) = -12$  y  $f'(8) = 90$  por lo que  $\bar{x} = 0$  y  $\bar{x} = 8$  son máximos locales. Por lo tanto,  $\bar{x} = 2$  es el único punto mínimo local estricto y por tanto el punto óptimo del problema.

¿Qué hubiera pasado si el dominio fuera  $-8 \leq x \leq 8$ ? Los puntos críticos serían:  $x = -8$ ;  $x = -1$ ;  $x = 2$ ;  $x = 8$ . De ellos  $x = -8$  y  $x = 2$  corresponden a mínimos locales. Para determinar el mínimo global es necesario evaluar ambos en la función objetivo y escoger el menor, en este caso  $x = -8$  es el óptimo del nuevo problema.

### Ejemplo 12

$$\begin{array}{ll} P) & \text{Min } \sqrt{x} \\ \text{s.a} & 1 \leq x \leq 2 \end{array}$$

En este caso,  $f'(x) = \frac{1}{2\sqrt{x}}$  por lo que no existe un punto en el dominio ( $1 \leq x \leq 2$ ) tal que la derivada se anule. Por lo tanto, es necesario analizar los bordes:

$$f'(1) = \frac{1}{2} > 0 \Rightarrow \bar{x} = 1 \text{ es punto mínimo local estricto de } P).$$

$$f'(2) = \frac{1}{2\sqrt{2}} > 0 \Rightarrow \bar{x} = 2 \text{ es punto máximo local estricto de } P).$$

Así, es claro que  $\bar{x} = 1$  es el punto mínimo global del problema.

Esta respuesta podría obtenerse sin analizar las condiciones en los bordes del dominio pues, como  $f'(x) > 0$  sobre  $[1, 2]$ ,  $f(x)$  es estrictamente creciente sobre  $[1, 2]$ . Así, el mínimo global debe encontrarse en el límite izquierdo del dominio, esto es  $\bar{x} = 1$ .

### Ejemplo 13

$$\begin{array}{ll} P) & \text{Min } -x^2 - x^3 \\ \text{s.a} & x^2 \leq 1 \quad \leftrightarrow -1 \leq x \leq 1 \\ & x \in \mathbb{R} \end{array}$$

Inicialmente se buscan puntos de derivada nula:  $f'(x) = -2x - 3x^2 = 0 \Rightarrow x = 0$  ó  $x = -\frac{2}{3}$  ambos en el dominio. Por lo tanto, los puntos críticos son:  $x = -1$ ;  $x = -\frac{2}{3}$ ;  $x = 0$ ;  $x = 1$ . A continuación se analizan uno por uno:

$f'(-1) = -1 < 0 \Rightarrow$  Por teorema de condición suficiente de 1<sup>er</sup> orden,  $\bar{x} = -1$  es máximo local estricto.

$f'(-\frac{2}{3}) = 0$ ;  $f''(-\frac{2}{3}) = 2 > 0 \Rightarrow$  Por teorema de condición suficiente de 2<sup>o</sup> orden,  $\bar{x} = -\frac{2}{3}$  es mínimo local estricto.

$f'(0) = 0$ ;  $f''(0) = -2 < 0 \Rightarrow$  Por teorema de condición suficiente de 2<sup>o</sup> orden,  $\bar{x} = 0$  es máximo local estricto.

$f'(1) = -5 < 0 \Rightarrow$  Por teorema de condición suficiente de 1<sup>er</sup> orden,  $\bar{x} = 1$  es mínimo local estricto.

Para identificar el mínimo global del problema se comparan los valores de la función objetivo de los puntos mínimos locales, es decir,  $f(-\frac{2}{3}) = -\frac{4}{27}$  y  $f(1) = -2$ . Así, se concluye que  $\bar{x} = 1$  es mínimo global del problema.

### 2.2.2 Caso 2: Restricciones de igualdad

En esta sección se vuelve a considerar un problema multivariable, pero esta vez sujeto a una o más restricciones de igualdad. En términos generales, el problema se formula del siguiente modo:

$$\begin{aligned} P) \quad & \text{Min } f(\mathbf{x}) \\ \text{s.a} \quad & h_i(\mathbf{x}) = 0 \quad \forall i = 1, \dots, m \end{aligned}$$

donde  $\mathbf{x} = (x_1, \dots, x_n)$  y  $m < n$  (¿qué sucede si  $m = n$ ?). Para abordar este problema, inicialmente se supondrá que  $n = 2$  y  $m = 1$ , es decir, el caso bidimensional con una restricción. En este caso, el problema  $P$ ) se puede expresar del siguiente modo:

$$\begin{aligned} P) \quad & \text{Min } f(x_1, x_2) \\ \text{s.a} \quad & h(x_1, x_2) = a \end{aligned}$$

donde se supondrá que  $f$  y  $h$  tienen derivadas continuas de 2° orden. En este caso particular ( $n = 2$ ,  $m = 1$ ), el problema puede representarse gráficamente en  $\mathbb{R}^2$  (ver figura). El dominio queda representado por los puntos  $(x_1, x_2)$  que satisfacen  $h(x_1, x_2) = a$ . Para representar la función objetivo se requeriría una tercera dimensión, sin embargo alternativamente se puede representar mediante curvas de nivel que grafican en el plano  $\mathbb{R}^2$  conjuntos continuos de puntos cuya función objetivo arroja un idéntico valor. Así, si en la figura asociamos a cada uno de los anillos un determinado valor constante sabremos aproximadamente la forma de la función objetivo (la precisión aumenta en la medida que se dibujan más curvas). Supongamos en este caso que los anillos más pequeños tienen asociado un valor de función objetivo cada vez más pequeño. En ese caso sabremos que, de continuar esa tendencia, existirá un punto interior al anillo más pequeño que corresponderá a un mínimo local (identificado con "\*" en la figura).

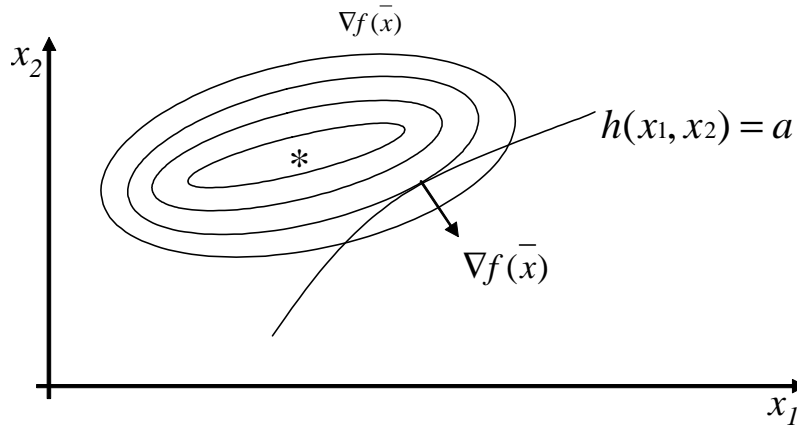


Figure 2.8:

En el caso no-restringido, un punto mínimo debía cumplir con:  $\nabla f(x) = 0$ . Sin embargo, ahora

las variables están enlazadas entre ellas en una restricción que las relaciona y las hace dependientes una de la otra, por lo que en el punto óptimo no necesariamente se satisfecerá esta condición.

En el caso de la figura vemos que el mínimo de la función  $f(x_1, x_2)$  no es factible en  $P$ ) ya que no satisface la restricción  $h(x_1, x_2) = a$ . El punto óptimo al problema deberá pertenecer a la curva  $h(x_1, x_2) = a$  y también a una curva de nivel (seguramente no dibujada)  $f(x_1, x_2) = z^*$ . Intuitivamente, vemos que en el óptimo de  $P$ ), la restricción y la curva de nivel correspondiente al punto óptimo en la función objetivo, son tangentes. Más adelante comprobaremos que esta será la nueva condición necesaria de primer orden para optimalidad de este problema.

Es posible imaginar que se pudiera expresar  $x_2$  en función de  $x_1$  a partir de la restricción  $h(x_1, x_2) = a$  como  $x_2 = r(x_1)$ . En ese caso, se puede reemplazar la restricción en la función objetivo redefiniéndola como:  $F(x_1) = f(x_1, r(x_1))$  y el problema quedará:

$$\begin{aligned} P) \quad & \text{Min } F(x_1) \\ & \text{irrestricto} \end{aligned}$$

En la sección 2.1, se enseñó cómo abordar problemas irrestrictos como el anterior: las condiciones necesarias de 1<sup>er</sup> orden para un punto mínimo de  $P$ ) son:

$$\frac{\partial F}{\partial x_1} = 0 \Rightarrow \frac{\partial f}{\partial x_1} + \frac{\partial f}{\partial x_2} \frac{dr}{dx_1} = 0$$

Pero  $x_2 = r(x_1)$  no es siempre identificable como función, por lo que para identificar  $\frac{\partial r}{\partial x_1}$  debemos recurrir a la restricción.

$$h(x_1, x_2) = a \rightarrow h(x_1, r(x_1)) = a.$$

Nos interesa determinar las condiciones de optimalidad de primer orden. Estas corresponden a condiciones que se deben satisfacer en el punto de interés en relación a las derivadas de primer orden. El punto óptimo local debe ser mejor que los puntos en su vecindad *que pertenezcan al dominio*. Es decir nos interesa analizar diferenciales de primer orden que garanticen la permanencia en el dominio. En nuestro caso debe suceder que:

$$h(x_1 + \Delta x_1, x_2 + \Delta x_2) = a$$

Podemos aproximar el lado izquierdo de esta expresión mediante una expansión de Taylor de primer orden en torno a  $(x_1, x_2)$ :

$$h(x_1 + \Delta x_1, x_2 + \Delta x_2) \simeq h(x_1, x_2) + \Delta x_1 \frac{\partial h}{\partial x_1} + \Delta x_2 \frac{\partial h}{\partial x_2}$$

Reemplazando en la expresión anterior y recordando que  $h(x_1, x_2) = a$  se obtiene:

$$\frac{\partial h}{\partial x_1} \Delta x_1 + \frac{\partial h}{\partial x_2} \Delta x_2 = 0$$

que corresponde al resultado que hubiéramos obtenido de diferenciar la restricción original:

$$\frac{\partial h}{\partial x_1} dx_1 + \frac{\partial h}{\partial x_2} dx_2 = 0$$

y despejando esta expresión nos queda

$$\frac{dx_2}{dx_1} = -\frac{\frac{\partial h}{\partial x_1}}{\frac{\partial h}{\partial x_2}} = \frac{dr}{dx_1}.$$

Reemplazando  $\frac{dr}{dx_1}$  de esta expresión en XX se obtiene:

$$\frac{\partial f}{\partial x_1} - \frac{\frac{\partial f}{\partial x_2}}{\frac{\partial h}{\partial x_2}} \frac{\partial h}{\partial x_1} = 0$$

Diremos que esta condición junto a  $h(x_1, x_2) = a$  son las condiciones necesarias de primer orden para nuestro problema. Sin embargo, si observamos la siguiente obvia identidad:

$$\frac{\partial f}{\partial x_2} - \frac{\frac{\partial f}{\partial x_2}}{\frac{\partial h}{\partial x_2}} \frac{\partial h}{\partial x_2} = 0$$

Si definimos  $\lambda = \frac{\partial f}{\partial x_2} / \frac{\partial h}{\partial x_2}$ , podemos expresar las condiciones necesarias de primer orden del siguiente modo:

$$\begin{aligned} \frac{\partial f}{\partial x_i} - \lambda \frac{\partial h}{\partial x_i} &= 0 & i = 1, 2 \\ h(x_1, x_2) &= a \end{aligned}$$

Esto permite reescribir el problema original como el siguiente problema equivalente irrestricto con una variable adicional ( $\lambda$ ):

$$\text{Min } L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda(a - h(x_1, x_2))$$

y abordarlo con las herramientas presentadas en la sección 2.1. Efectivamente, las condiciones de primer orden para este problema son:

$$\begin{aligned} \frac{\partial L}{\partial x_i} &= \frac{\partial f}{\partial x_i} - \lambda \frac{\partial h}{\partial x_i} = 0 & i = 1, 2 \\ \frac{\partial L}{\partial \lambda} &= a - h(x_1, x_2) = 0, \end{aligned}$$

lo cual corresponde a las condiciones derivadas anteriormente.

Note que este procedimiento asumió que  $\frac{\partial h}{\partial x_2} \neq 0$  en el punto de interés. En caso que  $\frac{\partial h}{\partial x_2} = 0$  y  $\frac{\partial h}{\partial x_1} \neq 0$ , se obtiene alternativamente una expresión para  $\frac{dx_1}{dx_2}$  mediante un procedimiento análogo



comenzando con  $x_1 = r(x_2)$ . Note que las condiciones de optimalidad serían las mismas. Sin embargo, este método no determina las condiciones de optimalidad local si el gradiente de la restricción en el punto de interés es nulo ( $\frac{\partial h}{\partial x_1} = \frac{\partial h}{\partial x_2} = 0$ ). A los puntos que cumplen con esta condición se les denomina puntos singulares y exigen un análisis alternativo.

Este método se conoce como "Método de Lagrange" (1790 aprox.), donde  $L$  es la función de Lagrange, y  $\lambda$  es el multiplicador de Lagrange. El método resulta especialmente útil cuando resulta complejo expresar  $x_2$  en función de  $x_1$  (y viceversa) a partir de la restricción del problema. Adicionalmente, este desarrollo provee un muy buen punto de partida para determinar condiciones de optimalidad para problemas más complejos.

Ahora probaremos lo que se desprendía gráficamente de la figura anterior, que (para  $n = 2$ ,  $m = 1$ ) en un punto óptimo local, la derivada de la curva de nivel que pasa por el punto y la derivada de la restricción deben ser iguales. Habíamos obtenido que:

$$\frac{\partial f}{\partial x_1} - \frac{\frac{\partial f}{\partial x_2}}{\frac{\partial h}{\partial x_2}} \frac{\partial h}{\partial x_1} = 0 \quad \Rightarrow \quad \frac{\frac{\partial f}{\partial x_1}}{\frac{\partial f}{\partial x_2}} = \frac{\frac{\partial h}{\partial x_1}}{\frac{\partial h}{\partial x_2}}.$$

De la curva de nivel  $f(x_1, x_2) = c$  (suponiendo  $c$  como el valor óptimo), tenemos que:

$$\frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 = 0$$

o bien,

$$\left. \frac{dx_2}{dx_1} \right|_{\text{en curva de nivel}} = - \frac{\frac{\partial f}{\partial x_1}}{\frac{\partial f}{\partial x_2}} \quad (2.5)$$

y análogamente en la restricción  $h(x_1, x_2) = a$  tenemos que:

$$\left. \frac{dx_2}{dx_1} \right|_{\text{en la restricción}} = - \frac{\frac{\partial h}{\partial x_1}}{\frac{\partial h}{\partial x_2}}. \quad (2.6)$$

Por lo tanto, de (2.5) y (2.6) vemos que en el punto óptimo ambas pendientes son iguales.

$$\left. \frac{dx_2}{dx_1} \right|_{\text{en curva de nivel}} = \left. \frac{dx_2}{dx_1} \right|_{\text{en la restricción}}$$

**Nota:** Esta condición se cumple para cada óptimo local (ver Figura 2.9).

**Ejemplo 14** Debemos comprar un terreno de una hectárea rectangular que colinde con el río y con el camino (que forman un ángulo de  $90^\circ$ , ver figura). El problema es definir las dimensiones del terreno, tal que su superficie útil sea máxima. Por cada metro de límite con el río se pierden  $3m^2$  de terreno útil. Por cada metro de límite con el camino se pierden  $2m^2$  de terreno útil. Por cada metro de límite con un vecino interior se pierde  $\frac{1}{2}m^2$  útil.

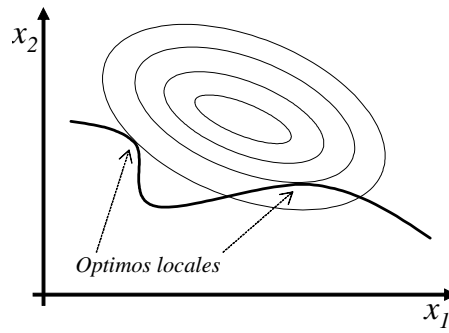


Figure 2.9: Óptimos locales

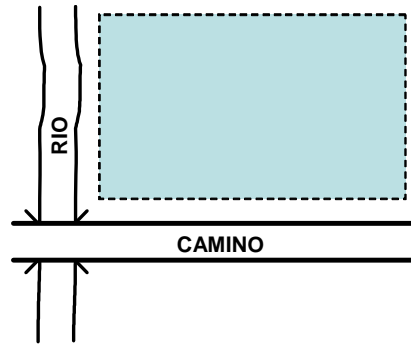


Figure 2.10:

Un problema de optimización equivalente sería:

$$\begin{aligned} P) \quad & \text{Max } (x - 3, 5)(y - 2, 5) \\ \text{s.a} \quad & xy = 10000 \end{aligned}$$

el cual es equivalente a uno en que se minimiza  $(3, 5 - x)(y - 2, 5)$  y puede ser resuelto mediante el método de Lagrange. La función de Lagrange sería:

$$L = (3, 5 - x)(y - 2, 5) + \lambda(10000 - xy)$$

y las condiciones de primer orden son:

$$\begin{aligned} \frac{\partial L}{\partial x} &= 2, 5 - y - \lambda y = 0 & \Rightarrow \lambda &= \frac{2, 5 - y}{y} \\ \frac{\partial L}{\partial y} &= 3, 5 - x - \lambda x = 0 & \Rightarrow \lambda &= \frac{3, 5 - x}{x} \\ \frac{\partial L}{\partial \lambda} &= 10000 - xy = 0. \end{aligned}$$

Reemplazando y resolviendo el sistema de ecuaciones, obtenemos la solución  $x = 118, 32$ ;  $y = 84, 5$ ;

$\lambda = -0,97$ . Por lo tanto, el valor óptimo es  $\hat{v} = -9.415,24 \text{ m}^2$ . Observar que esta solución corresponde al problema equivalente de minimización al de maximización del terreno útil.

**Nota:** El sistema de ecuaciones también admite la solución  $x = -118,32$ ;  $y = -84,5$ ; pero ésta se descarta ya que el problema no admite dimensiones negativas de terreno.

### Caso general con restricciones de igualdad

A continuación veremos las condiciones necesarias de optimalidad para un problema multidimensional ( $n$  variables) con muchas restricciones de igualdad ( $m$ ). Consideremos el siguiente problema:

$$\begin{array}{lll} P) & \text{Min } f(x) & x = (x_1, \dots, x_n) \\ \text{s.a} & h_j(x) = a_j & j = 1, \dots, m \quad \text{con } m \leq n \end{array}$$

Si hacemos un análisis similar al caso bidimensional representando  $m$  variables en función de las otras ( $n - m$ ), se puede demostrar el siguiente teorema:

**Teorema 33 (Lagrange)** *Considere el problema  $P$ ) en que todas las funciones involucradas son diferenciables. La condición necesaria para que el punto factible regular  $\bar{x}$  sea un mínimo local de  $P$ ), es que existan  $m$  escalares  $\lambda_j$  ( $j = 1, \dots, m$ ), denominados multiplicadores de Lagrange, tales que:*

$$\nabla f(\bar{x}) + \sum_{j=1}^m \lambda_j \nabla h_j(\bar{x}) = 0$$

**Nota:** El teorema de Lagrange indica que el gradiente de la función objetivo evaluado en  $\bar{x}$ , es decir,  $\nabla f(\bar{x})$ , se puede expresar como combinación lineal de los gradientes de las restricciones de  $P$ ) evaluados en  $\bar{x}$ , es decir,  $\{\nabla h_j(\bar{x})\}_{j \in J}$ .

Es decir las condiciones necesarias de primer orden para un óptimo son las siguientes:

$$\begin{array}{ll} \frac{\partial f(x)}{\partial x_i} - \sum_{j=1}^m \lambda_j \frac{\partial h_j(x)}{\partial x_i} = 0 & i = 1, \dots, n \\ a_j - h_j(x) = 0 & j = 1, \dots, m \end{array}$$

Es decir, si un punto cumple con satisfacer estas  $(n + m)$  restricciones, estamos frente a un punto candidato a mínimo o a máximo local. Es importante recordar que este análisis de primer orden es incapaz de distinguir entre un mínimo y un máximo local.

Así definiremos una función:

$$L(x_1, \dots, x_n, \lambda_1, \dots, \lambda_m) = f(x_1, \dots, x_n) + \sum_{j=1}^m \lambda_j [a_j - h_j(x_1, \dots, x_n)]$$

tal que las condiciones necesarias de 1<sup>er</sup> orden de nuestro problema  $P$ ) corresponden a las condiciones necesarias de primer orden del problema no-restringido:

$$\text{Min } L(x_1, \dots, x_n, \lambda_1, \dots, \lambda_m)$$

las cuales son:

$$\begin{aligned} \frac{\partial L}{\partial x_i} &= \frac{\partial f}{\partial x_i} - \sum_{j=1}^m \lambda_j \frac{\partial h_j}{\partial x_i} = 0 & i = 1, \dots, n \\ \frac{\partial L}{\partial \lambda_j} &= a_j - h_j(x_1, \dots, x_n) = 0 & j = 1, \dots, m \end{aligned}$$

Sin embargo, al igual que en el caso bidimensional, es importante tener presente que el método de Lagrange sólo identifica **puntos regulares**. Los puntos que no cumplan esta propiedad (singulares) deben ser analizados separadamente para determinar si son óptimos al problema.

**Definición 13** (*Punto regular*) Un punto es regular cuando el jacobiano de las restricciones (2.7), evaluado en el punto, es de rango máximo  $m$ , es decir, las  $m$  filas deben ser linealmente independientes (recordar que  $n > m$ ).

$$J(x) = \begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial h_m}{\partial x_1} & \cdots & \frac{\partial h_m}{\partial x_n} \end{bmatrix} \quad (2.7)$$

**Ejemplo 15** Consideremos el siguiente problema:

$$\begin{aligned} P) \text{ Min } x \\ \text{s.a. } -x^3 + y^2 = 0 \\ (x, y) \in \mathbb{R}^2 \end{aligned}$$

Las condiciones de primer orden del Lagrangeano  $L(x, y, \lambda) = x + \lambda(-x^3 + y^2)$  son:

$$\begin{aligned} \frac{\partial L}{\partial x} &= 1 - 3x^2\lambda = 0 \\ \frac{\partial L}{\partial y} &= 2y\lambda = 0 \\ \frac{\partial L}{\partial \lambda} &= -x^3 + y^2 = 0 \end{aligned}$$

de donde se observa que este sistema no tiene solución. Sin embargo, es fácil ver que la solución óptima es  $(0, 0)$ . Pero, ¿es  $\bar{x} = (0, 0)$  un punto regular? Observando el Jacobiano de la restricción  $h(x) : \nabla h(\bar{x}) = \begin{bmatrix} -3x^2 & 2y \end{bmatrix}_{(0,0)} = \begin{bmatrix} 0 & 0 \end{bmatrix}$  vemos que éste es nulo, por lo tanto  $\bar{x}$  es solución óptima no regular, por esta razón las condiciones necesarias de primer orden no lo detectan.

**Observación:** El método de Lagrange detecta todos los puntos óptimos locales regulares, pero no necesariamente detecta los singulares.

Condiciones de 2° orden:

Antes de describir las condiciones necesarias y suficientes de segundo orden para que un punto que, satisfaciendo las condiciones necesarias de primer orden, sea mínimo local, es necesario entender que estas condiciones poco tienen que ver con la convexidad de la función objetivo. Es decir, que la función objetivo sea convexa no indica que un punto que cumpla con las condiciones necesarias de primer orden sea efectivamente un mínimo local. Como ejemplo considere el problema bidimensional de minimizar una función objetivo paraboidal sujeto a que el punto deba satisfacer una curva que describe una elipse. Esto se ilustra en la figura YY<sup>1</sup>. En ella se observa que son dos los puntos que satisfacen la condición necesaria de primer orden, pero sólo uno de ellos corresponde a un mínimo.

Volviendo al caso más general, para que  $\bar{x}$  sea mínimo local debe suceder que:  $f(\bar{x} + \varepsilon \Delta x) \geq f(\bar{x})$ , para todo  $(\bar{x} + \varepsilon \Delta x)$  factible ( $\varepsilon \Delta x$  pequeño), esto es para todo  $(\bar{x} + \varepsilon \Delta x)$  que satisfaga todas las restricciones:  $h_j(\bar{x} + \varepsilon \Delta x) = a_j$ ,  $j = 1, \dots, m$ .

La función que interesa minimizar es  $L(x, \lambda) = f(x) + \lambda(a - h(x))$ . Es decir buscamos  $\bar{x}$  tal que para todo  $\bar{x} + \varepsilon \Delta x$  factible,  $L(\bar{x} + \varepsilon \Delta x, \lambda) - L(\bar{x}, \lambda) = f(\bar{x} + \varepsilon \Delta x) - f(\bar{x}) \geq 0$  (para que  $\bar{x}$  sea mínimo).

Obtendremos las condiciones de segundo orden por medio de expandir  $L$  en serie de Taylor en segundo grado en torno al punto  $(\bar{x}, \lambda)$ . Recordando que este punto satisface las condiciones de 1<sup>er</sup> orden, tendremos que:

$$\begin{aligned} L(\bar{x} + \varepsilon \Delta x, \lambda) - L(\bar{x}, \lambda) &= L(\bar{x}, \lambda) + \frac{(\bar{x} + \varepsilon \Delta x - \bar{x}) \nabla L(x^*, \lambda)}{1!} + \frac{(\varepsilon \Delta x) D^2 L(x^*, \lambda) (\varepsilon \Delta x)^T}{2!} + 0(\varepsilon^2) - L(\bar{x}, \lambda) \\ &= \frac{1}{2} \varepsilon^2 \Delta x \frac{\partial^2 L(\bar{x}, \lambda)}{\partial x^2} \Delta x^T + 0(\varepsilon^2) \end{aligned}$$

Considerando que  $\lim_{\varepsilon \rightarrow 0} \frac{0(\varepsilon^2)}{\varepsilon^2 \|\Delta x\|} = 0$ , si  $\bar{x}$  es un óptimo local entonces debe cumplirse que  $\Delta x \frac{\partial^2 L(\bar{x}, \lambda)}{\partial x^2} \Delta x^T \geq 0$ . Bastaría con que el Hessiano del Lagrangeano sea positivo semi definido (o positivo definido) en el punto para que el punto cumpla las condiciones necesarias (suficientes) de segundo orden para ser considerado mínimo local. Sin embargo las direcciones factibles de movimiento  $\Delta \vec{x}$  no son arbitrarias pues no pueden salirse del dominio. Así, es posible que el punto  $\bar{x}$  sea un óptimo local sin que el Hessiano del Lagrangeano sea positivo semi definido pues restringiendo los  $\Delta \vec{x}$  a direcciones factibles sí se cumpla la condición  $\Delta x \frac{\partial^2 L(\bar{x}, \lambda)}{\partial x^2} \Delta x^T \geq 0$ . Es decir interesa determinar cuáles son esas direcciones factibles para  $\Delta \vec{x}$  de modo de evaluar la condición de optimalidad de segundo grado. Mediante un desarrollo idéntico al realizado en la sección 2.2 se obtienen las condiciones que se deben satisfacer:

$$\sum_{i=1}^n \frac{\partial h_j(\bar{x})}{\partial x_i} \Delta x_i = 0 \quad \forall j = 1, \dots, m$$

---

<sup>1</sup> Agregar Figura

Así, la condición necesaria de 2° orden será que el Hessiano del Lagrangeano debe ser positivo semidefinido en el subespacio definido por las restricciones.

Análogamente, la condición suficiente de 2° orden, es que el Hessiano del Lagrangeano sea positivo definido en dicho subespacio. En este caso, el punto  $\bar{x}$  será un mínimo local estricto de  $P$ ).

### Ejemplo 16

$$\begin{aligned} \text{Min } f(x) &= x_1^2 + x_2^2 + x_3^2 \\ \text{s.a. } h_1(x) &= x_1 + x_2 + 3x_3 - 2 = 0 \\ h_2(x) &= 5x_1 + 2x_2 + x_3 - 5 = 0 \end{aligned}$$

Lo primero que se debe hacer es demostrar la existencia de solución para este problema. Este problema evidentemente satisface las condiciones del teorema de existencia por lo que tenemos garantía de que el problema tendrá solución.

$$L = x_1^2 + x_2^2 + x_3^2 + \lambda_1(2 - x_1 - x_2 - 3x_3) + \lambda_2(5 - 5x_1 - 2x_2 - x_3)$$

$$\left. \begin{aligned} \frac{\partial L}{\partial x_1} &= 2x_1 - \lambda_1 - 5\lambda_2 = 0 \\ \frac{\partial L}{\partial x_2} &= 2x_2 - \lambda_1 - 2\lambda_2 = 0 \\ \frac{\partial L}{\partial x_3} &= 2x_3 - 3\lambda_1 - \lambda_2 = 0 \end{aligned} \right\} \begin{aligned} 2x_1 - 2x_2 - 3\lambda_2 &= 0 \\ 6x_2 - 2x_3 - 5\lambda_2 &= 0 \end{aligned} \left\} \begin{aligned} 5x_1 - 14x_2 + 3x_3 &= 0 \end{aligned} \right.$$

$$\left. \begin{aligned} x_1 + x_2 + 3x_3 &= 2 \\ 5x_1 + 2x_2 + x_3 &= 5 \\ 5x_1 - 14x_2 + 3x_3 &= 0 \end{aligned} \right\} \begin{aligned} 3x_2 + 14x_3 &= 5 \\ 16x_2 - 2x_3 &= 5 \end{aligned} \left\} \begin{aligned} x_2^* &= \frac{8}{23} \quad \text{y} \quad x_3^* = \frac{13}{46} \end{aligned}$$

$$\Rightarrow x_1^* = \frac{37}{46} \quad \text{y} \quad \lambda_1 = \frac{2}{23}, \lambda_2 = \frac{7}{23}$$

¿Es el punto  $(\frac{37}{46}, \frac{8}{23}, \frac{13}{46})$  un mínimo? Para esto basta observar el Hessiano del Lagrangeano, esto es

$$H = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Dado que el Hessiano del Lagrangeano (que en este caso coincide con el Hessiano de la función objetivo pues todas las restricciones son lineales<sup>2</sup>) es positivo definido en todo el dominio y en particular en el punto  $(\frac{37}{46}, \frac{8}{23}, \frac{13}{46})$ , este punto es un mínimo único y global.

### Ejemplo 17

$$\begin{aligned} \text{Min } x_1^2 + x_2^2 \\ \text{s.a. } (x_1 - 1)^2 + (x_2 - 2)^2 &= 45 \end{aligned}$$

---

<sup>2</sup>Explicar mejor este parentesis

Lo primero que se debe hacer es demostrar la existencia de solución para este problema. Este problema evidentemente satisface las condiciones del teorema de Bolzano-Weierstrass por lo que tenemos garantía de que el problema tendrá solución.

El lagrangeano queda:

$$L = x_1^2 + x_2^2 + \lambda(45 - (x_1 - 1)^2 - (x_2 - 2)^2)$$

De las condiciones de primer orden podemos obtener:

$$\begin{aligned} \frac{\partial L}{\partial x_1} &= 2x_1 - 2\lambda(x_1 - 1) = 0 & \Rightarrow \lambda = \frac{x_1}{x_1 - 1}; \quad x_1 \neq 1 \quad (x_1 = 1 \text{ no satisface la condición}) \\ \frac{\partial L}{\partial x_2} &= 2x_2 - 2\lambda(x_2 - 2) = 0 & \Rightarrow \lambda = \frac{x_2}{x_2 - 2}; \quad x_2 \neq 2 \quad (x_2 = 2 \text{ no satisface la condición}) \\ \frac{x_1}{x_1 - 1} &= \frac{x_2}{x_2 - 2} \rightarrow 2x_1 = x_2 \end{aligned}$$

Utilizando esta condición en conjunto con la restricción tenemos que:

$$\begin{aligned} (x_1 - 1)^2 + (2x_1 - 2)^2 &= 45 \\ x_1^2 - 2x_1 + 1 + 4x_1^2 - 8x_1 + 4 &= 45 \\ 5x_1^2 - 10x_1 - 40 &= 0 \end{aligned}$$

Por lo tanto las soluciones posibles son:

$$\begin{aligned} x_1 &= 4 & x_2 &= 8 & \lambda &= \frac{4}{3} \\ x_1 &= -2 & x_2 &= -4 & \lambda &= \frac{2}{3} \end{aligned}$$

Estos dos puntos son los únicos puntos que satisfacen las condiciones necesarias de primer orden. Esto significa que, de no existir puntos singulares al problema, el punto óptimo al problema será uno de ellos. A continuación revisaremos las condiciones de segundo orden. La matriz Hessiana es:

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

Para caracterizar esta matriz debemos reemplazar los valores de  $\lambda$ :

$H(4, 8, \frac{4}{3}) = \begin{bmatrix} \frac{-2}{3} & 0 \\ 0 & \frac{-2}{3} \end{bmatrix} \Rightarrow$  Definida Negativa, por lo tanto,  $\bar{x} = (4, 8)$  es un máximo local estricto,  $f(\bar{x}) = 80$ .

$H(-2, -4, \frac{2}{3}) = \begin{bmatrix} \frac{2}{3} & 0 \\ 0 & \frac{2}{3} \end{bmatrix} \Rightarrow$  Definida positiva, por lo tanto,  $\bar{x} = (-2, -4)$  es un mínimo local estricto,  $f(\bar{x}) = 20$ .

¿Existen puntos singulares en el dominio?

$J(\bar{x}) = \left[ \frac{\partial h}{\partial x_1} \quad \frac{\partial h}{\partial x_2} \right] = \left[ 2(x_1 - 1) \quad 2(x_2 - 2) \right] \rightarrow$  El único punto no regular (que hace nulo el jacobiano) es  $(1, 2)$  pero este punto no es factible. Así, el punto extremo identificado es mínimo global del problema.

### Ejemplo 18

$$\begin{aligned} \text{Min} \quad & -x_1x_2 - x_2x_3 - x_3x_1 \\ \text{s.a.} \quad & x_1 + x_2 + x_3 = 3 \\ & (x_1, x_2, x_3) \in \mathbb{R}^3 \end{aligned}$$

Lo primero que se debe hacer es demostrar la existencia de solución para este problema. Este problema satisface las condiciones del teorema de existencia de soluciones óptimas por lo que tenemos garantía de que el problema tendrá solución. La función objetivo es continua, el dominio es cerrado y no vacío (el punto  $(1, 1, 1)$  es factible). Para mostrar que  $f(x) \rightarrow +\infty$  si  $\|x\| \rightarrow +\infty, x \in D$  se necesita que al menos una de las variables tienda a infinito. Para comenzar asignemos a una de las variables, digamos  $x_1$ , el valor  $M$ . En ese caso  $x_2 + x_3 = 3 - M$ . Y la función objetivo queda  $M(M - 3) - x_3x_2$ . Ahora bien, si  $M \rightarrow +\infty$ , entonces es necesario determinar qué sucede con la función objetivo. Es fácil demostrar que el máximo valor que toman  $x_3x_2$  restringidos a  $x_2 + x_3 = 3 - M$  es  $x_2 = x_3 = \frac{3-M}{2}$ . Es decir, la función objetivo tenderá como mínimo a  $M(M - 3) - \frac{(M-3)^2}{4}$ . Y esta función claramente tiende a infinito cuando  $M$  tiende a ello.

Sólo ahora podemos preocuparnos de la búsqueda de óptimos locales.

$$\begin{aligned} L &= -x_1x_2 - x_2x_3 - x_3x_1 + \lambda(3 - x_1 - x_2 - x_3) \\ \left. \begin{aligned} \frac{\partial L}{\partial x_1} &= -x_2 - x_3 - \lambda = 0 \\ \frac{\partial L}{\partial x_2} &= -x_1 - x_3 - \lambda = 0 \\ \frac{\partial L}{\partial x_3} &= -x_2 - x_1 - \lambda = 0 \end{aligned} \right\} x_1 = x_2 = x_3 = 1; \quad \lambda = 2 \end{aligned}$$

Dado que el dominio es lineal, no existen puntos singulares en el dominio. Ahora, ¿cómo saber si  $\bar{x}$  es un punto mínimo o no? Como hemos demostrado la existencia de una solución óptima, este punto debe corresponder a ese mínimo. A pesar de esta certidumbre, proseguiremos identificando el Hessiano del Lagrangeano:

$$H = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}$$

pero este Hessiano no es ni positivo semidefinido ni negativo semidefinido. Así, es necesario analizar si limitando las direcciones de movimiento desde el punto  $(1, 1, 1)$  sólo a aquéllas que satisfagan la restricción, es posible discernir si este punto corresponde a un mínimo o máximo local del problema. El punto  $(\bar{x}_1 + \Delta x_1, \bar{x}_2 + \Delta x_2, \bar{x}_3 + \Delta x_3)$  debe satisfacer la restricción por lo que debe cumplirse:  $\bar{x}_1 + \Delta x_1 + \bar{x}_2 + \Delta x_2 + \bar{x}_3 + \Delta x_3 = 3$ , y como  $\bar{x}_1 + \bar{x}_2 + \bar{x}_3 = 3$ , el vector  $\Delta \vec{x}$  debe satisfacer



$\Delta x_1 + \Delta x_2 + \Delta x_3 = 0$ . Así, para que el punto  $(1, 1, 1)$  sea mínimo local debe cumplirse que

$$\begin{bmatrix} \Delta x_1 & \Delta x_2 & \Delta x_3 \end{bmatrix} \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \\ \Delta x_3 \end{bmatrix} = -2(\Delta x_1 \Delta x_2 + \Delta x_2 \Delta x_3 + \Delta x_3 \Delta x_1) \geq 0$$

para todo  $\Delta \vec{x}$  que satisfaga  $\Delta x_1 + \Delta x_2 + \Delta x_3 = 0$ .

Despejando  $\Delta x_1 = -\Delta x_2 - \Delta x_3$  el lado izquierdo queda  $2(\Delta x_2^2 + \Delta x_2 \Delta x_3 + \Delta x_3^2)$ . Esta expresión es necesariamente positiva para todo  $(\Delta x_2, \Delta x_3) \neq (0, 0)$ . Esto se puede mostrar de dos formas:  $2(\Delta x_2^2 + \Delta x_2 \Delta x_3 + \Delta x_3^2) = (\Delta x_2 + \Delta x_3)^2 + \Delta x_2^2 + \Delta x_3^2$ . Y la suma de elementos positivos no puede más que ser positivo. Alternativamente se puede representar matricialmente la expresión del siguiente modo:

$$2(\Delta x_2^2 + \Delta x_2 \Delta x_3 + \Delta x_3^2) = [\Delta x_2 \ \Delta x_3] \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} \Delta x_2 \\ \Delta x_3 \end{bmatrix}.$$

Es fácil ver que este valor es siempre positivo  $\forall (\Delta x_2, \Delta x_3) \neq (0, 0)$  pues la matriz involucrada en la operación es positiva definida. Por lo tanto,  $\bar{x} = (1, 1, 1)$  es un punto mínimo global del problema.

### Interpretación de los multiplicadores $\lambda_j$

Proveeremos una interpretación para los multiplicadores de Lagrange  $\lambda_j$  comenzando del caso más simple, esto es  $(n = 2, m = 1)$ . En este caso las condiciones de primer orden indicaban:

$$\begin{aligned} \frac{\partial f}{\partial x_1} - \lambda \frac{\partial h}{\partial x_1} &= 0 \\ \frac{\partial f}{\partial x_2} - \lambda \frac{\partial h}{\partial x_2} &= 0 \\ a - h(x_1, x_2) &= 0 \end{aligned}$$

podemos expresar la solución óptima  $(x^*, \lambda^*)$  en función del parámetro  $a$ , como  $x_1^*(a)$ ,  $x_2^*(a)$ ,  $\lambda^*(a)$ . Así, el Lagrangeano sería:

$$L(a) = f(x_1^*(a), x_2^*(a)) + \lambda^*(a)(a - h(x_1^*(a), x_2^*(a))).$$

Luego, derivando con respecto al parámetro  $a$ , tenemos:

$$\frac{\partial L}{\partial a} = \left[ \left( \frac{\partial f}{\partial x_1} - \lambda \frac{\partial h}{\partial x_1} \right) \frac{dx_1}{da} + \left( \frac{\partial f}{\partial x_2} - \lambda \frac{\partial h}{\partial x_2} \right) \frac{dx_2}{da} + (a - h(x_1, x_2)) \frac{d\lambda}{da} + \lambda(a) \right]_{(x, \lambda) = (x^*(a), \lambda^*(a))},$$

Sin embargo, los tres primeros términos del lado derecho de esta expresión evaluados en un punto mínimo local son nulos (por condiciones de optimalidad). Así,

$$\left. \frac{\partial L}{\partial a} \right|_{\text{óptimo}} = \left. \frac{\partial f}{\partial a} \right|_{\text{óptimo}} = \lambda.$$

En resumen,  $\lambda$  es la sensibilidad del valor óptimo ( $\hat{v}$ ) frente a pequeñas variaciones en el parámetro  $a$ . Si el parámetro  $a$  es modificado en  $a' = a + \Delta a$ , entonces el valor óptimo puede aproximarse como  $\hat{v}' = \hat{v} + \Delta \hat{v} \simeq \hat{v} + \lambda \Delta a$ . El multiplicador  $\lambda$  puede tomar valores negativos o positivos, de hecho basta multiplicar la restricción por  $(-1)$  para que el valor del multiplicador cambie de signo (sin haber modificado el conjunto factible). Esto se debe a que al aumentar en una unidad el lado derecho de la restricción, es posible que la función objetivo aumente o disminuya pues se está modificando completamente los puntos del dominio. Para el caso general  $(n, m)$ , la interpretación de los multiplicadores es absolutamente análoga, es decir, si  $\bar{x}$  es el punto mínimo local,

$$\frac{\partial f(\bar{x})}{\partial a_j} = \lambda_j,$$

es decir cada multiplicador  $\lambda_j$  se interpreta como el impacto en el valor óptimo de un aumento en una unidad del lado derecho de la restricción respectiva. Esto puede ser muy útil al trabajar con un problema real y complejo, pues permite obtener rápidamente una sensibilidad respecto al impacto sobre el valor óptimo de la disponibilidad de recursos de cada tipo que limitan este valor. Por ejemplo, conocer los multiplicadores de lagrange (o “precios sombra”) asociados a restricciones de capacidad total de una máquina, de insumos disponibles, o de mano de obra contratada permite identificar en cuál de estos tres ítems resulta más conveniente invertir. Es importante destacar que esta predicción se desprende de una aproximación lineal del lagrangeano en torno al parámetro  $a$ . Así, sólo tiene validez para modificaciones leves de este valor.

Volviendo al ejemplo anterior del terreno a la orilla del río, ¿cuánto cambiará la función objetivo si ahora hubiese que comprar  $10.010 \text{ m}^2$  en vez de  $10.000 \text{ m}^2$ ?

Como  $\Delta \hat{v} = \lambda \Delta a$ , la función objetivo cambia  $\Delta \hat{v} \simeq (-0,97)(10) = -9,7$ . Volviendo al problema original de maximización, la función objetivo que representa los  $m^2$  útiles aumentará  $9,7 \text{ m}^2$ .

### 2.2.3 Caso 3: Restricciones de desigualdad

Para comenzar constatemus que todo problema de programación matemática continua (con funciones diferenciables) puede formularse bajo un formato estándar en que todas las restricciones son del tipo menor o igual y todas las variables deben ser no negativas. Es decir cualquier problema se puede formular del siguiente modo:

$$\begin{aligned} P) \quad & \text{Min } f(x_1, \dots, x_n) \\ \text{s.a} \quad & g_j(x_1, \dots, x_n) \leq b_j \quad j = 1, \dots, m \\ & x_1 \geq 0, \dots, x_n \geq 0 \end{aligned}$$

A continuación se describen algunas transformaciones necesarias para llevar cualquier problema a este formato:

- Cada restricción del tipo  $g(x) = b$ , puede transformarse en dos desigualdades del siguiente

modo:

$$\begin{aligned} g(x) &\leq b \\ -g(x) &\leq -b \end{aligned}$$

- Cada restricción del tipo  $g(x) \geq b$ , se puede reescribir acorde al formato estándar como:

$$-g(x) \leq -b$$

- Si el modelo exige que una variable sea no positiva, basta redefinir una nueva variable como el negativo de la anterior y reemplazarla en el problema. Esta nueva variable deberá cumplir la restricción de no negatividad.
- Si el modelo deja una variable libre (esto es que no está restringida a tomar valores no-negativos) se puede realizar las siguientes modificaciones que permitan modelar el problema bajo la estructura indicada: reemplazar  $x = x' - x''$  en el modelo y exigir que ambas variables nuevas ( $x'$  y  $x''$ ) sean no-negativas. Este cambio permite ajustarse a la estructura que exige que todas las variables sean no-negativas y no elimina soluciones. Note que cualquier valor positivo de  $x$  se representa por un par  $(x', x'')$  tal que  $x' > x''$ . Análogamente, cualquier valor negativo de  $x$  se representa por un par  $(x', x'')$  tal que  $x' < x''$ . Asimismo, resulta claro que cualquier punto factible en el modelo original (en que supongamos  $x = x_0$ ) puede representarse por una infinidad de puntos en el modelo nuevo puesto que habrá infinitas combinaciones de  $(x', x'')$  tal que  $x' - x'' = x_0$ .

En contraposición, es importante notar que toda restricción  $g(x) \leq b$ , puede convertirse en una restricción de igualdad introduciendo una nueva variable de *holgura*  $h$  no-negativa que represente la diferencia entre el lado derecho y el izquierdo de la desigualdad. De este modo, la restricción quedaría  $g(x) + h = b$ , con  $h \geq 0$ . Asimismo cada restricción del tipo  $g(x) \geq b$  puede convertirse en una restricción de igualdad introduciendo una variable de *exceso*  $e$  no negativa. De este modo, la restricción quedaría  $g(x) - e = b$ , con  $e \geq 0$ . Si  $h = 0$  (o  $e = 0$ ) para algún vector  $\bar{x}$ , se dice que la restricción correspondiente está *activa* en el punto  $\bar{x}$ . Si  $h \neq 0$  (o  $e \neq 0$ )  $\Rightarrow$  la restricción está *inactiva* en el punto  $\bar{x}$ .

### Análisis de optimalidad :

Al igual que en el caso unidimensional analizado en la sección 2.2.1, las condiciones de optimalidad difieren si se trata de un punto en el borde del dominio (en que al menos una restricción está activa) o si se trata de un punto interior (en que ninguna restricción está activa). Si supiéramos cuáles son las restricciones activas en el punto, bastaría con expresarlas como igualdades y, eliminando el resto (inactivas), emplear el método de Lagrange. Sin embargo típicamente se desconoce cuáles restricciones estarán activas en el punto óptimo del problema.

Para comenzar el análisis supongamos un problema unidimensional que cuenta con una sola restricción: no negatividad en la variable. Del análisis de la sección 2.2.1, observamos que si la derivada es estrictamente positiva (negativa) en  $x = 0$  entonces este punto es necesariamente un mínimo (máximo) local. Sin embargo, si la derivada es nula en  $x = 0$  es necesario analizar su segunda derivada para caracterizar el punto. Por otra parte, la derivada en cualquier punto  $x > 0$  que sea máximo o mínimo local debe ser nula. Así, se desprenden las siguientes condiciones necesarias de primer orden para caracterizar un punto mínimo  $\bar{x}$ :

$$\begin{aligned} f'(\bar{x}) &= 0 \text{ si } \bar{x} > 0 \\ f'(\bar{x}) &\geq 0 \text{ si } \bar{x} = 0, \end{aligned}$$

lo que matemáticamente puede expresarse del siguiente modo (note que ambas condiciones son satisfechas por un idéntico conjunto de puntos):

$$\begin{aligned} f'(\bar{x}) &\geq 0 \\ \bar{x}f'(\bar{x}) &= 0. \end{aligned}$$

Estas condiciones pueden extenderse directamente al caso de un problema multidimensional sin restricciones, salvo que todas sus variables deben ser no negativas. En ese caso las condiciones quedan:

$$\begin{aligned} \frac{\partial f(\bar{x})}{\partial x_i} &\geq 0 \quad \forall i \\ \frac{\partial f(\bar{x})}{\partial x_i} \bar{x}_i &= 0 \quad \forall i \end{aligned}$$

Al igual que en el caso unidimensional, estas condiciones son necesarias para optimalidad, pero no son suficientes. Si  $\frac{\partial f(\bar{x})}{\partial x_i} > 0$  el punto  $\bar{x}$  es un mínimo local, pero si  $\frac{\partial f(\bar{x})}{\partial x_i} = 0$  podría tratarse tanto de un máximo como de un mínimo local. Es importante destacar que estas condiciones de optimalidad rigen sólo para aquellas variables del problema que estén restringidas a tomar valores no negativos. Si algunas de las variables no están sujetas a estas restricciones, para esas variables las condiciones de optimalidad son simplemente  $\frac{\partial f(\bar{x})}{\partial x_i} = 0$ .

Consideremos ahora un problema un poco más general en que se agregan algunas restricciones de igualdad al problema. Es decir:

$$\begin{aligned} &P) \text{ Min } f(x) \\ \text{s.a} \quad &g_j(x) = b_j \quad \forall j \\ &x_i \geq 0 \quad \forall i \end{aligned}$$

Se podría definir un Lagrangeano que incorpore las restricciones complejas dejando sólo las

restricciones de no-negatividad en el dominio:

$$\begin{aligned} \text{Min } L(x, \mu) &= f(x) + \sum_{j=1}^m \mu_j (g_j(x) - b_j) \\ \text{s.a. } x_i &\geq 0 \quad \forall i \end{aligned}$$

Note que hemos optado por invertir el orden en que se incorpora cada restricción al Lagrangeano por motivos que más adelante serán aparentes. Por lo pronto esto tiene como único impacto invertir la interpretación que se le da al multiplicador  $\mu_j$ : ahora corresponde a cuánto cae la función objetivo si  $b_j$  aumenta en una unidad.

Note también que en este nuevo problema en que se minimiza el Lagrangeano sólo algunas de sus variables (i.e., los  $x_i$ ) están restringidas a tomar valores no negativos (los  $\mu_j$  están libres). Así, debemos considerar las condiciones apropiadas de optimalidad en cada caso. Estas serían:

$$\begin{aligned} \frac{\partial L}{\partial x_i} &\geq 0 \rightarrow \frac{\partial f}{\partial x_i} + \sum_{j=1}^m \mu_j \frac{\partial g_j}{\partial x_i} \geq 0 \quad \forall i = 1, \dots, n \\ x_i \frac{\partial L}{\partial x_i} &= 0 \rightarrow x_i \left( \frac{\partial f}{\partial x_i} + \sum_{j=1}^m \mu_j \frac{\partial g_j}{\partial x_i} \right) = 0 \quad \forall i = 1, \dots, n \\ \frac{\partial L}{\partial \mu_j} &= 0 \rightarrow g_j(\bar{x}) - b_j = 0 \quad \forall j = 1, \dots, m \end{aligned}$$

Además se debe cumplir con las restricciones de no negatividad que en este caso son:

$$x_i \geq 0 \quad \forall i = 1, \dots, n.$$

Estas cuatro familias de condiciones describen las condiciones necesarias de primer orden para optimalidad de puntos regulares en este caso. Esto nos deja a un paso de identificar las condiciones generales para un problema no lineal de programación matemática continua.

Considere el siguiente problema:

$$\begin{aligned} P) \text{ Min } f(x) \\ \text{s.a. } g_j(x) &\leq b_j \quad \forall j = 1, \dots, m \\ x_i &\geq 0 \quad \forall i = 1, \dots, n \end{aligned}$$

Para determinar las condiciones necesarias de primer orden para optimalidad en este caso general, consideraremos variables auxiliares de holgura  $y_j$  (no negativas) para cada una de las  $m$

restricciones que nos permitirán reformular el problema del siguiente modo equivalente:

$$\begin{aligned} & \tilde{P}) \text{ Min } f(x) \\ \text{s.a. } & g_j(x) + y_j = b_j \quad \forall j = 1, \dots, m \\ & x_i \geq 0, \quad \forall i = 1, \dots, n \\ & y_j \geq 0 \quad \forall j = 1, \dots, m \end{aligned}$$

Este problema es similar al explicado en los párrafos previos por lo que se puede identificar las condiciones que deberá satisfacer todo mínimo local a partir del siguiente problema:

$$\begin{aligned} \text{Min } L(x, y, \mu) &= f(x) + \sum_{j=1}^m \mu_j (g_j(x) + y_j - b_j) \\ \text{s.a. } & x_i \geq 0 \quad \forall i \\ & y_j \geq 0 \quad \forall j \end{aligned}$$

Las condiciones son:

$$\begin{aligned} \frac{\partial L}{\partial x_i} &= \frac{\partial f}{\partial x_i} + \sum_{j=1}^m \mu_j \frac{\partial g_j}{\partial x_i} \geq 0 & \forall i = 1, \dots, n \\ x_i \frac{\partial L}{\partial x_i} &= x_i \left( \frac{\partial f}{\partial x_i} + \sum_{j=1}^m \mu_j \frac{\partial g_j}{\partial x_i} \right) = 0 & \forall i = 1, \dots, n \\ \frac{\partial L}{\partial y_j} &= \mu_j \geq 0 & \forall j = 1, \dots, m \\ y_j \frac{\partial L}{\partial y_j} &= y_j \mu_j = 0 & \forall j = 1, \dots, m \\ \frac{\partial L}{\partial \mu_j} &= g_j(x) + y_j - b_j = 0 & \forall j = 1, \dots, m \\ x_i &\geq 0 & \forall i = 1, \dots, n \\ y_j &\geq 0 & \forall j = 1, \dots, m \end{aligned} \tag{2.8}$$

Ahora, de la quinta familia de condiciones, se puede eliminar la variable  $y_j$  de la formulación reemplazando  $y_j = b_j - g_j(x)$ . De este modo las familias de condiciones cuarta y séptima quedan:

$$\begin{aligned} [b_j - g_j(x)] \mu_j &= 0 & \forall j = 1, \dots, m \\ b_j - g_j(x) &\geq 0 & \forall j = 1, \dots, m \end{aligned} \tag{2.9}$$

Con lo anterior se puede plantear las condiciones generales de optimalidad desarrolladas por Karush (1939) y Kuhn y Tucker (1951).<sup>3</sup>

---

<sup>3</sup>Estas condiciones fueron conocidas por más de una década como las condiciones de Kuhn-Tucker. Esto debido a que en 1951 mientras Harold W. Kuhn desarrollaba su tesis doctoral en Princeton bajo la supervisión del profesor

**Proposición 34** (*Condiciones de Karush-Kuhn-Tucker (KKT)*) Supóngase que  $f(x)$ ,  $g_j(x)$  con  $j = 1, \dots, m$ , son funciones diferenciables que satisfacen ciertas condiciones de regularidad. Entonces,  $\bar{x}$  puede ser una solución óptima para el problema de programación no-lineal, sólo si existen  $m$  números  $\mu_1, \dots, \mu_m$  que satisfagan todas las condiciones (necesarias) siguientes:

$$\begin{aligned} \frac{\partial f(\bar{x})}{\partial x_i} + \sum_{j=1}^m \mu_j \frac{\partial g_j(\bar{x})}{\partial x_i} &\geq 0 & \forall i = 1, \dots, n \\ \bar{x}_i \left( \frac{\partial f(\bar{x})}{\partial x_i} + \sum_{j=1}^m \mu_j \frac{\partial g_j(\bar{x})}{\partial x_i} \right) &= 0 & \forall i = 1, \dots, n \\ g_j(\bar{x}) - b_j &\leq 0 & \forall j = 1, \dots, m \\ \mu_j (g_j(\bar{x}) - b_j) &= 0 & \forall j = 1, \dots, m \\ \bar{x}_i &\geq 0 & \forall i = 1, \dots, n \\ \mu_j &\geq 0 & \forall j = 1, \dots, m. \end{aligned}$$

Es importante notar que el cumplimiento de estas condiciones no garantiza que la solución sea óptima (todo mínimo local las satisface). Se necesitan ciertas suposiciones de convexidad adicionales para obtener esta garantía.

En este caso los multiplicadores  $\mu_j$  nos indican cuánto baja la función objetivo si  $b_j$  aumenta en una unidad. Como la función objetivo debe ser minimizada, esto indica que sólo se puede mejorar a través de aumentar  $b_j$ . Esto es natural puesto que aumentar  $b_j$  implica expandir el dominio, esto es expandir el espectro de posibilidades de donde escoger.

La cuarta condición se denomina *complementariedad de las holguras*, y exige que si la restricción de desigualdad está activa ( $g_j(\bar{x}) - b_j = 0$ ), su holgura sea cero y su multiplicador asociado ( $\mu_j$ ) tome cualquier valor no negativo. Indicando que si se expande el dominio relajando esa restricción, la solución óptima cambiaría y la función objetivo seguramente mejoraría. Por otra parte, si la restricción está inactiva ( $g_j(\bar{x}) - b_j < 0$ ), la holgura es estrictamente positiva y es el multiplicador asociado quien toma un valor nulo. Esto indica que si se relaja esta restricción no se modificará la optimalidad local del punto  $\bar{x}$  identificado.

Es interesante recordar que en el caso de problemas de optimización sólo con restricciones de igualdad, la condición necesaria de primer orden puede interpretarse de la siguiente manera: el gradiente de la función objetivo en el punto óptimo es igual a una combinación lineal de los gradientes de las restricciones en ese mismo punto. En el caso de problemas con restricciones de desigualdad, esto aún es cierto, pero en este caso el gradiente de la función objetivo en el punto óptimo debe equivaler a una combinación lineal de los gradientes de las restricciones *activas* ( $\mu_j > 0$ )

---

Albert W. Tucker, descubrió estas condiciones de optimalidad. En la década de los 60s, alguien encontró en una biblioteca una tesis de master desarrollada en la Universidad de Chicago en 1939 por William Karush, donde establecía exactamente las mismas condiciones. Este trabajo nunca fue publicado debido a la poca importancia que le dio su profesor guía en ese minuto. Apenas esto se supo, Kuhn y Tucker antepusieron a Karush en el nombre de estas condiciones y de ahí en adelante de habla de las condiciones KKT.

en el punto. Por ejemplo, en un punto interior al dominio todos los  $\mu_j$  son nulos (las restricciones están inactivas) y por tanto la condición de optimalidad se traduce en que el gradiente de la función objetivo sea nulo en el punto.

Mencionamos que si el problema original incorpora restricciones de igualdad o variables libres, el problema puede ser transformado en uno equivalente en que sólo haya restricciones de menor o igual y todas las variables deban tomar valores no negativos. Sin embargo esta transformación conlleva un costo ya que se expande el número de restricciones o de variables en el problema. Alternativamente se puede modificar levemente las condiciones de optimalidad de KKT para tratar directamente estas características del problema. Consideremos el siguiente problema:

$$\begin{aligned}
 P) \text{ Min } & f(x) \\
 \text{s.a } & g_j(x) \leq b_j \quad \forall j = 1, \dots, p \\
 & g_j(x) = b_j \quad \forall j = p+1, \dots, m \\
 & x_i \geq 0 \quad \forall i = 1, \dots, q \\
 x_i \text{ libre} & \quad \forall i = q+1, \dots, n
 \end{aligned}$$

En este caso se ha considerado un conjunto de restricciones de igualdad y algunas variables libres. En ese caso es fácil demostrar a partir del procedimiento desarrollado en este texto que las condiciones de optimalidad son las siguientes:

$$\begin{aligned}
 \frac{\partial f(\bar{x})}{\partial x_i} + \sum_{j=1}^m \mu_j \frac{\partial g_j(\bar{x})}{\partial x_i} & \geq 0 \quad \forall i = 1, \dots, q \\
 \bar{x}_i \left( \frac{\partial f(\bar{x})}{\partial x_i} + \sum_{j=1}^m \mu_j \frac{\partial g_j(\bar{x})}{\partial x_i} \right) & = 0 \quad \forall i = 1, \dots, q \\
 \frac{\partial f(\bar{x})}{\partial x_i} + \sum_{j=1}^m \mu_j \frac{\partial g_j(\bar{x})}{\partial x_i} & = 0 \quad \forall i = q+1, \dots, n \\
 g_j(\bar{x}) - b_j & \leq 0 \quad \forall j = 1, \dots, p \\
 \mu_j (g_j(\bar{x}) - b_j) & = 0 \quad \forall j = 1, \dots, p \\
 g_j(\bar{x}) - b_j & = 0 \quad \forall j = p+1, \dots, m \\
 \bar{x}_i & \geq 0 \quad \forall i = 1, \dots, q \\
 \mu_j & \geq 0 \quad \forall j = 1, \dots, p.
 \end{aligned}$$

Es decir sólo se agregan dos familias de restricciones asociadas a los cambios efectuados: para las variables libres sólo se exige que la derivada parcial del Lagrangeano con respecto a esas variables se a nulo y se exige que se cumplan las restricciones de igualdad. Note que los multiplicadores asociados a estas restricciones quedan libres.

Las condiciones de optimalidad de Karush-Kuhn-Tucker son condiciones necesarias y sólo garantizarían optimalidad global si se cumplen adicionalmente otras condiciones de convexidad. Con-



sidere el siguiente corolario:

**Corolario 35** *Supóngase que  $f(x)$  es una función convexa y diferenciable, y que  $g_1(x), g_2(x), \dots, g_m(x)$  también lo son en donde todas estas funciones satisfacen las condiciones de regularidad. Entonces  $\bar{x} = (\bar{x}_1, \dots, \bar{x}_n)$  es una solución óptima si y sólo si se satisfacen todas las condiciones del teorema.*

Es importante destacar que el método, derivado del método de Lagrange, identifica puntos óptimos locales que cumplan condiciones de regularidad (tal como lo exige el corolario anterior). En este caso, estas condiciones consisten en que los gradientes de las restricciones *activas* en el punto sean linealmente independientes. Es decir, si el punto óptimo al problema es tal que el jacobiano de los gradientes de las restricciones activas en el punto no es de rango máximo, entonces este método no garantiza que se encuentre el punto óptimo.

Es posible identificar ciertas condiciones de problemas que garanticen regularidad en todo el dominio. Considere las siguientes proposiciones.

**Proposición 36** (Slater) *Las condiciones de regularidad de Slater establecen que si hay un dominio  $D = \{x : g_j(x) \leq 0\}$  con  $g_j(x)$  funciones convexas, y existe un punto  $\bar{x} \in D$  tal que  $g_j(\bar{x}) < 0$   $\forall j = 1, \dots, m$ , entonces todo punto  $x \in D$  es regular.*

**Proposición 37** (Slater-Usawa) *Las condiciones de regularidad de Slater-Usawa establecen que si hay un dominio  $D = \{x : g_j(x) \leq 0\}$  con  $g_j(x)$  funciones convexas, y existe un punto  $\bar{x} \in D$  tal que  $g_j(\bar{x}) < 0$  para toda restricción no lineal, y  $g_j(\bar{x}) \leq 0$  para toda restricción lineal, entonces todo punto  $x \in D$  es regular.*

**Corolario 38** *En problemas lineales basta que haya un punto factible para decir que su dominio es regular.*

#### Solución a condiciones de Karush-Kuhn-Tucker :

Si bien las condiciones de KKT expresan un sistema finito de ecuaciones que basta resolver para identificar el conjunto de puntos que las satisface, esto suele exigir mucho tiempo y trabajo. Esto se debe a que para cada restricción de desigualdad o de no negatividad se tendrá distintas condiciones dependiendo de si la restricción se asume activa o inactiva. Esto motiva las condiciones del tipo  $xz = 0$  presentes en las condiciones KKT. La única forma de abordarlas es considerar dos casos separados:  $x = 0$  y  $z = 0$  y en la medida que se tenga muchas condiciones de este tipo se deberá analizar separadamente cada una de las combinaciones posibles que se pueden presentar. Así, resolver el problema exige recorrer un árbol binario de decisiones en que en cada nodo se debe abordar una de estas condiciones y tomar dos caminos separados (e.g.  $x = 0$  en una dirección y  $x \neq 0$  en la otra). Por lo tanto cada nodo del árbol lleva asociado un conjunto de decisiones para satisfacer las condiciones ya abordadas. Así, por ejemplo si se tiene un problema de tres variables ( $x_1, x_2$  y  $x_3$ ), todas restringidas a tomar valores no negativos, se deberá analizar (en el peor de los

casos) los siguientes ocho casos:

$$\begin{aligned}
 x_1 &= 0; x_2 \neq 0; x_3 \neq 0 \\
 x_1 &= 0; x_2 = 0; x_3 \neq 0 \\
 x_1 &= 0; x_2 \neq 0; x_3 = 0 \\
 x_1 &= 0; x_2 = 0; x_3 = 0 \\
 x_1 &\neq 0; x_2 \neq 0; x_3 \neq 0 \\
 x_1 &\neq 0; x_2 = 0; x_3 \neq 0 \\
 x_1 &\neq 0; x_2 \neq 0; x_3 = 0 \\
 x_1 &\neq 0; x_2 = 0; x_3 = 0.
 \end{aligned}$$

Sin embargo, al recorrer este árbol típicamente es posible descartar algunos escenarios antes de revisarlos en detalle como se verá en los ejemplos al final de esta sección.

**Ejemplo 19** *Considere el siguiente problema:*

$$\begin{aligned}
 P) \quad & \text{Max } f(x_1, x_2) = \ln(x_1 + 1) + x_2 \\
 s.a \quad & 2x_1 + x_2 \leq 3 \\
 & x_1, x_2 \geq 0
 \end{aligned}$$

El problema puede modificarse del siguiente modo:

$$\begin{aligned}
 P) \quad & \text{Min } -\ln(x_1 + 1) - x_2 \quad \rightarrow -f(x) \text{ es convexa} \\
 s.a \quad & 2x_1 + x_2 \leq 3 \quad \rightarrow g(x) \text{ es convexa} \\
 & x_1, x_2 \geq 0
 \end{aligned}$$

Resulta claro que el problema presenta características que permiten garantizar que el método de KKT identificará sólo un punto y que éste será el óptimo global. Además, el dominio está definido por restricciones lineales por lo que no habrá puntos singulares. El Lagrangeano del problema queda:

$$L = -\ln(x_1 + 1) - x_2 + \mu_1(2x_1 + x_2 - 3)$$

Resolviendo las condiciones de KKT es posible obtener una solución óptima.

$$\begin{aligned}
 1a. \quad & \frac{-1}{x_1 + 1} + 2\mu_1 \geq 0 \\
 2a. \quad & x_1 \left( \frac{-1}{x_1 + 1} + 2\mu_1 \right) = 0 \\
 1b. \quad & -1 + \mu_1 \geq 0 \\
 2b. \quad & x_2(-1 + \mu_1) = 0 \\
 3. \quad & 2x_1 + x_2 \leq 3 \\
 4. \quad & \mu_1(2x_1 + x_2 - 3) = 0 \\
 5. \quad & x_1 \geq 0, \ x_2 \geq 0 \\
 6. \quad & \mu_1 \geq 0
 \end{aligned}$$

De donde es fácil ver que:

- $\mu_1 \geq 1$  de la condición 1b.
- $x_1 \geq 0$  de la condición 5.
- Por lo tanto,  $\frac{-1}{x_1 + 1} + 2\mu_1 > 0$ .
- Por lo tanto,  $x_1 = 0$  de la condición 2a.
- $\mu_1 \neq 0$  implica que  $2x_1 + x_2 - 3 = 0$ , de la condición 4.
- Los pasos 3 y 4 implican que  $x_2 = 3$ .
- $x_2 \neq 0$  implica que  $\mu_1 = 1$ , de la condición 2b.
- Los valores  $x_1 = 0$ ,  $x_2 = 3$  y  $\mu_1 = 1$  no violan ninguna condición.
- Por lo tanto,  $\bar{x} = (0, 3)$  es solución óptima para este problema.

## 2.3 Problemas resueltos

### 2.3.1 Optimización de una función sin restricciones

**Problema 39 (Condiciones Necesarias y Suficientes (I2 1°03/Ex1°03))** ¿Qué condiciones deben cumplir los parámetros  $a$  y  $b$  para que un punto de la función  $f(x, y) = x^2 + axy + by^2 + x + y$  cumpla con las condiciones necesarias, **pero no suficientes** para ser mínimo?

**Solución:** Las condiciones necesarias son que los  $\Delta_i \geq 0$ . Por lo que habrá que construir la matriz Hessiana.

$$\begin{aligned}\frac{\partial f}{\partial x} &= 2x + ay + 1 \\ \frac{\partial f}{\partial y} &= ax + 2by + 1\end{aligned}$$

Luego,  $H = \begin{bmatrix} 2 & a \\ a & 2b \end{bmatrix}$ . Entonces, para que se cumpla con las condiciones necesarias, como  $\Delta_1 = 2 > 0$ ,  $\Delta_2 = 4b - a^2 \geq 0$ .

Las condiciones suficientes son que los  $\Delta_i > 0$ , pero queremos que estas condiciones no se cumplan, por lo que algún  $\Delta_i = 0$ . Como  $\Delta_1 = 2 > 0$ , entonces  $\Delta_2 = 4b - a^2 = 0$ .

Por lo tanto, se concluye que la condición es  $4b = a^2$ .

**Problema 40** Considere el siguiente problema de optimización:

$$\text{Min} f(x) = x^4 - 4x^3 - 48x^2 + 480x - 160$$

Realice 4 iteraciones con el método de Newton comenzando en  $x_0 = 5$  y luego repita pero con  $x_0 = -5$ .

**Solución:**

En primer lugar se ve que

$$\frac{\partial f}{\partial x} = 4x^3 - 12x^2 + 96x + 480y \quad \frac{\partial^2 f}{\partial x^2} = 12x^2 - 24x + 96$$

a) Punto inicial:  $x_0 = 5$

$$\begin{aligned}x_1 &= x_0 - \frac{f'(x_0)}{f''(x_0)} \text{ con } x_0 = 5 \Rightarrow x_1 = 2.62 \\ x_2 &= x_1 - \frac{f'(x_1)}{f''(x_1)} \Rightarrow x_2 = 5.4 \\ x_3 &= 3.566 \\ x_4 &= 9.308\end{aligned}$$

b) Punto inicial:  $x_0 = -5$

$$\begin{aligned}x_0 &= -5 \\ x_1 &= -5.494 \\ x_2 &= -5.4485 \\ x_3 &= -5.4481 \\ x_4 &= -5.4481\end{aligned}$$

**Problema 41** Considere los siguientes problemas de optimización:

a)

$$\text{Min } f(x, y) = x^2 - 4y - 2xy + 2y^2$$

b)

$$\text{Min } f(x, y) = 5x^2 + 5y^2 - xy - 11x + 11y + 11$$

Utilizando el método del gradiente realice dos iteraciones, tomando como punto de partida el punto  $(x, y) = (0, 0)$ .

**Solución:**

a) En primer lugar se ve que

$$\frac{\partial f}{\partial x} = 2x - 2y, \quad \frac{\partial f}{\partial y} = -4 - 2x + 4y$$

A partir de esto, se realizan las iteraciones:

i)

$$x_1 = x_0 - \alpha \vec{\nabla} f(x_0) \Rightarrow x_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \alpha \begin{pmatrix} 0 \\ -4 \end{pmatrix} = \begin{pmatrix} 0 \\ 4\alpha \end{pmatrix}$$

$$\begin{aligned} \text{Min}_{\alpha \geq 0} f(0, 4\alpha) &= -16\alpha + 32\alpha^2 \\ \frac{\partial f}{\partial \alpha} &= -16 + 64\alpha \equiv 0 \Rightarrow \alpha^* = 1/4, \\ \frac{\partial^2 f}{\partial \alpha^2} &= 64 > 0 \Rightarrow \alpha^* \text{ es mínimo global.} \\ &\Rightarrow x_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{aligned}$$

ii)

$$x_2 = x_1 - \alpha \vec{\nabla} f(x_1) \Rightarrow x_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \alpha \begin{pmatrix} -2 \\ 0 \end{pmatrix} = \begin{pmatrix} 2\alpha \\ 1 \end{pmatrix}$$

$$\begin{aligned} \text{Min}_{\alpha \geq 0} f(2\alpha, 1) &= 4\alpha^2 - 4 - 4\alpha + 2 \\ \frac{\partial f}{\partial \alpha} &= 8\alpha - 4 \equiv 0 \Rightarrow \alpha^* = 1/2, \\ \frac{\partial^2 f}{\partial \alpha^2} &= 8 > 0 \Rightarrow \alpha^* \text{ es mínimo global.} \\ &\Rightarrow x_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \end{aligned}$$

Se necesita seguir iterando para llegar al punto óptimo.

b) Del mismo modo, se aprecia que:

$$\frac{\partial f}{\partial x} = 10x - y - 11, \quad \frac{\partial f}{\partial y} = 10y - x + 11$$

Nuevamente, a partir de estas derivadas, se realizan las iteraciones:

i)

$$x_1 = x_0 - \alpha \vec{\nabla} f(x_0) \Rightarrow x_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \alpha \begin{pmatrix} -11 \\ 11 \end{pmatrix} = \begin{pmatrix} 11\alpha \\ -11\alpha \end{pmatrix}$$

$$\text{Min}_{\alpha \geq 0} f(11\alpha, -11\alpha) = 5(11\alpha)^2 + 5(-11\alpha)^2 + 121\alpha^2 - 121\alpha - 121\alpha + 11$$

$$\frac{\partial f}{\partial \alpha} = 2 \times 110 \times 11\alpha + 242\alpha - 242 \equiv 0 \Rightarrow \alpha^* = 1/11,$$

$$\frac{\partial^2 f}{\partial \alpha^2} > 0 \Rightarrow \alpha^* \text{ es mínimo global.}$$

$$\Rightarrow x_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

ii)

$$x_2 = x_1 - \alpha \vec{\nabla} f(x_1) \Rightarrow x_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} - \alpha \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Por lo que  $x_2 = x_1 \Rightarrow$  se obtuvo el punto óptimo  $(1, -1)$ .

**Problema 42** Determine el mínimo de la función

$$f(x, y, z) = x^2 + y^2 + 2z^2$$

utilizando para ello el método de Newton y el método del gradiente, y como punto inicial  $(2, -2, 1)$ . Compare ambos resultados.

**Solución:**

a) Método de Newton

En primer lugar se ve que

$$\frac{\partial f}{\partial x} = 2x, \quad \frac{\partial f}{\partial y} = 2y, \quad \frac{\partial f}{\partial z} = 4z$$

El hessiano estará dado por

$$H = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix} \Rightarrow H^{-1} = \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/4 \end{bmatrix}, \quad \forall x, y, z$$

Y las iteraciones estarán dadas por

$$\begin{aligned} x_1 &= x_0 - H^{-1}(f(x_0)) \cdot \nabla f(x_0) \text{ con } x_0 = (2, -2, 1) \\ \Rightarrow x_1 &= \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} - \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/4 \end{bmatrix} \begin{pmatrix} 4 \\ -4 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} x_2 &= x_1 - H^{-1}(f(x_1)) \cdot \nabla f(x_1) \text{ con } x_0 = (0, 0, 0) \\ \Rightarrow x_2 &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} - \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/4 \end{bmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \end{aligned}$$

Por lo que  $x_2 = x_1 \Rightarrow$  llegamos al punto extremo  $(0, 0, 0)$ .

b) Método del Gradiente

Las iteraciones estarán dadas por,

i)

$$x_1 = x_0 - \alpha \vec{\nabla} f(x_0) \Rightarrow x_1 = \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} - \alpha \begin{pmatrix} 4 \\ -4 \\ 4 \end{pmatrix} = \begin{pmatrix} 2 - 4\alpha \\ -2 + 4\alpha \\ 1 - 4\alpha \end{pmatrix}$$

$$\begin{aligned} \text{Min}_{\alpha \geq 0} f(x_1) &= (2 - 4\alpha)^2 + (-2 + 4\alpha)^2 + 2(1 - 4\alpha)^2 \\ \frac{\partial f}{\partial \alpha} &= -8(2 - 4\alpha) + 8(-2 + 4\alpha) - 16(1 - 4\alpha) \equiv 0 \Rightarrow \alpha^* = 3/8, \\ \frac{\partial^2 f}{\partial \alpha^2} &= 8 \cdot 4 + 8 \cdot 4 + 16 \cdot 4 > 0 \Rightarrow \alpha^* \text{ es mínimo global.} \\ \Rightarrow x_1 &= \begin{pmatrix} 1/2 \\ -1/2 \\ -1/2 \end{pmatrix} \end{aligned}$$

ii)

$$x_2 = x_1 - \alpha \vec{\nabla} f(x_1) \Rightarrow x_2 = \begin{pmatrix} 1/2 \\ -1/2 \\ -1/2 \end{pmatrix} - \alpha \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix} = \begin{pmatrix} 1/2 - \alpha \\ -1/2 + \alpha \\ -1/2 + 2\alpha \end{pmatrix}$$

$$\begin{aligned}
\text{Min}_{\alpha \geq 0} f(x_2) &= (1/2 - \alpha)^2 + (-1/2 + \alpha)^2 + 2(-1/2 + 2\alpha)^2 \\
\frac{\partial f}{\partial \alpha} &= -2(1/2 - \alpha) + 2(-1/2 + \alpha) + 8(-1/2 + 2\alpha) \equiv 0 \Rightarrow \alpha^* = 3/10, \\
\frac{\partial^2 f}{\partial \alpha^2} &= 2 + 2 + 16 = 20 > 0 \Rightarrow \alpha^* \text{ es mínimo global.} \\
\Rightarrow x_2 &= \begin{pmatrix} 1/5 \\ -1/5 \\ 1/10 \end{pmatrix}
\end{aligned}$$

iii)

$$x_3 = x_2 - \alpha \vec{\nabla} f(x_2) \Rightarrow x_3 = \begin{pmatrix} 1/5 \\ -1/5 \\ 1/10 \end{pmatrix} - \alpha \begin{pmatrix} 2/5 \\ -2/5 \\ 2/5 \end{pmatrix} = \begin{pmatrix} 1/5 - 2/5\alpha \\ -1/5 + 2/5\alpha \\ 1/10 - 2/5\alpha \end{pmatrix}$$

$$\begin{aligned}
\text{Min}_{\alpha \geq 0} f(x_2) &= \frac{1}{25} [(1 - 2\alpha)^2 + (-1 + 2\alpha)^2 + 2(1/2 - 2\alpha)^2] \\
\frac{\partial f}{\partial \alpha} &= \frac{1}{25} [-4(1 - 2\alpha) + 4(-1 + 2\alpha) - 8(1/2 - 2\alpha)] \equiv 0 \Rightarrow \alpha^* = 3/8, \\
\frac{\partial^2 f}{\partial \alpha^2} &= \frac{1}{25} [8 + 8 + 16] > 0 \Rightarrow \alpha^* \text{ es mínimo global.} \\
\Rightarrow x_3 &= \begin{pmatrix} 1/5 - 3/20 \\ -1/5 + 3/20 \\ 1/10 - 3/20 \end{pmatrix} = \begin{pmatrix} 1/20 \\ -1/20 \\ -1/20 \end{pmatrix} \quad \dots etc.
\end{aligned}$$

Ambos métodos llegarán al mismo punto óptimo, pero a través del método de Newton sólo requirió 2 iteraciones, mientras que por el método del gradiente se hicieron 3 iteraciones, y aún no se logra llegar al óptimo. La función objetivo representa una elipse, por lo que el método del gradiente converge muy lentamente.

**Problema 43** Determine la relación que debe existir entre los parámetros  $a$  y  $b$ , para que la función

$$f(x) = e^{ax} \cos(bx)$$

llegue a su punto óptimo en la segunda iteración a través del método de Newton, utilizando como punto inicial  $x_0 = 2\pi/b$ .

**Solución:**

En primer lugar, se ve que

$$\begin{aligned}
\frac{\partial f}{\partial x} &= ae^{ax} \cos(bx) - be^{ax} \sin(bx) \\
\frac{\partial^2 f}{\partial x^2} &= a^2 e^{ax} \cos(bx) - 2abe^{ax} \sin(bx) - b^2 e^{ax} \cos(bx)
\end{aligned}$$



Luego

$$\begin{aligned}
 x_1 &= x_0 - \frac{f'(x_0)}{f''(x_0)} \text{ con } x_0 = 2\pi/b \\
 \Rightarrow x_1 &= x_0 - \frac{a \cos(bx_0) - b \sin(bx_0)}{(a^2 - b^2) \cos(bx_0) - 2ab \sin(bx_0)} = \frac{2\pi}{b} - \frac{a}{a^2 - b^2} \\
 \\
 x_2 &= x_1 - \frac{f'(x_1)}{f''(x_1)} \\
 \Rightarrow x_2 &= x_1 - \frac{a \cos\left(-\frac{ab}{a^2 - b^2}\right) - b \sin\left(-\frac{ab}{a^2 - b^2}\right)}{(a^2 - b^2) \cos\left(-\frac{ab}{a^2 - b^2}\right) - 2ab \sin\left(-\frac{ab}{a^2 - b^2}\right)} \equiv x_1
 \end{aligned}$$

Así, para encontrar el óptimo en la segunda iteración, se necesita

$$a \cos\left(-\frac{ab}{a^2 - b^2}\right) - b \sin\left(-\frac{ab}{a^2 - b^2}\right) = 0 \Rightarrow \frac{a}{b} = \tan\left(\frac{-ab}{a^2 - b^2}\right)$$

Se puede corroborar esta solución buscando el punto crítico con el gradiente,  $x^* = \frac{1}{b} \arctan\left(\frac{a}{b}\right)$ . Al introducir la condición que deben cumplir los parámetros a y b, el punto crítico toma el valor de  $x_1$ .

**Problema 44** Determine el mínimo de la función

$$f(x, y, z) = 5x^2 + 2y^2 + 2z^2 + 2xy + 2yz - 2xz - 6z$$

utilizando el método del gradiente, y considerando como punto de inicio  $x_0 = (0, 0, 0)$ .

**Solución:**

En primer lugar se obtienen las derivadas parciales, resultando

$$\frac{\partial f}{\partial x} = 10x + 2y - 2z, \quad \frac{\partial f}{\partial y} = 4y + 2x + 2z, \quad \frac{\partial f}{\partial z} = 4z + 2y - 2x - 6$$

Luego se comienzan a realizar las iteraciones:

i)

$$x_1 = x_0 - \alpha \vec{\nabla} f(x_0) \Rightarrow x_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} - \alpha \begin{pmatrix} 0 \\ 0 \\ -6 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 6\alpha \end{pmatrix}$$

$$\begin{aligned}
\text{Min}_{\alpha \geq 0} f(x_1) &= 2(6\alpha)^2 - 36\alpha \\
\frac{\partial f}{\partial \alpha} &= 24(6\alpha) - 36 \equiv 0 \Rightarrow \alpha^* = 1/4, \\
\frac{\partial^2 f}{\partial \alpha^2} &= 24 \cdot 6 > 0 \Rightarrow \alpha^* \text{ es mínimo global.} \\
\Rightarrow x_1 &= \begin{pmatrix} 0 \\ 0 \\ 3/2 \end{pmatrix}
\end{aligned}$$

ii)

$$x_2 = x_1 - \alpha \vec{\nabla} f(x_1) \Rightarrow x_2 = \begin{pmatrix} 0 \\ 0 \\ 3/2 \end{pmatrix} - \alpha \begin{pmatrix} -3 \\ 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 3\alpha \\ -3\alpha \\ 3/2 \end{pmatrix}$$

$$\begin{aligned}
\text{Min}_{\alpha \geq 0} f(x_2) &= 5(3\alpha)^2 + 2(3\alpha)^2 + 9/2 - 18\alpha^2 - 9\alpha - 9\alpha - 9 \\
\frac{\partial f}{\partial \alpha} &= 30 \cdot 3\alpha - 18 \equiv 0 \Rightarrow \alpha^* = 1/5, \\
\frac{\partial^2 f}{\partial \alpha^2} &= 90 > 0 \Rightarrow \alpha^* \text{ es mínimo global.} \\
\Rightarrow x_2 &= \begin{pmatrix} 3/5 \\ -3/5 \\ 3/2 \end{pmatrix}
\end{aligned}$$

iii)

$$x_3 = x_2 - \alpha \vec{\nabla} f(x_2) \Rightarrow x_3 = \begin{pmatrix} 3/5 \\ -3/5 \\ 3/2 \end{pmatrix} - \alpha \begin{pmatrix} 9/5 \\ 9/5 \\ -12/5 \end{pmatrix} = \begin{pmatrix} 3/5 - 9\alpha/5 \\ -3/5 - 9\alpha/5 \\ 3/2 + 12\alpha/5 \end{pmatrix}$$

$$\begin{aligned}
\text{Min}_{\alpha \geq 0} f(x_3) &= 5(3/5 - 9\alpha/5)^2 + 2(-3/5 - 9\alpha/5)^2 + 2(3/2 + 12\alpha/5)^2 \\
&\quad + 2(3/5 - 9\alpha/5)(-3/5 - 9\alpha/5) - 2(-3/5 - 9\alpha/5)(3/2 + 12\alpha/5) \\
&\quad - 2(3/5 - 9\alpha/5)(3/2 + 12\alpha/5) - 6(3/2 + 12\alpha/5)
\end{aligned}$$

Como se aprecia, un problema puede complicarse bastante. En este caso, por el tipo de función, el método no logra encontrar fácilmente la solución. Por otro lado, si se estudia el hessiano de esta función se obtiene fácilmente el punto (1,-2,3) como mínimo global del problema.

### 2.3.2 Optimización de una función con restricciones

**Problema 45 (Lagrange (Ex.1°03))** Dado el siguiente problema:

$$\begin{aligned}
P) \quad & \text{Min } 2xy - 6y - 5x + 15 \\
s.a. \quad & 500 = xy
\end{aligned}$$

(a) ¿Cuáles son las condiciones de Lagrange?

(b) Si cambio la restricción a  $510 = xy$ , sin resolver el problema, ¿cuánto varía la función objetivo y cuánto la solución óptima? ¿Es exacta aquella variación? ¿Es positiva o negativa la variación?

**Solución:** (a) Las Condiciones de Lagrange son:

$$L(x, y, z) = 2xy - 6y - 5x + 15 + \lambda(500 - xy)$$

$$\begin{aligned}\frac{\partial L}{\partial x} &= 2y - 5 - \lambda y = 0 \\ \frac{\partial L}{\partial y} &= 2x - 6 - \lambda x = 0 \\ \frac{\partial L}{\partial \lambda} &= 500 - xy = 0\end{aligned}$$

(b) Si la restricción es  $510 = xy$ :

- La función objetivo varía en  $10\lambda$ .
- **No** se puede saber sin resolver si la solución óptima  $(x, y)$  varía.
- La variación no es exacta, es sólo una estimación.
- No se puede saber el signo de la variación. El multiplicador de Lagrange en este caso puede tener cualquier signo.

**Problema 46 (KKT (I2 1°03))** Considere el siguiente problema de optimización restringida y escriba las condiciones de KKT. Luego, encuentre los puntos de KKT candidatos a óptimo.

$$\begin{aligned}P) \quad & \text{Min } 3x^2 + 2y \\ \text{s.a.} \quad & 3x + 4y^2 = 8 \\ & -x^2 + 5y \geq -5 \\ & 0 \leq x \leq 30\end{aligned}$$

**Solución:** Reescribiendo el problema se obtiene:

$$\begin{aligned}P) \quad & \text{Min } 3x^2 + 2y \\ \text{s.a.} \quad & 3x + 4y^2 = 8 \\ & x^2 - 5y \leq 5 \\ & x \leq 30 \\ & x \geq 0\end{aligned}$$

Luego el lagrangeano viene dado por:

$$L(x, y, \mu_1, \mu_2, \mu_3) = 3x^2 + 2y + \mu_1(3x + 4y^2 - 8) + \mu_2(x^2 - 5y - 5) + \mu_3(x - 30)$$

Y las condiciones de KKT:

$$\frac{\partial L}{\partial x} = 6x + 3\mu_1 + 2\mu_2x + \mu_3 \geq 0 \quad (2.10)$$

$$x \frac{\partial L}{\partial x} = x(6x + 3\mu_1 + 2\mu_2x + \mu_3) = 0 \quad (2.11)$$

$$\frac{\partial L}{\partial y} = 2 + 8\mu_1y - 5\mu_2 = 0 \quad (2.12)$$

$$\frac{\partial L}{\partial \mu_1} = 3x + 4y^2 - 8 = 0 \quad (2.13)$$

$$\frac{\partial L}{\partial \mu_2} = x^2 - 5y - 5 \leq 0 \quad (2.14)$$

$$\mu_2 \frac{\partial L}{\partial \mu_2} = \mu_2(x^2 - 5y - 5) = 0 \quad (2.15)$$

$$\frac{\partial L}{\partial \mu_3} = x - 30 \leq 0 \quad (2.16)$$

$$\mu_3 \frac{\partial L}{\partial \mu_3} = \mu_3(x - 30) = 0 \quad (2.17)$$

$$x \geq 0 \quad (2.18)$$

$$\mu_2 \geq 0 \quad (2.19)$$

$$\mu_3 \geq 0 \quad (2.20)$$

Una vez expresadas estas condiciones corresponde resolver el problema. La resolución consistirá en analizar un árbol de ocho casos diferentes pues para  $x$ ,  $\mu_2$  y  $\mu_3$  se debe analizar los casos en que cada una de estas variables es cero o estrictamente positiva. Es importante notar que cada uno de estos ocho casos conlleva un sistema de ecuaciones distinto. Se procurará realizar este proceso de un modo eficiente que evite resolver estos ocho sistemas de ecuaciones.

$$\text{Si } x = 0 \begin{cases} y = 0 \text{ Se contradice con (2.13).} \\ y \neq 0 \text{ Por (2.13) } y = +\sqrt{2} \text{ o } y = -\sqrt{2} \quad (*) \end{cases}$$

En este caso, por 2.17 se debe cumplir que  $\mu_3 = 0$  y por 2.14 que  $y \geq -1$  (esto descarta la

solución  $y = -\sqrt{2}$ .

Es decir, si  $x = 0$  se debe cumplir  $y = \sqrt{2}$ ,  $\mu_3 = 0$ .

Luego, considerando la condición (2.12):  $\begin{cases} \mu_2 = 0 \implies \mu_1 = \frac{-1}{4y}. \text{ Por 2.10, } \mu_1 \geq 0, \text{ lo que no se satisface para } y \\ \mu_2 \neq 0 \implies 5y = -5 \Rightarrow y = -1, \rightarrow \leftarrow \text{ con } (*). \end{cases}$

Por lo tanto se puede concluir que  $x > 0$ . En ese caso:

$$\begin{aligned}
 6x + 3\mu_1 + 2\mu_2x + \mu_3 &= 0 \\
 2 + 8\mu_1y - 5\mu_2 &= 0 \\
 3x + 4y^2 - 8 &= 0
 \end{aligned} \tag{2.21}$$

Si  $y = 0 \rightarrow x = \frac{8}{3}$ , lo que no cumple con la condición (2.14)  $x^2 \leq 5$ . Por lo tanto, hay contradicción. Por lo tanto,  $y \neq 0$ .

Si  $\mu_3 \neq 0$ :

$$x = 30 \Rightarrow y^2 = \frac{-82}{4} \rightarrow \leftarrow \text{Imaginario.}$$

Por lo que  $\mu_3 = 0$ . Así, el sistema de ecuaciones queda:

$$\begin{aligned}
 6x + 3\mu_1 + 2\mu_2x &= 0 \\
 2 + 8\mu_1y - 5\mu_2 &= 0 \\
 3x + 4y^2 - 8 &= 0
 \end{aligned}$$

Si  $\mu_2 = 0$  el sistema se simplifica a:

$$\begin{aligned}
 6x + 3\mu_1 &= 0 \\
 2 + 8\mu_1y &= 0 \\
 3x + 4y^2 - 8 &= 0
 \end{aligned}$$

con lo que se obtiene  $\mu_1 = -2x$ ,  $x = -\frac{-1}{\frac{4y}{2}} = \frac{1}{8y}$ .  $\therefore y^3 - 2y + \frac{3}{32} = 0$ . Serán puntos de KKT aquéllos cuyo  $x$  sea positivo ( $x > 0$ ). Es decir es necesario resolver esta función cúbica e identificar los puntos  $y$  positivos. Estos tendrán asociado un  $x$  positivo.

Y si  $\mu_2 \neq 0$ :

$$\begin{aligned}
 x^2 - 5y - 5 &= 0 \\
 3x + 4y^2 - 8 &= 0
 \end{aligned}$$

$\therefore 16y^4 - 64y^2 - 45y + 19 = 0$ . Serán puntos de KKT aquéllos que tengan un  $x > 0$  y un  $\mu_2 > 0$ .

**Problema 47 (KKT (I2 1°03))** Considere el siguiente problema de optimización. Llévelo a una formulación estándar y encuentre mediante las condiciones de KKT todos los puntos candidatos a mínimo que hayan. Luego, utilizando condiciones de segundo orden verifique cuáles de estos candidatos son mínimos locales, y cuál(es) es(son) el(los) mínimo(s) global(es).

$$\begin{aligned}
 P) \quad & \text{Max } x^2 + x^3 \\
 \text{s.a.} \quad & x^2 \leq 1
 \end{aligned}$$

**Solución:** Primero planteamos el problema equivalente:

$$\begin{aligned} \tilde{P}) \quad & \text{Min} \quad -x^2 - x^3 \\ \text{s.a.} \quad & x^2 \leq 1 \end{aligned}$$

Para llevarlo a formulación estándar, todas las variables deben ser no negativas. Sea  $x = x_1 - x_2$ , donde  $x_1 \geq 0$  y  $x_2 \geq 0$ . Reemplazando obtenemos el siguiente modelo en formulación estándar:

$$\begin{aligned} \tilde{P}) \quad & \text{Min} \quad -(x_1 - x_2)^2 - (x_1 - x_2)^3 \\ \text{s.a.} \quad & (x_1 - x_2)^2 \leq 1 \\ & x_1, x_2 \geq 0 \end{aligned}$$

Luego el lagrangeano sería:

$$L = -(x_1 - x_2)^2 - (x_1 - x_2)^3 + \mu((x_1 - x_2)^2 - 1)$$

y las condiciones de K.K.T.:

$$\frac{\partial L}{\partial x_1} = -2(x_1 - x_2) - 3(x_1 - x_2)^2 + 2\mu(x_1 - x_2) \geq 0 \quad (2.22)$$

$$x_1 \frac{\partial L}{\partial x_1} = x_1(-2(x_1 - x_2) - 3(x_1 - x_2)^2 + 2\mu(x_1 - x_2)) = 0 \quad (2.23)$$

$$x_1 \geq 0$$

$$\frac{\partial L}{\partial x_2} = 2(x_1 - x_2) + 3(x_1 - x_2)^2 - 2\mu(x_1 - x_2) \geq 0 \quad (2.24)$$

$$x_2 \frac{\partial L}{\partial x_2} = x_2(2(x_1 - x_2) + 3(x_1 - x_2)^2 - 2\mu(x_1 - x_2)) = 0 \quad (2.25)$$

$$x_2 \geq 0$$

$$\frac{\partial L}{\partial \mu} = (x_1 - x_2)^2 \leq 1 \quad (2.26)$$

$$\mu \frac{\partial L}{\partial \mu} = \mu((x_1 - x_2)^2 - 1) = 0 \quad (2.27)$$

$$\mu \geq 0$$

El dominio no admite soluciones singulares

De (2.22) y (2.24), obtenemos que o bien  $x_1 = x_2 = 0$ , o bien:

$$3(x_1 - x_2)^2 + (2 - 2\mu)(x_1 - x_2) = 0 \quad (2.28)$$

El punto  $(0, 0)$  satisface todas las condiciones por lo que constituye un primer punto candidato.

Ahora si  $\mu \neq 0$ , entonces de (2.27) tenemos que  $(x_1 - x_2)^2 = 1$ .

Así de (2.28):

$$\text{i) } 3 + (2 - 2\mu)(1) = 0 \Rightarrow \mu = \frac{5}{2}$$

$$\text{ii) } 3 + (2 - 2\mu)(-1) = 0 \Rightarrow \mu < 0, \text{ que viola la restricción de no negatividad en } \mu.$$

Por lo tanto,  $x_1 - x_2 = x = 1$ , con  $\mu = \frac{5}{2}$  constituye un segundo punto candidato.

Ahora si  $\mu = 0$  entonces de (2.28) y retornando a la formulación original en términos de  $x$ , se obtiene:  $3(x_1 - x_2)^2 + 2(x_1 - x_2) = 0 \Rightarrow (x_1 - x_2)(3(x_1 - x_2) + 2) = 0 \Rightarrow$

$$\begin{cases} x_1 - x_2 = 0 \\ x_1 - x_2 = -\frac{2}{3} \end{cases}.$$

Es decir los puntos críticos del problema original detectados por KKT son  $x = -\frac{2}{3}$ ,  $x = 0$  y  $x = 1$ . En términos de las variables  $x_1$  y  $x_2$  existe infinitas combinaciones para obtener cada uno de estos tres puntos.

Para determinar el óptimo global del problema basta reemplazar cada uno de estos tres puntos y escoger el mejor. Sin embargo se realizará previamente un análisis de convexidad de cada punto. La primera y segunda derivada de la función objetivo son:

$$\begin{aligned} f'(x) &= -2x - 3x^2 \\ f''(x) &= -2 - 6x \end{aligned}$$

Así se puede ver que:

$$f'(1) = -5 < 0 \Rightarrow \text{Mínimo de borde (derecho) estricto. } x = 1.$$

$$\left. \begin{aligned} f'(0) &= 0 \\ f''(0) &= -2 < 0 \end{aligned} \right\} \text{ Por condición suficiente de 2º orden } x = 0 \text{ es un máximo local estricto (punto interior del dominio con gradiente nulo y en que el hessiano es negativo definido).}$$

$$\left. \begin{aligned} f'(-\frac{2}{3}) &= 0 \\ f''(-\frac{2}{3}) &= 2 > 0 \end{aligned} \right\} \text{ Por condición suficiente de 2º orden } x = -\frac{2}{3} \text{ es un mínimo local estricto (punto interior del dominio con gradiente nulo y en que el hessiano es positivo definido).}$$

Por lo tanto,  $x = 1$  y  $x = -\frac{2}{3}$  son mínimos locales y  $f(1) = -2 < f(-\frac{2}{3}) = -0,148$ .

$\therefore x^* = 1$  es el mínimo global.

**Problema 48 (KKT (I2 1º03))** *Responda:*

(a) ¿Por qué para el problema de minimización estándar las primeras condiciones de KKT son del tipo  $\frac{\partial L}{\partial x} \geq 0$  y  $x \frac{\partial L}{\partial x} = 0$ ? Explique.

(b) “El gradiente de la función objetivo es una combinación lineal de los gradientes de las restricciones activas sólo en el punto óptimo global.” Comente.

**Solución:** (a) Porque en el problema de minimización estándar las variables son no-negativas,

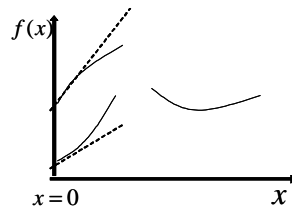


Figure 2.11:

por lo que en el caso de estar en los bordes ( $x = 0$ ) las derivadas deben ser positivas, y en el interior deben ser cero como siempre. Lo que se ilustra en la Figura 2.11.

(b) Esto es falso solamente por el “sólo”, ya que esto ocurre con todo punto extremo que identifica el método de Lagrange.

**Problema 49 (KKT(Ex1°03, pregunta especial I2))** *Resuelva el siguiente problema mediante el método de Karush-Kuhn-Tucker:*

$$\begin{aligned}
 P) \quad & \text{Min } xy^2 - x^3 \\
 s.a. \quad & y + x^2 \leq 4 \\
 & y + 2 \leq x \\
 & y + 2 \geq x^2 \\
 & x \geq 0
 \end{aligned}$$

**Solución:** El lagrangeano está dado por:



$$L = -x^3 + xy^2 + \lambda_1(y + x^2 - 4) + \lambda_2(y - x + 2) + \lambda_3(x^2 - 2 - y) \quad (2.29)$$

$$\frac{\partial L}{\partial x} = -3x^2 + y^2 + 2x\lambda_1 - \lambda_2 + 2x\lambda_3 \geq 0 \quad (2.30)$$

$$x \frac{\partial L}{\partial x} = x(-3x^2 + y^2 + 2x\lambda_1 - \lambda_2 + 2x\lambda_3) = 0 \quad (2.31)$$

$$\frac{\partial L}{\partial y} = 2xy + \lambda_1 + \lambda_2 - \lambda_3 = 0 \quad (2.32)$$

$$\lambda_1(y + x^2 - 4) = 0$$

$$y + x^2 - 4 \leq 0$$

$$\lambda_1 \geq 0$$

$$\lambda_2(y - x + 2) = 0$$

$$y - x + 2 \leq 0$$

$$\lambda_2 \geq 0$$

$$\lambda_3(x^2 - 2 - y) = 0 \quad (2.33)$$

$$x^2 - 2 - y \leq 0$$

$$\lambda_3 \geq 0$$

$$x \geq 0 \quad (2.34)$$

Considerando el caso en que  $x = 0$ , de (2.30) obtenemos:

$$y^2 - \lambda_2 \geq 0 \quad (2.35)$$

y de (2.32):

$$\lambda_1 + \lambda_2 = \lambda_3 \quad (2.36)$$

Luego, analizando por casos, tenemos:

i)  $\lambda_1 = 0$ ,  $\lambda_2 = 0$  y  $\lambda_3 = 0$ .

De (2.32):  $2xy = 0 \begin{cases} \text{Si } x = 0, \text{ de (2.33), } y \leq -2, y \geq -2, y \geq 4 \Rightarrow (0, -2) \text{ es un punto de KKT.} \\ \text{Si } x \neq 0 \rightarrow y = 0; \text{ y de (2.31), } -3x^2 + y^2 = 0, \rightarrow x = 0 \text{ contradice } x \neq 0. \end{cases}$

ii)  $\lambda_1, \lambda_2 = 0$  y  $\lambda_3 \neq 0$

De (2.33),  $x^2 = 2 + y \begin{cases} \text{Si } x = 0, \text{ contradice (2.32).} \\ \text{Si } x \neq 0, -3x^2 + y^2 + 2x\lambda_3 = 0. \text{ Y de (2.32), tenemos adem\'as } 2xy = \lambda_3. \end{cases}$

Por lo tanto,  $-3x^2 + y^2 + 4x^2y = 0$ , pero  $x^2 = 2 + y \Rightarrow -3(2 + y) + y^2 + 4y(2 + y) = 0$ . Entonces,  $5y^2 + 5y - 6 = 0$ , con lo que obtenemos los valores:  $y_1 = 0,704 \rightarrow x_1 = \sqrt{2,704}$ ;  $y_2 = -1,704 \rightarrow x_2 = 0,29$ . Ninguno de estos puntos satisface todas las condiciones. Mientras el primero de estos puntos no satisface ( $y - x + 2 \leq 0$ ), el segundo implica  $2x_2y_2 < 0$  y por lo tanto  $\lambda_3 < 0$ .

iii)  $\lambda_1 = 0$ ,  $\lambda_2 \neq 0$  y  $\lambda_3 = 0$ .

$$\Rightarrow y = x - 2 \begin{cases} x = 0 \text{ contradice (2.32).} \\ x \neq 0 \rightarrow -3x^2 + y^2 = \lambda_2. \\ \lambda_2 = -2xy \end{cases}$$

Luego,  $-3x^2 + y^2 + 2xy = 0$ , ya que  $y = x - 2 \Rightarrow -3x^2 + (x - 2)^2 + 2x(x - 2) = 0$ . Por lo que

$-8x + 4 = 0$ , entonces  $x = \frac{1}{2} \rightarrow y = \frac{-3}{2}$ . Y como  $\lambda_2 = -2xy$ ,  $\lambda_2 = \frac{3}{2} > 0$ . Por lo tanto,  $(x, y)$  es un punto de KKT.

iv)  $\lambda_1 = 0$ ,  $\lambda_2 \neq 0$  y  $\lambda_3 \neq 0$ .

$$\left. \begin{array}{l} x = 2 + y \\ x^2 = 2 + y \end{array} \right\} x^2 = x \left\{ \begin{array}{l} x = 1 \rightarrow y = -1 \\ x = 0 \rightarrow y = -2 \end{array} \right.$$

Para el primer punto  $(1, -1)$ , de (2.31) y de (2.32) se tiene:  $\left\{ \begin{array}{l} -3 + 1 - \lambda_2 + 2\lambda_3 = 0 \\ -2 + \lambda_2 - \lambda_3 = 0 \end{array} \right.$  Lo que implica que  $\lambda_3 = 4$  y  $\lambda_2 = 6$ . Con lo que el punto  $(x = 1, y = -1)$  es un punto de KKT.

El segundo punto  $(0, -2)$ , ya fue detectado como punto KKT previamente.

v)  $\lambda_1 \neq 0$ ,  $\lambda_2, \lambda_3 = 0$ .

$$\Rightarrow y = 4 - x^2 \left\{ \begin{array}{l} x = 0, \text{ contradice (2.32).} \\ x \neq 0, 2xy + \lambda_1 = 0 \text{ y } -3x^2 + y^2 + 2x\lambda_1 = 0 \end{array} \right.$$

Reemplazando y resolviendo se obtiene  $x_1 = 0,823 \rightarrow y_1 = 3,322 \Rightarrow \lambda_1 < 0$ , lo que contradice la condición. Por otra parte, se obtiene además la solución  $x_2 = 2,173 \rightarrow y_2 = -0,721$ , pero este punto no cumple con la restricción  $x^2 - 2 - y \leq 0$ .

vi)  $\lambda_1 \neq 0$ ,  $\lambda_2 \neq 0$  y  $\lambda_3 = 0$ .

$$\left. \begin{array}{l} y = 4 - x^2 \\ y = x - 2 \end{array} \right\} x^2 + x - 6 = 0. \text{ Obteniéndose los valores } x_1 = 2 \rightarrow y_1 = 0 \text{ y } x_2 = -3 \rightarrow y_2 = -5.$$

Pero ninguno de estos puntos cumple las condiciones para ser un punto de KKT: el primero no satisface la restricción  $(x^2 - 2 - y \leq 0)$  mientras el segundo viola la no negatividad en  $x$ .

vii)  $\lambda_1 \neq 0$ ,  $\lambda_2 = 0$  y  $\lambda_3 \neq 0$ .

$\left. \begin{array}{l} y = 4 - x^2 \\ y = x^2 - 2 \end{array} \right\} 2x^2 = 6$ . Por lo que  $x = \pm\sqrt{3} \Rightarrow y = 1$ . Pero ninguno de estos dos puntos cumple con la condición  $y - x + 2 \leq 0$ .

viii)  $\lambda_1, \lambda_2, \lambda_3 \neq 0$ .

$$\left. \begin{array}{l} y + x^2 - 4 = 0 \rightarrow x^2 = 4 - y \\ y - x + 2 = 0 \rightarrow x = y + 2 \\ x^2 - 2 - y = 0 \rightarrow x^2 = 2 + y \end{array} \right\} \text{Ningún punto satisface estas tres condiciones.}$$

Por lo tanto, los tres puntos de KKT son:

$x$	$y$	$f(x, y)$
0	-2	0
$\frac{1}{2}$	$\frac{-3}{2}$	1
1	-1	0

**Problema 50** Considere los siguientes problemas de optimización:

a)

$$\begin{aligned} \text{Min } & \left(x_1 - \frac{13}{3}\right)^2 + \left(x_2 - \frac{1}{2}\right)^2 - x_3 \\ \text{s.a. } & x_1 + \frac{5}{3}x_2 = 10 \\ & (x_2 - 2)^2 + x_3 = 4 \end{aligned}$$

b)

$$\begin{aligned} \text{Min } & -xy \\ \text{s.a. } & x^2 + y^2 = 1 \end{aligned}$$

Encuentre las soluciones óptimas a través del método de Lagrange y caracterícelas. ¿Cuál sería aproximadamente el valor óptimo del problema b) si la constante de la restricción fuera 1.05?

**Solución:**

a)

$$L = \left(x_1 - \frac{13}{3}\right)^2 + \left(x_2 - \frac{1}{2}\right)^2 - x_3 + \lambda_1 \left[x_1 + \frac{5}{3}x_2 - 10\right] + \lambda_2 [(x_2 - 2)^2 + x_3 - 4]$$

$$\begin{aligned} \Rightarrow \frac{\partial L}{\partial x_1} &= 2\left(x_1 - \frac{13}{3}\right) + \lambda_1 \equiv 0 \\ \frac{\partial L}{\partial x_2} &= 2\left(x_2 - \frac{1}{2}\right) + 2\lambda_2(x_2 - 2) + \frac{5}{3}\lambda_1 \equiv 0 \\ \frac{\partial L}{\partial x_3} &= -1 + \lambda_2 \equiv 0 \\ \frac{\partial L}{\partial \lambda_1} &\equiv 0 \\ \frac{\partial L}{\partial \lambda_2} &\equiv 0 \end{aligned}$$

Resolviendo este sistema se obtiene

$$x^* = \left(\frac{35}{6}, \frac{5}{2}, \frac{15}{4}\right) \quad \text{con } \lambda_1 = -3, \lambda_2 = 1$$

También se debe estudiar la existencia de puntos singulares, utilizando el Jacobiano

$$J = \begin{bmatrix} 1 & 5/3 & 0 \\ 0 & 2(x_2 - 2) & 1 \end{bmatrix}$$

Así, el Jacobiano es linealmente independiente  $\forall x_2 \Rightarrow$  es regular  $\forall x_2 \Rightarrow \vec{x}^*$  es regular.

Para caracterizar este punto, se estudia en primer lugar el hessiano

$$H = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 + 2\lambda_2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Que resulta ser semi-definido positivo en  $\vec{x}^*$ , por lo que se deben estudiar las direcciones para caracterizar el punto, que son

$$\begin{aligned} \Delta h_1 &= \Delta x_1 + \frac{5}{3} \Delta x_2 \equiv 0 \\ \Delta h_2 &= 2(x_2^* - 2) \Delta x_2 + \Delta x_3 \equiv 0 \end{aligned}$$

$$\Rightarrow \Delta \vec{x}^T H \Delta \vec{x} = \begin{pmatrix} \frac{-5}{3} \Delta x_2 & \Delta x_2 & -\Delta x_2 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{-5}{3} \Delta x_2 \\ \Delta x_2 \\ -\Delta x_2 \end{pmatrix} = \left( \frac{50}{9} + 4 \right) \Delta x_2^2 \geq 0$$

Por lo que el punto encontrado es un mínimo local.

b)

$$L = -xy + \lambda(x^2 + y^2 - 1)$$

$$\begin{aligned} \Rightarrow \frac{\partial L}{\partial x} &= -y + 2\lambda x \equiv 0 \\ \frac{\partial L}{\partial y} &= -x + 2\lambda y \equiv 0 \\ \frac{\partial L}{\partial \lambda} &\equiv 0 \end{aligned}$$

$$\Rightarrow x^2 = y^2, \lambda = \frac{y}{2x}$$

Resolviendo este sistema se obtiene

$$x = \pm \frac{1}{\sqrt{2}}, y = \pm \frac{1}{\sqrt{2}}, \lambda = \pm \frac{1}{2}$$

Casos posibles:

i)  $x = y \Rightarrow \lambda = 1/2$

ii)  $x \neq y \Rightarrow \lambda = -1/2$

Para estudiar la regularidad del dominio, se obtiene el Jacobiano

$$J = \begin{bmatrix} 2x & 2y \end{bmatrix}$$

Así se aprecia que el Jacobiano es l.i. para todo punto perteneciente al dominio, por lo que éste es regular (El punto (0,0) no pertenece al dominio).

Para caracterizar los 4 puntos candidatos se estudia el hessiano

$$H = \begin{bmatrix} 2\lambda & -1 \\ -1 & 2\lambda \end{bmatrix}$$

Caso i)

$$H = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

no es definido positivo por lo que estudiaremos las direcciones de movimiento para obtener las condiciones de 2do orden:

$$\begin{aligned} \Delta h &= \frac{\partial h}{\partial x} \Delta x + \frac{\partial h}{\partial y} \Delta y = 0 \implies \Delta x = -\Delta y \\ \implies \begin{pmatrix} \Delta x & -\Delta x \end{pmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{pmatrix} \Delta x \\ -\Delta x \end{pmatrix} &= 4\Delta x^2 \geq 0 \end{aligned}$$

Así los puntos  $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$  y  $\left(\frac{-1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right)$  son mínimos locales.

Caso ii)

$$H = \begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix}$$

no es definido positivo por lo que estudiaremos las direcciones de movimiento para obtener las condiciones de 2do orden:

$$\begin{aligned} \Delta h &= \frac{\partial h}{\partial x} \Delta x + \frac{\partial h}{\partial y} \Delta y = 0 \implies \Delta x = \Delta y \\ \implies \begin{pmatrix} \Delta x & \Delta x \end{pmatrix} \begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix} \begin{pmatrix} \Delta x \\ \Delta x \end{pmatrix} &= -4\Delta x^2 \leq 0 \end{aligned}$$

Así los puntos  $\left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right)$  y  $\left(\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$  son máximos locales.

El valor óptimo de este problema será -1/2. Para estudiar la sensibilidad del valor óptimo al cambiar la constante de la restricción, se utiliza que

$$\frac{\partial L}{\partial a} \Big|_{\text{óptimo}} = \frac{\partial f}{\partial a} \Big|_{\text{óptimo}} = -\lambda \implies \Delta f = -\lambda \Delta a = -0.5 \cdot 0.05 = -0.025$$

Se utilizó  $\lambda = 1/2$ , pues con este valor del multiplicador se encuentran los mínimos. Así el valor óptimo decrecería en 0.025 unidades, quedando aproximadamente en -0.525.

**Problema 51** Considere una empresa que produce mesas de madera utilizando dos factores productivos: mano de obra y madera. Los costos estimados de dichos factores son \$10/hora-hombre y

\$6/m^3\$ de madera. Para la producción de las mesas se considera la función de producción

$$f(x, y) = x^a y^b \quad x, y \geq 0,$$

donde  $x$  representa las horas hombres dedicadas e  $y$  los metros cúbicos de madera utilizados;  $f(x, y)$  representa el número de mesas producidas con dichos factores (asuma un valor real). (a) Formule un modelo de minimización de costos para producir  $q$  mesas; y (b) ¿cuáles son las condiciones sobre  $a$  y  $b$  para que exista solución al problema anterior?

**Solución:**

a)

$$\begin{aligned} \text{Min } 10x + 6y \\ \text{s.a. } x^a y^b &= q \\ x, y &\geq 0 \end{aligned}$$

b) Para que exista solución óptima se debe cumplir el teorema de existencia. Se puede apreciar que la función objetivo es continua sobre  $D$ . Además  $D$  es cerrado y no vacío, pues

$$x^* = \sqrt[a]{q}, y^* = 1$$

pertenece a  $D$ . También se observa que

$$f \longrightarrow \infty \text{ si } \|\vec{x}\| \longrightarrow \infty, \vec{x} \in D,$$

pues  $y = \sqrt[b]{\frac{q}{x^a}}$  y reemplazando esto en la función objetivo se obtiene que ésta tiende a infinito cuando  $x$  tiende a infinito, para todo  $a, b$ .

Si se resolviera este problema por Lagrange, se obtendría

$$x = \frac{-a\lambda q}{10}, y = \frac{-b\lambda q}{6}, \lambda = -\left(\frac{10^a 6^b}{a^a b^b q^{a+b-1}}\right)^{\frac{1}{a+b}}$$

**Problema 52** Considere el siguiente problema de optimización

$$\begin{aligned} \text{Max } x + y \\ \text{s.a. } 12y - x^2 - y^2 &\leq 11 \\ x &\leq 3 \\ x, y &\geq 0 \end{aligned}$$

- Establezca las condiciones de KKT.
- Verifique si se cumplen las condiciones anteriores en los puntos  $(3, 2)$  y  $(0, 1)$ .
- Determine si estos puntos son máximos o mínimos.

d) Si la segunda restricción se sustituyera por  $-x + 3y \leq 3$ , estudie las condiciones de KKT en los mismos puntos.

**Solución:**

a)

$$L = -x - y + \mu_1(12y - x^2 - y^2 - 11) + \mu_2(x - 3)$$

$$\begin{aligned} i) \quad & \frac{\partial L}{\partial x} = -1 - 2x\mu_1 + \mu_2 \geq 0 \\ ii) \quad & x \frac{\partial L}{\partial x} = x(-1 - 2x\mu_1 + \mu_2) = 0 \\ iii) \quad & \frac{\partial L}{\partial y} = -1 + 12\mu_1 - 2y\mu_1 \geq 0 \\ iv) \quad & y \frac{\partial L}{\partial y} = y(-1 + 12\mu_1 - 2y\mu_1) = 0 \\ v) \quad & \frac{\partial L}{\partial \mu_1} = 12y - x^2 - y^2 - 11 \leq 0 \\ vi) \quad & \mu_1 \frac{\partial L}{\partial \mu_1} = \mu_1(12y - x^2 - y^2 - 11) = 0 \\ vii) \quad & \frac{\partial L}{\partial \mu_2} = x - 3 \leq 0 \\ viii) \quad & \mu_2 \frac{\partial L}{\partial \mu_2} = \mu_2(x - 3) = 0 \\ ix) \quad & x, y \geq 0 \quad \mu_1, \mu_2 \geq 0 \end{aligned}$$

b) Sólo el primer punto es candidato a mínimo pues el segundo no cumple las condiciones de KKT. En el primer punto, ambas restricciones están activas y las derivadas del lagrangeano con respecto a las variables se anulan.

c) Para esto se estudiará el hessiano

$$H = \begin{bmatrix} -2\mu_1 & 0 \\ 0 & -2\mu_2 \end{bmatrix}$$

evaluado en el punto, pues para que el punto sea óptimo, éste debe ser definido positivo, dado que las derivadas parciales se anulan en el punto. Pero  $\mu_1 > 0$  (primera restricción está activa), por lo que el hessiano es definido negativo en el punto. Así éste es un máximo local estricto, y no es solución óptima para el problema.

d) Al cambiar la restricción, ninguno de los dos puntos cumple las condiciones de KKT.

**Problema 53** Suponga que se sabe que, en el siguiente problema, la primera restricción se activa

en el óptimo. Determine una solución óptima utilizando el método de KKT.

$$\begin{aligned} \text{Min } & (x-3)^2 + (y-5)^2 \\ \text{s.a. } & x+y \leq 6 \\ & x, y \geq 0 \end{aligned}$$

**Solución:**

Si la primera restricción está activa:  $x+y=6 \implies \mu$  libre (no necesariamente mayor que cero). Al introducir esto en las condiciones de KKT se obtiene el punto óptimo (2,4). Este problema podría haber sido resuelto directamente utilizando Lagrange, pues se sabe que la restricción está activa. Luego se escogerían los puntos candidatos tales que  $x, y = 0$ .

Las condiciones de KKT serían

$$L = (x-3)^2 + (y-5)^2 + \mu(x+y-6)$$

$$\begin{aligned} i) \quad & \frac{\partial L}{\partial x} = 2(x-3) + \mu \geq 0 \\ ii) \quad & x \frac{\partial L}{\partial x} = x[2(x-3) + \mu] = 0 \\ iii) \quad & \frac{\partial L}{\partial y} = 2(y-5) + \mu \geq 0 \\ iv) \quad & y \frac{\partial L}{\partial y} = y[2(y-5) + \mu] = 0 \\ v) \quad & \frac{\partial L}{\partial \mu} = x+y-6 = 0 \\ vi) \quad & x, y \geq 0, \quad \mu \text{ libre} \end{aligned}$$

Además el hessiano es definido positivo en todo el dominio, por lo que el punto encontrado será un mínimo global.

**Problema 54** Considere le problema

$$\begin{aligned} \text{Min } & y \\ \text{s.a. } & x^2 + y^2 \geq 4 \\ & x^2 - y \leq 0 \\ & x^2 - y \leq 0 \\ & x^2 - y \leq 0 \end{aligned}$$

Identifique todos los puntos que satisfacen las condiciones de KKT.

**Solución:**



$$L = y + \mu_1(-x^2 - y^2 + 4) + \mu_2(x^2 - y)$$

$$\begin{aligned} i) \quad & \frac{\partial L}{\partial x} = -2x\mu_1 + 2x\mu_2 \geq 0 \\ ii) \quad & x \frac{\partial L}{\partial x} = x(-2x\mu_1 + 2x\mu_2) = 0 \\ iii) \quad & \frac{\partial L}{\partial y} = 1 - 2y\mu_1 - \mu_2 = 0 \\ iv) \quad & \frac{\partial L}{\partial \mu_1} = -x^2 - y^2 + 4 \leq 0 \\ v) \quad & \mu_1 \frac{\partial L}{\partial \mu_1} = \mu_1(-x^2 - y^2 + 4) = 0 \\ vi) \quad & \frac{\partial L}{\partial \mu_2} = x^2 - y \leq 0 \\ vii) \quad & \mu_2 \frac{\partial L}{\partial \mu_2} = \mu_2(x^2 - y) = 0 \\ viii) \quad & x \geq 0 \quad \mu_1, \mu_2 \geq 0 \end{aligned}$$

a) Sea  $x, y \neq 0$ :

$$\text{En } ii) \implies \mu_1 = \mu_2$$

Caso I:  $\mu_1, \mu_2 \neq 0$

$$\text{En } v) \implies x^2 + y^2 = 4$$

$$\text{En } vii) \implies x^2 = y$$

El punto candidato será así  $(x, y) = (2.44, 1.56)$  con  $\mu_1 = \mu_2 = 0.24$ .

Caso II:  $\mu_1 = \mu_2 = 0$

$$\text{En } iii) \rightarrow \leftarrow$$

b) Sea  $x, y = 0$ :

$$\text{En } iv) \rightarrow \leftarrow$$

c) Sea  $x \neq 0, y = 0$ :

$$\text{En } vi) \rightarrow \leftarrow$$

d) Sea  $x = 0, y \neq 0$ :

$$\text{En } iv) \implies y^2 \geq 4 \implies y \leq -2 \text{ o } y \geq 2$$

$$\text{En } vii) \implies \mu_2 = 0$$

$$\text{En } vi) \implies y \geq 0$$

$$\text{Casos: } i) y = 2 \implies \mu_1 = 1/4$$

$$\text{Candidato: } (x, y) = (0, 2) \text{ con } \mu_1 = 1/4, \mu_2 = 0$$

$$ii) y > 2 \implies \mu_1 = 0. \text{ En } iii) \rightarrow \leftarrow$$

**Problema 55** Suponga que se sabe que en el siguiente problema, la primera restricción se activa

en el óptimo. Determine una solución óptima utilizando el método de KKT.

$$\begin{aligned} \text{Min } (x-3)^2 + (y-5)^2 \\ \text{s.a. } x+y &\leq 6 \\ x, y &\geq 0 \end{aligned}$$

**Solución:**

Las condiciones de KKT serían

$$L = (x-3)^2 + (y-5)^2 + \mu(x+y-6)$$

$$\begin{aligned} i) \quad & \frac{\partial L}{\partial x} = 2(x-3) + \mu \geq 0 \\ ii) \quad & x \frac{\partial L}{\partial x} = x[2(x-3) + \mu] = 0 \\ iii) \quad & \frac{\partial L}{\partial y} = 2(y-5) + \mu \geq 0 \\ iv) \quad & y \frac{\partial L}{\partial y} = y[2(y-5) + \mu] = 0 \\ v) \quad & \frac{\partial L}{\partial \mu} = x+y-6 = 0 \\ vi) \quad & x, y \geq 0, \mu \text{ libre} \end{aligned}$$

a) Sea  $x, y \neq 0$ :

$$\text{En } ii), iv) \text{ y } v) \Rightarrow \begin{cases} 2(x-3) + \mu = 0 \\ 2(y-5) + \mu = 0 \\ x+y = 6 \end{cases}$$

Resolviendo este sistema, vemos que el candidato será (2,4) con  $\mu = 2$ , que cumple con todas las condiciones de KKT.

b) Sea  $x, y = 0$ :

$$\text{En } v) \rightarrow \leftarrow$$

c) Sea  $x \neq 0, y = 0$ :

$$\text{En } v) \Rightarrow x = 6$$

$$\text{En } ii) \Rightarrow \mu = -6$$

$$\text{En } iii) \rightarrow \leftarrow$$

d) Sea  $x = 0, y \neq 0$ :

$$\text{En } v) \Rightarrow y = 6$$

$$\text{En } iv) \Rightarrow \mu = -2$$

$$\text{En } i) \rightarrow \leftarrow$$

Si la primera restricción está activa:  $x+y=6 \Rightarrow \mu$  libre (no necesariamente mayor que cero).

Al introducir esto en las condiciones de KKT se obtiene el punto óptimo (2,4). Este problema podría haber sido resuelto directamente utilizando Lagrange, pues se sabe que la restricción está activa. Luego se escogerían los puntos candidatos tales que  $x, y \geq 0$ .

Además el hessiano es definido positivo en todo el dominio, por lo que el punto encontrado será un mínimo global.

**Problema 56** Considere le problema

$$\begin{aligned} \text{Min } & y \\ \text{s.a. } & x^2 + y^2 \geq 4 \\ & x^2 - y \leq 0 \\ & x \geq 0, y \text{ libre} \end{aligned}$$

Identifique todos los puntos que satisfacen las condiciones de KKT.

**Solución:**

$$L = y + \mu_1(-x^2 - y^2 + 4) + \mu_2(x^2 - y)$$

$$\begin{aligned} i) \quad & \frac{\partial L}{\partial x} = -2\mu_1 x + 2\mu_2 x \geq 0 \\ ii) \quad & x \frac{\partial L}{\partial x} = x[-2\mu_1 x + 2\mu_2 x] = 0 \\ iii) \quad & \frac{\partial L}{\partial y} = 1 - 2\mu_1 y - \mu_2 = 0 \\ iv) \quad & \frac{\partial L}{\partial \mu_1} = -x^2 - y^2 + 4 \leq 0 \\ v) \quad & \mu_1 \frac{\partial L}{\partial \mu_1} = \mu_1(-x^2 - y^2 + 4) = 0 \\ vi) \quad & \frac{\partial L}{\partial \mu_2} = x^2 - y \leq 0 \\ vii) \quad & \mu_2 \frac{\partial L}{\partial \mu_2} = \mu_2(x^2 - y) = 0 \\ viii) \quad & x \geq 0; \mu_1, \mu_2 \geq 0 \end{aligned}$$

a) Sea  $x, y \neq 0$ :

$$\text{En } ii) \Rightarrow \mu_1 = \mu_2$$

Caso I:  $\mu_1, \mu_2 \neq 0$

$$\text{En } v) \Rightarrow x^2 + y^2 = 4$$

$$\text{En } vii) \Rightarrow x^2 = y$$

El punto candidato será así  $(x, y) = (1.25, 1.56)$  con  $\mu_1 = \mu_2 = 0.24$ .

Caso II:  $\mu_1, \mu_2 = 0$

En *iii*)  $\rightarrow \leftarrow$

b) Sea  $x, y = 0$ :

En *iv*)  $\rightarrow \leftarrow$

c) Sea  $x \neq 0, y = 0$ :

En *vi*)  $\rightarrow \leftarrow$

d) Sea  $x = 0, y \neq 0$ :

En *iv*)  $\Rightarrow y^2 \geq 4 \Rightarrow y \leq -2$  o  $y \geq 2$

En *vii*)  $\Rightarrow \mu_2 = 0$

En *vi*)  $\Rightarrow y \geq 0$

Casos: i)  $y = 2 \Rightarrow$  de *iii*)  $\mu_1 = 1/4$

Candidato:  $(x, y) = (0, 2)$  con  $\mu_1 = 1/4, \mu_2 = 0$

ii)  $y > 2 \Rightarrow$  de *v*)  $\mu_1 = 0$  En *iii*)  $\rightarrow \leftarrow$

## Chapter 3

# Programación Lineal

### 3.1 Introducción

Un problema de optimización se dice lineal si las variables son todas continuas y tanto la función objetivo como las restricciones del problema son funciones lineales en las variables.

Todo problema de Programación Lineal puede expresarse mediante el siguiente formato estándar:

$$\begin{array}{ll} \text{Min} & c_1x_1 + c_2x_2 + \dots + c_nx_n \\ \text{s.a} & a_{11}x_1 + \dots + a_{1n}x_n \leq b_1 \\ & \vdots \\ & a_{m1}x_1 + \dots + a_{mn}x_n \leq b_m \end{array}$$

o bien en notación matricial:

$$\begin{array}{ll} \text{Min} & cx \\ \text{s.a} & Ax \leq b \end{array}$$

en que denominamos  $x^T = (x_1 \dots x_n)$  como las variables de decisión,  $c = (c_1 \dots c_n)$  el vector de costos,  $b^T = (b_1 \dots b_m)$  el vector de insumos disponibles, y

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$$

la matriz estructural de las restricciones.

Como todo problema de optimización continua, este problema de Programación Lineal es abordable via KKT con la particularidad que al tratarse de un problema convexo, cualquier solución que cumpla con las condiciones de KKT será un óptimo global al problema (KKT es condición suficiente en vez de necesaria). No podemos garantizar que esta solución óptima será única, pues

las funciones objetivo lineales no son estrictamente convexas.

En este caso las condiciones de KKT serán:  $L(x, \mu) = cx + \mu(Ax - b)$

$$\begin{aligned}\frac{\partial L}{\partial x} &= c + \mu A = 0 \\ \frac{\partial L}{\partial \mu} &= Ax - b \leq 0 \\ \mu \frac{\partial L}{\partial \mu} &= \mu(Ax - b) = 0 \\ \mu &\geq 0\end{aligned}$$

En caso que el problema incorpore restricciones de signo en todas las variables  $x_i \geq 0$ , las condiciones KKT serían:

$$\begin{aligned}\frac{\partial L}{\partial x} &= c + \mu A \geq 0 \\ x \frac{\partial L}{\partial x} &= x(c + \mu A) = 0 \\ \frac{\partial L}{\partial \mu} &= Ax - b \leq 0 \\ \mu \frac{\partial L}{\partial \mu} &= \mu(Ax - b) = 0 \\ \mu &\geq 0 \\ x &\geq 0\end{aligned}$$

Si bien este problema puede ser abordado mediante las condiciones generales de KKT, posee condiciones y características propias que hacen propicio desarrollar un método específico para enfrentarlo.

Para introducirnos en las características propias de la Programación Lineal, veamos un ejemplo simple:

**Ejemplo 20** *Una fábrica produce dos tipos de planchas de aluminio pintado y requiere determinar la cantidad a producir de cada tipo. Producir una plancha del tipo 1 requiere 7m<sup>2</sup> de aluminio bruto, 0,3lts de pintura y 15 min de trabajo. El costo por plancha (en aluminio y pintura) para el fabricante es de \$400 y el precio unitario de venta es de \$1200. Producir una plancha del tipo 2 requiere 14m<sup>2</sup> de aluminio bruto, 0,3lts de pintura y 5 min de trabajo. El costo por plancha es \$900 y el precio unitario de venta es de \$1500. El fabricante maneja un stock diario máximo de 630m<sup>2</sup> de aluminio bruto y 15lts de pintura. Trabaja solo y dispone de 10hrs cada día. El fabricante no dispone de un trabajo alternativo para las horas no utilizadas en fabricar planchas de aluminio ¿Cuánto es lo óptimo a producir de modo de maximizar la utilidad?*

Para esto definiremos dos variables  $x_1$  y  $x_2$ , que representan el número de planchas pintadas

diarias a producir de cada tipo.

$$\begin{aligned}
 P) \quad & \text{Max } 800x_1 + 600x_2 \\
 s.a \quad & 15x_1 + 5x_2 \leq 600 \text{ (minutos disponibles)} \\
 & 7x_1 + 14x_2 \leq 630 \text{ (m}^2 \text{ de aluminio)} \\
 & 0,3x_1 + 0,3x_2 \leq 15 \text{ (lts de pintura)} \\
 & x_1, x_2 \geq 0
 \end{aligned}$$

El objetivo es identificar la combinación de  $x_1$  y  $x_2$  que, satisfaciendo las restricciones del modelo, maximiza las utilidades para la empresa. Podemos ver que el problema tiene solución óptima ya que en un problema de programación lineal basta determinar una solución factible, por ejemplo producir 20 planchas tipo 1 y 20 del tipo 2. Esto es factible ya que consume sólo 6hrs 40 min, 420m<sup>2</sup> de aluminio y 12lts pintura de las disponibles, lo que alcanza. Su utilidad sería de \$28.000 diarios, pero claramente no sería óptimo (sobran insumos de todo tipo!).

Este problema tiene la particularidad de ser bidimensional. Esto permite visualizar gráficamente el problema y mediante esta herramienta identificar la o las soluciones óptimas. En la Figura 3.1 se observa que la solución óptima corresponde a la combinación de  $x_1$  y  $x_2$  en que la primera y tercera restricción están activas. Es decir, se utiliza toda la pintura y las horas disponibles. Esta solución corresponde a  $x_1^* = 35$  y  $x_2^* = 15$ , con una utilidad de \$37.000.

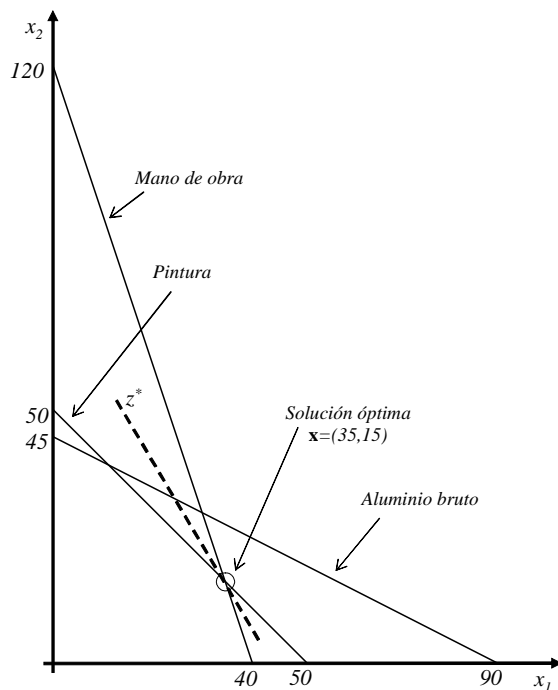


Figure 3.1: Ejemplo Planchas de Aluminio

**Observación:** Evidentemente cuando  $n$  crece (número de variables), el problema ya no puede

ser abordado gráficamente.

Planteemos este problema como uno KKT y abordémoslo genéricamente.

$$L = -800x_1 - 600x_2 + \mu_1(15x_1 + 5x_2 - 600) + \mu_2(7x_1 + 14x_2 - 630) + \mu_3(0,3x_1 + 0,3x_2 - 15)$$

Por lo tanto, un punto máximo del problema debe cumplir con:

$$\begin{aligned} -800 + 15\mu_1 + 7\mu_2 + 0,3\mu_3 &\geq 0 \\ x_1(-800 + 15\mu_1 + 7\mu_2 + 0,3\mu_3) &= 0 \\ x_1 &\geq 0 \\ -600 + 5\mu_1 + 14\mu_2 + 0,3\mu_3 &\geq 0 \\ x_2(-600 + 5\mu_1 + 14\mu_2 + 0,3\mu_3) &= 0 \\ x_2 &\geq 0 \\ 15x_1 + 5x_2 - 600 &\leq 0 \\ \mu_1(15x_1 + 5x_2 - 600) &= 0 \end{aligned}$$

$$\begin{aligned} \mu_1 &\geq 0 \\ 7x_1 + 14x_2 - 630 &\leq 0 \\ \mu_2(7x_1 + 14x_2 - 630) &= 0 \\ \mu_2 &\geq 0 \\ 0,3x_1 + 0,3x_2 - 15 &\leq 0 \\ \mu_3(0,3x_1 + 0,3x_2 - 15) &= 0 \\ \mu_3 &\geq 0 \end{aligned}$$

Lo primero que haremos será identificar los multiplicadores  $\mu_1, \mu_2$  y  $\mu_3$  asociados al punto  $x_1^* = 35$  y  $x_2^* = 15$  y verificar que esta solución satisface estas condiciones de KKT.

Las condiciones de complementaridad de las holguras en este punto exigen:

$$\begin{aligned} -800 + 15\mu_1 + 7\mu_2 + 0,3\mu_3 &= 0 \\ -600 + 5\mu_1 + 14\mu_2 + 0,3\mu_3 &= 0 \\ \mu_2 &= 0 \end{aligned}$$

Así obtenemos:

$$\begin{aligned} \mu_1 &= 20 \\ \mu_3 &= 1666,\bar{6} \end{aligned}$$

Es fácil comprobar que el punto  $x_1^* = 35$ ,  $x_2^* = 15$ ,  $\mu_1 = 20$ ,  $\mu_2 = 0$ ,  $\mu_3 = 1666,\bar{6}$  satisface todas



las condiciones de KKT y por lo tanto corresponde a un óptimo global del problema.

Como primera observación es interesante destacar que los multiplicadores obtenidos no dependen de los insumos disponibles. De hecho, cambios menores (y no tan menores también) en el vector de insumos  $\mathbf{b}$  modifican la combinación óptima de planchas ( $x_1^*$  y  $x_2^*$ ), pero no modifican los multiplicadores óptimos. Esto es consistente con la interpretación que se dio a estos multiplicadores en el capítulo anterior. Se indicó que el multiplicador  $\mu_i$  entrega una estimación de primer orden respecto a cuánto cambia el valor óptimo si el insumo disponible del recurso  $i$ , esto es  $b_i$ , aumenta en una unidad. Al respecto podemos hacer dos observaciones:

- Como se trata de un problema de programación lineal, el impacto en el valor óptimo de aumentar en una unidad el insumo disponible de una restricción activa será el mismo independiente de el valor del lado derecho (en la medida que el conjunto restante de restricciones activas en el punto óptimo no cambie). Asimismo si la restricción está inactiva, el multiplicador permanecerá nulo ante cambios en el total de insumos disponibles que la mantengan inactiva.
- Como se trata de un problema de programación lineal, la estimación de primer orden del impacto en el valor óptimo es exacta para perturbaciones menores.

Así, el valor de los multiplicadores asociado a un punto del dominio depende de los coeficientes estructurales de las restricciones activas en el punto y del gradiente de la función objetivo, pero es independiente del vector de insumos disponibles (una vez definidas las restricciones activas). Geométricamente, el gradiente de la función objetivo debe corresponder a una combinación lineal (con coeficientes positivos) de los gradientes de las restricciones activas.

Por ejemplo, si revisamos las condiciones de KKT para el punto factible  $x_1^* = 10$ ,  $x_2^* = 40$ , observamos que el punto satisface todas las condiciones de KKT a excepción de  $\mu_2 \geq 0$ . Esto indica que la segunda restricción está activa en el punto, y que la solución mejoraría si la restricción dejara de estarlo.

Ahora bien, sabemos que los multiplicadores óptimos satisfacen  $\mu_j = \frac{\partial L}{\partial b_j}$  y por lo tanto representan el valor económico marginal del recurso  $j$ .

Supongamos ahora que se decide aumentar los recursos disponibles de uno de los insumos. ¿Qué posibilidad se estudiará primero? ¿La de tener más pintura, mayor mano de obra, o más planchas de aluminio bruto? Por una parte, el aluminio bruto no es restricción activa por lo que realmente sobra. Al comparar una expansión en pintura o tiempo disponible, se debe tener cuidado, ya que el multiplicador  $\mu$  está referido a *una unidad* de  $b_j$ .

Si hacemos una comparación porcentual equivalente (digamos en un 1%) veremos que 6 minutos adicionales de trabajo significan \$120 más de utilidad, mientras que 0,15 litros de pintura significan \$250 adicionales. Como decíamos anteriormente esta estimación no contempla error (es decir es exacta) en la medida que las restricciones activas siguen siendo las mismas. Si esto no sucede entonces el impacto en el valor óptimo variaría.

Habíamos dicho que esta interpretación de los multiplicadores cabe sólo para variaciones de los recursos disponibles que no altere el conjunto de restricciones activas. A continuación identificare-

mos para cada insumo los rangos en que esto sucede manteniendo los demás insumos en su valor actual.

### Restricción de mano de obra

¿En qué rango de disponibilidad diaria de tiempo de trabajo la afirmación “El valor de un minuto del tiempo del fabricante es \$20” es cierta? Será cierta mientras las restricciones activas sigan siendo las de tiempo y pintura disponible. Sin embargo, si la disponibilidad de tiempo (actualmente en 10 horas) varía demasiado puede suceder que en el óptimo se active la restricción de aluminio o alguna de las variables se haga cero (se activa la restricción de no negatividad). A través de un análisis gráfico se puede observar que si la disponibilidad diaria de tiempo es mayor a 5hr 50 min y menor a 12hr 30 min., el valor del multiplicador  $\mu_1$  se mantiene en \$20.

¿Qué sucede si dispone de menos de 5hr 50 min? Resolviendo las ecuaciones de KKT se obtienen los siguientes multiplicadores  $\mu_1 = 40$ ,  $\mu_2 = 28.5$ ,  $\mu_3 = 0$ , por lo tanto el valor de un minuto de su tiempo aumenta a \$40 el minuto. El límite inferior de disponibilidad diaria de tiempo para este valor del tiempo es 3hr 45 min. Este aumento en el valor del tiempo es esperable pues en la medida que se cuenta con menor cantidad de un insumo, se destina a tareas más productivas.

¿Y si trabaja menos de 3hr 45 min? En este caso los multiplicadores serán  $\mu_3 = 0$  y  $\mu_2 = 0$ , por lo tanto  $x_1 = 0$  y  $\mu_1 = 120$ . Así el valor de su tiempo aumenta a \$120 el minuto.

Es interesante notar que si se obvia la restricción de aluminio y pintura y se supone un problema exclusivamente con restricción de tiempo se observa que la solución óptima consiste en producir la mayor cantidad de planchas tipo 2 que se pueda con el tiempo disponible. Esto se debe a que en las planchas tipo 1 se ganaría \$800 por cada 15 minutos de trabajo, versus \$600 por cada 5 minutos en las planchas tipo 2. Así, en este caso la productividad de un minuto extra sería \$120 que es exactamente el resultado obtenido si se cuenta con menos de 3hr 45 min al día. En este caso el aluminio y la pintura abundan para poder sacar un máximo provecho a los escasos minutos disponibles. Sin embargo, en la medida que se dispone de más de 3hr 45 min ya no se cuenta con todos los recursos de aluminio y pintura para poder producir la solución que maximiza la productividad de los minutos. Y en ese caso estos otros insumos comienzan a ser importantes y a definir la combinación óptima. En ese caso el valor del tiempo cae de los \$120 a \$40 y luego a \$20 el minuto y el valor del aluminio y pintura progresivamente dejan de tener valor económico nulo. Es interesante notar también que cuando se dispone más de 3hr 45 min, pero menos de 5hr 50 min el aluminio tiene un valor económico, pero no así la pintura. Sin embargo, cuando se cuenta con más de 5hr 50 min el aluminio deja de tener un valor económico y sí lo tiene la pintura.

También es interesante notar que de los 600 minutos que el fabricante dispone, los primeros 225 tienen un valor de \$120/min, los siguientes 125 un valor de \$40/min mientras los últimos 250 un valor de \$20/min. Si sumamos el valor de cada uno de estos minutos obtenemos la utilidad de la empresa. Es decir:  $225 \text{ min} * \$120/\text{min} + 125 \text{ min} * \$40/\text{min} + 250 \text{ min} * \$20/\text{min} = \$37.000$ .

¿Qué pasa si se trabaja más de 12hr 30 min? En este caso la restricción tiempo deja de ser activa y por lo tanto  $\mu_1 = 0$  en este rango.

¿Qué debiera hacer este fabricante si alguien le ofrece un trabajo por \$100/minuto? En este

caso el debiera trabajar 3hr 45 min en la fábrica de pintado (con una rentabilidad de \$120/minuto) y emplearse por el resto de su tiempo disponible en el trabajo que le ofrecen.

### Restricción de aluminio

¿Qué valor tiene una plancha de aluminio bruto de más o de menos? El valor actual es cero. Para que tenga algún valor debe convertirse en una restricción que incida en el resultado de la fábrica. Del gráfico se observa que esto sólo pasaría si el stock diario fuese sólo  $455m^2$  de aluminio bruto. Por lo tanto mientras  $455 \leq b_2 \leq \infty$ ,  $\mu_2 = 0$ .

### Restricción de pintura

La restricción 3 tiene un multiplicador asociado (o precio sombra) de \$1666,7. Del gráfico observamos que éste tiene validez para  $12 \leq b_3 \leq 18$ . En el caso de que se disponga de menos de 12 litros deja de ser activa la restricción 1 y no se produce  $x_2$ . Por otro lado, si se dispone de más de 18 lts., comienza a sobrar pintura en vez de planchas de aluminio ( $\mu_3 = 0$  y  $\mu_2 \neq 0$ ).

¿Qué pasa si se dispone de 16 lts.? El precio sombra (utilidad económica) nos indica que la nueva utilidad debiera ser  $\$37.000 + 1(1.666,7) = \$38.666,7$ . Hemos dicho que esta aproximación es exacta. Comprobémoslo. Las restricciones activas siguen siendo las mismas:

$$\left. \begin{array}{l} 15x_1 + 5x_2 = 600 \\ 0,3x_1 + 0,3x_2 = 16 \end{array} \right\} x_1 = 33, \bar{3}; x_2 = 20$$

y este corresponde a un punto factible para restricción 2. La utilidad es  $800(33, \bar{3}) + 600(20) = \$38.666,7$ . Como podemos ver la estimación proporciona un valor exacto y no aproximado. Esto se debe a que tanto la restricción como la función objetivo son lineales.

### Función objetivo

Qué pasa si el fabricante súbitamente sufre un cambio en la utilidad unitaria de sus productos terminados (ya sea por un aumento o disminución en el precio o en sus costos) ¿Varía la producción? ¿Varían los precios sombra?

Hemos dicho que en el óptimo el gradiente de la función objetivo debe constituir una combinación lineal (con coeficientes positivos) de los gradientes de las restricciones activas. Gráficamente esto equivale a que el gradiente de la función objetivo apunte a una dirección que esté entre los gradientes de las restricciones activas. En nuestro caso los gradientes de las restricciones activas apuntan en las direcciones  $(3,1)$  y  $(1,1)$ . Así, manteniendo la utilidad de las planchas tipo 2 constante en \$600, la solución óptima no cambiará si la utilidad de las planchas tipo 1 se ubica en el intervalo:  $\$600 \leq U_1 \leq \$1800$ . En este rango el punto óptimo de producción sigue siendo el mismo, incluso  $\mu_2$  sigue siendo cero, pero  $\mu_1$  y  $\mu_3$  varían con cada alternativa pues el ángulo que forman los gradientes de esas restricciones con el nuevo gradiente de la función objetivo ha cambiado. Para determinarlos se debe resolver el siguiente sistema:

$$\begin{aligned} -U_1 + 15\mu_1 + 7\mu_2 + 0,3\mu_3 &= 0 \\ -U_2 + 5\mu_1 + 14\mu_2 + 0,3\mu_3 &= 0 \end{aligned}$$

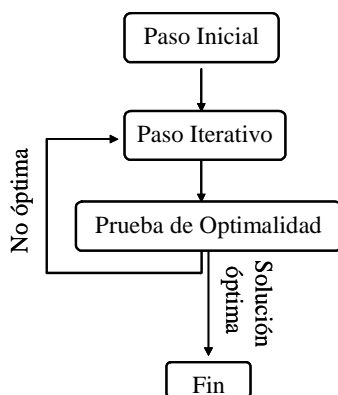


Figure 3.2: Estructura del método Simplex

Es interesante observar que en los casos límites la solución óptima no es única. Esto sucede en los casos en que el gradiente de la función objetivo puede expresarse como una combinación lineal de las restricciones activas, pero en que al menos uno de los coeficientes involucrados sea nulo. Es decir en este caso el multiplicador asociado a una de las restricciones activas será nulo. Geométricamente este fenómeno puede interpretarse como el caso en que el gradiente de la función objetivo es perpendicular a una de las aristas del poliedro y por lo tanto todos los puntos de la arista son igualmente atractivos y óptimos. En el caso anterior esto sucede si  $U_1$  vale exactamente \$600 o exactamente \$1800.

Este tipo de análisis de sensibilidad que hemos realizado para este problema es sumamente útil en el caso de una empresa cuando un modelo como éste es utilizado. Sin embargo, hemos descansado nuestro procedimiento en el análisis gráfico que no está disponible en problemas de tres o más dimensiones. En la próxima sección introduciremos el método Simplex que permite resolver problemas de programación lineal y desarrollar un análisis de sensibilidad como el anterior usando un número mucho mayor de variables y restricciones.

## 3.2 El Método Simplex

El método Simplex es un procedimiento general para resolver problemas de programación lineal. Fue desarrollado por George Dantzig en 1947, y ha probado ser un método extraordinariamente eficiente que se usa en forma rutinaria para resolver problemas grandes en las computadoras de hoy en día.

En realidad el método Simplex es un algoritmo, es decir, es simplemente un proceso en el que se repite (se itera) un procedimiento sistemático una y otra vez hasta obtener el resultado deseado. Con este algoritmo se sustituye un problema difícil por una serie de problemas más simples.

La estructura general de este algoritmo se puede observar en la Figura 3.2. El paso inicial consiste en la preparación para iniciar las iteraciones; el paso iterativo, es la realización de una iteración; y la prueba de optimalidad, es analizar si la solución obtenida es óptima. En el caso en que sí lo sea, termina el proceso, si no, se vuelve a la etapa iterativa.

Desde una perspectiva geométrica, los problemas de programación lineal se pueden interpretar como la búsqueda del mejor punto perteneciente a un poliedro (conjunto de oportunidades) de  $m$  caras en  $n$  dimensiones. Al resolver nuestro ejemplo pudimos ver que los posibles candidatos a ser soluciones óptimas de nuestro problema se encuentran en el borde del conjunto de oportunidades, más específicamente en los vértices.

Las tres propiedades de las soluciones factibles en un vértice que forman el fundamento del método Simplex son las siguientes:

1. Si existe exactamente una solución óptima, entonces debe ser una solución factible en un vértice. Si existe soluciones óptimas múltiples, al menos dos de ellas deben ser soluciones factibles en vértices adyacentes.
2. Existe sólo un número finito de soluciones factibles en los vértices.
3. Si una solución en un vértice es igual o mejor que todas las soluciones factibles en los vértices adyacentes a ella, entonces es igual o mejor que todas las demás soluciones en los vértices, es decir, es óptima.

Adicionalmente es importante observar que cada vértice del dominio queda definido por un conjunto distinto de restricciones que se activan. En un problema de  $n$  variables y  $m$  restricciones, un vértice del poliedro (dominio) queda definido por la intersección de  $n$  restricciones. Sin embargo, puede suceder que en un vértice coincidan más de  $n$  restricciones (por ejemplo en el vértice superior de una pirámide de base cuadrada). Asimismo, dos vértices adyacentes de un poliedro se puede decir que comparten  $n - 1$  restricciones activas. Así, cada iteración del método consiste en identificar un conjunto de  $n$  restricciones cuya intersección defina un vértice del poliedro. Estos puntos definidos como la intersección de  $n$  restricciones se les denominará *soluciones básicas*. Las soluciones básicas que además son factibles para el problema se les denominará *soluciones básicas factibles*. Evidentemente, es posible que una solución básica no sea factible, pues la intersección de las  $n$  restricciones podría encontrarse fuera del dominio.

En definitiva, el método Simplex comienza en una solución básica factible (vértice del poliedro), y evalúa si la solución es óptima comparándola con las soluciones de los vértices adyacentes. Si no es óptima, el método se desplaza a un vértice adyacente que ojalá sea mejor (no puede ser peor) que la solución básica factible vigente. Para esto, el método mantiene  $n - 1$  de las  $n$  restricciones activas que definían la solución básica anterior y permuta la restante por otra restricción que antes se mantenía (muy probablemente) inactiva. Una vez determinadas las  $n$  restricciones activas, identificar la solución básica asociada es muy simple pues basta resolver el sistema de  $n \times n$  ecuaciones resultante (en Simplex esto se hace de un modo muy eficiente). Este procedimiento se repite hasta alcanzar una solución óptima.

Es importante destacar que Simplex exige que todas las variables sean no negativas y que todas las restricciones sean ecuaciones (igualdad). Esto exige agregar variables de holgura no negativas en

casos de desigualdades menor o igual y restar variables de exceso no negativas en las restricciones de mayor o igual.

En su forma más simple (más adelante se discutirá cómo abordar problemas que no satisfagan este formato), el método Simplex opera sobre problemas presentados bajo el siguiente formato en que se contemplan  $n$  actividades (con variables no negativas asociadas a cada una) y  $m$  recursos (el vector de recursos disponibles  $\mathbf{b}$  contiene sólo elementos no negativos):

$$\begin{aligned}
 P) \quad & \text{Min } c_1x_1 + \dots + c_nx_n \\
 \text{s.a.} \quad & a_{11}x_1 + \dots + a_{1n}x_n \leq b_1 \\
 & \vdots \\
 & a_{m1}x_1 + \dots + a_{mn}x_n \leq b_m \\
 & x_i \geq 0 \quad \forall i \in \{1, \dots, n\}
 \end{aligned}$$

<sup>1</sup>En este formato estándar se tiene  $n$  variables originales y  $m$  holguras, es decir  $n + m$  variables en total; además se tiene  $n + m$  restricciones (recordar las restricciones de no negatividad). Una solución básica factible queda definida por  $n$  restricciones activas y (a lo más)  $m$  inactivas. Es decir, cada solución básica factible tiene asociadas  $n$  variables nulas y a lo más  $m$  variables estrictamente positivas (podrían ser menos de  $m$ , recordar la pirámide de base cuadrada). Definiremos como variables básicas a estas  $m$  variables que no corresponden a las holguras de las restricciones definidas como activas y como variables no básicas a las variables de holgura de las restricciones activas (esto incluye a las variables originales que estén en cero pues la restricción de no negatividad respectiva es parte de las restricciones activas).

El formato de la formulación anterior calza con el de nuestro ejemplo de las planchas de aluminio pintado. En él hay 2 actividades: producir planchas de tipo 1, y producir planchas de tipo 2. En cuanto a los recursos, contamos con tres, las horas hombre, las planchas de aluminio bruto y la pintura. Para mantener el orden es importante considerar la siguiente matriz de datos:

Recurso	Actividad				Recursos disponibles
	1	2	...	$n$	
1	$a_{11}$	$a_{12}$	...	$a_{1n}$	$b_1$
2	$a_{21}$	$a_{22}$	...	$a_{2n}$	$b_2$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$m$	$a_{m1}$	$a_{m2}$	...	$a_{mn}$	$b_m$
Beneficio unitario	$c_1$	$c_2$	...	$c_n$	
Nivel de Actividad	$x_1$	$x_2$	...	$x_n$	

---

<sup>1</sup>REVIDAR con mucha calma este párrafo

Así, nuestro ejemplo se verá de la siguiente forma:

Recurso	Tipo 1	Tipo 2	Rec.Disp.
<i>Min</i> /hombre	15	5	600
Aluminio( $m^2$ )	7	14	700
Pintura ( $lt$ )	0,3	0,3	15
Benef/unitario	-800	-600	
Nivel de Producción	$x_1$	$x_2$	

### 3.2.1 Pasos en el método Simplex

El método Simplex está compuesto por tres pasos que se pueden describir resumidamente del siguiente modo:

1. *Paso inicial*: Determinar una solución factible en un vértice.
2. *Prueba de optimalidad*: La solución factible en un vértice es óptima cuando ninguna de las soluciones en vértices adyacentes a ella sean mejores.
3. *Paso iterativo*: Traslado a una mejor solución factible en un vértice adyacente (repetir las veces que sea necesario).

A continuación explicaremos cada uno de estos pasos en forma detallada, ilustrándolos con nuestro ejemplo de las planchas de aluminio. El problema inicial es:

$$\begin{aligned}
 P) \quad & \text{Min} \quad -800x_1 - 600x_2 \\
 \text{s.a} \quad & 15x_1 + 5x_2 \leq 600 \\
 & 7x_1 + 14x_2 \leq 630 \\
 & 0,3x_1 + 0,3x_2 \leq 15 \\
 & x_1, x_2 \geq 0
 \end{aligned}$$

Antes de comenzar se debe crear 3 variables de holgura ( $x_3, x_4, x_5$ ), una para cada restricción. De esta forma el modelo nos queda:

$$\begin{aligned}
 P) \quad & \text{Min} \quad -800x_1 - 600x_2 \\
 \text{s.a} \quad & 15x_1 + 5x_2 + x_3 = 600 \\
 & 7x_1 + 14x_2 + x_4 = 630 \\
 & 0,3x_1 + 0,3x_2 + x_5 = 15 \\
 & x_1, x_2 \geq 0 \\
 & x_3, x_4, x_5 \geq 0
 \end{aligned}$$

1. *Paso Inicial:* Determinar una solución inicial factible puede ser no trivial. Sin embargo, para el caso simple presentado en esta sección (todas las restricciones son desigualdades de menor o igual, todas las variables son no negativas y todos los recursos disponibles son no negativos) existe una solución factible evidente que consiste en asignar un valor cero a cada una de las variables originales. Es decir se asume que las  $n$  restricciones activas serán las restricciones de no negatividad de las variables originales del problema. Así, todas las variables de holgura tomarían valores no negativos.

En nuestro nuevo lenguaje de variables básicas y no básicas, recordemos que siempre se tiene exactamente  $m$  (en este caso  $m = 3$ ) variables básicas que corresponde a las variables que no se asocian a las restricciones activas. En este caso las variables básicas son  $x_3, x_4, x_5$ . Por consiguiente las variables no básicas son  $x_1, x_2$ .

Por lo tanto, podemos presentar la solución básica factible inicial en el siguiente formato (al que se denomina tableau) de  $m + 1$  filas y  $m + n + 1$  columnas:

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	
15	5	1	0	0	600
7	14	0	1	0	630
0,3	0,3	0	0	1	15
-800	-600	0	0	0	0

En este diagrama las primeras  $m + n$  columnas corresponden a una variable cada una mientras la última columna de la derecha corresponde al vector de recursos disponibles. Las  $m$  primeras filas corresponden a las restricciones originales mientras la fila inferior corresponde a la función objetivo. Los coeficientes de esta fila se denominan costos reducidos. La casilla inferior derecha del tableau indica el inverso aditivo del valor de la función objetivo.

El método Simplex representará cada una de las soluciones básicas factibles del proceso iterativo mediante tableaux como el anterior. Este formato permitirá identificar inmediatamente las variables básicas, las variables no básicas, la solución básica, el valor de la función objetivo asociado a la solución básica y la eventual optimalidad de la solución. Esto se debe a que se mantendrá una estructura en que  $m$  columnas (las asociadas a las variables básicas) formarán una matriz identidad en las  $m$  filas superiores y tendrán valores nulos en la función objetivo y a que se procurará que la columna de recursos disponibles (al extremo derecho) contenga sólo valores no negativos. Esto permitirá identificar esas  $m$  variables como básicas y las restantes como no básicas. Al asignar valores nulos a las variables no básicas, el conjunto de restricciones se transforma en una asignación directa de las variables básicas a los recursos disponibles correspondientes. En este caso al hacer  $x_1 = x_2 = 0$  nos queda el sistema de ecuaciones  $x_3 = 600, x_4 = 630, x_5 = 15$  es decir identifica completamente la solución básica. Como decíamos anteriormente además identifica el valor de la función objetivo que corresponde al inverso aditivo del valor de la celda inferior derecha del tableau, en este caso 0. Adicionalmente el tableau permite identificar la optimalidad de la solución básica como



se verá en el paso siguiente del algoritmo.

2. *Prueba de Optimalidad:* Se determina si la solución es óptima. Se verifica si el valor de  $Z$  puede mejorar (en este caso disminuir) al hacer que una de las variables no básicas crezca. Esto se puede verificar observando la última fila del tableau ( $Z$ ). Si todos los valores son positivos o cero, estamos en el óptimo, de lo contrario regresamos al paso iterativo. Para comprender la lógica de esta aseveración considere el siguiente problema:

$$\begin{aligned} \Rightarrow \quad & P) \text{ Min } c_1x_1 + \dots + c_nx_n + v \\ \text{s.a} \quad & \mathbf{x} \in D: \\ & x_i \geq 0 \quad \forall i \in \{1, \dots, n + m\} \end{aligned}$$

en que todos los coeficientes de la función objetivo ( $c_i$ ) son no negativos. En ese caso el valor óptimo al problema no podrá ser menor a  $v$ . Esta cota inferior se alcanzaría si se encuentra una solución factible en que sean cero todas las variables  $x_1$  a  $x_n$ . Y esta es justamente la característica que cumple la solución básica. Todas las variables que aparecen en la función objetivo tienen valor cero lo que permite garantizar que si los coeficientes de esas variables son no negativos se habrá alcanzado una solución óptima.

En nuestro ejemplo, podemos observar que tanto la variable  $x_1$  como  $x_2$  (no básicas por tanto actualmente nulas) mejoran la función objetivo si aumentan en una unidad ya que sus coeficientes en la fila de la función objetivo son estrictamente negativos ( $-800$  y  $-600$  respectivamente), por lo que nos convendrá que una de ellas (cualquiera) entre a la base (conjunto de variables básicas).

3. *Paso Iterativo:* posee tres partes.

3.1 Parte I: Se determina la variable no básica que entra a la base. Típicamente se escoge la variable cuyo coeficiente en la función objetivo sea el más negativo, pues este coeficiente indica cuánto mejora la función objetivo por unidad de la variable que se aumente. Sin embargo, podría escogerse cualquier variable con costo reducido negativo y el método igual podría converger. De hecho escoger la variable con costo reducido más negativo no garantiza que se alcance la mejor próxima solución básica (de entre todas las soluciones asociadas a vértices adyacentes en el poliedro) pues no se cuenta con un indicativo de a cuánto se podrá aumentar la variable entrante sin salirse del dominio. Así, en el caso de nuestro ejemplo, la variable que más aporta a la función objetivo es  $x_1$  (su costo reducido es  $-800$ ), por lo que ésta será la variable entrante.

3.2 Parte II: Se determina la variable básica que sale de la base. Se elige la variable básica que primero alcanza el valor cero cuando se incrementa la variable básica entrante. Para comprender el procedimiento que se describirá es necesario visualizar que lo que se hace es aumentar una variable no básica que tomaba valor cero manteniendo las otras  $n - 1$  variables no básicas en cero. Es decir, nos desprenderemos de sólo una

restricción activa en cada iteración. El cambio en la variable no básica lo absorben las otras variables básicas modificando su valor. En esas circunstancias, cada fila del tableau (que representa a una restricción de igualdad) contendrá sólo dos variables que están cambiando su valor: la variable entrante y la variable básica que figura en la restricción (dado por la matriz identidad del tableau). Es decir, la restricción  $j$  del tableau se puede escribir del siguiente modo:

$$a_{j1}x_1 + a_{j2}x_2 + \dots + a_{jn}x_n + x_{n+j} = b_j$$

en que hemos ordenado las variables de modo que las  $n$  primeras variables correspondan a las variables no básicas y que la restricción  $j$  contenga a la variable básica  $x_{n+j}$  (note que el coeficiente asociado a esta variable es 1 pues la variable es básica). Supongamos que la variable entrante es  $x_i$ , entonces: ¿cuál es el valor más alto que podría tomar esta variable sin salirse del dominio? Dado que las demás variables no básicas no cambian (se mantienen en cero) el alza en  $x_i$  debe ser absorbido por  $x_{n+j}$ . Es decir se debe mantener la siguiente igualdad:

$$a_{ji}x_i + x_{n+j} = b_j$$

En este contexto la única restricción que puede violarse al aumentar el valor de  $x_i$  es la no negatividad de  $x_{n+j}$ . Es decir:

$$x_{n+j} = b_j - a_{ji}x_i \geq 0 \rightarrow \begin{array}{l} \text{Si } a_{ji} > 0 \text{ entonces } x_i \leq \frac{b_j}{a_{ji}} \\ \text{Si } a_{ji} \leq 0 \text{ entonces se cumple } \forall x_i \end{array}$$

Así, hemos encontrado una cota para el alza de la variable  $x_i$  (recordar que  $b_j \geq 0$ ). ¿Qué significa que  $a_{ji} < 0$ ? Significa que en la medida que  $x_i$  crece,  $x_{n+j}$  también lo hace. ¿Qué significa que  $a_{ji} = 0$ ? Significa que  $x_{n+j}$  no se ve afectada por el valor que tome  $x_i$ . Así, si consideramos simultáneamente las  $m$  restricciones del tableau podemos determinar lo siguiente.

La variable saliente se debe escoger el menor valor entre las fracciones  $\frac{b_j}{a_{ji}}, \forall j$  (en que  $i$  es el índice de la variable entrante), considerando sólo las filas en que el coeficiente  $a_{ji}$  sea estrictamente positivo. El mínimo de entre estos coeficientes indica el valor que tomará la variable entrante en la solución básica a la que migrará el método en la siguiente iteración. En nuestro ejemplo debemos realizar el siguiente cálculo:  $\min\{\frac{600}{15}, \frac{630}{7}, \frac{15}{0.3}\} = 40$ . De él observamos que la variable saliente está dictada por la primera restricción, ya que es la que determina el mínimo cociente y que la variable entrante ( $x_1$ ) tomará el valor 40 en la próxima solución básica del método. Por lo tanto,  $x_3$  es la variable básica que sale (note que al hacer  $x_1 = 40 \Rightarrow x_3 = 0$ ).

- 3.3 Parte III: Se determina la nueva solución básica factible una vez determinado que será la variable  $x_{n+j}$  la que dejará la base. Para esto se toma la restricción  $j$  y se reemplaza el valor de  $x_i$  en dicha restricción en el resto de las filas (esto es, en las demás ecuaciones

y en la función objetivo). De este modo, se garantiza que el coeficiente asociado a  $x_i$  en esas filas sea cero. Además se multiplica la fila  $j$  por  $\frac{1}{a_{ji}}$  para transformar en 1 el coeficiente asociado a la variable  $x_i$ . De este modo, se obtiene un nuevo tableau cuya solución básica factible asociada es evidente.

Veamos nuestro ejemplo. Como observamos, la variable entrante era  $x_1$  y la variable saliente  $x_3$ . Así, fue la primera restricción la que limitó el aumento en la variable entrante. Por lo tanto reemplazaremos el valor de  $x_1$  correspondiente a esta restricción en las demás filas del tableau. Esto es:

$$x_1 = \frac{600 - 5x_2 - x_3}{15}$$

Nos queda:

$$\begin{aligned} P) \text{ Min } & -800 * \frac{600 - 5x_2 - x_3}{15} - 600x_2 \\ & \text{s.a.} \quad 15x_1 + 5x_2 + x_3 = 600 \\ & 7 * \frac{600 - 5x_2 - x_3}{15} + 14x_2 + x_4 = 630 \\ & 0,3 * \frac{600 - 5x_2 - x_3}{15} + 0,3x_2 + x_5 = 15 \\ & x_1, x_2 \geq 0 \\ & x_3, x_4, x_5 \geq 0 \end{aligned}$$

Reordenando y dividiendo la primera restricción por 15 nos queda:

$$\begin{aligned} P) \text{ Min } & -\frac{1000}{3}x_2 + \frac{800}{15}x_3 - 32.000 \\ & \text{s.a.} \quad x_1 + \frac{1}{3}x_2 + \frac{1}{15}x_3 = 40 \\ & \frac{35}{3}x_2 - \frac{7}{15}x_3 + x_4 = 350 \\ & \frac{1}{5}x_2 - \frac{1}{50}x_3 + x_5 = 3 \\ & x_1, x_2 \geq 0 \\ & x_3, x_4, x_5 \geq 0 \end{aligned}$$

En que resulta evidente desprender la solución básica asociada. En este caso las variables no básicas son  $x_2 = x_3 = 0$  mientras que las variables básicas son  $x_1 = 40$ ,  $x_4 = 350$ , y  $x_5 = 3$ . Esta solución básica contempla un valor de función objetivo igual a  $-32.000$ . Esto es consistente con que la variable entrante tomó un valor de 40 y su costo reducido era  $-800$ , por lo que la función objetivo debía caer en  $-32.000$ ; como el valor anterior de la función objetivo era 0, se esperaba que ahora fuera  $-32.000$ . Como se observa se ha generado un sistema que permite transitar entre soluciones básicas factibles de modo muy expedito mediante simples reemplazos de variables cuidadosamente seleccionadas

en algunas restricciones y función objetivo.

El método Simplex no trabaja con la formulación formal del problema sino directamente con los tableau. En este caso, para determinar el nuevo tableau se hubiera seguido una formalidad matemática distinta pero que es absolutamente equivalente al procedimiento recientemente desarrollado. El procedimiento consistiría en dividir la 1ª fila por 15, y restar a las demás filas un ponderado de la 1ª fila de modo que en la columna asociada a  $x_1$  el coeficiente se haga cero (operaciones fila). Haciendo eso nos queda:

$$\begin{array}{ccccc|c} 1 & \frac{1}{3} & \frac{1}{15} & 0 & 0 & 40 \\ 0 & \frac{35}{3} & \frac{-7}{15} & 1 & 0 & 350 \\ 0 & \frac{1}{5} & \frac{-1}{50} & 0 & 1 & 3 \\ \hline 0 & -\frac{1000}{3} & \frac{800}{15} & 0 & 0 & 32.000 \end{array}$$

que es el tableau correspondiente al problema de minimización anteriormente obtenido. Es importante notar que el valor de la función objetivo aparece en el tableau con signo cambiado. Esto es consistente pues la columna de la derecha del tableau contiene números que se encuentran al otro lado del signo de igualdad de cada restricción. Así, si se quisiera expresar estos números como parte del lado izquierdo de cada restricción, habría que multiplicarlos por  $-1$ .

Para acentuar el aprendizaje del método continuemos el ejemplo. Observamos del tableau que aún existe una variable no básica cuyo costo reducido es negativo:  $x_2$ . Esta variable (actualmente en cero) reduce en  $\frac{1000}{3}$  la función objetivo por cada unidad que aumente su valor, por lo que nos convendrá que entre a la base. El lector podrá preguntarse porqué una unidad extra del producto 2 agrega sólo  $\$ \frac{1000}{3}$  a la función objetivo siendo que su utilidad era  $\$600$ . Lo que sucede es que dada la combinación actual de productos, y nula disponibilidad de recursos del tipo 1 (tiempo), cada unidad extra del producto 2 exige dejar de producir  $\frac{1}{3}$  de unidad del producto 1. Es decir la utilidad neta es  $\$600 - \frac{1}{3} * \$800$ , esto es  $\$ \frac{1000}{3}$ . Para determinar la variable que sale de la base realizamos el siguiente cálculo:

$$\text{Min} \left\{ \frac{40}{\frac{1}{3}}, \frac{350}{\frac{25}{3}}, \frac{3}{\frac{1}{5}} \right\} = \text{Min}\{120, 30, 15\} = 15$$

Por lo tanto, la variable que sale de la base es  $x_5$ . Realizando las operaciones filas correspondientes obtenemos el siguiente tableau:

$$\begin{array}{ccccc|c} 1 & 0 & \frac{1}{10} & 0 & \frac{-5}{3} & 35 \\ 0 & 0 & \frac{7}{10} & 1 & \frac{-175}{3} & 175 \\ 0 & 1 & \frac{-1}{10} & 0 & 5 & 15 \\ \hline 0 & 0 & 20 & 0 & \frac{5000}{3} & 37.000 \end{array}$$

Realizando la prueba de optimalidad podemos ver que todos los costos reducidos no básicos son positivos, por lo que estamos en una solución óptima. Por lo tanto, la solución obtenida es la

siguiente:

$$\begin{aligned}
 x_1 &= 35 \\
 x_2 &= 15 \\
 x_3 &= 0 \text{ (Restricción activa)} \\
 x_4 &= 175 \text{ (holgura de aluminio)} \\
 x_5 &= 0 \text{ (Restricción activa)}
 \end{aligned}$$

Y el valor óptimo es  $-\$37.000$  (este valor calza con lo que se esperaba:  $-32.000 - 15 \cdot \frac{1000}{3}$ ). La optimalidad de esta solución es evidente si uno transforma este tableau en el problema de minimización correspondiente:

$$\begin{aligned}
 P) \text{ Min } & 20x_3 + \frac{5000}{3}x_5 - 37.000 \\
 \text{s.a.} \quad & x_1 + \frac{1}{10}x_3 - \frac{5}{3}x_5 = 35 \\
 & \frac{7}{10}x_3 + x_4 - \frac{175}{3}x_5 = 175 \\
 & x_2 - \frac{1}{10}x_3 + 5x_5 = 15 \\
 & x_1, x_2 \geq 0 \\
 & x_3, x_4, x_5 \geq 0
 \end{aligned}$$

Dado que las variables deben ser no negativas, el valor óptimo no puede ser inferior a  $-\$37.000$ . Si observamos, basta hacer  $x_3 = x_5 = 0$  para obtener una solución factible que alcanza la cota mínima para el valor óptimo antes mencionada ( $-\$37.000$ ). Así, esta solución debe ser óptima.

Por último, es interesante observar que el tableau final entrega los multiplicadores asociados a cada uno de los recursos necesarios para producir planchas de aluminio. Los costos reducidos indican que si  $x_3$  aumenta en una unidad, la función objetivo empeora en  $\$20$ . Aumentar una variable de holgura en una unidad equivale a disponer de una unidad menos del recurso respectivo. Es decir contar con una unidad más de un recurso equivale a que la holgura respectiva se reduzca en una unidad. Podemos observar que si  $x_3$  cae en una unidad (tomaría el valor  $-1$ ) la función objetivo mejoraría en  $\$20$ . Es decir,  $\lambda_1 = 20$ . Análogamente podemos concluir que  $\lambda_2 = 0$ ;  $\lambda_3 = 1666, \bar{6}$  lo que es consistente con nuestros resultados previos.

### 3.2.2 Solución inicial factible básica

El método Simplex, tal como lo hemos planteado aquí, requiere de una solución inicial factible básica (SIFB) para comenzar a iterar. En cualquier problema de programación lineal en forma estándar con  $\mathbf{b} \geq \mathbf{0}$ , resulta sencillo identificar una SIFB. Basta definir el conjunto de variables básicas como el conjunto de holguras de las restricciones. En ese caso se obtiene:  $\mathbf{x}_{\text{holguras}} = \mathbf{b}$ , y las demás variables originales (no básicas) iguales a cero.

Evidentemente esta situación es bastante particular. Muchos problemas lineales contienen restricciones de igualdad (ecuaciones) o desigualdades tales que al asignar un valor cero a las variables originales del problema no se obtiene un punto factible del dominio.

Volvamos al ejemplo inicial de este capítulo, en que debíamos determinar la producción de dos tipos de planchas de aluminio. Supongamos ahora que por alguna razón estamos forzados a utilizar al menos 15 *lts.* de pintura. En este caso, el problema en forma estándar sería:

$$\begin{aligned}
 \text{Min} \quad & -800x_1 - 600x_2 \\
 \text{s.a.} \quad & 15x_1 + 5x_2 + x_3 = 600 \\
 & 7x_1 + 14x_2 + x_4 = 630 \\
 & 0,3x_1 + 0,3x_2 - x_5 = 15 \\
 & x_i \geq 0, \quad i \in \{1, 2, 3, 4, 5\}
 \end{aligned}$$

en que  $x_3$  y  $x_4$  son variables de holgura y  $x_5$  es de exceso.

Como podemos ver, no poseemos una SIFB. Ante estas situaciones se procederá a resolver el problema en *dos fases*. En la *primera fase* se procurará determinar una SIFB mientras que en la *segunda fase* se procederá con el método Simplex a partir de la SIFB encontrada.

En la primera fase se procurará inventar una solución básica al problema que sea aparente. Para esto se agregará una *variable artificial*  $y_i$  no negativa en cada una de las restricciones que no cumpla con el formato estándar; esto es en las restricciones de igualdad y en las restricciones que requirieron de variables de exceso. Agregar estas variables permite generar inmediatamente una solución básica al nuevo problema: el conjunto de variables no básicas quedaría compuesto por las variables originales y las variables de exceso, mientras las variables básicas serían las artificiales y las de holgura. Esta solución básica sería factible para el nuevo problema, pero no para el problema original. De hecho cualquier solución a este nuevo problema sólo será factible en el problema original si todas las variables artificiales son nulas. Por lo tanto en este nuevo problema se procurará identificar una solución factible básica en que todas las variables artificiales sean no básicas, es decir nulas (en este caso se habrá encontrado una SIFB al problema original). Por este motivo, se reemplazará (sólo durante la primera fase del algoritmo) la función objetivo del problema original por  $\text{Min} \sum y_i$ .

Supongamos que el problema original es el siguiente:

$$\begin{aligned}
 \text{Min} \quad & \mathbf{c}^T \mathbf{x} \\
 \text{s.a.} \quad & \mathbf{A}_1 \mathbf{x} \leq \mathbf{b}_1 \\
 & \mathbf{A}_2 \mathbf{x} \geq \mathbf{b}_2 \\
 & \mathbf{A}_3 \mathbf{x} = \mathbf{b}_3 \\
 & \mathbf{x} \geq 0
 \end{aligned}$$

en que  $b_1, b_2$  y  $b_3$  son vectores no negativos. En ese caso se lleva el problema a formato estándar

(agregando variables de holgura y restando variables de exceso) y a continuación se prepara el problema para la primera fase agregando variables artificiales donde corresponda y modificando la función objetivo. En este caso el problema queda:

$$\begin{aligned}
 \text{Min} \quad & \sum_{y_i \in \mathbf{a}_2} y_i + \sum_{y_i \in \mathbf{a}_3} y_i \\
 \text{s.a.} \quad & \mathbf{A}_1 \mathbf{x} + \mathbf{h}_1 = \mathbf{b}_1 \\
 & \mathbf{A}_2 \mathbf{x} - \mathbf{e}_2 + \mathbf{a}_2 = \mathbf{b}_2 \\
 & \mathbf{A}_3 \mathbf{x} + \mathbf{a}_3 = \mathbf{b}_3 \\
 & \mathbf{x}, \mathbf{h}_1, \mathbf{e}_2, \mathbf{a}_2, \mathbf{a}_3 \geq 0
 \end{aligned}$$

Como se observa, este nuevo problema tiene una SIFB evidente, pues basta considerar los vectores  $\mathbf{h}_1$ ,  $\mathbf{a}_2$  y  $\mathbf{a}_3$  como el conjunto de variables básicas. Así, es posible aplicar Simplex directamente comenzando en dicha solución.

Si en la solución óptima a este problema todos los elementos  $y_i$  tanto de  $\mathbf{a}_2$  como de  $\mathbf{a}_3$  son nulos, entonces basta eliminar las variables artificiales del problema y utilizar la solución óptima como SIFB para el problema original. Si en cambio, la solución óptima contempla algún  $y_i > 0$  significa que no es posible encontrar una solución en que todos los  $y_i$  sean nulos. Esto indicaría que el dominio del problema original no admite soluciones factibles, es decir, es vacío.

A continuación procederemos a aplicar esta metodología al problema del ejemplo.

### Primera fase

En una primera fase, nos interesa resolver el siguiente problema:

$$\begin{aligned}
 \text{Min } y_3 \\
 \text{s.a.} \quad & 15x_1 + 5x_2 + x_3 = 600 \\
 & 7x_1 + 14x_2 + x_4 = 630 \\
 & 0,3x_1 + 0,3x_2 - x_5 + y_3 = 15 \\
 & x_i \geq 0, \quad i \in \{1, 2, 3, 4, 5\} \\
 & y_3 \geq 0
 \end{aligned}$$

Su tableau asociado es el siguiente:

15	5	1	0	0	0	600
7	14	0	1	0	0	630
0,3	0,3	0	0	-1	1	15
0	0	0	0	0	1	0
-800	-600	0	0	0	0	0

Se opta por incluir una fila adicional en la parte inferior del tableau con la función objetivo

original. Esto permitirá facilitar la transición entre la fase 1 y la fase 2 en que se deberá volver a la función objetivo original. Sin embargo, se debe tener claridad que la fila en que se mantiene la función objetivo del problema de la fase 1 es la penúltima fila. Los elementos de dicha fila indicarán la eventual optimalidad de las soluciones alcanzadas. En este tableau, la SIFB corresponde a  $h_1 = 600, h_2 = 630, y_3 = 15$ , mientras las demás variables (no básicas) son cero. Sin embargo, resta ajustar la función objetivo para que contenga sólo ceros en las columnas asociadas a las variables básicas. Para lograrlo es necesario restar la tercera fila a la cuarta fila obteniéndose el siguiente tableau:

15	5	1	0	0	0	600
7	<b>14</b>	0	1	0	0	630
0,3	0,3	0	0	-1	1	15
-0,3	-0,3	0	0	1	0	-15
-800	-600	0	0	0	0	0

La solución no es óptima pues  $x_1$  y  $x_2$  tienen costo reducido negativo (en ambos casos  $-0,3$ ). Realizando una iteración de simplex en el pivote destacado se alcanza el siguiente tableau:

$\frac{25}{2}$	0	1	$-\frac{5}{14}$	0	0	375
$\frac{1}{2}$	1	0	$\frac{1}{14}$	0	0	45
$\frac{3}{20}$	0	0	$-\frac{3}{140}$	-1	1	1,5
$-\frac{3}{20}$	0	0	$\frac{3}{140}$	1	0	-1,5
-500	0	0	$\frac{600}{14}$	0	0	27000

La solución básica alcanzada aún no es factible para el problema original pues  $y_3$  aún es positiva. Observamos también que la solución no es óptima para la fase 1 pues el costo reducido de la variable  $x_1$  es negativo ( $-\frac{3}{20}$ ). Por lo tanto realizamos una iteración adicional y obtenemos:

0	0	1	$\frac{20}{14}$	$\frac{250}{3}$	$-\frac{250}{3}$	250
0	1	0	$\frac{1}{7}$	$\frac{10}{3}$	$-\frac{10}{3}$	40
1	0	0	$-\frac{1}{7}$	$-\frac{20}{3}$	$\frac{20}{3}$	10
0	0	0	0	0	1	0
0	0	0	$-\frac{200}{7}$	$-\frac{10000}{3}$	$\frac{10000}{3}$	32000

La solución básica alcanzada en este caso es óptima para la fase 1 pues todas las variables artificiales (en este caso sólo una) han salido de la base. Es decir a través de operaciones fila hemos



transformado el problema de la fase 1 en el siguiente:

$$\begin{array}{rcl}
 & & \text{Min } y_3 \\
 \text{s.a} & x_3 + \frac{20}{14}x_4 + \frac{250}{3}x_5 - \frac{250}{3}y_3 & = 250 \\
 & x_2 + \frac{1}{7}x_4 + \frac{10}{3}x_5 - \frac{10}{3}y_3 & = 40 \\
 & x_1 - \frac{1}{7}x_4 - \frac{20}{3}x_5 + \frac{20}{3}y_3 & = 10 \\
 & x_i & \geq 0, \quad i \in \{1, 2, 3, 4, 5\} \\
 & y_3 & \geq 0
 \end{array}$$

cuya solución óptima prescinde de la variable  $y_3$ . Así, la solución básica factible alcanzada ( $x_1 = 10, x_2 = 40, x_3 = 250, x_4 = x_5 = 0$ ) es factible para el problema original. Así, basta eliminar la variable artificial ( $y_3$ ) y reemplazar la función objetivo por la original para poder comenzar la segunda fase del problema. Es importante notar que las variables básicas de esta última solución aparecen en la función objetivo original, por lo que es necesario realizar operaciones adicionales para que la función objetivo quede expresada sólo en función de las variables no básicas. Es por este motivo que durante la primera fase mantuvimos la última fila del tableau con la función objetivo original. Para corroborar esta afirmación basta reemplazar el valor de las variables básicas en la función objetivo original:

$$\begin{aligned}
 f(x_1, x_2) &= -800x_1 - 600x_2 \\
 &= -800\left(10 + \frac{1}{7}x_4 + \frac{20}{3}x_5 - \frac{20}{3}y_3\right) - 600\left(40 - \frac{1}{7}x_4 - \frac{10}{3}x_5 + \frac{10}{3}y_3\right) \\
 &= -32.000 - \frac{200}{7}x_4 - \frac{10000}{3}x_5 + \frac{10000}{3}y_3
 \end{aligned}$$

que es exactamente la expresión que aparece en la última fila del tableau. Es interesante notar también que en la penúltima fila correspondiente a la función objetivo artificial el único costo reducido distinto de cero es el de la variable artificial con coeficiente 1. Esto es esperable dado que la variable artificial está fuera de la base y la función objetivo consistía en minimizarla.

Para continuar con el problema basta eliminar la penúltima fila y la penúltima columna. De este modo se obtendrá un tableau inicial para la segunda fase que nos permite comenzar a iterar de inmediato en búsqueda de la solución óptima al problema.

### Segunda fase

Eliminando la cuarta fila y sexta columna del último tableau de la primera fase, obtenemos el tableau inicial para esta segunda fase:

$$\begin{array}{ccccc|c}
 0 & 0 & 1 & \frac{20}{14} & \frac{250}{3} & 250 \\
 0 & 1 & 0 & \frac{1}{7} & \frac{10}{3} & 40 \\
 1 & 0 & 0 & \frac{-1}{7} & \frac{-20}{3} & 10 \\
 \hline
 0 & 0 & 0 & \frac{-200}{7} & \frac{-10000}{3} & 32000
 \end{array}$$

Así comenzamos desde la S.I.F.B.:

$$\begin{aligned}
 x_1 &= 10 \\
 x_2 &= 40 \\
 x_3 &= 250 \\
 x_4 &= 0 \\
 x_5 &= 0
 \end{aligned}$$

Es decir, estamos comenzando desde un vértice del dominio del problema original. Al hacer una iteración del algoritmo obtenemos el siguiente tableau:

$$\begin{array}{ccccc|c}
 0 & 0 & \frac{3}{250} & \frac{3}{175} & 1 & 3 \\
 0 & 1 & \frac{-1}{25} & \frac{15}{175} & 0 & 30 \\
 1 & 0 & \frac{2}{25} & \frac{-5}{175} & 0 & 30 \\
 \hline
 0 & 0 & 40 & \frac{5000}{175} & 0 & 42000
 \end{array}$$

que corresponde al óptimo del problema.

### 3.2.3 Análisis Matricial del Método Simplex

El problema que estamos analizando (una vez agregadas las variables de holgura y de exceso necesarias) se puede describir en su forma estándar o canónica del siguiente modo:

$$\begin{aligned}
 \text{Min} \quad & \mathbf{c}^T \mathbf{x} \\
 \text{s.a.} \quad & \mathbf{Ax} = \mathbf{b} \\
 & \mathbf{x} \geq 0
 \end{aligned}$$

donde debemos notar que todos los componentes del vector  $\mathbf{b}$  deben ser mayor o igual a cero. En cada iteración del algoritmo, los componentes del vector  $\mathbf{x}$  se dividen en dos grupos: las variables básicas a las que asociaremos un vector  $\mathbf{x}_B$  y las variables no básicas a las que asociaremos un vector  $\mathbf{x}_D$ . Así, podremos expresar nuestro problema en función de estos vectores del siguiente

modo:

$$\begin{aligned} \text{Min} \quad & \mathbf{c}_B^\top \mathbf{x}_B + \mathbf{c}_D^\top \mathbf{x}_D \\ \text{s.a.} \quad & \mathbf{B}\mathbf{x}_B + \mathbf{D}\mathbf{x}_D = \mathbf{b} \\ & \mathbf{x}_B, \mathbf{x}_D \geq 0 \end{aligned}$$

en que hemos dividido el vector de costos en  $\mathbf{c}_B^\top$  y  $\mathbf{c}_D^\top$  y la matriz de coeficientes de las restricciones en  $B$  y  $D$  de modo que se ajusten a los dos grupos de variables respectivos. Es importante destacar que la matriz  $B$  es una matriz cuadrada de  $m \times m$  que debe ser de rango máximo (de otro modo el sistema de ecuaciones  $\mathbf{B}\mathbf{x}_B = \mathbf{b}$  podría no tener solución factible).

Si multiplicamos la restricción en ambos lado por  $\mathbf{B}^{-1}$  (para esto no pueden haber filas l.d.) nos queda:

$$\mathbf{x}_B = \mathbf{B}^{-1}\mathbf{b} - \mathbf{B}^{-1}\mathbf{D}\mathbf{x}_D \quad (3.1)$$

y reemplazando en la función objetivo queda:

$$\hat{v} = \mathbf{c}_B^\top \mathbf{B}^{-1}\mathbf{b} + (\mathbf{c}_D^\top - \mathbf{c}_B^\top \mathbf{B}^{-1}\mathbf{D})\mathbf{x}_D \quad (3.2)$$

Así, se alcanza el siguiente problema equivalente al original:

$$\begin{aligned} \text{Min} \quad & \mathbf{c}_B^\top \mathbf{B}^{-1}\mathbf{b} + (\mathbf{c}_D^\top - \mathbf{c}_B^\top \mathbf{B}^{-1}\mathbf{D})\mathbf{x}_D \\ \text{s.a.} \quad & \mathbf{x}_B + \mathbf{B}^{-1}\mathbf{D}\mathbf{x}_D = \mathbf{B}^{-1}\mathbf{b} \\ & \mathbf{x}_B, \mathbf{x}_D \geq 0 \end{aligned}$$

En la medida que todos los elementos de  $\mathbf{B}^{-1}\mathbf{b}$  sean no negativos, la siguiente solución corresponde a una solución básica factible del problema:

$$\begin{aligned} \mathbf{x}_B &= \mathbf{B}^{-1}\mathbf{b} \\ \mathbf{x}_D &= \mathbf{0}. \end{aligned}$$

en que su valor de la función objetivo asociado es:

$$\hat{v} = \mathbf{c}_B^\top \mathbf{B}^{-1}\mathbf{b}$$

Es decir, si el tableau inicial se expresa en forma gruesa del siguiente modo:

$\mathbf{B}$	$\mathbf{D}$	$\mathbf{b}$
$\mathbf{c}_B^\top$	$\mathbf{c}_D^\top$	$\mathbf{0}$

el tableau correspondiente a una solución básica factible será:

<b>I</b>	<b>B<sup>-1</sup>D</b>	<b>B<sup>-1</sup>b</b>
<b>0</b>	<b>c<sub>D</sub><sup>T</sup> - c<sub>B</sub><sup>T</sup>B<sup>-1</sup>D</b>	<b>-c<sub>B</sub><sup>T</sup>B<sup>-1</sup>b</b>

Si además resulta que todos los elementos del vector de **costos reducidos**  $\mathbf{r}_D^T = \mathbf{c}_D^T - \mathbf{c}_B^T \mathbf{B}^{-1} \mathbf{D}$  son no negativos, podemos afirmar que la solución básica factible anterior es también óptima. Observar en (3.2) que cada elemento del vector de costos reducidos puede expresarse como  $c_j - \mathbf{c}_B^T \mathbf{B}^{-1} \mathbf{D}_j$  en que  $c_j$  y  $\mathbf{D}_j$  corresponden al coeficiente asociado a la variable  $x_j$  (no básica) en la función objetivo original y la columna en  $A$  asociada a la variable  $x_j$  respectivamente.

Cabe mencionar que en el caso en que el problema no tenga una SIFB inmediata y haya que recurrir al método de las dos fases para su resolución, este análisis matricial funcionará sin problemas. No importa si hubo que hacer la primera fase o no, solo importa la base actual y referirse a la información del problema original para el análisis.

Veamos el siguiente ejemplo:

$$\begin{aligned}
 \text{Min} \quad & -x_1 - 2x_2 - 3x_3 \\
 & -3x_1 + 2x_2 + x_3 \leq 2 \\
 & -x_1 + 2x_2 - x_3 \leq 4 \\
 & x_1 + x_2 + 4x_3 \leq 5 \\
 & x_1, x_2, x_3 \geq 0
 \end{aligned}$$

Claramente vemos que para transformar el problema a formato estándar habrá que agregar 3 variables de holgura (con sus respectivas restricciones de no negatividad), una para cada restricción. En este caso, las variables de holgura formarán la base inicial. Así,  $\mathbf{x}_B = \{x_4, x_5, x_6\}$  y  $\mathbf{x}_D = \{x_1, x_2, x_3\}$  y las matrices iniciales son:

$$\begin{aligned}
 \mathbf{B} &= \mathbf{B}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, & \mathbf{D} &= \begin{bmatrix} -3 & 2 & 1 \\ -1 & 2 & -1 \\ 1 & 1 & 4 \end{bmatrix}, \\
 \mathbf{b} &= \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}, & \mathbf{c}_B &= \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, & \mathbf{c}_D &= \begin{bmatrix} -1 \\ -2 \\ -3 \end{bmatrix}.
 \end{aligned}$$

Es decir el tableau inicial es el siguiente:

$$\begin{array}{cccccc|c}
 -3 & 2 & 1 & 1 & 0 & 0 & 2 \\
 -1 & 2 & -1 & 0 & 1 & 0 & 4 \\
 1 & 1 & 4 & 0 & 0 & 1 & 5 \\
 \hline
 -1 & -2 & -3 & 0 & 0 & 0 & 0
 \end{array}$$

Dado que el vector  $\mathbf{c}_B$  es nulo,  $\mathbf{r}_D^T = \mathbf{c}_D^T$ , y así vemos que todas las variables no básicas tienen costo reducido negativo. Es decir, si incorporamos cualquiera de ellas a la base, la función objetivo mejora (lo que es bastante obvio en este caso). Escogeremos la variable con costo reducido más negativo, esto es  $\min\{-1, -2, -3\} = -3$ ; es decir la variable  $x_3$  entra a la base. Para determinar la variable que sale de la base debemos realizar el cociente entre los elementos de la columna del extremo derecho ( $\mathbf{B}^{-1}\mathbf{b}$ ) del tableau y los elementos positivos de la columna de  $D$  asociada a  $x_3$ . Es decir  $\min\{\frac{2}{1}, \cdot, \frac{5}{4}\} = \frac{5}{4}$ . Esto indica que la variable básica que aparezca en la tercera restricción es la que sale de la base, ésta es  $x_6$ . Esto se debe a que cuando  $x_3$  crece,  $x_6$  es la primera variable básica que se vuelve cero.

El próximo tableau lo podemos obtener pivoteando (mediante operaciones filas) en la posición de  $x_3$  de la tercera fila, o bien simplemente redefinir las matrices del problema acorde a la nueva base  $\mathbf{x}_B = \{x_4, x_5, x_3\}$ . Definiremos el conjunto de variables no básicas como:  $\mathbf{x}_D = \{x_1, x_2, x_6\}$ . Las estructuras matriciales asociadas a esta nueva base son:

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 4 \end{bmatrix} \Rightarrow \mathbf{B}^{-1} = \begin{bmatrix} 1 & 0 & -\frac{1}{4} \\ 0 & 1 & \frac{1}{4} \\ 0 & 0 & \frac{1}{4} \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} -3 & 2 & 0 \\ -1 & 2 & 0 \\ 1 & 1 & 1 \end{bmatrix},$$

$$\mathbf{b} = \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}, \quad \mathbf{c}_B = \begin{bmatrix} 0 \\ 0 \\ -3 \end{bmatrix}, \quad \mathbf{c}_D = \begin{bmatrix} -1 \\ -2 \\ 0 \end{bmatrix}$$

En este desarrollo matricial es importante mantener claridad respecto al orden en que las variables forman el conjunto  $\mathbf{x}_B$  y  $\mathbf{x}_D$  de modo de garantizar consistencia entre las distintas estructuras matriciales del sistema. Por ejemplo, si alternativamente definimos  $\mathbf{x}_B = \{x_3, x_4, x_5\}$  las matrices asociadas a las variables no básicas deben absorber esta modificación. Por ejemplo,  $\mathbf{D}$  en este caso sería:

$$\mathbf{D} = \begin{bmatrix} 0 & -3 & 2 \\ 0 & -1 & 2 \\ 1 & 1 & 1 \end{bmatrix}$$

Es importante destacar que el orden que se les da a las variables dentro de la base o del conjunto de variables no básicas es irrelevante (el resultado será idéntico), lo que importa es mantener la consistencia. Realizando las operaciones necesarias para determinar el nuevo tableau obtenemos:

$$\mathbf{B}^{-1}\mathbf{D} = \begin{bmatrix} \frac{-13}{4} & \frac{7}{4} & \frac{-1}{4} \\ \frac{-3}{4} & \frac{9}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}, \quad \mathbf{B}^{-1}\mathbf{b} = \begin{bmatrix} \frac{3}{4} \\ \frac{21}{4} \\ \frac{5}{4} \end{bmatrix},$$

$$\mathbf{r}_D = (\mathbf{c}_D^T - \mathbf{c}_B^T \mathbf{B}^{-1} \mathbf{D})^T = \begin{bmatrix} -\frac{1}{4} \\ -\frac{5}{4} \\ \frac{3}{4} \end{bmatrix}, \quad \mathbf{c}_B^T \mathbf{B}^{-1} \mathbf{b} = -\frac{15}{4}$$

es decir, obtenemos el siguiente tableau:

$$\begin{array}{cccccc|c}
 \frac{-13}{4} & \mathbf{\frac{7}{4}} & 0 & 1 & 0 & \frac{-1}{4} & \frac{3}{4} \\
 \frac{-3}{4} & \frac{9}{4} & 0 & 0 & 1 & \frac{1}{4} & \frac{21}{4} \\
 \frac{1}{4} & \frac{1}{4} & 1 & 0 & 0 & \frac{1}{4} & \frac{5}{4} \\
 \hline
 \frac{-1}{4} & \frac{-5}{4} & 0 & 0 & 0 & \frac{3}{4} & \frac{15}{4}
 \end{array}$$

Evidentemente este mismo tableau se hubiera alcanzado si se hubiera realizado operaciones fila pivotando en el casillero de la tercera fila y tercera columna (destacado en negrita).

¿Es esta solución óptima? Evidentemente no, ya que aún hay costos reducidos negativos. Decidimos que  $x_2$  entra a la base pues  $\min\{-\frac{1}{4}, -\frac{5}{4}\} = -\frac{5}{4}$ . La variable saliente de la base es  $x_4$  pues  $\min\{\frac{3}{7}, \frac{21}{9}, 5\} = \frac{3}{7}$  lo que indica que es la variable de la base asociada a la primera restricción la que deja la base.

Pivoteando obtenemos:

$$\begin{array}{cccccc|c}
 \frac{-13}{7} & 1 & 0 & \frac{4}{7} & 0 & \frac{-1}{7} & \frac{3}{7} \\
 \mathbf{\frac{24}{7}} & 0 & 0 & \frac{-9}{7} & 1 & \frac{4}{7} & \frac{30}{7} \\
 \frac{5}{7} & 0 & 1 & \frac{-1}{7} & 0 & \frac{2}{7} & \frac{8}{7} \\
 \hline
 \frac{-18}{7} & 0 & 0 & \frac{5}{7} & 0 & \frac{4}{7} & \frac{30}{7}
 \end{array}$$

con  $\mathbf{x}_B = \{x_5, x_3, x_2\}$  y  $\mathbf{x}_D = \{x_1, x_6, x_4\}$ . ¿Estamos en el óptimo? Aún no, ya que queda un costo reducido negativo ( $\frac{-18}{7}$ )  $\Rightarrow x_1$  entra a la base.

Así,  $\min\{\cdot, \frac{30}{24}, \frac{8}{5}\} = \frac{5}{4} \Rightarrow x_5$  sale de la base.

Pivoteando

$$\begin{array}{cccccc|c}
 0 & 1 & 0 & \frac{-1}{8} & \frac{13}{24} & \frac{1}{6} & \frac{11}{4} \\
 1 & 0 & 0 & \frac{-3}{8} & \frac{7}{24} & \frac{1}{6} & \frac{5}{4} \\
 0 & 0 & 1 & \frac{1}{8} & \frac{-5}{24} & \frac{1}{6} & \frac{1}{4} \\
 \hline
 0 & 0 & 0 & \frac{-1}{4} & \frac{3}{4} & 1 & \frac{15}{2}
 \end{array}$$

Como se puede ver, podemos pivotar nuevamente. Así obtenemos:

$$\begin{array}{cccccc|c}
 0 & 1 & 1 & 0 & \frac{1}{3} & \frac{1}{3} & 3 \\
 1 & 0 & 3 & 0 & \frac{-1}{3} & \frac{2}{3} & 2 \\
 0 & 0 & 8 & 1 & \frac{-5}{3} & \frac{4}{3} & 2 \\
 \hline
 0 & 0 & 2 & 0 & \frac{1}{3} & \frac{4}{3} & 8
 \end{array}$$

donde todos los costos reducidos son mayores o iguales a cero. Por lo tanto, la solución óptima es:

$$\begin{array}{rcl}
 x_1 & = & 2 \quad x_4 = 2 \\
 x_2 & = & 3 \quad x_5 = 0 \\
 x_3 & = & 0 \quad x_6 = 0
 \end{array}$$

que alcanza en la función objetivo un valor óptimo  $\hat{v} = -8$ .

Como se puede observar, la variable  $x_4$  estaba inicialmente en la base, luego salió de ésta, y

en la última iteración volvió a entrar. Esto nos hace pensar que escogimos el camino largo para llegar al óptimo. Le dejamos al lector que resuelva este problema nuevamente pero en la segunda iteración incorporando la variable  $x_1$  en vez de  $x_2$ .

### 3.2.4 Casos especiales en el desarrollo de Simplex

#### (a) Solución degenerada

¿Qué pasa si al buscar la variable que sale de la base me encuentro un empate, es decir, hay dos cocientes  $\frac{b_i}{a_{ij}}$  de idéntico valor? En dicho caso, observamos que la solución será degenerada (alguna variable básica tomará valor cero). Por ejemplo, supongamos que el tableau inicial del ejemplo anterior ahora tiene  $b_3 = 8$  en vez de  $b_3 = 5$ , de la siguiente manera:

$$\begin{array}{cccccc|c} -3 & 2 & 1 & 1 & 0 & 0 & 2 \\ -1 & 2 & -1 & 0 & 1 & 0 & 4 \\ 1 & 1 & 4 & 0 & 0 & 1 & 8 \\ \hline -1 & -2 & -3 & 0 & 0 & 0 & 0 \end{array}$$

Al determinar la variable saliente, asumiendo que  $x_3$  entra a la base, tenemos que  $\min\{\frac{2}{1}, \cdot, \frac{8}{4}\} = 2$ . Pivoteando en alguna de las dos alternativas tenemos:

$$\begin{array}{cccccc|c} -3 & 2 & 1 & 1 & 0 & 0 & 2 \\ -4 & 4 & 0 & 1 & 1 & 0 & 6 \\ 13 & -7 & 0 & -4 & 0 & 1 & 0 \\ \hline -10 & 4 & 0 & 3 & 0 & 0 & 6 \end{array}$$

En esta caso la solución alcanzada no corresponde a un óptimo, sin embargo el vértice corresponde a una **solución factible básica degenerada**, donde  $\mathbf{x}_B = \{x_3, x_5, x_6\}$ , con  $x_3 = 2$ ,  $x_5 = 6$ , y  $x_6 = 0$ . Geométricamente una solución básica degenerada se crea cuando confluye en un punto un número mayor de restricciones al mínimo necesario para definir un vértice. Así, si el espacio de soluciones es  $\mathbb{R}^2$  entonces se crea una solución degenerada cuando coinciden tres restricciones en un punto. Un ejemplo de solución degenerada en  $\mathbb{R}^3$  sería el vértice superior de una pirámide cuya base tenga cuatro o más esquinas. En estos casos el vértice geométrico queda definido por varias bases distintas. Esto causa un problema, pues si la solución degenerada es óptima muy posiblemente no todas esas bases lo serán. Así, es posible haber alcanzado la solución óptima y no haberlo detectado. Si la solución degenerada no es óptima, entonces el algoritmo podrá requerir varias iteraciones en el vértice cambiando en cada una de base (pero no de vértice) para por fin poder moverse a una solución con mejor valor de la función objetivo.

#### (b) Dominio no acotado

¿Qué pasa si al buscar la variable que sale de la base, me encuentro con que todos los  $a_{ij}$  son negativos? En dicho caso, observamos que la solución es no acotada, es decir, el dominio es no acotado. Por ejemplo, si la última restricción hubiese sido  $-x_1 + x_2 + 4x_3 \leq 8$ , y de haber escogido

que entrase  $x_1$  a la base, hubiéramos tenido que escoger entre valores negativos. Esto es razonable, pues  $x_1$  aumenta la función objetivo y disminuye las restricciones. Las restricciones serían:

$$\begin{aligned} -3x_1 + 2x_2 + x_3 + x_4 &= 2 \\ -x_1 + 2x_2 - x_3 + x_5 &= 4 \\ -x_1 + x_2 + 4x_3 + x_6 &= 8 \end{aligned}$$

Ahora si  $x_2$  y  $x_3$  permanecen en valor cero, tenemos que

$$\begin{aligned} x_1 &= (x_4 - 2)/3 \\ x_1 &= x_5 - 4 \\ x_1 &= x_6 - 8 \end{aligned}$$

donde no hay un valor máximo para  $x_1$ .

### (c) Múltiples soluciones

¿Qué pasa si al analizar la solución final observamos que hay un costo reducido igual a cero asociado a una variable no básica? Significa que si esa variable se incorpora a la base, el valor de la función objetivo no se verá alterado. En dicho caso, observamos que hay dos vértices con idéntico valor objetivo. Entonces cualquier combinación convexa de ambos, sería óptima también.

**Ejemplo:** Utilicemos el análisis matricial para determinar cuánto debió haber sido  $c_3$  para que

hubiese sido atractivo de incorporar a la base. El costo reducido final de las variables no básicas es:  $\mathbf{c}_D^T - \mathbf{c}_B^T \mathbf{B}^{-1} \mathbf{D}$ , y en este caso,  $\mathbf{x}_D = \{x_3, x_5, x_6\}$ ,  $\mathbf{x}_B = \{x_1, x_2, x_4\}$ .

$$\left. \begin{aligned} \mathbf{c}_D^T &= [-3 \quad 0 \quad 0] \\ \mathbf{c}_B^T &= [-1 \quad -2 \quad 0] \\ \mathbf{B} &= \begin{bmatrix} 3 & 2 & 1 \\ -1 & 2 & 0 \\ 1 & 1 & 0 \end{bmatrix} \\ \mathbf{D} &= \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix} \end{aligned} \right\} \begin{aligned} \mathbf{c}_B^T \mathbf{B}^{-1} \mathbf{D} &= [-5 \quad \frac{-1}{3} \quad \frac{-4}{3}] \\ \text{Por lo tanto, } \mathbf{c}_D^T - \mathbf{c}_B^T \mathbf{B}^{-1} \mathbf{D} &= [2 \quad \frac{1}{3} \quad \frac{4}{3}] \end{aligned}$$

Entonces, para que  $x_3$  sea atractivo de incorporar a la base,  $c_3 - 5 \geq 0 \Rightarrow c_3 \geq 5$ . Cualquier valor mayor que 5 resulta atractivo. Supongamos que  $c_3 = 5$ , así la solución:

$$\begin{aligned} x_1 &= 2 & x_4 &= 2 \\ x_2 &= 3 & x_5 &= 0 \\ x_3 &= 0 & x_6 &= 0 \end{aligned}$$



sigue siendo óptima.

Ahora podríamos incorporar  $x_3$  a la base.

$$\begin{array}{cccc|cc|c} 0 & 1 & 1 & 0 & \frac{1}{3} & \frac{1}{3} & 3 \\ 1 & 0 & 3 & 0 & \frac{1}{3} & \frac{2}{3} & 2 \\ 0 & 0 & \mathbf{8} & 1 & \frac{-5}{3} & \frac{4}{3} & 2 \\ \hline 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{4}{3} & 8 \end{array}$$

Pivoteando para que salga  $x_4$  de la base, tenemos

$$\begin{array}{cccc|cc|c} 0 & 1 & 0 & \frac{-1}{8} & \frac{13}{24} & \frac{1}{6} & \frac{11}{4} \\ 1 & 0 & 0 & \frac{-3}{8} & \frac{7}{24} & \frac{1}{6} & \frac{5}{4} \\ 0 & 0 & 1 & \frac{1}{8} & \frac{-5}{24} & \frac{2}{3} & \frac{1}{4} \\ \hline 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{4}{3} & 8 \end{array},$$

es decir, otra solución de vértice es:

$$\begin{aligned} x_1 &= \frac{5}{4} & x_4 &= 0 \\ x_2 &= \frac{11}{4} & x_5 &= 0 \\ x_3 &= \frac{1}{4} & x_6 &= 0 \end{aligned}$$

la cual también nos da el mismo valor de la función objetivo  $\hat{v} = -8$ . Por lo tanto, también es solución óptima. Ahora, cualquier combinación convexa también será solución óptima, es decir, todo  $\mathbf{x}$  tal que

$$\mathbf{x} = \alpha \begin{bmatrix} \frac{5}{4} \\ \frac{11}{4} \\ \frac{1}{4} \\ 0 \\ 0 \\ 0 \end{bmatrix} + (1 - \alpha) \begin{bmatrix} 2 \\ 3 \\ 0 \\ 2 \\ 0 \\ 0 \end{bmatrix} \quad \text{donde } 0 \leq \alpha \leq 1.$$

Supongamos que a última hora se nos ocurre incorporar una nueva variable que podría ser atractiva en la solución óptima. Para verificar esto tomamos el tableau final y vemos cuáles son las variables básicas y cuáles las no básicas.

Podremos estimar el costo reducido que hubiera tenido esta variable a esta altura, si nunca hubiese sido incorporada a la base.

Por ejemplo, si una nueva actividad  $x_7$  aportara  $-6$  por unidad a la función objetivo y consum-

iera una unidad de cada recurso en las restricciones, el costo reducido respectivo sería:

$$r_7 = -6 - [-1 \quad -2 \quad 0] \begin{bmatrix} -3 & 2 & 1 \\ -1 & 2 & 0 \\ 1 & 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{-13}{3} < 0$$

Por lo tanto, sí nos interesaría incorporar  $x_7$  a la base y así mejorar la solución actual.

### 3.3 Análisis de Sensibilidad de los Resultados

En nuestro análisis hemos asumido un conocimiento cabal de cada uno de los parámetros que influyen en el problema. Sin embargo, a menudo la realidad es muy distinta. Por una parte, el modelador suele no conocer perfectamente los valores exactos que toman los parámetros. Por otra, aunque los conociera, muchas veces los parámetros tienen una naturaleza estocástica que impide asignarles un valor único. Es decir, muchas veces problemas de naturaleza estocástica son modelados como si fueran determinísticos, pues esta modelación permite abordarlos con mayor facilidad (la optimización estocástica escapa del alcance de este texto). Por este motivo, cuando se modela un problema en forma determinística (en que se asume que se conoce el valor de cada parámetro con total precisión), a menudo interesa saber cuán robusto es el resultado frente a perturbaciones en algunos parámetros del modelo.

A continuación se presenta un análisis de sensibilidad de los resultados de problemas de programación lineal continua como los que se han tratado en este texto con respecto a los parámetros de la función objetivo y de los niveles de cada uno de los recursos disponibles. En cada uno de los casos analizados se perturbará sólo un parámetro a la vez. Un análisis de sensibilidad más complejo en que se perturban varios parámetros simultáneamente no será analizado en las secciones que siguen, sin embargo, responde a la misma lógica que se explicará a continuación.

#### 3.3.1 Rango de variación de los costos

En esta sección realizaremos un análisis de sensibilidad de la solución óptima al problema respecto de los parámetros de su función objetivo. Evidentemente modificar la función objetivo no cambia en absoluto el dominio del problema. Una función objetivo diferente sólo podría modificar su vértice óptimo. Por lo tanto, el análisis de sensibilidad aquí presentado pretenderá identificar el rango de las perturbaciones en uno de los coeficientes de la función objetivo que no altera la solución óptima alcanzada.

Como hemos visto, el vector de costos de la función objetivo lo podemos dividir en dos grupos: los costos asociados a las variables básicas y los asociados a las variables no básicas. A medida que la base cambia en búsqueda de una base óptima, los costos reducidos también lo hacen.

Analicemos primero variaciones en algún costo asociado a una variable no básica de la solución óptima. Para poder preservar la optimalidad de la solución alcanzada, los costos reducidos deben

permanecer positivos o nulos (el caso de un problema de minimización). Es decir,

$$\mathbf{r}_D = \mathbf{c}_D - \mathbf{z} \geq \mathbf{0}$$

donde  $\mathbf{z}^\top = \mathbf{c}_B^\top \mathbf{B}^{-1} \mathbf{D}$ . Por lo tanto,

$$r_k = c_k - z_k \geq 0, \quad \forall k \text{ no básico.}$$

Ahora, supongamos que el costo  $c_j$  asociado a la variable no básica  $x_j$  varía en  $\Delta c_j$ . Para que esta variable siga siendo no básica deberá cumplirse que su respectivo costo reducido siga siendo no negativo, es decir,

$$c_j + \Delta c_j - z_j \geq 0 \rightarrow r_j + \Delta c_j \geq 0$$

Por lo tanto, el rango de esta variación  $\Delta c_j$  es

$$-r_j \leq \Delta c_j \leq \infty$$

Este resultado es muy simple de interpretar. Basta conocer el costo reducido de una variable no básica para saber el rango en que se puede mover su coeficiente en la función objetivo sin que la solución óptima al problema se vea afectada.

Por otro lado, respecto a costos asociados a variables básicas, el análisis es un poco más complejo. Esto se debe a que una perturbación en algún elemento de  $\mathbf{c}_B$ , digamos  $c_j$ , alterará todos los componentes del vector  $\mathbf{z}^\top = \mathbf{c}_B^\top \mathbf{B}^{-1} \mathbf{D}$ . En la medida que ante esta perturbación los costos reducidos no básicos sigan siendo no negativos (los básicos seguirán siendo nulos por construcción), la base óptima permanecerá inalterada. Sin embargo, basta que uno de los costos reducidos no básicos se torne negativo para que la base (y por tanto la solución) óptima cambie. Definamos  $z_k = \sum_{i=1}^r c_i \alpha_{ik}$ , donde  $\alpha_{ik}$  es el producto de la fila  $i$  de la matriz inversa de la base con la columna  $k$  de la matriz  $\mathbf{D}$ , es decir,  $\alpha_{ik} = B_{\text{fila } i}^{-1} D_k$ . De esta forma si modificamos el costo  $c_j$  en  $\Delta c_j$ , los costos reducidos  $\hat{r}_k$  de todas la variables no básicas quedarán:

$$\begin{aligned} \hat{r}_k &= c_k - z_k \\ &= c_k - \left( \sum_{i \in \text{Base}} c_i \alpha_{ik} \right) - \Delta c_j \alpha_{jk} \\ &= r_k - \Delta c_j \alpha_{jk} \geq 0, \quad \forall k \text{ no básico} \end{aligned}$$

Por lo tanto, en la medida que  $r_k \geq \Delta c_j \alpha_{jk}$  para todos los costos reducidos, la base óptima no cambiará. Para que esto suceda, dado que los  $\alpha_{jk}$  pueden ser positivos, negativos o nulos, deberá

cumplirse que:

$$\begin{aligned} \text{Si } \alpha_{jk} > 0 &\Rightarrow \Delta c_j \leq \frac{r_k}{\alpha_{jk}} \\ \text{Si } \alpha_{jk} < 0 &\Rightarrow \Delta c_j \geq \frac{r_k}{\alpha_{jk}} \\ \text{Si } \alpha_{jk} = 0 &\Rightarrow \Delta c_j \text{ puede tomar cualquier valor} \end{aligned}$$

Recordemos que  $r_k \geq 0$ . Como estas condiciones deben cumplirse  $\forall k$ , el rango de la variación  $\Delta c_j$  que mantiene la base óptima inalterada será:

$$\max_{\forall k: \alpha_{jk} < 0} \left\{ \frac{r_k}{\alpha_{jk}} \right\} \leq \Delta c_j \leq \min_{\forall k: \alpha_{jk} > 0} \left\{ \frac{r_k}{\alpha_{jk}} \right\}. \quad (3.3)$$

Es importante notar que dentro de este rango el valor óptimo cambiará en  $\Delta c_j x_j$ .

¿Qué ocurre si  $\Delta c_j$  llega a un límite? ¿Cuál es la nueva base?

En ambos casos de perturbaciones en el coeficiente  $c_j$  de la función objetivo de una variable  $x_j$  (sea ésta no básica o básica), si  $\Delta c_j$  llega a uno de los bordes de los intervalos derivados en esta sección, se tendrá que el costo reducido de (al menos) una variable no básica se vuelve nulo. Esto indicaría que en ese caso la solución óptima no sería única pues de incorporar la variable de costo reducido nulo a la base se obtendría una nueva solución básica con idéntico valor en la función objetivo. Ahora bien,

- i) si  $c_j$  está asociada a una variable no-básica, entonces en la solución óptima alternativa esta variable pasará a ser básica.
- ii) si  $c_j$  está asociada a una variable básica, entonces en la solución óptima alternativa la variable de índice  $k$  que establece el límite entrará a la base, y dejará la base la variable de índice  $i$  tal que:

$$\min_i \left\{ \frac{x_i}{\alpha_{ik}} : \alpha_{ik} \geq 0 \right\}.$$

En ambos casos la solución óptima alternativa corresponderá a un vértice adyacente al vértice asociado a la solución óptima original en el dominio del problema.

### 3.3.2 Rango de variación del nivel de recursos

En esta sección realizaremos un análisis de sensibilidad de la solución óptima respecto de los parámetros del vector de recursos del sistema. En este caso la función objetivo se mantiene inalterada, pero se modifica el dominio de las soluciones. Si se modifica el nivel de recursos de una restricción activa, la solución óptima (a menos que ésta sea degenerada) inmediatamente cambiará. Sin embargo, si la solución óptima continúa siendo definida por la misma base (el mismo conjunto de restricciones activas) entonces los costos reducidos permanecerán inalterables (no dependen de  $b$ ). Es decir, perturbaciones pequeñas en uno de los recursos disponibles no cambiaría el conjunto de variables básicas, pero sí (levemente) el valor que ellas toman en la solución óptima. Sin embargo,

perturbaciones mayores a un determinado umbral convertirían la base en una base infactible (el valor que toma una de las variables básicas se torna negativo). Así, interesará identificar el rango de variación en el recurso considerado que mantiene la base óptima actual dentro del dominio. El análisis para recursos de restricciones inactivas en el óptimo es similar. Perturbaciones dentro de un cierto rango no modificarán la factibilidad de la base (la holgura de la restricción modificada seguirá siendo positiva). Sin embargo, si se restringe el nivel de recursos sobre un determinado umbral, la base pasará a ser infactible debido a que la holgura del recurso en cuestión será negativa (el recurso no será suficiente para los requerimientos de la solución óptima).

La factibilidad de una base queda definida por el lado derecho del tableau asociado a dicha base, es decir  $\mathbf{x}_B = \mathbf{B}^{-1}\mathbf{b}$  que debe ser no negativo. Así, la solución óptima puede expresarse como:

$$\mathbf{x}^* = \mathbf{B}^{-1}\mathbf{b} = \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{m1} & \beta_{m2} & \cdots & \beta_{mm} \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} = \begin{bmatrix} b_1\beta_{11} + b_2\beta_{12} + \cdots + b_m\beta_{1m} \\ \vdots \\ b_1\beta_{m1} + b_2\beta_{m2} + \cdots + b_m\beta_{mm} \end{bmatrix}$$

Por lo tanto, si el nivel del recurso  $j$  varía en  $\Delta b_j$ , se deberá cumplir que

$$\hat{\mathbf{x}}^* = \begin{bmatrix} b_1\beta_{11} + b_2\beta_{12} + \cdots + b_m\beta_{1m} + \Delta b_j\beta_{1j} \\ \vdots \\ b_1\beta_{m1} + b_2\beta_{m2} + \cdots + b_m\beta_{mm} + \Delta b_j\beta_{mj} \end{bmatrix} \geq 0$$

Es decir,  $x_i^* + \beta_{ij}\Delta b_j \geq 0$ ,  $\forall i = 1, \dots, m$ . De esta condición se desprende que debe cumplirse:

$$\max_{\forall i: \beta_{ij} > 0} \left\{ \frac{-x_i^*}{\beta_{ij}} \right\} \leq \Delta b_j \leq \min_{\forall i: \beta_{ij} < 0} \left\{ \frac{-x_i^*}{\beta_{ij}} \right\}$$

para que la base óptima no cambie.

Al realizar cambios en los  $b_j$ , dentro de este rango, la nueva solución óptima será  $x_i^* + \beta_{ij}\Delta b_j$  para todo  $i$ , y el cambio en la función objetivo será:  $\lambda_j\Delta b_j$ . Observar que dentro de este rango de variación, los costos reducidos no se alterarán.

### 3.3.3 Ejemplo

Volvamos a nuestro ejemplo de las planchas de aluminio, y veamos cuánto podría variar el beneficio de las planchas de tipo 1 sin que cambie la política de producción. El tableau inicial era:

15	5	1	0	0	600
7	14	0	1	0	630
0,3	0,3	0	0	1	15
800	600	0	0	0	0

Y el tableau final era:

$$\begin{array}{ccccc|c} 1 & 0 & \frac{1}{10} & 0 & \frac{-5}{3} & 35 \\ 0 & 0 & \frac{7}{10} & 1 & \frac{-175}{3} & 175 \\ 0 & 1 & \frac{-1}{10} & 0 & 5 & 15 \\ \hline 0 & 0 & -20 & 0 & \frac{-5000}{3} & -37000 \end{array}$$

Como podemos ver:

$$\mathbf{B}^{-1} = \begin{bmatrix} \frac{1}{10} & 0 & \frac{-5}{3} \\ \frac{7}{10} & 1 & \frac{-175}{3} \\ \frac{-1}{10} & 0 & 5 \end{bmatrix} \quad \text{y} \quad \mathbf{D} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Como  $c_1$  está asociado a una variable básica, entonces su rango de variación está determinado por (3.3). Aquí tenemos  $k = 1$  y  $k = 2$ , son las variables no-básicas. Tenemos que calcular los  $\alpha_{1k}$ , es decir,  $\alpha_{11}$  y  $\alpha_{12}$ .

$$\alpha_{11} = \begin{bmatrix} \frac{1}{10} & 0 & \frac{-5}{3} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \frac{1}{10} > 0,$$

$$\alpha_{12} = \begin{bmatrix} \frac{1}{10} & 0 & \frac{-5}{3} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \frac{-5}{3} < 0. \quad r_2 = \frac{-5000}{3}.$$

Además sabemos que  $r_1 = -20$  y  $r_2 = \frac{-5000}{3}$ . Así,

$$\max_{\alpha_{1k} > 0} \left\{ \frac{-20}{\frac{1}{10}} \right\} \leq \Delta c_1 \leq \min_{\alpha_{1k} < 0} \left\{ \frac{\frac{-5000}{3}}{\frac{-5}{3}} \right\},$$

por lo tanto,  $-200 \leq \Delta c_1 \leq 1000$  para que no cambie la base.

Es fácil comprobarlo con los gradientes de las restricciones activas en el punto (35, 15). Si  $\Delta c_1 = 1000 \Rightarrow \nabla f = (1800, 600)$ , y si  $\Delta c_1 = -200 \Rightarrow \nabla f = (600, 600)$ , que corresponden a los gradientes de las restricciones activas en el punto óptimo.

### 3.4 Dualidad

<sup>2</sup>En la presente sección se analiza una característica muy particular de los problemas de progra-

mación lineal continua. Para cada problema de este tipo existe un problema asociado que a simple vista parece completamente diferente, pero que comparte múltiples de sus características y de cuya solución se pueden desprender elementos de la solución del problema original. A este nuevo problema le denominaremos el problema dual del problema original (o primal).

---

<sup>2</sup> Agregar dibujo de soluciones optimas y suboptimas en PP y PD

Consideremos el siguiente problema general:

$$\begin{aligned} PP) \quad & \text{Max } \mathbf{c}^\top \mathbf{x} \\ & \text{s.a. } \mathbf{Ax} \leq \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned}$$

Podemos escribir este problema como uno equivalente de minimización de manera de dejarlo en nuestro formato estándar para establecer las condiciones KKT de la forma vista en este libro.

$$\begin{aligned} PP) \quad & \text{Min } -\mathbf{c}^\top \mathbf{x} \\ & \text{s.a. } \mathbf{Ax} \leq \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned}$$

El Lagrangeano de este problema es  $L = -\mathbf{c}^\top \mathbf{x} + \boldsymbol{\lambda}^\top (\mathbf{Ax} - \mathbf{b})$ , y las respectivas condiciones de KKT son:

$$\begin{aligned} \frac{\partial L}{\partial x_i} = -\mathbf{c}_i + \boldsymbol{\lambda}^\top \mathbf{A}_i &\geq \mathbf{0} & \forall i = 1, \dots, n & \quad \frac{\partial L}{\partial \lambda_j} = \mathbf{A}^j \mathbf{x} - \mathbf{b}_j &\leq \mathbf{0} & \quad \forall j = 1, \dots, m \\ x_i \frac{\partial L}{\partial x_i} = x_i(-\mathbf{c}_i + \boldsymbol{\lambda}^\top \mathbf{A}_i) &= \mathbf{0} & \forall i = 1, \dots, n & \quad \lambda_j \frac{\partial L}{\partial \lambda_j} = \lambda_j(\mathbf{A}^j \mathbf{x} - \mathbf{b}_j) &= \mathbf{0} & \quad \forall j = 1, \dots, m \\ x_i &\geq \mathbf{0} & \forall i = 1, \dots, n & \quad \lambda_j &\geq \mathbf{0} & \quad \forall j = 1, \dots, m \end{aligned}$$

En que  $\mathbf{A}_i$  y  $\mathbf{A}^j$  corresponden al vector  $i$ -ésima columna y  $j$ -ésima fila de la matriz  $\mathbf{A}$  respectivamente. El primer vector es de dimensiones  $m * 1$ , mientras el segundo es de  $1 * n$ .

Así, para que un vector  $\mathbf{x}$  sea óptimo debe existir también otro vector  $\boldsymbol{\lambda}$  tal que en conjunto satisfagan estas condiciones. Además debe suceder que los gradientes de las restricciones activas sean l.i. para que exista regularidad.

De las condiciones de KKT observamos que la solución óptima debe satisfacer:

$$\mathbf{c}^\top \mathbf{x} \leq \boldsymbol{\lambda}^\top \mathbf{Ax} \leq \boldsymbol{\lambda}^\top \mathbf{b}$$

Esto se cumple ya que la solución óptima  $(\mathbf{x}, \boldsymbol{\lambda})$  debe cumplir:  $\boldsymbol{\lambda} \geq \mathbf{0}$ ,  $\boldsymbol{\lambda}^\top \mathbf{A} \geq \mathbf{c}^\top$ ,  $\mathbf{x} \geq \mathbf{0}$  y  $\mathbf{Ax} \leq \mathbf{b}$ . Adicionalmente, se observa que en el óptimo (si éste existe), debe cumplir también que:

$$\begin{aligned} \sum_{i=1}^n x_i(-\mathbf{c}_i + \boldsymbol{\lambda}^\top \mathbf{A}_i) &= \mathbf{0} \Rightarrow \mathbf{c}^\top \mathbf{x} = \boldsymbol{\lambda}^\top \mathbf{Ax} \\ \sum_{j=1}^m \lambda_j(\mathbf{A}^j \mathbf{x} - \mathbf{b}_j) &= \mathbf{0} \Rightarrow \boldsymbol{\lambda}^\top \mathbf{Ax} = \boldsymbol{\lambda}^\top \mathbf{b} \\ \therefore \mathbf{c}^\top \mathbf{x} &= \boldsymbol{\lambda}^\top \mathbf{b} \end{aligned}$$

Es decir estos dos problemas de maximización y minimización comparten el mismo valor óptimo. Si llamamos  $v$  a este valor entonces el valor en la función objetivo de cualquier solución factible no óptima a  $PP$ ) será necesariamente inferior a  $v$ . Asimismo, veremos que el valor de la función

objetivo de cualquier solución factible no óptima en el problema dual será necesariamente superior a  $v$ .

De la estructura del problema y de esta observación surge la idea de reescribir un nuevo problema equivalente al problema primal ( $PP$ ). A este nuevo problema equivalente le denominaremos **Problema Dual** ( $PD$ ) :

$$\begin{aligned} PD) \quad & \text{Min } \boldsymbol{\lambda}^T \mathbf{b} \\ \text{s.a} \quad & \boldsymbol{\lambda}^T \mathbf{A} \geq \mathbf{c}^T \\ & \boldsymbol{\lambda} \geq \mathbf{0} \end{aligned}$$

Es fácil observar que las condiciones de KKT de este problema son exactamente las mismas que las del problema original en que las variables del vector  $\mathbf{x}$  toman el rol de los multiplicadores de las restricciones del problema dual. Así, los siguientes problemas son sustancialmente equivalentes:

$$\begin{array}{ll} PP) \quad \text{Max } \mathbf{c}^T \mathbf{x} & PD) \quad \text{Min } \boldsymbol{\lambda}^T \mathbf{b} \\ \text{s.a} \quad \mathbf{A} \mathbf{x} \leq \mathbf{b} & y \quad \text{s.a} \quad \boldsymbol{\lambda}^T \mathbf{A} \geq \mathbf{c}^T \\ \mathbf{x} \geq \mathbf{0} & \boldsymbol{\lambda} \geq \mathbf{0} \end{array}$$

Diremos que cada uno de ellos corresponde al problema dual del otro problema. Es posible generalizar esta relación primal-dual para problemas que no necesariamente satisfacen el formato estándar del problema originalmente considerado en esta sección. Así, la siguiente tabla ilustra la equivalencia de dos problemas genéricos en que cada uno de ellos contiene todo tipo de restricciones (igualdades, desigualdades, no-negatividad, o libres):

$\begin{aligned} & \text{Max } \mathbf{c}^T \mathbf{x} \\ \text{s.a} \quad & \mathbf{A}^j \mathbf{x} \geq b_j \quad j \in M_1 \\ & \mathbf{A}^j \mathbf{x} \leq b_j \quad j \in M_2 \\ & \mathbf{A}^j \mathbf{x} = b_j \quad j \in M_3 \\ & x_i \geq 0 \quad i \in N_1 \\ & x_i \leq 0 \quad i \in N_2 \\ & x_i \text{ libre} \quad i \in N_3 \end{aligned}$	$\begin{aligned} & \text{Min } \boldsymbol{\lambda}^T \mathbf{b} \\ \text{s.a} \quad & \lambda_j \leq 0 \quad j \in M_1 \\ & \lambda_j \geq 0 \quad j \in M_2 \\ & \lambda_j \text{ libre} \quad j \in M_3 \\ & \boldsymbol{\lambda}^T \mathbf{A}_i \geq c_i \quad i \in N_1 \\ & \boldsymbol{\lambda}^T \mathbf{A}_i \leq c_i \quad i \in N_2 \\ & \boldsymbol{\lambda}^T \mathbf{A}_i = c_i \quad i \in N_3 \end{aligned}$
--	--

A modo de ilustración volvamos a nuestro ejemplo de las planchas de aluminio y establezcamos los problemas primal y dual.

$$\begin{aligned} PP) \quad & \text{Max } 800x_1 + 600x_2 \\ \text{s.a} \quad & 15x_1 + 5x_2 \leq 600 \\ & 7x_1 + 14x_2 \leq 700 \\ & 0,3x_1 + 0,3x_2 \leq 15 \\ & x_i \geq 0 \quad \forall i = 1, \dots, 2 \end{aligned}$$



En este caso el problema dual asociado sería:

$$\begin{aligned}
 PD) \quad & \text{Min } 600\lambda_1 + 700\lambda_2 + 15\lambda_3 \\
 s.a \quad & 15\lambda_1 + 7\lambda_2 + 0,3\lambda_3 \geq 800 \\
 & 5\lambda_1 + 14\lambda_2 + 0,3\lambda_3 \geq 600 \\
 & \lambda_j \geq 0 \quad \forall j = 1, \dots, 3
 \end{aligned}$$

### 3.4.1 Relaciones Primal-Dual

A continuación formalizaremos algunas propiedades de la relación entre un problema primal y su problema dual que fueron esbozadas en la sección previa.

1. **Propiedad de dualidad débil:** Si  $\mathbf{x}$  es una solución factible para el problema primal, y  $\boldsymbol{\lambda}$  es una solución factible para el problema dual, entonces  $\mathbf{c}^T \mathbf{x} \leq \boldsymbol{\lambda}^T \mathbf{b}$ .
2. **Propiedad de dualidad fuerte:** Si  $\mathbf{x}^*$  es una solución óptima para el problema primal y  $\boldsymbol{\lambda}^*$  es solución óptima para el problema dual, entonces:  $\mathbf{c}^T \mathbf{x}^* = \boldsymbol{\lambda}^{*T} \mathbf{b}$ .
3. **Propiedades de soluciones complementarias:** En cada iteración, el método simplex identifica simultáneamente una solución factible en un vértice,  $\mathbf{x}$  para el  $PP$  y una solución complementaria  $\boldsymbol{\lambda}$  para el  $PD$ , donde se cumple que  $\mathbf{c}^T \mathbf{x} = \boldsymbol{\lambda}^T \mathbf{b}$ . Si  $\mathbf{x}$  no es óptimo para el  $PP$ , entonces  $\boldsymbol{\lambda}$  no es factible para el  $PD$ .
4. **Propiedad de soluciones complementarias óptimas:** En la última iteración del método simplex se identifica simultáneamente una solución óptima  $\mathbf{x}^*$  para el  $PP$  y una solución óptima complementaria  $\boldsymbol{\lambda}^*$  para el  $PD$ , en donde  $\mathbf{c}^T \mathbf{x}^* = \boldsymbol{\lambda}^{*T} \mathbf{b}$ . En este caso los valores de  $\boldsymbol{\lambda}^*$  corresponden a los precios sombra del  $PP$ .
5. **Propiedades de simetría:** Para cualquier  $PP$  y su  $PD$ , las relaciones entre ellas deben ser simétricas debido a que el problema dual de  $PD$  corresponde a  $PP$ .

Todas las propiedades anteriores se cumplen sin importar cuál de los dos problemas se etiquete como problema primal o dual.

Veamos otro ejemplo:

$$\begin{aligned}
 & \text{Max } 5x_1 + 6x_2 \\
 s.a \quad & x_1 + 2x_2 = 5 \\
 & -x_1 + 5x_2 \geq 3 \\
 & 4x_1 + 7x_2 \leq 8 \\
 & x_2 \geq 0
 \end{aligned}$$

Dejemos este problema en forma estándar incorporando variables de holgura y de exceso según corresponda, y haciendo el cambio de variable  $x_1 = x'_1 - x''_1$  con  $x'_1 \geq 0$  y  $x''_1 \geq 0$ . Así, el problema

nos queda:

$$\begin{aligned}
 &Max \ 5x'_1 - 5x''_1 + 6x_2 \\
 &s.a \quad x'_1 - x''_1 + 2x_2 = 5 \\
 &\quad -x'_1 + x''_1 + 5x_2 - x_3 = 3 \\
 &\quad 4x'_1 - 4x''_1 + 7x_2 + x_4 = 8 \\
 &\quad x'_1, x''_1, x_2, x_3, x_4 \geq 0
 \end{aligned}$$

Ahora veamos como queda el respectivo problema dual.

$$\begin{aligned}
 &Min \ 5\lambda_1 + 3\lambda_2 + 8\lambda_3 \\
 &s.a \quad \lambda_1 - \lambda_2 + 4\lambda_3 \geq 5 \\
 &\quad -\lambda_1 + \lambda_2 - 4\lambda_3 \geq -5 \\
 &\quad 2\lambda_1 + 5\lambda_2 + 7\lambda_3 \geq 6 \\
 &\quad -\lambda_2 \geq 0 \\
 &\quad \lambda_3 \geq 0 \\
 &\quad \lambda_1 \text{ irrestricta} \\
 &\quad \lambda_2, \lambda_3 \text{ irrestrictas (redundantes).}
 \end{aligned}$$

### 3.4.2 Análisis Matricial del problema dual

De acuerdo a la propiedad de dualidad fuerte, en el óptimo debe satisfacerse que  $\mathbf{c}^T \mathbf{x}^* = \boldsymbol{\lambda}^{*T} \mathbf{b}$ . Como  $\mathbf{x}_B^* = \mathbf{B}^{-1} \mathbf{b}$ ,  $\mathbf{x}_D^* = 0$ , observamos que  $\hat{\boldsymbol{\lambda}} = \mathbf{c}_B^T \mathbf{B}^{-1}$  satisface esta propiedad y por lo tanto podría constituir la representación matricial de la solución óptima al problema dual. Para esto será necesario verificar la factibilidad de  $\hat{\boldsymbol{\lambda}}$  al problema dual. Con este propósito observamos que el problema primal presenta restricciones del tipo  $\mathbf{Ax} \leq \mathbf{b}$ , a las que es posible agregar variables de holgura (una por restricción). El costo reducido asociado a cada una de estas  $m$  variables corresponderá a  $\mathbf{r}_h = \mathbf{c}_h - \mathbf{c}_B^T \mathbf{B}^{-1} \mathbf{D}_h$ , en que  $\mathbf{c}_h = 0$  y  $\mathbf{D}_h = \mathbf{I}$  (donde  $h$  corresponde exclusivamente al conjunto de variables de holgura). Así, el costo reducido a estas variables de holgura es exactamente  $-\mathbf{c}_B^T \mathbf{B}^{-1} = -\hat{\boldsymbol{\lambda}}$ . Si la solución al problema primal (de maximización) es óptima, entonces estos costos reducidos deben ser todos negativos. Así, la solución óptima al problema primal debe satisfacer  $\hat{\boldsymbol{\lambda}} \geq 0$ . Adicionalmente se debe comprobar que  $\hat{\boldsymbol{\lambda}}^T \mathbf{A} \geq \mathbf{c}^T$ . Habíamos visto que el vector de costos reducidos del problema primal puede escribirse matricialmente como:  $\mathbf{r}_D^T = \mathbf{c}_D^T - \mathbf{c}_B^T \mathbf{B}^{-1} \mathbf{D}$ . Como consideramos que el problema primal es de maximización, entonces en el óptimo debe verificarse que  $\mathbf{r}_D \leq 0$ , es decir,  $\mathbf{c}_D^T \leq \mathbf{c}_B^T \mathbf{B}^{-1} \mathbf{D}$ . Si ahora estudiamos el valor de  $\hat{\boldsymbol{\lambda}}^T \mathbf{A} = \mathbf{c}_B^T \mathbf{B}^{-1} \mathbf{A}$ , observamos

que:

$$\begin{aligned}\widehat{\boldsymbol{\lambda}}^T \mathbf{A} &= [\widehat{\boldsymbol{\lambda}}^T \mathbf{B} : \widehat{\boldsymbol{\lambda}}^T \mathbf{D}] \\ &= [\mathbf{c}_B^T : \mathbf{c}_B^T \mathbf{B}^{-1} \mathbf{D}] \\ &\geq [\mathbf{c}_B^T : \mathbf{c}_D^T]\end{aligned}$$

es decir,  $\widehat{\boldsymbol{\lambda}}$  satisface la restricción  $\boldsymbol{\lambda}^T \mathbf{A} \geq \mathbf{c}^T$  del problema dual. Así, concluimos que esta solución es óptima para dicho problema.

Procedamos a reescribir la restricción del problema dual  $\boldsymbol{\lambda}^T \mathbf{A} \geq \mathbf{c}^T$  del siguiente modo:  $\boldsymbol{\lambda}^T \mathbf{A} - \boldsymbol{\mu}^T \mathbf{I} = \mathbf{c}^T$ , en que las variables  $\boldsymbol{\mu}$  representan variables no negativas de exceso. De las condiciones de optimalidad (KKT) observamos que  $x_i(-\mathbf{c}_i + \boldsymbol{\lambda}^T \mathbf{A}_i) = 0$  ( $\forall i = 1, \dots, n$ ) lo que nos permite concluir que  $\mu_i x_i = 0$ . Es decir, si en el óptimo al problema primal,  $x_i > 0$ , entonces  $\mu_i = 0$  mientras que si  $x_i = 0$ , entonces  $\mu_i \geq 0$ . Esta observación nos permite intuir que el vector  $\boldsymbol{\mu}$  representa el vector de multiplicadores asociados a las restricciones de no negatividad del problema primal. De hecho si descomponemos el vector  $\boldsymbol{\mu}^T$  en los valores asociados a las variables básicas ( $\boldsymbol{\mu}_B$ ) y en los asociados a las variables no básicas ( $\boldsymbol{\mu}_D$ ), observamos que  $\boldsymbol{\mu}_B^T = \mathbf{c}_B^T \mathbf{B}^{-1} \mathbf{B} - \mathbf{c}_B^T = \mathbf{0}$ , mientras que  $\boldsymbol{\mu}_D^T = \mathbf{c}_B^T \mathbf{B}^{-1} \mathbf{D} - \mathbf{c}_D^T = -\mathbf{r}_D^T$  que corresponden precisamente a los costos reducidos de las variables respectivas (con signo inverso).

Asimismo, observamos que en el óptimo debe cumplirse que  $\mathbf{c}^T = \boldsymbol{\lambda}^T \mathbf{A} - \boldsymbol{\mu}^T \mathbf{I}$  lo que confirma que el gradiente de la función objetivo (en este caso del problema primal) constituye una combinación lineal de los gradientes de las restricciones activas.

Finalmente es interesante observar que el costo reducido asociado a una variable de holgura es  $r_j = c_j - \mathbf{c}_B^T \mathbf{B}^{-1} \mathbf{D}_j$ , pero como  $c_j = 0$  y  $\mathbf{D}_j$  es un vector unitario,  $r_j = -\mathbf{c}_B^T \mathbf{B}^{-1} \mathbf{e}_j = -\lambda_j$ . Es decir, el costo reducido de una variable de holgura que no está en la base, es igual al multiplicador de Lagrange de la restricción con signo contrario. Si la variable de holgura no está en la base del problema primal, entonces ésta toma un valor nulo, la restricción asociada debe estar activa, y su costo reducido asociado es también nulo. Si por otra parte la variable de holgura es básica, su costo reducido corresponde a la variable dual asociado a la restricción con signo cambiado. Esto debiera ser intuitivo pues el efecto en la función objetivo de un aumento en la variable de holgura es equivalente al efecto de disminuir en la misma cantidad el recurso asociado.

Es interesante mencionar que de las propiedades antes mencionadas se desprende que el problema dual de un problema no acotado debe ser tal que su dominio sea vacío, de otro modo es imposible satisfacer la propiedad de dualidad débil. Asimismo, el problema dual de un problema con soluciones múltiples debe ser un problema cuya solución óptima sea degenerada ya que los costos reducidos básicos nulos del problema primal corresponderán a las variables básicas del problema dual.

## 3.5 Problemas resueltos

### 3.5.1 Geometría de PL

**Problema 57 (Modelo (I2 1°03))** Una fábrica produce dos tipos de planchas de aluminio pintado, y necesita definir su plan de producción diario de manera que su utilidad sea maximizada. Producir una plancha del tipo 1 requiere de 15 minutos-hombre, de 7 m<sup>2</sup> de aluminio bruto, y de 0.3 litros de pintura. Cada plancha tipo 1 tiene un costo para el fabricante de \$400 y éste la vende a \$1200. Producir una plancha del tipo 2 requiere de 5 minutos-hombre, de 14 m<sup>2</sup> de aluminio bruto, y de 0.3 litros de pintura. Cada plancha tipo 2 tiene un costo para el fabricante de \$900 y éste la vende a \$1500. El fabricante maneja un stock diario de 630 m<sup>2</sup> de aluminio bruto y de 15 litros de pintura. Él trabaja sólo y está dispuesto a trabajar 10 horas cada día.

(a) Formule un modelo de optimización para lograr el objetivo del fabricante. Asuma que las variables de producción son continuas.

(b) En forma muy ordenada haga una tabla (planilla) de  $4 \times 4$  celdas señalando las columnas desde la A a la D, y las filas desde la 1 a la 4. Llénela con el modelo formulado en la parte (a) tal como lo haría en Excel. Fuera de la tabla indique cuales celdas corresponden a las variables de decisión, función objetivo, y restricciones.

(c) Especifique los cuatro “parámetros” (datos de entrada) del solver de Excel que se necesitan para resolver su problema.

(d) Resuelva este problema gráficamente indicando el dominio, curvas de nivel y la solución óptima.

(e) En programación lineal vimos que al menos una de las soluciones óptimas pertenecerá a un conjunto “especial” de soluciones factibles. Sea  $\Omega \subset \mathbb{R}^2$  ese conjunto de soluciones factibles. Diga cuáles son todos los puntos que pertenecen a  $\Omega$ , y cómo se llaman.

(f) Considere un modelo equivalente al formulado en (a) en el formato estándar usado en el curso. Escriba el lagrangeano y las condiciones de Karush-Kuhn-Tucker (KKT).

(g) Del gráfico en (d) usted puede ver cuales restricciones son activas en el óptimo. Utilice esta información para encontrar la solución del problema a partir de las condiciones de KKT establecidas en (f).

(h) Sin resolver nuevamente el problema, ¿cuánto cambiaría la utilidad del fabricante si éste tuviese 10 m<sup>2</sup> más de aluminio bruto al día? ¿Y cuánto cambiaría si éste trabajara 10 minutos más al día?

**Solución:** (a) La variable de decisión es:

$x_i$  = número de planchas tipo  $i$ ;  $i = 1, 2$ .

Como el fabricante quiere maximizar su utilidad, la función objetivo es:

$$P) \text{ Max } 800x_1 + 600x_2$$



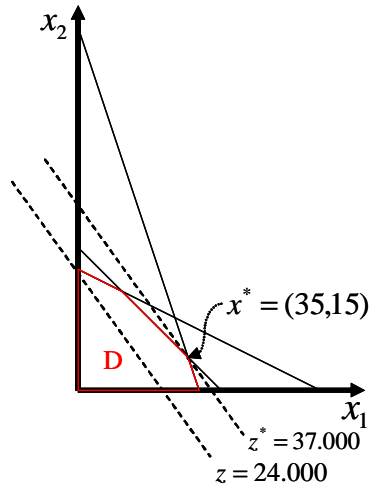


Figure 3.3:

(f) Un modelo equivalente es:

$$\begin{aligned}
 \tilde{P}) \quad & \text{Min} \quad -800x_1 - 600x_2 \\
 \text{s.a.} \quad & 15x_1 + 5x_2 \leq 600 \\
 & 7x_1 + 14x_2 \leq 630 \\
 & 0,3x_1 + 0,3x_2 \leq 15 \\
 & x_i \geq 0 \quad \forall i = 1, 2.
 \end{aligned}$$

El lagrangeano viene dado por:

$$L = -800x_1 - 600x_2 + \mu_1(15x_1 + 5x_2 - 600) + \mu_2(7x_1 + 14x_2 - 630) + \mu_3(0,3x_1 + 0,3x_2 - 15)$$

Por lo tanto, un punto mínimo de  $\tilde{P}$ ) debe cumplir con las condiciones de K.K.T.:

$$\begin{aligned}
 -800 + 15\mu_1 + 7\mu_2 + 0,3\mu_3 & \geq 0 \\
 x_1(-800 + 15\mu_1 + 7\mu_2 + 0,3\mu_3) & = 0 \\
 x_1 & \geq 0
 \end{aligned}$$

$$\begin{aligned}
 -600 + 5\mu_1 + 14\mu_2 + 0,3\mu_3 & \geq 0 \\
 x_2(-600 + 5\mu_1 + 14\mu_2 + 0,3\mu_3) & = 0 \\
 x_2 & \geq 0
 \end{aligned}$$

$$\begin{aligned} 15x_1 + 5x_2 - 600 &\leq 0 \\ \mu_1(15x_1 + 5x_2 - 600) &= 0 \\ \mu_1 &\geq 0 \end{aligned}$$

$$\begin{aligned} 7x_1 + 14x_2 - 630 &\leq 0 \\ \mu_2(7x_1 + 14x_2 - 630) &= 0 \\ \mu_2 &\geq 0 \end{aligned}$$

$$\begin{aligned} 0,3x_1 + 0,3x_2 - 15 &\leq 0 \\ \mu_3(0,3x_1 + 0,3x_2 - 15) &= 0 \\ \mu_3 &\geq 0 \end{aligned}$$

(g) En (d) vemos que las restricciones activas son las de horas-hombre y de pintura, y que hay holgura de aluminio bruto. Lo anterior nos dice que  $\mu_1 > 0$  y  $\mu_3 > 0$ , y además que  $\mu_2 = 0$  (restricción inactiva). Así, algunas condiciones de K.K.T. se reducen a:

$$\begin{aligned} 15x_1 + 5x_2 &= 600 \\ 0,3x_1 + 0,3x_2 &= 15 \end{aligned}$$

donde se obtiene, al resolver el sistema de  $2 \times 2$  que  $x_1 = 35$  y  $x_2 = 15$  en el óptimo.

(h) Como hay holgura de aluminio ( $\mu_2 = 0$ ), el tener  $10 \text{ m}^2$  más no afecta la utilidad del fabricante.

En el caso de que se trabajara 10 minutos más al día, la utilidad aumentará aproximadamente  $10\mu_1$ .  $\mu_1$  se obtiene resolviendo el siguiente sistema:

$$\begin{aligned} 15\mu_1 + 0,3\mu_3 &= 800 \\ 5\mu_1 + 0,3\mu_3 &= 600 \end{aligned}$$

de donde  $\mu_1 = 20$ . Por lo tanto, la utilidad aumentaría a \$37.200 aproximadamente.

**Problema 58** *Una empresa ha violado la ley ambiental y ha recibido la pena de plantar árboles en una plaza de uso público. Los árboles disponibles son robles y abedules. Se le exige además que por cada roble, se deben plantar a lo más 4 abedules, y que se deben plantar como mínimo 2 robles o un abedul. Las horas-hombre necesarias para la plantación de un abedul y un roble son 1 y 3 horas-hombre respectivamente, y los 3 trabajadores sólo tienen 3 horas para realizar el trabajo (c/u). Resuelva los siguientes problemas gráficamente. Considere variables reales.*

a) Si la empresa, arrepentida por su falta, buscara maximizar el beneficio social de la plantación de árboles, siendo éste de 4 para los abedules y 1 para los robles. ¿Cuál sería el óptimo?

b) Si la empresa buscara minimizar sus costos, y estos fueran de 4 para los abedules y 1 para los robles. ¿Cuál sería el óptimo?

c) Si la empresa, arrepentida por su falta, buscara maximizar el beneficio social de la plantación de árboles, pero éste fuera de 1 para los abedules y 3 para los robles. ¿Cuál sería el óptimo?

d) Si no existiese la restricción de la mano de obra, ¿cuál sería el óptimo para los casos descritos en a) y b)?

### Solución:

Las variables de decisión serán:

$x$ : número de robles que se plantarán

$y$ : número de abedules que se plantarán

a) El problema descrito será

$$\begin{aligned} \text{Max } 4y + x \\ \text{s.a. } 4x &\geq y \\ 2y + x &\geq 2 \\ y + 3x &\leq 9 \\ x, y &\geq 0 \end{aligned}$$

Graficando este problema y las curvas de nivel asociadas, se obtiene el Figura 3.4. Las curvas de nivel se obtienen con la función objetivo, es decir,

$$4y + x = z \implies y = \frac{-1}{4}x + \frac{z}{4}$$

Así, la pendiente de las curvas de nivel será  $-1/4$ , y el valor óptimo ( $z$ ) estará relacionado con la intersección de las curvas de nivel con el eje  $y$ . En este caso, a medida que aumenta el valor óptimo, la intersección con el eje  $y$  será más "arriba". De este modo se puede conocer hacia "que lado" buscar el óptimo, dependiendo de si se intenta maximizar o minimizar la función objetivo.

En este caso, se puede apreciar que el máximo se encuentra en el punto donde  $4x = y$  y  $y + 3x = 9$ , es decir, en  $(x, y) = (9/7, 36/7)$ .

b) La minimización tendrá como solución óptima el punto  $(x, y) = (2, 0)$ , es decir, se plantarían sólo 2 robles. Para obtener esta solución también se utilizó la Figura 3.4, pero esta vez se buscó la curva de nivel más "baja" (de menor valor óptimo) que todavía se intersectara con el dominio.

c) Al graficar las curvas de nivel de la nueva función objetivo se aprecia que éstas son paralelas a la restricción de la mano de obra, por lo que existen múltiples soluciones óptimas (todos los puntos que pertenecen al dominio y hacen que ésta restricción se active). Ver Figura 3.5.



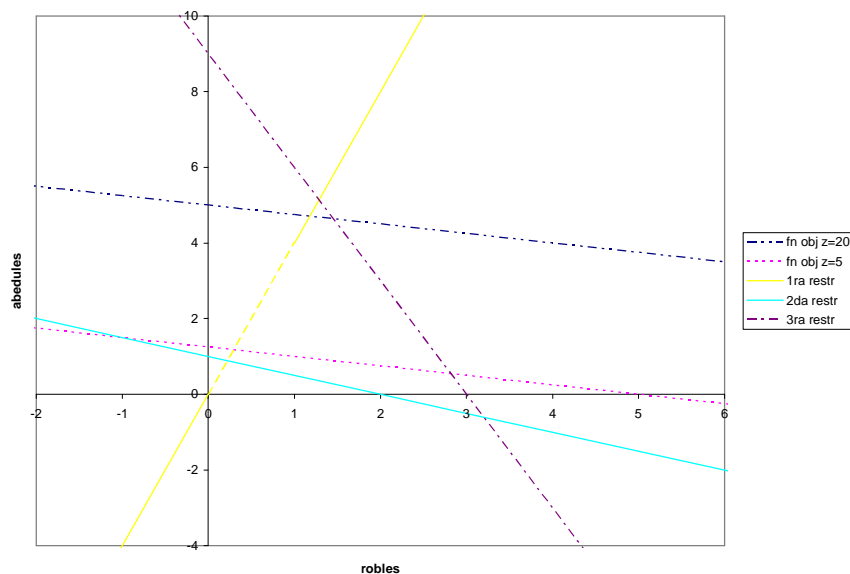


Figure 3.4: Plantación de robles y abedules - Partes a) y b)

d) Si no estuviese la restricción de mano de obra, entonces se tendría un dominio no acotado. Ver Figura 3.6. Así para el caso a) de maximización, el problema no tendría solución óptima pues puedo "subir" eternamente por las curvas de nivel, mientras que para el caso b) de minimización, el problema mantendría como solución óptima el punto  $(2,0)$ , pese a tener un dominio no acotado.

**Problema 59 (I21°03)** ¿Qué condición debe cumplir  $a$  y  $b$  en el siguiente problema de modo que la solución óptima sea  $(\frac{2}{3}, \frac{2}{3})$ ?

$$\begin{aligned}
 P) \quad & \text{Max } ax + 2by \\
 \text{s.a.} \quad & x + 2y \leq 2 \\
 & 2x + y \leq 2 \\
 & x, y \geq 0
 \end{aligned}$$

**Solución:** El gradiente de  $f(x, y)$  debe estar entre los gradientes de las restricciones, es decir:

$$\begin{aligned}
 \nabla f &= \alpha(1, 2) + \beta(2, 1) \text{ con } \alpha, \beta \geq 0 \\
 &= (\alpha + 2\beta, 2\alpha + \beta)
 \end{aligned}$$

esto implica que:

$$\begin{aligned}
 \alpha + 2\beta &= a \\
 2\alpha + \beta &= 2b
 \end{aligned}$$

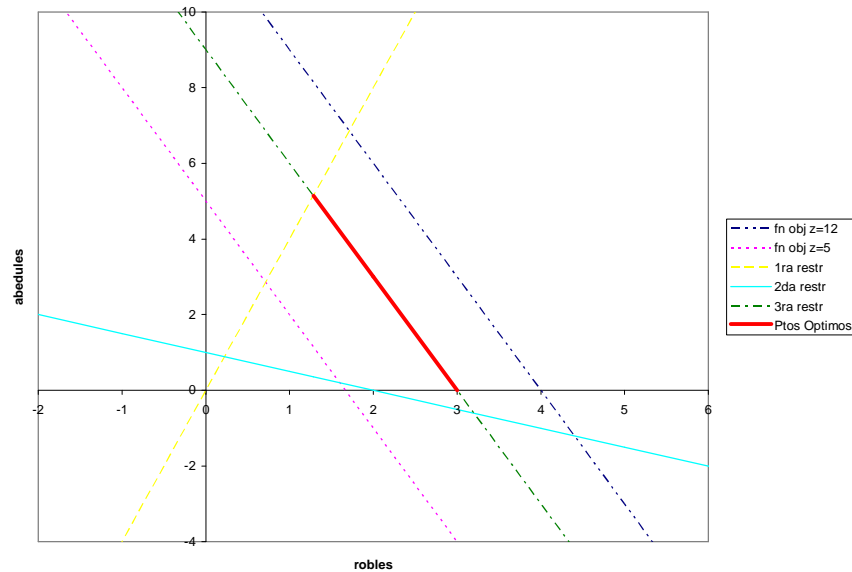


Figure 3.5: Plantación de robles y abedules - Parte c)

0	3	1	c	d
1	a	0	2	e
0	b	0	-1	-10

Table 3.1: Tableau

Luego,

$$\begin{aligned}\beta &= \frac{2}{3}(a-b) \geq 0 \implies \frac{a}{b} \geq 1 \\ \alpha &= \frac{-1}{3}a + \frac{4}{3}b \geq 0 \implies \frac{a}{b} \leq 4\end{aligned}$$

Por lo tanto,  $a$  y  $b$  deben cumplir con  $1 \leq \frac{a}{b} \leq 4$ .

### 3.5.2 Método Simplex

**Problema 60 (Simplex (I3 1°03))** Responda a las siguientes preguntas en base al tableau correspondiente a un problema de maximización que se muestra en la Tabla 1.

(a) ¿Qué debe suceder con los parámetros para que éste sea un tableau óptimo? ¿Cuál sería el valor óptimo del problema?

(b) Si  $a = 2$ ,  $b = 1$ ,  $d = 6$ , ¿qué debe suceder para que  $x_3$  salga de la base?

(c) ¿Qué significa que  $b = 0$ ? ¿Y qué significa que  $d = 0$ ?

(d) Si se asume que  $d = 0$ , ¿qué implicancia tiene en el problema dual?

**Solución:** (a)  $b \leq 0$  y  $d, e \geq 0$ . El valor óptimo sería  $z = 10$ .

Figure 3.6: Plantación de robles y abedules - Parte d)

(b)  $e > 4 \implies \text{Min}\{\frac{6}{3}, \frac{e}{2}\} = \frac{6}{3}$ . Por lo tanto,  $x_3$  sale de la base.

(c) Si  $b = 0$ , hay múltiples soluciones que hacen  $z = 10$ . Si  $d = 0$ , estamos ante una solución básica degenerada.

(d) Si  $d = 0 \implies$  en el dual hay múltiples soluciones.

**Problema 61 (Simplex (I3 1°03))** Resuelva el siguiente problema de optimización utilizando el método Simplex. Indique la solución óptima y el valor óptimo.

$$\begin{aligned}
 P) \quad & \text{Max} \quad x_1 + 2x_2 \\
 \text{s.a.} \quad & -x_1 + 2x_2 \leq 2 \\
 & 0 \leq x_1 \leq 2 \\
 & x_2 \geq 1
 \end{aligned}$$

**Solución:** En formato estandar  $P)$  nos queda:

$$\begin{array}{rcll}
 P) \quad & \text{Max} & x_1 & + 2x_2 \\
 \text{s.a.} & & x_1 & + x_3 = 2 \\
 & & x_2 & - x_4 + x_6 = 1 \\
 & & -x_1 + 2x_2 & + x_5 = 2 \\
 & & & x_i \geq 0
 \end{array}$$

Como  $x_6$  es una variable artificial, habrá que eliminarla con la Fase I y así encontrar una solución inicial posible básica a  $P)$ .



(b) La solución óptima sería:

$$\begin{aligned}x_1 &= 2 \\x_2 &= 0 \\x_3 &= e \\x_4 &= 0\end{aligned}$$

y el valor óptimo:  $z^* = 220$ .

(c) Las condiciones que deben darse son:  $b = 0$  y/o  $c = 0$ , y que  $e \geq 0$ .

(d) Para estar en el óptimo y que éste consista en una solución degenerada se debe dar:  $b \leq 0$ ,  $c \leq 0$  y  $e = 0$ .

(e) Sin estar en el óptimo, para que la solución sea no acotada, debe darse:

$$\begin{aligned}e &\geq 0 \\d &< 0 \\c &> 0 \\b &< 0\end{aligned}$$

**Problema 63 (Simplex (Ex 1°03))** *Suponga el siguiente problema de maximización lineal:*

$$\begin{aligned}P) \quad & \text{Max } x_1 + 2x_2 \\s.a. \quad & 3x_1 + x_2 \leq 210 \\& x_1 + 3x_2 \leq 150 \\& x_1, x_2 \geq 0\end{aligned}$$

*y que el tableau final es:*

$x_1$	$x_2$	Holg.1	Holg.2	
1	0	$\frac{3}{8}$	$-\frac{1}{8}$	60
0	1	$-\frac{1}{8}$	$\frac{3}{8}$	30
0	0	$-\frac{1}{8}$	$-\frac{5}{8}$	-120

*Indique en qué rango pueden moverse los valores  $c_2 = 2$  y  $b_2 = 150$  para que la base no cambie.*

**Solución:**

$$\begin{aligned}
r &= c_D^T - c_B^T B^{-1} D \\
\vec{r} &= \vec{0} - \begin{bmatrix} 1 & c_2 \end{bmatrix} \begin{bmatrix} \frac{3}{8} & \frac{-1}{8} \\ \frac{-1}{8} & \frac{3}{8} \end{bmatrix} = - \begin{bmatrix} \frac{3}{8} - \frac{c_2}{8} & \frac{-1}{8} + \frac{3c_2}{8} \end{bmatrix} \leq \vec{0} \\
\frac{-3}{8} + \frac{c_2}{8} &\leq 0 \Rightarrow c_2 \leq 3 \\
\frac{1}{8} - \frac{3c_2}{8} &\leq 0 \Rightarrow c_2 \geq \frac{1}{3} \\
\therefore \frac{1}{3} &\leq c_2 \leq 3 \\
B^{-1}b &= \begin{bmatrix} \frac{3}{8} & \frac{-1}{8} \\ \frac{-1}{8} & \frac{3}{8} \end{bmatrix} \begin{bmatrix} 210 \\ b_2 \end{bmatrix} = \begin{bmatrix} \frac{630}{8} - \frac{b_2}{8} \\ \frac{-210}{8} + \frac{3b_2}{8} \end{bmatrix} \geq \vec{0} \\
\therefore 70 &\leq b_2 \leq 630.
\end{aligned}$$

**Problema 64 (Simplex (Ex 1°03, Preg. Especial I1))** Considere el siguiente problema de optimización:

$$\begin{aligned}
P) \quad & \text{Max } 6x + 4y \\
s.a. \quad & x + 2y \leq 2 \\
& 2x + y \leq 2 \\
& x, y \geq 0
\end{aligned}$$

- (a) Resuelva utilizando Simplex.
- (b) ¿Cuál es la base y solución factible inicial?
- (c) ¿Cuál es la base y solución óptima?
- (d) ¿En qué rango puede variar  $b_1$  sin que cambie la solución óptima?
- (e) ¿Qué sucede si  $b_2 = 5$ ?

**Solución:** (a) Reescribimos el problema:

$$\begin{aligned}
P) \quad & \text{Max } 6x + 4y \\
s.a. \quad & x + 2y + z_1 = 2 \\
& 2x + y + z_2 = 2 \\
& x, y, z_1, z_2 \geq 0
\end{aligned}$$

Luego, planteamos el *tableau* y resolvemos:

$$\begin{array}{c|ccc|c} x & y & z_1 & z_2 & b \\ \hline 1 & 2 & 1 & 0 & 2 \\ 2 & 1 & 0 & 1 & 2 \\ \hline 6 & 4 & 0 & 0 & 0 \end{array} \rightarrow \begin{array}{c|ccc|c} x & y & z_1 & z_2 & b \\ \hline 0 & \frac{3}{2} & 1 & \frac{-1}{2} & 1 \\ 1 & \frac{1}{2} & 0 & \frac{1}{2} & 1 \\ \hline 0 & 1 & 0 & -3 & -6 \end{array} \rightarrow \begin{array}{c|ccc|c} x & y & z_1 & z_2 & b \\ \hline 0 & 1 & \frac{2}{3} & \frac{-1}{3} & \frac{2}{3} \\ 1 & 0 & \frac{-1}{3} & \frac{2}{3} & \frac{2}{3} \\ \hline 0 & 0 & \frac{-2}{3} & \frac{-8}{3} & \frac{-20}{3} \end{array}$$

$\uparrow$  Sale  $z_2$ 
 $\uparrow$  Sale  $z_1$

Por lo tanto,  $x = \frac{2}{3}$ ,  $y = \frac{2}{3}$ ,  $z_1 = 0$ ,  $z_2 = 0$  y  $v^* = \frac{20}{3}$ .

(b) Base inicial:  $B = \{z_1 \ z_2\} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ .

Solución factible inicial:  $x = 0$ ,  $y = 0$ ,  $z_1 = 2$  y  $z_2 = 2$ .

(c) Base óptima:  $B = \{x \ y\} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ .

Solución óptima:  $x = \frac{2}{3}$ ,  $y = \frac{2}{3}$ ,  $z_1 = 0$  y  $z_2 = 0$ .

(d)  $x_B = B^{-1}b = \begin{bmatrix} \frac{-1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{-1}{3} \end{bmatrix} \begin{bmatrix} b_1 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{-b_1}{3} + \frac{4}{3} \\ \frac{2b_1}{3} - \frac{2}{3} \end{bmatrix}$

$$\frac{-b_1}{3} + \frac{4}{3} \geq 0 \Rightarrow -b_1 + 4 \geq 0 \Rightarrow 4 \geq b_1.$$

$$\frac{2b_1}{3} - \frac{2}{3} \geq 0 \Rightarrow 2b_1 - 2 \geq 0 \Rightarrow b_1 - 1 \geq 0 \Rightarrow b_1 \geq 1.$$

$$\therefore 1 \leq b_1 \leq 4.$$

(e) Si  $b_2 = 5$ :

$$\frac{-2}{3} + \frac{2}{3}b_2 \geq 0 \Rightarrow -1 + b_2 \geq 0 \Rightarrow b_2 \geq 1.$$

$$\frac{4}{3} - \frac{b_2}{3} \geq 0 \Rightarrow 4 - b_2 \geq 0 \Rightarrow 4 \geq b_2.$$

$$\therefore 1 \leq b_2 \leq 4.$$

Si  $b_2 = 5$ , varía la solución óptima, y por lo tanto, cambia la base óptima y valor óptimo.

**Problema 65 (Análisis Matricial (I3 1°03))** Considere el siguiente problema de optimización.

$$\begin{aligned} P) \quad & \text{Max} \quad 2x_1 + 4x_2 - 3x_3 \\ \text{s.a.} \quad & x_1 + 2x_2 + x_3 \leq 3 \\ & 2x_1 - x_2 + 2x_3 \leq 2 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

(a) Utilizando exclusivamente análisis matricial justifique si la base  $\{x_2; x_3\}$  es óptima o no.

(b) Resuelva el problema P) como más le acomode.

**Solución:** (a) Base =  $\{x_2; x_3\}$

$$r = c_D^T - c_B^T B^{-1} D$$

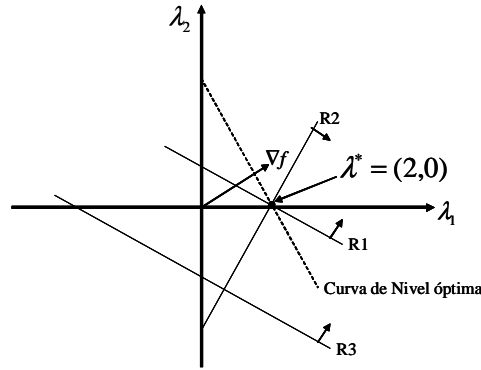


Figure 3.7: Solución Gráfica

$$r = \begin{bmatrix} 2 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 4 & -3 \end{bmatrix} \begin{bmatrix} \frac{2}{5} & \frac{-1}{5} \\ \frac{1}{5} & \frac{2}{5} \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$$

$$r = \begin{bmatrix} 5 & -1 & 2 \end{bmatrix}$$

Como no todos los costos reducidos son negativos, entonces la base  $\{x_2; x_3\}$  no es óptima.

(b) Escribiendo el dual de  $P$ ) quedamos en dos dimensiones:

$$\begin{aligned} D) \quad & \text{Min } 3\lambda_1 + 2\lambda_2 \\ \text{s.a.} \quad & \lambda_1 + 2\lambda_2 \geq 2 \end{aligned} \tag{3.4}$$

$$2\lambda_1 - \lambda_2 \geq 4 \tag{3.5}$$

$$\lambda_1 + 2\lambda_2 \geq -3 \tag{3.6}$$

$$\lambda_1, \lambda_2 \geq 0$$

Gráficamente, como  $\lambda_1 \neq 0$ , la primera restricción de  $P$ ) está activa.  $x_1 + 2x_2 + x_3 = 3$ .

Se puede ver en  $P$ ) que  $x_3^* = 0 \Rightarrow x_1 + 2x_2 = 3$ . Como el gradiente de esta restricción es igual al de la función objetivo, cualquier combinación convexa de  $x^* = (\frac{7}{5}, \frac{4}{5})$  y  $x^* = (0, \frac{3}{2})$  será una solución óptima.

**Problema 66** Considere el siguiente problema

$$\begin{aligned} \text{Max } & 2x_1 + 4x_2 + 3x_3 \\ \text{s.a. } & x_1 + 2x_2 \leq 4 \\ & x_2 + 3x_3 \leq 6 \\ & 2x_1 + x_2 + 2x_3 \leq 10 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

a) Resuelva mediante el algoritmo Simplex, especificando la solución y el valor óptimo.

b) Considere ahora que la tercera restricción fuese  $2x_1 + x_2 + 2x_3 \geq 10$ . Resuelva nuevamente



considerando el método de las dos fases.

c) Resuelva el problema considerando que la segunda restricción fuese  $x_2 + 3x_3 \leq 2$ . ¿Nota algo especial? Interprete.

d) Si las constantes que acompañan a la variable en las restricciones fuesen todas negativas en lugar de positivas, ¿qué estaría sucediendo?

e) Resuelva el problema considerando que la tercera restricción fuese  $2x_1 + 4x_2 + 3x_3 \leq 10$ . ¿Nota algo especial? Interprete.

### Solución:

El criterio de parada de las iteraciones de Simplex en la búsqueda de un máximo, es que los costos reducidos en el óptimo deben ser todos menores o iguales a cero. En el caso de un mínimo, los costos reducidos en el óptimo serán mayores o iguales a cero.

a) Tras añadir las variables de holgura, el problema será

$$\begin{aligned} \text{Max } & 2x_1 + 4x_2 + 3x_3 \\ \text{s.a. } & x_1 + 2x_2 + x_4 = 4 \\ & x_2 + 3x_3 + x_5 = 6 \\ & 2x_1 + x_2 + 2x_3 + x_6 = 10 \\ & x_1, x_2, x_3, x_4, x_5, x_6 \geq 0 \end{aligned}$$

Llevando este problema a un tableau se obtiene

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$b_i$
1	2	0	1	0	0	4
0	1	3	0	1	0	6
2	1	2	0	0	1	10
2	4	3	0	0	0	0

donde la base inicial es  $x_B = \{x_4, x_5, x_6\}$  y las variables no-básicas serán  $x_N = \{x_1, x_2, x_3\}$ . Dado que  $\max\{2, 4, 3\} = 4$ ,  $x_2$  entra a la base. Luego  $\min\{4/2, 6/1, 10/1\} = 2$ , por lo que  $x_4$  sale de la base. Pivoteando se obtiene el tableau

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$b_i$
1/2	1	0	1/2	0	0	2
-1/2	0	3	-1/2	1	0	4
3/2	0	2	-1/2	0	1	8
0	0	3	-2	0	0	-8

Se observa que aún queda un costo reducido positivo, por lo que  $x_3$  entra ahora a la base. Para buscar la variable saliente:  $\min\{\bullet, 4/3, 8/2\} = 4/3$ , por lo que  $x_5$  sale de la base. Pivoteando

nuevamente, ahora se obtiene

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$b_i$
1/2	1	0	1/2	0	0	2
-1/6	0	1	-1/6	1/3	0	4/3
11/6	0	0	-1/6	-2/3	1	16/3
1/2	0	0	-3/2	-1	0	-12

Ahora  $x_1$  entra a la base.  $\min \left\{ \frac{2}{1/2}, \bullet, \frac{16/3}{11/6} \right\} = \frac{32}{11}$ , por lo que  $x_6$  sale de la base. Pivoteando se obtiene

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$b_i$
0	1	0	6/11	2/11	-3/11	6/11
0	0	1	-2/11	3/11	1/11	20/11
1	0	0	-1/11	-4/11	6/11	32/11
0	0	0	-16/11	-9/11	-3/11	-148/11

Todos los costos reducidos son menores o iguales a cero, por lo que se ha llegado al óptimo, que estará dado por:  $(x_1, x_2, x_3, x_4, x_5, x_6) = (\frac{32}{11}, \frac{6}{11}, \frac{20}{11}, 0, 0, 0)$  y tendrá valor óptimo  $v^* = -\frac{148}{11}$ . Ya que todas las holguras toman valor cero, las tres restricciones estarán activas en el óptimo.

b) Ya que la restricción es mayor o igual a cero, y no se puede invertir multiplicando por (-1) ya que se requiere  $b_i \geq 0$ ,  $\forall i$ , se debe agregar una variable de exceso y luego una variable artificial, para poder formar una base inicial. La restricción será entonces

$$2x_1 + x_2 + 2x_3 - x_6 + y = 10$$

En la primera fase se busca *Min*  $y$  (*Max*  $-y$ ), sujeto a las mismas restricciones, buscando idealmente que  $y$  adquiera valor cero. El tableau inicial para la primera fase será

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$y$	$b_i$
1	2	0	1	0	0	0	4
0	1	3	0	1	0	0	6
2	1	2	0	0	-1	1	10
0	0	0	0	0	0	-1	0
2	4	3	0	0	0	0	0

Reescribiendo este tableau sumando la 4ta y 5ta fila, se obtiene

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$y$	$b_i$
1	2	0	1	0	0	0	4
0	1	3	0	1	0	0	6
2	1	2	0	0	-1	1	10
2	1	2	0	0	-1	0	10
2	4	3	0	0	0	0	0

Observando los costos reucidos de la primera fase, se aprecia una empate entre el de  $x_1$  y  $x_3$ . Se escoje arbitrariamente a  $x_3$  para entrar a la base.  $\min\{\bullet, \frac{6}{3}, \frac{10}{2}\} = 2$ , por lo que  $x_5$  sale de la base. Pivoteando se obtiene

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$y$	$b_i$
1	2	0	1	0	0	0	4
0	1/3	1	0	1/3	0	0	2
2	1/3	0	0	-2/3	-1	1	6
2	1/3	0	0	-2/3	-1	0	6
2	3	0	0	-1	0	0	-6

Ahora  $x_1$  entra a la base.  $\min\{\frac{4}{1}, \bullet, \frac{6}{2}\} = 3$ , por lo que  $y$  sale de la base. Pivoteando se obtiene

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$y$	$b_i$
0	11/6	0	1	1/3	1/2	-1/2	1
0	1/3	1	0	1/3	0	0	2
1	1/6	0	0	-1/3	-1/2	1/2	3
0	0	0	0	0	0	-1	0
0	8/3	0	0	-1/3	1	-1	-12

Ya que  $y$  ha salido de la base adquiriendo valor nulo, puede ser eliminada. Así se pasa a la Fase II donde se vuelve a la función objetivo inicial y se comienza con el tableau

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$b_i$
0	11/6	0	1	1/3	1/2	1
0	1/3	1	0	1/3	0	2
1	1/6	0	0	-1/3	-1/2	3
0	8/3	0	0	-1/3	1	-12

Ahora  $x_2$  entra a la base.  $\min\{\frac{1}{11/6}, \frac{2}{1/3}, \frac{3}{1/6}\} = \frac{6}{11}$ , por lo que  $x_4$  sale de la base. Pivoteando

se obtiene

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$b_i$
0	1	0	6/11	2/11	3/11	6/11
0	0	1	-2/11	3/11	-1/11	20/11
1	0	0	-1/11	-4/11	-6/11	32/11
0	0	0	-16/11	-9/11	3/11	-148/11

Ahora  $x_6$  entra a la base.  $\min \left\{ \frac{6/11}{3/11}, \bullet, \bullet \right\} = 2$ , por lo que  $x_2$  sale de la base. Pivoteando se obtiene

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$b_i$
0	11/3	0	2	2/3	1	2
0	1/3	1	0	1	0	2
1	2	0	1	0	0	4
0	-1	0	-22/11	-1	0	-14

Luego el punto óptimo será  $(x_1, x_2, x_3, x_4, x_5, x_6) = (4, 0, 2, 0, 0, 2)$  y tendrá valor óptimo  $v^* = 14$ . Las dos primera restricciones están activas en el óptimo, no así la tercera.

c) El problema será

$$\begin{aligned}
 &Max \ 2x_1 + 4x_2 + 3x_3 \\
 &s.a. \ x_1 + 2x_2 + x_4 = 4 \\
 &\quad \quad x_2 + 3x_3 + x_5 = 2 \\
 &\quad \quad 2x_1 + x_2 + 2x_3 + x_6 = 10 \\
 &\quad \quad x_1, x_2, x_3, x_4, x_5, x_6 \geq 0
 \end{aligned}$$

El tableau inicial pasaría a ser

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$b_i$
1	2	0	1	0	0	4
0	1	3	0	1	0	2
2	1	2	0	0	1	10
2	4	3	0	0	0	0

donde la base inicial es  $x_B = \{x_4, x_5, x_6\}$  y las variables no-básicas serán  $x_N = \{x_1, x_2, x_3\}$ . Dado que  $\max \{2, 4, 3\} = 4$ ,  $x_2$  entra a la base. Luego  $\min \{4/2, 2/1, 10/1\} = 2$ , pero se tiene un "empate" entre  $x_4$  y  $x_5$ . Esto quiere decir que el dominio determinado por las restricciones está sobredeterminado (hay una restricción "de más"). A esto se le llama una solución degenerada. Se escoje

arbitrariamente que  $x_4$  salga de la base, donde pivoteando se obtiene

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$b_i$
1/2	1	0	1/2	0	0	2
-1/2	0	3	-1/2	1	0	0
3/2	0	2	-1/2	0	1	8
0	0	3	-2	0	0	-8

Se puede observar que la variable  $x_5$ , pese a estar dentro de la base, toma un valor nulo. Este no es el óptimo. Estamos ante una solución factible degenerada, que será  $(x_1, x_2, x_3, x_4, x_5, x_6) = (0, 2, 0, 0, 0, 8)$  y tendrá valor NO-óptimo  $v = 8$ .

d) El problema será

$$\begin{aligned}
 &Max \ 2x_1 + 4x_2 + 3x_3 \\
 &s.a. \ x_1 - 2x_2 + x_4 = 4 \\
 &\quad -x_2 + 3x_3 + x_5 = 6 \\
 &\quad 2x_1 - x_2 + 2x_3 + x_6 = 10 \\
 &\quad x_1, x_2, x_3, x_4, x_5, x_6 \geq 0
 \end{aligned}$$

El tableau inicial pasaría a ser

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$b_i$
1	-2	0	1	0	0	4
0	-1	3	0	1	0	6
2	-1	2	0	0	1	10
2	4	3	0	0	0	0

Dado que  $\max\{2, 4, 3\} = 4$ ,  $x_2$  entra a la base. Luego se tendría  $\min\{\bullet, \bullet, \bullet\}$ , pues todos los valores son negativos. Un caso como este señala que la solución óptima no es acotada, pues  $x_2$  puede seguir creciendo ilimitadamente (mejorando así la función objetivo), sin que se dejen de cumplir las restricciones

e) El problema será

$$\begin{aligned}
 &Max \ 2x_1 + 4x_2 + 3x_3 \\
 &s.a. \ x_1 + 2x_2 + x_4 = 4 \\
 &\quad x_2 + 3x_3 + x_5 = 6 \\
 &\quad 2x_1 + 4x_2 + 3x_3 + x_6 = 10 \\
 &\quad x_1, x_2, x_3, x_4, x_5, x_6 \geq 0
 \end{aligned}$$

Llevando este problema a un tableau se obtiene

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$b_i$
1	2	0	1	0	0	4
0	1	3	0	1	0	6
2	4	3	0	0	1	10
2	4	3	0	0	0	0

Dado que  $\max\{2, 4, 3\} = 4$ ,  $x_2$  entra a la base. Luego  $\min\{4/2, 6/1, 10/4\} = 2$ , por lo que  $x_4$  sale de la base. Pivoteando se obtiene el siguiente tableau

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$b_i$
1/2	1	0	1/2	0	0	2
-1/2	0	3	-1/2	1	0	4
0	0	3	-2	0	1	2
0	0	3	-2	0	0	-8

Se observa que aún queda un costo reducido positivo, por lo que  $x_3$  entra ahora a la base. Para buscar la variable saliente:  $\min\{\bullet, 4/3, 2/3\} = 2/3$ , por lo que  $x_6$  sale de la base. Pivoteando nuevamente, ahora se obtiene

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$b_i$
1/2	1	0	1/2	0	0	2
-1/2	0	0	3/2	1	-1	2
0	0	1	-2/3	0	1/3	2/3
0	0	0	0	0	-1	-10

Se puede apreciar que los costos reducidos de las variables  $x_1$  y  $x_4$ , pese a que éstas están fuera de la base, tienen valor nulo. La solución óptima será  $(x_1, x_2, x_3, x_4, x_5, x_6) = (0, 2, 2/3, 0, 2, 0)$  y tendrá valor óptimo  $v^* = 10$ . Interesante es notar que sucede si se introduce  $x_1$  y/o  $x_4$  a la base. Veamos que sucede al introducir  $x_1$ , sacando por ejemplo a  $x_2$ . Pivoteando se obtiene

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$b_i$
1	2	0	1	0	0	4
0	1	0	2	1	-1	4
0	0	1	-2/3	0	1/3	2/3
0	0	0	0	0	-1	-10

Se aprecia una solución óptima alternativa, que tiene el mismo valor óptimo. Así este es un problema de múltiples soluciones. La solución óptima general será la combinación lineal de todos los vértices óptimos que se encuentren a través de Simplex.

**Problema 67** Considere el siguiente problema

$$\begin{aligned}
 \text{Min } & 3x_1 + 2x_2 - 3x_3 - 6x_4 + 10x_5 - 5x_6 \\
 \text{s.a. } & x_1 + 2x_2 + x_4 - x_6 = 11 \\
 & x_2 + x_3 + 3x_4 - 2x_5 - x_6 = 6 \\
 & x_1 + 2x_2 + x_3 + 3x_4 - x_5 - 5x_6 = 13 \\
 & x_i \geq 0, \forall i = 1 \dots 6
 \end{aligned}$$

Suponga que las variables básicas en el óptimo son:  $x_1, x_2, x_3$  y que

$$B^{-1} = \begin{bmatrix} -1 & -2 & 2 \\ 1 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix}$$

Calcule los valores de las variables básicas en el óptimo. Indique cómo puede comprobar que ésta es efectivamente la solución óptima.

a) Suponga que se desea introducir una nueva variable al problema, cuyo costo inicial es -7 y cuya columna traspuesta es  $\begin{pmatrix} 1 & 2 & -3 \end{pmatrix}$ . Indique si la solución óptima obtenida se modifica por este cambio.

b) Obtenga el rango del costo inicial de la variable  $x_5$  para el cual la solución óptima obtenida sigue siendo óptima.

c) Repita lo anterior para la variable  $x_1$ .

d) Obtenga el rango en el que puede variar la constante de la primera restricción para que se mantenga la base óptima. Expresé el valor óptimo de la función objetivo en función de esa constante, para el caso en que ésta se mantiene dentro de ese rango.

e) Analice el rango de variación de  $a_{2,5}$  para que la solución siga siendo óptima.

### Solución:

En primer lugar, recordemos Simplex en su notación matricial. El problema se puede escribir como

$$\begin{aligned}
 \text{Min } & c^T x \\
 \text{s.a. } & Ax = b \\
 & x \geq 0
 \end{aligned}$$

Luego separamos las variables básicas de las no-básicas tomando  $x = x_B + x_D$ .

$$\begin{aligned}
 \Rightarrow \quad \text{Min } & c_B^T x_B + c_D^T x_D \\
 \text{s.a. } & Bx_B + Dx_D = b \\
 & x_B, x_D \geq 0
 \end{aligned}$$

$$\begin{aligned}\Rightarrow x_B + B^{-1}Dx_D &= B^{-1}b \\ \Rightarrow x_B &= B^{-1}b - B^{-1}Dx_D\end{aligned}$$

Donde las variables no-básicas tomarán valor nulo ( $x_D = 0$ ). Así el valor de las variables básicas en el óptimo será

$$x_B = B^{-1}b = \begin{pmatrix} -1 & -2 & 2 \\ 1 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 11 \\ 6 \\ 13 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ 2 \end{pmatrix}$$

Para comprobar si ésta es una solución óptima se pueden calcular los costos reducidos de las variables no básicas comprobando si éstos son mayores o iguales a cero.

a) Para esto debemos evaluar cuánto sería el costo reducido de esta variable en el óptimo actual. El valor óptimo está dado por

$$\begin{aligned}v^* &= c^T x = c_B^T x_B + c_D^T x_D \\ \Rightarrow v^* &= c_B^T (B^{-1}b - B^{-1}Dx_D) + c_D^T x_D \\ \Rightarrow v^* &= c_B^T B^{-1}b + \underbrace{(c_D^T - c_B^T B^{-1}D)}_{\text{Costo reducido}} x_D\end{aligned}$$

Así,

$$r_7 = c_7 - c_B^T B^{-1}D_7 = -7 - \begin{pmatrix} 3 & 2 & -3 \end{pmatrix} \begin{pmatrix} -1 & -2 & 2 \\ 1 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} = 2 \geq 0$$

Como se está buscando un mínimo, y el costo reducido de la séptima variable es positivo, ésta se mantendrá fuera de la base y el punto óptimo no cambiará.

b) Nuevamente se analizará el costo reducido de la variable  $x_5$  que está fuera de la base. Suponiendo que se permite una variación de  $\Delta$  en el costo inicial  $c_5$ , el costo reducido de ésta variable será

$$\begin{aligned}r'_5 &= (c_5 + \Delta) - c_B^T B^{-1}D_5 \\ \Rightarrow r'_5 &= 10 + \Delta - \begin{pmatrix} 3 & 2 & -3 \end{pmatrix} \begin{pmatrix} -1 & -2 & 2 \\ 1 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ -2 \\ -1 \end{pmatrix} \\ \Rightarrow r'_5 &= 3 + \Delta \geq 0 \\ \Rightarrow \Delta &\geq -3\end{aligned}$$

$\Rightarrow c_5$  puede variar entre  $(7, \infty)$  sin que varíe la solución óptima.



c) Ya que  $x_1$  es una variable básica, se debe estudiar cómo un cambio en  $c_1$  afectará el costo reducido de las variables no-básicas.

Caso 1)

$$\begin{aligned} r'_4 &= c_4 - c_B^T B^{-1} D_4 \\ \Rightarrow r'_4 &= -6 - \begin{pmatrix} 3 + \Delta & 2 & -3 \end{pmatrix} \begin{pmatrix} -1 & -2 & 2 \\ 1 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ 3 \end{pmatrix} \\ \Rightarrow r'_4 &= 1 + \Delta \geq 0 \Rightarrow \Delta \geq -1 \end{aligned}$$

Caso 2)

$$\begin{aligned} r'_5 &= c_5 - c_B^T B^{-1} D_5 \\ \Rightarrow r'_5 &= 10 - \begin{pmatrix} 3 + \Delta & 2 & -3 \end{pmatrix} \begin{pmatrix} -1 & -2 & 2 \\ 1 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ -2 \\ -1 \end{pmatrix} \\ \Rightarrow r'_5 &= 3 - 2\Delta \geq 0 \Rightarrow \Delta \leq 3/2 \end{aligned}$$

Caso 3)

$$\begin{aligned} r'_6 &= c_6 - c_B^T B^{-1} D_6 \\ \Rightarrow r'_6 &= -5 - \begin{pmatrix} 3 + \Delta & 2 & -3 \end{pmatrix} \begin{pmatrix} -1 & -2 & 2 \\ 1 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ -1 \\ -5 \end{pmatrix} \\ \Rightarrow r'_6 &= -2 + 7\Delta \geq 0 \Rightarrow \Delta \geq 2/7 \end{aligned}$$

Así para que la solución óptima no cambie se requiere  $2/7 \leq \Delta \leq 3/2$ .

d) El valor de las variables básicas en el óptimo debe ser positivo, por lo que

$$\begin{aligned} x_B &= B^{-1}b \geq 0 \\ \Rightarrow x_B + B^{-1}\Delta b &\geq 0 \\ \Rightarrow \begin{pmatrix} 3 \\ 4 \\ 2 \end{pmatrix} + \begin{pmatrix} -1 & -2 & 2 \\ 1 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta \\ 0 \\ 0 \end{pmatrix} &= \begin{pmatrix} 3 - \Delta \\ 4 + \Delta \\ 2 - \Delta \end{pmatrix} \geq 0 \\ \Rightarrow -4 &\leq \Delta \leq 2 \end{aligned}$$

El cambio en el valor óptimo (asumiendo que  $\Delta$  se mantiene en este rango) será

$$\begin{aligned}\Delta v^* &= c_B^T B^{-1} \Delta b \\ \Rightarrow \Delta v^* &= \begin{pmatrix} 3 & 2 & -3 \end{pmatrix} \begin{pmatrix} -1 & -2 & 2 \\ 1 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta \\ 0 \\ 0 \end{pmatrix} = 2\Delta\end{aligned}$$

e) Esta variación afectará el costo reducido de  $x_5$ , por lo que se debe buscar el rango tal que logre que esta variable permanezca fuera de la base. Así se ve que

$$\begin{aligned}r'_5 &= c_5 - c_B^T B^{-1} D_5 \\ \Rightarrow r'_5 &= 10 - \begin{pmatrix} 3 & 2 & -3 \end{pmatrix} \begin{pmatrix} -1 & -2 & 2 \\ 1 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ -2 + \Delta \\ -1 \end{pmatrix} \\ \Rightarrow r'_5 &= 27 - 8\Delta \geq 0 \Rightarrow \Delta \leq 27/8\end{aligned}$$

**Problema 68 (I3 1°03)** Considere el problema de optimización

$$\begin{aligned}P) \quad & \text{Max} \quad 2x + 5y \\ & \text{s.a.} \quad (x, y) \in D \\ & \quad \quad x, y \in \mathbb{Z}^+\end{aligned}$$

donde  $D$  es un dominio definido por desigualdades lineales. Suponga que al resolver el problema relajado la solución es  $x^* = 5/3$ ,  $y^* = 2/3$ . Con la información dada encuentre la mejor cota superior para  $P$ ), y justifique su respuesta.

**Solución:** El valor óptimo del problema relajado es  $z = \frac{20}{3}$ , por lo que  $z = 6$  sería la mejor cota superior ya que los coeficientes de la función objetivo son números enteros.

### 3.5.3 Dualidad

**Problema 69 (Problema Dual (I3 1°03))** Considere el problema primal

$$\begin{aligned}\text{Min} \quad & \mathbf{c}^\top \mathbf{x} \\ & \text{s.a.} \quad \mathbf{Ax} \geq \mathbf{b} \\ & \quad \quad \mathbf{x} \geq \mathbf{0}\end{aligned}$$

donde  $\mathbf{x} \in \mathbb{R}^n$ . Formule el problema dual y luego conviértalo en uno equivalente de minimización. Derive un conjunto de condiciones para la matriz  $\mathbf{A}$  y los vectores  $\mathbf{b}$  y  $\mathbf{c}$ , bajo las cuales el problema dual sea idéntico al primal. Por último, construya un ejemplo donde estas condiciones sean satisfechas.

**Solución:**

$$\begin{array}{lll} P) \text{ Min } c^T x & D) \text{ Max } \lambda^T b & \tilde{D}) \text{ Min } -\lambda^T b \\ s.a. Ax \geq b & \longrightarrow s.a. A^T \lambda \leq c & \longrightarrow s.a. A^T \lambda \leq c \\ x \geq 0 & \lambda \geq 0 & \lambda \geq 0 \end{array}$$

Condiciones:

- i)  $c = -b$  y ambos de igual dimensión.
- ii)  $A$  es una matriz cuadrada tal que:  $A = -A^T$ , es decir,  $A$  tiene ceros en la diagonal y  $a_{ij} = -a_{ji} \forall (i, j) \notin \text{diag}(A)$ .

Ejemplo:

$$A = \begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -3 \\ 2 & 3 & 0 \end{bmatrix} \quad c = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad b = \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix}$$

$$\begin{array}{ll} P) \text{ Min } x_1 + x_2 + x_3 & \\ s.a. & -x_2 - 2x_3 \geq -1 \\ & x_1 - 3x_3 \geq -1 \\ & 2x_1 + 3x_2 \geq -1 \\ & x_1, x_2, x_3 \geq 0 \end{array}$$

$$\begin{array}{ll} D) \text{ Min } \lambda_1 + \lambda_2 + \lambda_3 & \\ s.a. & \lambda_2 + 2\lambda_3 \leq 1 \\ & -\lambda_1 + 3\lambda_3 \leq 1 \\ & -2\lambda_1 - 3\lambda_2 \leq 1 \\ & \lambda_1, \lambda_2, \lambda_3 \geq 0 \end{array}$$

$$P) \equiv D).$$

**Problema 70** Escriba el problema dual de los siguientes problemas:

a)

$$\begin{aligned}
 & \text{Min } 2x_1 + 3x_2 \\
 & \text{s.a. } x_1 + 2x_2 = 3 \\
 & \quad 3x_1 + 4x_2 = 7 \\
 & \quad x_1 + x_2 = 2 \\
 & \quad x_i \geq 0, \forall i = 1, 2
 \end{aligned}$$

b)

$$\begin{aligned}
 & \text{Max } 2x_1 + 3x_2 + x_3 \\
 & \text{s.a. } x_1 + 4x_2 + x_3 \geq 15 \\
 & \quad x_1 + x_2 + x_3 = 13 \\
 & \quad 3x_1 - x_2 - x_3 \leq 3 \\
 & \quad x_1 \geq 0, x_2 \leq 0, x_3 \text{ libre}
 \end{aligned}$$

**Solución:**

2

a)

$$\begin{aligned}
 & \text{Max } 3\lambda_1 + 7\lambda_2 + 2\lambda_3 \\
 & \text{s.a. } \lambda_1 + 3\lambda_2 + \lambda_3 \leq 2 \\
 & \quad 2\lambda_1 + 4\lambda_2 + \lambda_3 \leq 3 \\
 & \quad \lambda_i \text{ libres, } \forall i = 1, 2, 3
 \end{aligned}$$

b)

$$\begin{aligned}
 & \text{Min } 15\lambda_1 + 13\lambda_2 + 3\lambda_3 \\
 & \text{s.a. } \lambda_1 + \lambda_2 + 3\lambda_3 \geq 2 \\
 & \quad 4\lambda_1 + \lambda_2 - \lambda_3 \leq 3 \\
 & \quad \lambda_1 + \lambda_2 - \lambda_3 = 1 \\
 & \quad \lambda_1 \leq 0, \lambda_2 \text{ libre}, \lambda_3 \geq 0
 \end{aligned}$$

**Problema 71** Considere el siguiente problema

$$\begin{aligned} \text{Max} \quad & -5x_1 + 5x_2 + 13x_3 \\ \text{s.a.} \quad & -x_1 + x_2 + x_3 \leq 20 \\ & 12x_1 + 4x_2 + 10x_3 \leq 90 \\ & x_i \geq 0, \forall i = 1, \dots, 3 \end{aligned}$$

a) Resuelva el problema utilizando análisis gráfico del dual, y luego análisis matricial para obtener la solución del primal.

b) Desarrolle un análisis de sensibilidad de la solución óptima del problema, evaluando independientemente el impacto de cada uno de los siguientes cambios:

- i) Cambio del lado derecho derecho de la primera restricción al valor 30.
- ii) Cambio en el lado derecho de la segunda restricción al valor 70.
- iii) Cambio del coeficiente de la tercera variable de la F.O. al valor 8.
- iv) Cambio del coeficiente de la segunda variable de la F.O. al valor 6.
- v) Introducción de una nueva variable al modelo con coeficiente 10 en la F.O. y con coeficientes 3 y 5 en la primera y segunda restricción respectivamente.
- vi) Introducción de una nueva restricción al modelo:  $2x_1 + 3x_2 + 5x_3 \leq 100$ .

**Solución:**

a) El problema dual será

$$\begin{aligned} \text{Min } w \quad & = 20\lambda_1 + 90\lambda_2 \\ \text{s.a.} \quad & -\lambda_1 + 12\lambda_2 \geq -5 \\ & \lambda_1 + 4\lambda_2 \geq 5 \\ & \lambda_1 + 10\lambda_2 \geq 13 \\ & \lambda_i \geq 0, \forall i = 1, 2 \end{aligned}$$

La ventaja de buscar la solución gráfica del dual, es que este problema es de dos dimensiones. Realizando el análisis gráfico del problema dual se obtiene la solución óptima:  $\lambda_1 = 0$ ,  $\lambda_2 = 1.3$ ,  $w^* = 117$ .

Así el tableau del primal será

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$b_i$
			0	-1.3	-117

Esto indica que  $x_4$  está en la base, por lo que  $x_4 > 0$ . Debemos buscar una variable más

que pertenezca a la base entre  $x_1, x_2, x_3$ , pues se sabe que  $x_5$  no está en la base. Como la tercera restricción del problema dual está activa (ver gráfico), entonces  $x_3 > 0$ , y como la 1ra y 2da restricción están inactivas,  $x_1, x_2 = 0$ . Así la base será  $x_B = \{x_3, x_4\}$ , por lo que  $B = \begin{pmatrix} 1 & 1 \\ 10 & 0 \end{pmatrix}$ . Entonces, la base en el óptimo tomará el valor

$$x = B^{-1}b = \begin{pmatrix} 0 & 1/10 \\ 1 & -1/10 \end{pmatrix} \begin{pmatrix} 20 \\ 90 \end{pmatrix} = \begin{pmatrix} 9 \\ 11 \end{pmatrix} \rightarrow \begin{matrix} x_3 \\ x_4 \end{matrix}$$

y  $x_1, x_2, x_5 = 0$ .

b)

i)

$$\begin{aligned} x &= B^{-1}b > \vec{0} \text{ con } b' = \begin{pmatrix} 30 \\ 90 \end{pmatrix} \\ \Rightarrow x &= B^{-1}b' = \begin{pmatrix} 0 & 1/10 \\ 1 & -1/10 \end{pmatrix} \begin{pmatrix} 30 \\ 90 \end{pmatrix} = \begin{pmatrix} 9 \\ 21 \end{pmatrix} > \vec{0} \quad \checkmark\checkmark \end{aligned}$$

ii)

$$\begin{aligned} x &= B^{-1}b > \vec{0} \text{ con } b' = \begin{pmatrix} 20 \\ 70 \end{pmatrix} \\ \Rightarrow x &= B^{-1}b' = \begin{pmatrix} 0 & 1/10 \\ 1 & -1/10 \end{pmatrix} \begin{pmatrix} 20 \\ 70 \end{pmatrix} = \begin{pmatrix} 7 \\ 13 \end{pmatrix} > \vec{0} \quad \checkmark\checkmark \end{aligned}$$

iii)

$$r_j = c_j^T - c_B^T B^{-1} D_j$$

$$\begin{aligned} \bullet r'_1 &= c_1^T - c_B^T B^{-1} D_1 \\ &= -5 - \begin{pmatrix} 8 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1/10 \\ 1 & -1/10 \end{pmatrix} \begin{pmatrix} -1 \\ 12 \end{pmatrix} = -5 - \frac{812}{10} < 0 \quad \checkmark\checkmark \\ \bullet r'_2 &= c_2^T - c_B^T B^{-1} D_2 \\ &= 5 - \begin{pmatrix} 8 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1/10 \\ 1 & -1/10 \end{pmatrix} \begin{pmatrix} 1 \\ 4 \end{pmatrix} = 5 - \frac{32}{10} > 0 \\ &\Rightarrow x_2 \text{ entra a la base.} \\ \bullet r'_3 &= c_3^T - c_B^T B^{-1} D_3 \\ &= -0 - \begin{pmatrix} 8 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1/10 \\ 1 & -1/10 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\frac{8}{10} < 0 \quad \checkmark\checkmark \end{aligned}$$

Así la base cambia, entrando a ella la variable  $x_2$ .

iv)

$$r_j = c_j^T - c_B^T B^{-1} D_j$$

$$\begin{aligned} \bullet r'_2 &= c_2^T - c_B^T B^{-1} D_2 \\ &= 6 - \begin{pmatrix} 13 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1/10 \\ 1 & -1/10 \end{pmatrix} \begin{pmatrix} 1 \\ 4 \end{pmatrix} = 6 - \frac{52}{10} > 0 \\ &\Rightarrow \text{cambia la base, entrando } x_2. \end{aligned}$$

v)

$$c_6 = 10, a_{i6} = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$$

$$\begin{aligned} \bullet r'_6 &= c_6^T - c_B^T B^{-1} D_6 \\ &= 10 - \begin{pmatrix} 13 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1/10 \\ 1 & -1/10 \end{pmatrix} \begin{pmatrix} 3 \\ 5 \end{pmatrix} = 10 - \frac{65}{10} > 0 \\ &\Rightarrow \text{cambia la base, entrando } x_6. \end{aligned}$$

vi) La solución óptima cumple la nueva restricción, por lo que la solución óptima no cambia.

**Problema 72** Considere le siguiente problema de optimización

$$\begin{aligned} \text{Max } & 40x_1 + 60x_2 \\ \text{s.a. } & 2x_1 + x_2 \leq 3 \\ & x_1 + x_2 \leq 40 \\ & x_1 + 3x_2 \leq 90 \\ & x_i \geq 0, \forall i = 1, 2 \end{aligned}$$

a) Resuelva mediante el algoritmo Simplex

b) Plantee el problema dual y obtenga la solución óptima a partir de la solución del primal.

**Solución:**

a) El tableau inicial será

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$b_i$
2	1	1	0	0	70
1	1	0	1	0	40
1	3	0	0	1	90
40	60	0	0	0	0

En la primera iteración, entrará  $x_2$  a la base, y saldrá  $x_5$ . El tableau quedará

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$b_i$
5/3	0	1	0	-1/3	40
2/3	0	0	1	-1/3	10
1/3	1	0	0	1/3	30
20	0	0	0	-20	-1800

Ahora entrará  $x_1$  y saldrá  $x_4$  de la base. El tableau final será

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$b_i$
0	0	1	-5/2	1/2	15
1	0	0	3/2	-1/2	15
0	1	0	-1/2	1/2	25
0	0	0	-30	-10	-2100

Así la solución óptima será  $(x_1, x_2, x_3, x_4, x_5) = (15, 25, 15, 0, 0)$ , donde la segunda y tercera restricción están activas y el valor óptimo es  $v^* = 2100$ .

b) El problema dual será

$$\begin{aligned}
 \text{Min } & 70y_1 + 40y_2 + 90y_3 \\
 \text{s.a. } & 2y_1 + y_2 + y_3 \geq 40 \\
 & y_1 + y_2 + 3y_3 \geq 60 \\
 & y_i \geq 0, \forall i = 1, 2, 3
 \end{aligned}$$

Observando el tableau final del problema primal, se puede observar que los valores que tomarán las variables duales serán  $(y_1, y_2, y_3) = (0, 30, 10)$ . Además se sabe que el valor óptimo no cambia  $v^* = 2100$ . Ya que las variables asociadas a las restricciones del problema dual  $(x_1, x_2)$  son diferentes de cero en la solución óptima del primal, las dos restricciones del dual estarán activas en el óptimo.

**Problema 73** *Una fábrica de juguetes produce camiones tolva, muñecas, pelotas y conejitos inflables. Para ello utiliza una máquina inyectora de plástico, una máquina sopladora de plástico, mano de obra especializada para decorara los juguetes y plástico. La Tabla 3.2 muestra el uso de recursos por cada unidad de juguete fabricado y la capacidad disponible de cada recurso en un mes.*

La solución óptima se obtuvo computacionalmente; sin embargo, por problemas en el archivo correspondiente se perdió parte de ella y sólo se ha rescatado la siguiente información.

- El valor óptimo es \$2.527.000
- Las máquinas inyectora y sopladora se utilizan a plena capacidad.
- Se tienen 11.240 minutos-hombre (mh) de mano de obra ociosos.
- Se tienen 944 kg de plástico ociosos (sin usar).



	Camiones Tolva	Muñecas	Pelotas	Conejitos	Disponibilidad
Inyectora (min)	10	6	0	1	3840
Sopladora (min)	0	1	0.5	5	4200
Mano de obra (mh)	2	5	0.5	3	18000
Plástico (kg)	2	0.5	0.3	0.6	3400
Beneficios (\$)	3000	3500	350	1500	

Table 3.2: Utilización de recursos y capacidad disponible

- Los valores de las variables duales son:  $\pi_1 = 466.67$ ,  $\pi_2 = 700$  y  $\pi_3 = \pi_4 = 0$ .
- La base óptima es

$$B^{-1} = \begin{pmatrix} 1/6 & 0 & 0 & 0 \\ -1/3 & 2 & 0 & 0 \\ -2/3 & -1 & 1 & 0 \\ 1/60 & -6/10 & 0 & 1 \end{pmatrix}$$

- Reconstruya la solución óptima a partir de la información rescatada.
- Explique la solución óptima: ¿qué juguetes se fabrican?, ¿cuál es el beneficio total?, ¿qué recursos se utilizan completamente y cuales no?, etc.
- ¿En cuánto debe aumentar el beneficio de los camiones tolva para que sea conveniente fabricarlos?
- ¿Cuánto estaría dispuesto a pagar por tiempo adicional en la máquina inyectora? ¿y en la sopladora?
- Se está estudiando la posibilidad de reemplazar la máquina inyectora por una más moderna que tiene mayor confiabilidad. El precio de ella es \$4.000.000 y permite aumentar la capacidad de inyección en 20% ¿cuántos meses se requieren para recuperar la inversión?

**Solución:**

- El problema primal será

$$\begin{aligned}
 \text{Max } & 3000x_1 + 3500x_2 + 350x_3 + 1500x_4 \\
 \text{s.a. } & 10x_1 + 6x_2 + x_4 \leq 3840 \\
 & x_2 + 0.5x_3 + 5x_4 \leq 4200 \\
 & 2x_1 + 5x_2 + 0.5x_3 + 3x_4 \leq 18000 \\
 & 2x_1 + 0.5x_2 + 0.3x_3 + 0.6x_4 \leq 3400 \\
 & x_i \geq 0, \forall i = 1, \dots, 4
 \end{aligned}$$

Se asocian variables de holgura  $x_5, x_6, x_7, x_8$  a cada una de las restricciones.

Con los datos entregados se puede deducir que:

- el valor óptimo es \$2.527.000.
- $x_5, x_6 = 0$  en el óptimo.

- $x_7 = 11240$  en el óptimo.
- $x_8 = 944$  en el óptimo.
- $x_5, x_6$  están fuera de la base, mientras que  $x_7, x_8$  están dentro de ella, en el óptimo.
- Falta buscar las otras 2 variables que pertenecen a la base. Para esto, invertiremos la matriz  $B^{-1}$

$$\Rightarrow B = \begin{pmatrix} 6 & 0 & 0 & 0 \\ 1 & 0.5 & 0 & 0 \\ 5 & 0.5 & 1 & 0 \\ 0.5 & 0.3 & 0 & 1 \end{pmatrix}$$

Así, las variables que pertenecen a la base son:  $x_2, x_3, x_7, x_8$ .

Para obtener el valor de éstas variables en el óptimo, se calcula

$$\begin{pmatrix} x_2 \\ x_3 \\ x_7 \\ x_8 \end{pmatrix} = B^{-1}b = \begin{pmatrix} 1/6 & 0 & 0 & 0 \\ -1/3 & 2 & 0 & 0 \\ -2/3 & -1 & 1 & 0 \\ 1/60 & -6/10 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3840 \\ 4200 \\ 18000 \\ 3400 \end{pmatrix} = \begin{pmatrix} 640 \\ 820 \\ 11240 \\ 944 \end{pmatrix}$$

con  $x_1, x_4, x_5, x_6 = 0$  y un valor óptimo de \$2.527.000.

b) Se fabricarán 640 muñecas y 820 pelotas, y ningún camión tolva o conejito. Se utilizará toda la capacidad de las máquinas inyectoras y sopladoras (restricciones activas), pero la capacidad de mano de obra y plástico no será utilizada en su totalidad (restricciones inactivas). El beneficio total estará dado por el valor óptimo de \$2.527.000.

c) Para estudiar la sensibilidad del costo de los camiones tolva, se estudia el costo reducido de la variable  $x_1$ . Para esto se busca  $\Delta$  tal que el costo reducido de la variable se vuelva positivo, y así la variable entre a la base:

$$\begin{aligned} r'_1 &= c_1^T - c_B^T B^{-1} D_1 \\ &= 3000 + \Delta - \begin{pmatrix} 3500 & 350 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1/6 & 0 & 0 & 0 \\ -1/3 & 2 & 0 & 0 \\ -2/3 & -1 & 1 & 0 \\ 1/60 & -6/10 & 0 & 1 \end{pmatrix} \begin{pmatrix} 10 \\ 0 \\ 2 \\ 2 \end{pmatrix} \\ &= -1666.67 + \Delta > 0 \Rightarrow \Delta > 1666.67 \end{aligned}$$

Así el beneficio aportado por cada camión tolva debería ser mayor que 4666.67 para que fuera conveniente fabricarlos.

d) Para esto, se estudia en cuánto afecta un cambio en las disponibilidad del recurso a la función objetivo: la ganancia generada en este caso. En primer lugar se estudiará el rango en el que puede

variar la disponibilidad de la máquina inyectora:

$$\begin{aligned} \begin{pmatrix} x_2 \\ x_3 \\ x_7 \\ x_8 \end{pmatrix} = B^{-1}b &= \begin{pmatrix} 1/6 & 0 & 0 & 0 \\ -1/3 & 2 & 0 & 0 \\ -2/3 & -1 & 1 & 0 \\ 1/60 & -6/10 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3840 + \Delta_1 \\ 4200 \\ 18000 \\ 3400 \end{pmatrix} \\ &= \begin{pmatrix} 640 + \Delta_1/6 \\ 820 - \Delta_1/3 \\ 11240 - 2\Delta_1/3 \\ 944 + \Delta_1/60 \end{pmatrix} > 0 \end{aligned}$$

$$\Rightarrow \Delta_1 > -3840, \Delta_1 < 2460, \Delta_1 < 16860, \Delta_1 > -56640$$

$$\Rightarrow -3840 < \Delta_1 < 2460$$

Ahora, por una unidad de tiempo adicional,  $\Delta_1 = 1$ , el aumento de beneficio generado será:  $\Delta(\text{beneficio}) = 3500 \times 1/6 + 350 \times -1/3 = 466.67$ , por lo tanto, éste será el máximo que estaría dispuesto a pagar por una unidad de tiempo adicional en la máquina inyectora.

Siguiendo el mismo proceso para la máquina sopladora, se tiene

$$\begin{aligned} \begin{pmatrix} x_2 \\ x_3 \\ x_7 \\ x_8 \end{pmatrix} = B^{-1}b &= \begin{pmatrix} 1/6 & 0 & 0 & 0 \\ -1/3 & 2 & 0 & 0 \\ -2/3 & -1 & 1 & 0 \\ 1/60 & -6/10 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3840 \\ 4200 + \Delta_2 \\ 18000 \\ 3400 \end{pmatrix} \\ &= \begin{pmatrix} 640 \\ 820 + 2\Delta_2 \\ 11240 - \Delta_2 \\ 944 - 6\Delta_2/10 \end{pmatrix} > 0 \end{aligned}$$

$$\Rightarrow \Delta_2 > -410, \Delta_2 < 11240, \Delta_2 < 1573.33$$

$$\Rightarrow -410 < \Delta_2 < 1573.33$$

Ahora, por una unidad de tiempo adicional,  $\Delta_2 = 1$ , el aumento de beneficio generado será:  $\Delta(\text{beneficio}) = 350 \times 2 = 700$ , por lo tanto, éste será el máximo que estaría dispuesto a pagar por una unidad de tiempo adicional en la máquina sopladora.

**Notar** que el aumento en los beneficios generado por un aumento, en una unidad, de la capacidad de alguno de los recursos que restringen el problema (restricciones activas), coincide con los costos reducidos de las variables de holgura de la restricción correspondiente. Esto concuerda con la idea de que los costos reducidos representan al multiplicador de Lagrange de la restricción, que indica cuánto varía la función objetivo al variar en una unidad la disponibilidad del recurso.

e) Si se aumena la capacidad de inyección en un 20%  $\implies \Delta_1 = 768$  . El cambio en el beneficio generado mensualmente será:  $\Delta(beneficio) = 3500 \times \Delta_1/6 + 350 \times -\Delta_1/3 = 358400$ . Así el número de meses necesario para recuperar la inversión será:  $\$4.000.000/\$358.400 = 11.16$  meses.