



UNIVERSIDAD TECNOLÓGICA DE LA MIXTECA

**“UN RESULTADO GENÉRICO SOBRE UNICIDAD EN  
PROGRAMACIÓN LINEAL”**

TESIS  
PARA OBTENER EL TÍTULO DE:  
LICENCIADO EN MATEMÁTICAS APLICADAS

PRESENTA:  
DAVID LÓPEZ OSORIO

DIRECTOR DE TESIS:  
DR. GUILLERMO ARTURO LANCHO ROMERO

*Huajuapán de León, Oaxaca*

*Diciembre de 2010*



UN RESULTADO GENÉRICO SOBRE UNICIDAD EN  
PROGRAMACIÓN LINEAL

David López Osorio

Diciembre de 2010



*A mis padres, Joel y Eudelia.*



# Agradecimientos

En primer lugar agradezco a mi familia. A mis padres Joel y Eudelia les agradezco el apoyo y la confianza que me han brindado. A mis hermanos Joel y Daniel les agradezco su apoyo incondicional. Gracias a su ejemplo y motivación he concluido mis estudios.

Agradezco a mi asesor, Dr. Guillermo Arturo Lancho Romero, por guiarme en todas las etapas del desarrollo de este trabajo y por sus observaciones y sugerencias que han sido de provecho en mi desarrollo profesional.

A todos mis profesores y compañeros a quienes debo la formación académica y humana que conseguí durante mi estancia en esta universidad.

Agradezco el apoyo que recibí del Conacyt como parte del proyecto con número de registro 80993.

*David*





# Prefacio

La programación lineal es una rama de la optimización que estudia problemas que consisten en minimizar o maximizar una función lineal donde la solución se busca en el conjunto solución de un sistema de desigualdades lineales llamadas restricciones. La programación lineal ordinaria trata del estudio de los problemas lineales con un número finito de restricciones mientras que la programación lineal semi-infinita se ocupa de los problemas lineales en los que el número de restricciones es infinito.

En este trabajo consideramos el conjunto de los problemas de programación lineal ordinaria que tienen el mismo número de variables y restricciones. Equipando a este conjunto con una topología adecuada, demostraremos que el conjunto de problemas que tienen una única solución óptima es denso en el conjunto de problemas que tienen al menos una solución óptima, es decir, para cada problema lineal que tiene al menos una solución óptima existe un problema lineal que tiene una única solución óptima arbitrariamente próximo a él. Para esto usaremos caracterizaciones de optimalidad y unicidad así como condiciones de estabilidad de problemas lineales consistentes. Un problema lineal es consistente si el sistema de restricciones tiene al menos una solución. Un problema consistente es estable si al hacer modificaciones pequeñas en sus datos se obtiene un problema nuevo que también es consistente.

Veremos que el Teorema de Karush-Kuhn-Tucker es una condición necesaria y suficiente de optimalidad en programación lineal ordinaria. De esta condición de optimalidad se deriva una caracterización de los problemas que tienen una única solución óptima. Demostraremos algunas caracterizaciones del interior del conjunto de problemas consistentes usando la conocida condición de Slater y la continuidad de la llamada multifunción conjunto factible, y veremos que tales condiciones caracterizan a los problemas lineales consistentes que son estables. Para demostrar que el conjunto de problemas lineales que tienen una única solución óptima es denso en el conjunto de problemas lineales que tienen al menos una solución óptima, seguiremos los tres pasos siguientes. Primero, dado un problema de programación lineal ordinaria que tiene al menos una solución óptima, su correspondiente conjunto factible puede no ser estable. Haciendo modificaciones pequeñas en sus datos obtenemos un problema nuevo

arbitrariamente “cercano” al problema original que tiene conjunto factible estable, sin embargo este problema nuevo puede no tener solución óptima. Segundo, haciendo algunas perturbaciones convenientes al problema construido en el primer paso obtenemos un problema nuevo arbitrariamente “cercano” al problema construido en el primer paso y que tiene al menos una solución óptima. Tercero, modificando los datos del problema construido en el segundo paso obtenemos un problema nuevo con una única solución óptima arbitrariamente “cercano” al problema construido en el segundo paso y por lo tanto este último problema está arbitrariamente “cercano” al problema original.

En la referencia [6] está demostrado un resultado análogo para cierta clase de problemas de programación lineal semi-infinita. Las hipótesis empleadas en la demostración propuesta en [6] son sofisticadas y algunas no se requieren para enunciar el resultado en programación lineal ordinaria. Por ejemplo, el cono convexo generado por un conjunto finito de vectores es un cono cerrado mientras que el cono convexo generado por un número infinito de vectores puede no ser cerrado. Además en programación lineal ordinaria en cada punto de la frontera existe al menos una restricción activa, si un problema tiene solución óptima entonces existe un punto extremo que es solución óptima, lo cual no sucede en general en programación lineal semi-infinita.

La estructura de la tesis es la siguiente. En el capítulo 1 presentamos una reseña histórica del desarrollo de la programación matemática y en particular de la programación lineal. En la sección 1.1, presentamos la notación usual e introducimos una métrica en el conjunto de problemas de programación lineal ordinaria para dotar a este conjunto con una topología. En la sección 1.2 mencionamos algunos ejemplos que ilustran aplicaciones de la programación lineal.

En el Capítulo 2 presentamos una compilación de herramientas de análisis convexo y funciones convexas que utilizaremos en este trabajo. La sección 2.1 consta de definiciones y propiedades básicas de los conjuntos convexas. En la sección 2.2 se presentan las propiedades geométricas de separación y soporte de conjuntos convexas, mismas que son importantes para establecer los criterios de optimalidad y unicidad en los problemas de optimización. En la sección 2.3 estudiamos algunas propiedades del conjunto factible de un problema de programación lineal ordinaria, el cual pertenece a la clase de conjunto convexas llamados poliedros. En la sección 2.4 hablaremos de las funciones convexas, de algunas propiedades de estas funciones y de criterios que caracterizan a las funciones de este tipo.

En el Capítulo 3 presentamos las condiciones necesarias y suficientes de optimalidad y unicidad. La sección 3.1 tiene como resultado principal al Teorema que, para problemas de optimización cuya función objetivo es diferenciable, caracteriza las direcciones de descenso

en cada punto factible en términos del gradiente de la función objetivo en dicho punto. Este resultado permite derivar la condición necesaria de primer orden para optimalidad, la cual exponemos en la sección 3.2. En la sección 3.4 veremos que el Teorema de Karush-Kuhn-Tucker se deduce de la condición necesaria para optimalidad local usando el concepto de restricciones activas. Este Teorema es también una condición suficiente para problemas en los que las funciones involucradas además de ser diferenciables satisfacen propiedades de convexidad, tal como lo demostraremos en el Teorema 3.15. En la sección 3.5 presentamos una caracterización de unicidad en programación lineal ordinaria.

En la primera sección del Capítulo 4 demostramos criterios de estabilidad de los problemas consistentes y resolubles. En la sección 4.2 demostramos el resultado principal de esta tesis el cual mencionamos anteriormente.

La demostración de la suficiencia del Teorema de Karush-Kuhn-Tucker, enunciada en el Teorema 3.15, así como las demostraciones de los resultados presentados en el Capítulo 4 para programación lineal ordinaria son pruebas originales propuestas en esta tesis.



# Índice general

<b>Prefacio</b>	<b>vii</b>
<b>1. Introducción</b>	<b>1</b>
1.1. Preliminares . . . . .	1
1.2. Programación lineal . . . . .	3
1.3. Ejemplos . . . . .	6
<b>2. Análisis convexo</b>	<b>11</b>
2.1. Conjuntos convexos . . . . .	11
2.2. Separación y soporte de conjuntos convexos . . . . .	21
2.3. Conjunto factible de un problema de programación lineal . . . . .	24
2.4. Funciones convexas . . . . .	31
<b>3. Criterios de optimalidad y unicidad</b>	<b>39</b>
3.1. Direcciones de descenso . . . . .	40
3.2. Condición necesaria de primer orden . . . . .	42
3.3. Restricciones activas . . . . .	44
3.4. Teorema de Karush-Kuhn-Tucker . . . . .	48
3.5. Unicidad . . . . .	55
<b>4. Resultados principales</b>	<b>59</b>
4.1. Estabilidad . . . . .	59
4.2. Resultado principal . . . . .	65
<b>5. Conclusiones</b>	<b>71</b>
<b>Topología</b>	<b>73</b>

Diferenciabilidad	77
Bibliografía	79

# Capítulo 1

## Introducción

En este Capítulo mencionamos una reseña del desarrollo de la programación lineal y una introducción a la teoría de la optimización.

### 1.1. Preliminares

La teoría de optimización o programación matemática es la rama de las matemáticas aplicadas que se ocupa del estudio de los problemas que consisten en la minimización o maximización de alguna función, llamada función objetivo. La solución buscada debe ser solución de un sistema de ecuaciones o de desigualdades llamadas restricciones. Así, un problema de programación matemática tiene la forma siguiente:

$$(P) \quad \text{Minimizar } f(\mathbf{x}) \\ \text{sujeto a } g_i(\mathbf{x}) \geq 0, \text{ para cada } i = 1, 2, \dots, m.$$

donde  $\mathbf{x} \in \mathbb{R}^n$  es el vector de variables,  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es llamada función objetivo y  $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$  son las restricciones del problema. A cada  $\mathbf{x} \in \mathbb{R}^n$  que satisface las restricciones se le denomina punto factible.

En la teoría de optimización se persiguen varios objetivos, uno de ellos es la caracterización de soluciones óptimas, esto es, se buscan condiciones necesarias y suficientes para que un punto factible sea la solución óptima de un problema dado. Otro objetivo es el diseño de métodos iterativos que ayuden a encontrar la solución óptima. Finalmente mencionamos como otro objetivo la construcción de modelos de programación matemática que ayuden a resolver problemas prácticos, que de manera natural surgen de campos muy diversos como planeación, asignación, administración de recursos naturales, economía, etc.

Si  $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ ,  $\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$ , entonces con la expresión  $\mathbf{x}'$  denotaremos el vector transpuesto de  $\mathbf{x}$ , es decir,  $\mathbf{x}' = (x_1, x_2, \dots, x_n)$ ;  $\mathbf{x}'\mathbf{y}$  representa el producto interno usual en  $\mathbb{R}^n$  definido por  $\mathbf{x}'\mathbf{y} = \sum_{j=1}^n x_j y_j$ .

Si el objetivo y las restricciones del problema ( $P$ ) son funciones lineales entonces se dice que el problema ( $P$ ) es un problema de programación lineal. Si el objetivo es una función cuadrática y las restricciones son lineales entonces tenemos un problema de programación cuadrática. Así un problema de programación cuadrática tiene la forma siguiente:

$$\begin{aligned} & \text{Minimizar } \mathbf{x}'\mathbf{H}\mathbf{x} + \mathbf{a}'\mathbf{x} \\ & \text{sujeto a} \\ & \quad a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \geq b_1 \\ & \quad \vdots \\ & \quad a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \geq b_m. \end{aligned}$$

donde  $\mathbf{H}$  es una matriz  $n \times n$  con entradas reales,  $\mathbf{a} \in \mathbb{R}^n$  y  $b_i \in \mathbb{R}$ ,  $i = 1, \dots, m$  son constantes. Si las restricciones o la función objetivo son no lineales, entonces el problema ( $P$ ) se conoce como problema de programación no lineal.

La programación lineal se creó como rama independiente de la teoría de optimización gracias al impulso que se dió a la investigación y desarrollo de modelos para el uso eficiente de los recursos en el contexto de la segunda guerra mundial ([7], [1]). El estudio de la programación lineal recibió un gran impulso con el desarrollo del método simplex creado por George Dantzig en 1947. El incremento en la capacidad de procesamiento y disponibilidad de las computadoras desde esta época permitieron que el método simplex fuera capaz de resolver problemas involucrando gran cantidad de datos, con esto, la programación lineal atrajo la atención de quienes vieron en ésta una herramienta poderosa y eficiente. Posteriormente, la publicación del teorema de dualidad por Gale, Khun y Tucker en 1951 introdujo la teoría de dualidad, la cual ha sido una rama importante en el desarrollo de esta teoría.

Posteriormente al inicio de la programación lineal, el desarrollo de la teoría de optimización ha sido notable: En 1959, Wolfe desarrolló el método simplex para programación cuadrática y en 1963 desarrolló el método del gradiente reducido para problemas con restricciones lineales y función objetivo no lineal. Posteriormente se desarrollaron los métodos numéricos para programación no lineal, en 1970 diversos autores desarrollaron métodos cuasi-Newton para optimizar funciones no cuadráticas sin restricciones y en 1969 Abadie generalizó el método del gradiente reducido para problemas de programación no lineal.



## 1.2. Programación lineal

Un problema de programación lineal semi-infinita consiste en minimizar o maximizar una función objetivo, satisfaciendo un conjunto de restricciones dadas en términos de igualdades y desigualdades lineales donde el número de variables o el de restricciones, pero no ambos a la vez, puede ser infinito (ver [8], Capítulo 1). Si el número de variables es finito, entonces el problema tiene la forma siguiente:

$$\text{Minimizar} \quad \mathbf{c}'\mathbf{x} \quad (1.1)$$

$$\text{sujeto a} \quad \mathbf{a}'_t\mathbf{x} \geq b_t, \quad t \in T, \quad (1.2)$$

donde  $T$  es un conjunto de índices posiblemente infinito;  $\mathbf{x} \in \mathbb{R}^n$  es el *vector de variables*,  $\mathbf{c} \in \mathbb{R}^n$  es el *vector de costos*,  $a : T \rightarrow \mathbb{R}^n$  y  $b : T \rightarrow \mathbb{R}$  son funciones cuyas imágenes se representan con  $\mathbf{a}_t := a(t)$  y  $b_t := b(t)$ , respectivamente. Cualquier problema lineal con finitas variables puede ser reformulado en un problema de la forma anterior mediante algunas manipulaciones convenientes.

Evidentemente un problema de programación lineal semi-infinita está determinado por el vector  $\mathbf{c}$  y por los mapeos  $a$  y  $b$ , llamados los datos del problema; por lo cual el problema se puede identificar con el parámetro  $\sigma = (\mathbf{c}, a, b)$ . Se denota con  $\Pi$  al conjunto de todos los parámetros  $\sigma$  con el mismo conjunto de índices  $T$ . Cuando se consideran varios parámetros o problemas a la vez, éstos se distinguen mediante el uso de subíndices y sus datos se distinguen con superíndices por ejemplo  $\sigma_1 = (\mathbf{c}^1, a^1, b^1)$ ,  $\sigma_2 = (\mathbf{c}^2, a^2, b^2)$ .

**Definición 1.1.** Un punto  $\mathbf{x} \in \mathbb{R}^n$  se llama *punto factible* para  $\sigma$  si  $\mathbf{x}$  satisface el sistema de desigualdades asociado a  $\sigma$ , es decir,  $\mathbf{a}'_t\mathbf{x} \geq b_t$ , para cada  $t \in T$ . El conjunto de puntos factibles del problema  $\sigma$  se denota con  $F$  y se llama *conjunto factible*. Si  $F \neq \emptyset$  se dice que  $\sigma$  es un *problema consistente* y en caso contrario se denomina *problema inconsistente*.

**Definición 1.2.** Sea  $\bar{\mathbf{x}} \in F$ ,  $\bar{\mathbf{x}}$  es llamado *solución óptima* del problema  $\sigma$  si  $\mathbf{c}'\bar{\mathbf{x}} \leq \mathbf{c}'\mathbf{x}$ , para todo  $\mathbf{x} \in F$ . En este caso diremos que  $\mathbf{c}'\bar{\mathbf{x}}$  es el *valor objetivo óptimo* y lo denotaremos con  $V$ . El *conjunto de soluciones óptimas* del problema  $\sigma$  denotado con  $F^*$  está definido como sigue

$$F^* = \{\mathbf{x} \in F : \mathbf{c}'\mathbf{x} = V\}.$$

Se dice que el problema representado por  $\sigma$  tiene *solución óptima única* si  $F^* = \{x^*\}$ . Se dice que el problema definido por  $\sigma$  es *resoluble* si tiene al menos una solución óptima.

Si  $T$  es un conjunto finito, por ejemplo  $T = \{1, 2, \dots, m\}$ , (1.1) y (1.2) definen un problema

de programación lineal ordinaria de la forma siguiente:

$$\begin{array}{ll} \text{Minimizar} & c_1x_1 + c_2x_2 + \cdots + c_nx_n \\ \text{sujeto a} & \\ & a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \geq b_1 \\ & \vdots \\ & a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \geq b_m. \end{array}$$

Supondremos que las restricciones de no negatividad de las variables,  $x_i \geq 0$ ,  $i = 1, \dots, n$ , están incluidas en el sistema de restricciones anterior, de aquí,  $m \geq n$ .

Si  $T$  es un espacio topológico Hausdorff compacto y las funciones  $a$ ,  $b$  son continuas entonces se dice que el problema definido por  $\sigma = (c, a, b)$  es *continuo*. Se dice que  $\sigma$  es *acotado* si  $\sup\{\|a_t\| : t \in T\} < \infty$ . Claramente los problemas de programación lineal ordinaria son continuos considerando a  $T$  como espacio topológico equipado con la topología discreta. De acuerdo con el conocido Teorema de Weierstrass, si las funciones  $a : T \rightarrow \mathbb{R}^n$  y  $b : T \rightarrow \mathbb{R}$  son continuas sobre el compacto  $T$  entonces son acotadas, por lo cual, cada problema continuo es acotado.

Una de las áreas en el estudio de la programación lineal es el análisis de estabilidad. En el estudio de la estabilidad se trata de caracterizar a los problemas con ciertas propiedades de interés tales que los problemas “cercanos” conservan las mismas propiedades. Para este propósito en la literatura ([6], [8]) se ha dotado al espacio  $\Pi$  de la topología inducida por la métrica extendida  $d : \Pi \times \Pi \rightarrow [0, \infty]$  definida mediante:

$$d(\sigma_1, \sigma) := \max \left\{ \|c^1 - c\|, \sup_{t \in T} \left\| \begin{pmatrix} a_t^1 \\ b_t^1 \end{pmatrix} - \begin{pmatrix} a_t \\ b_t \end{pmatrix} \right\| \right\},$$

donde  $\|\cdot\|$  es cualquier norma definida en  $\mathbb{R}^n$ . Esto proporciona a  $(\Pi, d)$  la estructura de espacio topológico Hausdorff que satisface el primer axioma de numerabilidad, por lo cual la convergencia en este espacio se puede caracterizar mediante sucesiones.

Un problema de programación lineal ordinaria con  $m$  restricciones puede ser representado en forma matricial si se define la *matriz de coeficientes tecnológicos*, denotada con  $A$ , como sigue:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

y el vector de recursos disponibles, denotado con  $\mathbf{b}$ , se define como sigue:

$$\mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix},$$

con lo cual, el problema queda formulado como sigue

$$\begin{aligned} & \text{Minimizar } \mathbf{c}'\mathbf{x} \\ & \text{sujeito a } \mathbf{A}\mathbf{x} \geq \mathbf{b}. \end{aligned}$$

El problema anterior se identifica con el parámetro  $\sigma = (\mathbf{c}, \mathbf{A}, \mathbf{b})$ . Denotaremos con  $\theta$  al conjunto de parámetros  $\sigma$  que representan problemas de programación lineal ordinaria con  $m$  restricciones. Claramente  $\theta \equiv \mathbb{R}^n \times \mathbb{R}^{mn} \times \mathbb{R}^m$ .

En este trabajo estudiaremos algunas propiedades topológicas de subconjuntos especiales del conjunto  $\theta$ , tales como el subconjunto de problemas resolubles y el conjunto de problemas consistentes, para esto dotaremos a  $\theta$  con la topología inducida por la métrica definida por la función  $d : \theta \times \theta \rightarrow \mathbb{R}$  tal que, para cada  $\sigma = (\mathbf{c}, \mathbf{A}, \mathbf{b}), \sigma_1 = (\mathbf{c}^1, \mathbf{A}^1, \mathbf{b}^1) \in \theta$ ,

$$d(\sigma, \sigma_1) := \|\mathbf{A} - \mathbf{A}^1\| + \|\mathbf{b} - \mathbf{b}^1\|_\infty + \|\mathbf{c} - \mathbf{c}^1\|_\infty, \quad (1.3)$$

donde  $\|\mathbf{A}\| := \max_{i,j} |a_{ij}|$ .

**Proposición 1.3.** *La función  $d : \theta \times \theta \rightarrow \mathbb{R}$  definida mediante (1.3) es una métrica.*

*Demostración.* Sean  $\sigma, \sigma_1, \sigma_2 \in \theta$ , entonces

(i)  $d(\sigma, \sigma_1) \geq 0$  es evidente. Además

$$\begin{aligned} d(\sigma, \sigma_1) = 0 & \Leftrightarrow \|\mathbf{A} - \mathbf{A}^1\| = 0, \|\mathbf{b} - \mathbf{b}^1\|_\infty = 0, \|\mathbf{c} - \mathbf{c}^1\|_\infty = 0 \\ & \Leftrightarrow \mathbf{A} = \mathbf{A}^1, \mathbf{b} = \mathbf{b}^1, \mathbf{c} = \mathbf{c}^1 \\ & \Leftrightarrow \sigma = \sigma_1. \end{aligned}$$

(ii) Es claro que  $d(\sigma, \sigma_1) = d(\sigma_1, \sigma)$ .

(iii) Tenemos que

$$\begin{aligned} d(\sigma, \sigma_2) &= \|\mathbf{A} - \mathbf{A}^2\| + \|\mathbf{b} - \mathbf{b}^2\|_\infty + \|\mathbf{c} - \mathbf{c}^2\|_\infty \\ &\leq \|\mathbf{A} - \mathbf{A}^1\| + \|\mathbf{A}^1 - \mathbf{A}^2\| + \|\mathbf{b} - \mathbf{b}^1\|_\infty + \|\mathbf{b}^1 - \mathbf{b}^2\|_\infty + \\ &\quad \|\mathbf{c} - \mathbf{c}^1\|_\infty + \|\mathbf{c}^1 - \mathbf{c}^2\|_\infty \\ &= d(\sigma, \sigma_1) + d(\sigma_1, \sigma_2). \end{aligned}$$

De lo anterior,  $d(\sigma, \sigma_2) \leq d(\sigma, \sigma_1) + d(\sigma_1, \sigma_2)$ . Por lo tanto  $d$  es una métrica.

□

### 1.3. Ejemplos

Las aplicaciones de la programación lineal son diversas. Existen problemas que son formulables directamente como problemas de programación lineal en la planificación de actividades, diseño de políticas públicas, etc. El ejemplo que sigue muestra un problema de planificación de la producción que puede ser modelado mediante la programación lineal.

**Ejemplo 1.4.** *En Suchixtlahuaca el gobierno municipal local ha impulsado dos programas paralelos de empleo para sus ciudadanos, los cuales consisten en la producción de jitomate en un ambiente controlado de invernadero y la tradicional producción de jitomate y trigo como cultivos de temporal. La localidad, bajo el régimen de propiedad comunal, posee  $H$  hectáreas de tierra disponibles para su aprovechamiento, de las cuales se han usado  $K$  hectáreas para la construcción de invernaderos, y  $L$  horas de trabajo. Supondremos que los costos de producción unitarios del jitomate y trigo en campo son  $c_1$  y  $c_2$ , en cada caso. Además el costo de producción del jitomate en invernadero es  $c_3$ . De acuerdo con su experiencia en el comercio, los pobladores esperan los siguientes precios para la siguiente temporada*

Producto	Precio esperado	Costo
Jitomate en campo	$p_1$	$c_1$
Trigo	$p_2$	$c_2$
Jitomate en invernadero	$p_3$	$c_3$

Para fines de simplificación, supondremos que los costos de producción son constantes. La tecnología de producción se resume en la siguiente tabla

Producto	Tierra	Trabajo
Jitomate en campo	$a_{11}$	$a_{21}$
Trigo	$a_{12}$	$a_{22}$
Jitomate en invernadero	$a_{13}$	$a_{23}$

donde la tierra se mide en hectáreas y el trabajo en horas. El gobierno local desea planificar su producción para el siguiente período de forma que se maximizen sus ganancias.

Denotemos, de forma respectiva, con  $x_1$  y  $x_2$  a las cantidades a producir de jitomate y trigo en campo y con  $x_3$  la cantidad de producción de jitomate en invernadero. El problema a resolver es

$$\text{Max } x_1(p_1 - c_1) + x_2(p_2 - c_2) + x_3(p_3 - c_3),$$

donde las restricciones en cuanto al uso de la tierra son

$$a_{11}x_1 + a_{12}x_2 \leq H - K, \quad (1.4)$$

$$a_{13}x_3 \leq K, \quad (1.5)$$

y para el uso del trabajo

$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \leq L. \tag{1.6}$$

Aunque no se menciona explícitamente, los valores negativos de las variables de decisión no tienen sentido. Esto restringe el conjunto factible a estar contenido en el primer octante. La restricción (1.4) determina que el conjunto factible está contenido en el semiespacio negativo determinado por el hiperplano que contiene a los puntos  $C, D, F, G$ . La restricción (1.5) acota superiormente el conjunto factible por el hiperplano que contiene los puntos  $A, B, C, D$ . Por último la restricción (1.6) determina un hiperplano que contiene a los puntos  $A, B, E$  tal que su semiespacio negativo contiene al conjunto factible. El conjunto factible está representado por la Figura 1.1.

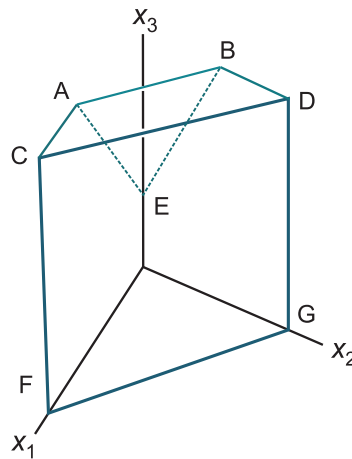


Figura 1.1: Conjunto factible del ejemplo 1.4.

Notemos que  $c = (p_1 - c_1, p_2 - c_2, p_3 - c_3)'$ ,  $c'x$  es una función continua en un compacto por lo cual existen soluciones óptimas para este problema. En los capítulos posteriores veremos que el conjunto solución puede ser descrito en términos de la pendiente del vector  $c$ .

Algunos problemas de programación matemática pueden ser aproximados, mediante técnicas de discretización, por problemas de programación lineal ordinaria, tal como lo muestra el ejemplo siguiente tomado de la sección 1.2 de [1] en el cual se aproxima un problema de control óptimo mediante un problema de programación lineal ordinaria.

**Ejemplo 1.5.** Una compañía desea determinar el nivel de producción de las  $T$  semanas siguientes de forma que se satisfagan las demandas, que son conocidas, y se minimicen los costos de producción y almacenamiento. La demanda al tiempo  $t$  se denota con  $g(t)$ . El nivel de producción e inventario al tiempo  $t$  se denotan con  $x(t)$  y  $y(t)$ , respectivamente. Se supone

que el nivel de inventario en el tiempo inicial 0 es  $y_0$  y se desea tener al final del horizonte de planificación el nivel  $y_T$ . Suponiendo que el costo de almacenamiento es proporcional al nivel de inventario, entonces se puede expresar como  $c_1 \int_0^T y(t)dt$ , donde la constante  $c_1 > 0$  se supone conocida. De la misma forma se supone que el costo de producción es proporcional al nivel de producción, y por lo tanto está dada por  $c_2 \int_0^T x(t)dt$ . Entonces el costo total es  $\int_0^T [c_1 y(t) + c_2 x(t)]dt$ . El nivel de inventario al tiempo  $t$  está dado por

$$y(t) = y_0 + \int_0^t [x(\tau) - g(\tau)]d\tau, \quad t \in [0, T].$$

Se supone que no se permiten atrasos, es decir, que la demanda debe ser satisfecha totalmente. Además la capacidad de manufactura restringe el nivel de producción a no exceder  $b_1$  en cualquier tiempo y la capacidad de almacenamiento permite un nivel de almacenamiento máximo  $b_2$ . Este problema puede ser formulado como sigue:

$$\begin{aligned} & \text{Minimizar} && \int_0^T [c_1 y(t) + c_2 x(t)]dt \\ & \text{sujeto a} && y(t) = y_0 + \int_0^t [x(\tau) - g(\tau)]d\tau, \quad t \in [0, T] \\ & && y(T) = y_T \\ & && 0 \leq x(t) \leq b_1, \quad t \in [0, T] \\ & && 0 \leq y(t) \leq b_2, \quad t \in [0, T] \end{aligned}$$

Este problema es un problema de control óptimo, donde la variable de control es el nivel de producción  $x(t)$  y la variable de estado es el nivel de inventario  $y(t)$ . El problema puede ser aproximado por un problema lineal discretizando las variables  $x(t)$  y  $y(t)$ . Primero se divide el horizonte de planificación  $[0, T]$  en  $n$  períodos  $[0, \Delta], [\Delta, 2\Delta], \dots, [(n-1)\Delta, n\Delta]$  donde  $T = n\Delta$ . El nivel de producción, el de inventarios y la demanda se asumen constantes en cada período  $j$  denotándolos con  $x_j, y_j$ , y  $g_j$ , respectivamente. Entonces el problema de control óptimo anterior puede ser aproximado por el problema de programación lineal ordinaria siguiente:

$$\begin{aligned} & \text{Minimizar} && \sum_{j=1}^n (c_1 \Delta) y_j + \sum_{j=1}^n (c_2 \Delta) x_j \\ & \text{sujeto a} && y_j = y_{j-1} + (x_j - g_j) \Delta, \quad j = 1, \dots, n \\ & && y_n = y_T \\ & && 0 \leq x_j \leq b_1, \quad j = 1, \dots, n \\ & && 0 \leq y_j \leq b_2, \quad j = 1, \dots, n. \end{aligned}$$

El ejemplo que sigue muestra la importancia de la programación lineal ordinaria como herramienta en la resolución de problemas lineales semi-infinitos.

**Ejemplo 1.6.** El método iterativo de discretización por mallas permite resolver problemas de programación lineal semi-infinita discretizando el conjunto de índices  $T$  y resolviendo, en cada

iteración, un problema lineal finito aproximado (ver Capítulo 11 de [8]). Este método consiste en construir una sucesión  $\{T_r\}_{r \in \mathbb{N}}$  de subconjuntos finitos de  $T$ , tales que  $\bigcap_{r=0}^{\infty} F_r = F$ , aquí  $F_r$  denota al conjunto solución del sistema de desigualdades  $\{\mathbf{a}'_t \mathbf{x} \geq b_t, t \in T_r\}$ . Dada la constante  $\epsilon > 0$  se resuelve sucesivamente el problema aproximado que consiste en optimizar el objetivo sobre el conjunto factible relajado  $F_r$  hasta que se obtiene una solución  $\bar{\mathbf{x}}$  la cual es  $\epsilon$ -factible para el problema original, es decir, tal que para todo  $t \in T$  satisfaga  $\mathbf{a}'_t \bar{\mathbf{x}} \geq b_t - \epsilon$ .

Una clase de problemas para los cuales este método de discretización es fácilmente aplicable está formada por los problemas para los cuales  $T$  es numerable, los cuales tienen la forma siguiente:

$$\begin{aligned} & \text{Minimizar} && \mathbf{c}'\mathbf{x} \\ & \text{sujeto a} && \mathbf{a}'_t \mathbf{x} \geq b_t, \quad t \in T, \end{aligned}$$

donde  $T = \{t_1, t_2, \dots\}$ .

En este caso la sucesión  $\{T_r\}$  con término  $r$ -ésimo  $T_r = \{t_1, t_2, \dots, t_r\}$  satisface  $\bigcap_{r=1}^{\infty} F_r = F$ . El algoritmo consiste en aplicar iterativamente los pasos siguientes

*Paso 1: Resolver el subproblema*

$$\begin{aligned} (P_r) \quad & \text{Minimizar} && \mathbf{c}'\mathbf{x} \\ & \text{sujeto a} && \mathbf{a}'_t \mathbf{x} \geq b_t, \quad t \in T_r \end{aligned}$$

Si  $(P_r)$  es inconsistente entonces parar, pues  $(P)$  es inconsistente.

De otra forma, calcular una solución óptima de  $(P_r)$ ,  $\mathbf{x}^r$ , e ir al Paso 2.

*Paso 2: Calcular  $s_r = \inf_{t \in T} \{\mathbf{a}'_t \mathbf{x}^r - b_t\}$ .*

Si  $s_r \geq -\epsilon$  entonces parar. La solución obtenida es  $\mathbf{x}^r$ .

Si  $s_r < -\epsilon$  entonces reemplazar  $r$  por  $r + 1$  y volver al Paso 1.

Para asegurar que el subproblema  $(P_r)$  a resolver en el Paso 1 tenga solución óptima se pide que la función objetivo  $\mathbf{c}'\mathbf{x}$  tenga conjuntos de nivel no vacíos y acotados sobre el conjunto  $F_r$ , lo cual sucede, por ejemplo, si  $F^*$  es no vacío y acotado.





## Capítulo 2

# Análisis convexo

El análisis convexo ha sido una herramienta fundamental en el desarrollo de la teoría de la programación matemática. Se dice que un problema de optimización es de programación convexa si la función objetivo es convexa y la solución óptima se busca en un conjunto convexo. En la programación convexa se han usado las propiedades geométricas de las funciones convexas y los conjuntos convexos para establecer condiciones de optimalidad y unicidad de las soluciones. Estos resultados se satisfacen particularmente en programación lineal, como caso particular de la programación convexa.

A continuación presentamos las herramientas necesarias para sustentar los teoremas de optimalidad y unicidad. La mayor parte de esta teoría está tomada de los primeros capítulos de [11]. En la sección 2.1 presentamos la definición de conjunto convexo y algunos ejemplos. En la sección 2.2 estudiamos dos de sus principales propiedades, la de separación y soporte de conjuntos convexos. En la sección 2.3 presentamos algunas propiedades del conjunto factible de un problema de programación lineal. Finalmente, en la sección 2.4, mencionamos las funciones convexas y algunos criterios para evaluar la convexidad de una función.

### 2.1. Conjuntos convexos

Geoméricamente, se dice que un conjunto es convexo si el segmento que une cualesquiera dos de sus puntos está contenido totalmente en el conjunto. La definición formal es como sigue:

**Definición 2.1.** Sea  $X \subseteq \mathbb{R}^n$ . Se dice que  $X$  es *convexo* si para cualesquiera dos puntos  $x_1, x_2 \in X$  se tiene  $\lambda x_1 + (1 - \lambda)x_2 \in X$ , para cada  $\lambda \in (0, 1)$ , ver la Figura 2.1.

Trivialmente  $\emptyset$  es convexo por vacuidad y  $\mathbb{R}^n$  es convexo por ser un espacio vectorial y por lo tanto cerrado bajo la suma y producto por escalar.

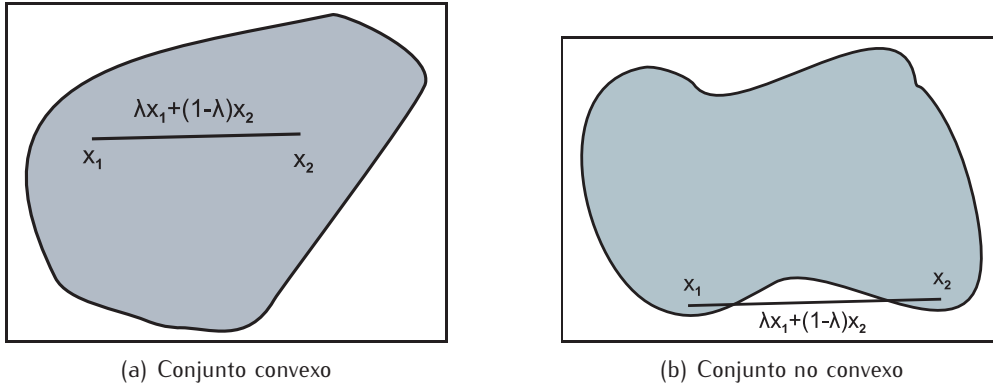


Figura 2.1: Ejemplos de un conjunto convexo y un conjunto no convexo.

**Definición 2.2.** Sean  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k \in \mathbb{R}^n$  y  $\lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{R}$ . Una expresión de la forma

$$\sum_{i=1}^k \lambda_i x_i \quad (2.1)$$

es llamada *combinación lineal* de  $x_1, \dots, x_k$ . Si  $\lambda_i \geq 0$ , se dice que (2.1) es una *combinación cónica*. Si  $\sum_{i=1}^k \lambda_i = 1$  entonces (2.1) es una *combinación afín*. Si  $\sum_{i=1}^k \lambda_i = 1$  y  $\lambda_i \geq 0$ , entonces la expresión (2.1) es una *combinación convexa* de  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ .

**Definición 2.3.**  $V \subseteq \mathbb{R}^n$  es un *conjunto afín* o *variedad afín* si para todo  $\mathbf{x}, \mathbf{y} \in V$

$$\mathbf{x} + \lambda(\mathbf{y} - \mathbf{x}) \in V, \quad \text{para cada } \lambda \in \mathbb{R}.$$

De la definición anterior, es claro que cualquier variedad afín es un conjunto convexo puesto que una combinación convexa es también combinación afín, de acuerdo con la Definición 2.2.

Si  $V$  es un conjunto afín, entonces existe un único subespacio de  $\mathbb{R}^n$ ,  $S$ , tal que  $V = \mathbf{a} + S$  para cierto  $\mathbf{a} \in \mathbb{R}^n$  fijo (Teorema 1.2 de [11]). Se define la dimensión de  $V$  como  $\dim(V) := \dim(S)$ .

**Definición 2.4.** Un *hiperplano* es una variedad afín en  $\mathbb{R}^n$  de dimensión  $n - 1$ .

De acuerdo con el Teorema 1.3 de [11],  $H$  es un hiperplano si y sólo si existen  $\mathbf{p} \in \mathbb{R}^n$  y  $\alpha \in \mathbb{R}$  tales que  $H = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{p}'\mathbf{x} = \alpha\}$ . Dada esta caracterización es fácil comprobar que el hiperplano  $H$  es convexo. El hiperplano  $H$  determina dos subconjuntos cerrados de  $\mathbb{R}^n$  llamados *semiespacios*, dados por  $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{p}'\mathbf{x} \geq \alpha\}$  y  $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{p}'\mathbf{x} \leq \alpha\}$  los cuales son conjuntos convexos. Similarmente los semiespacios abiertos  $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{p}'\mathbf{x} > \alpha\}$  y  $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{p}'\mathbf{x} < \alpha\}$  son conjuntos convexos también.

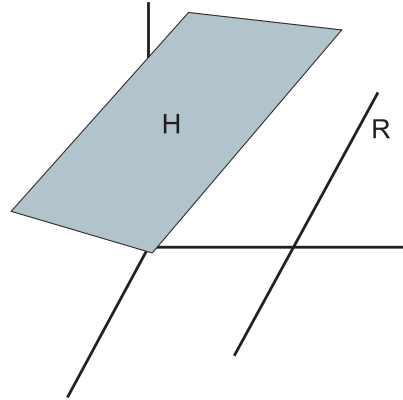


Figura 2.2: Variedades lineales.

En la Figura 2.2 están representados un plano  $H$  y una recta  $R$  como subconjuntos de  $\mathbb{R}^3$ , ambos son variedades afines. Sin embargo  $H$  es un hiperplano y  $R$  no lo es.

Los conjuntos convexos así como las variedades afines tienen varias propiedades que facilitan su estudio. Aunque un conjunto no esté en alguna de las clases anteriores siempre es posible considerar el mínimo conjunto convexo, lineal o afín que los contiene.

**Definición 2.5.** Sea  $X \subset \mathbb{R}^n$ . La *envoltura convexa (afín, lineal)* de  $X$ , denotada por  $convX$  ( $affX$ ,  $spanX$ ) es el mínimo conjunto convexo (variedad afín, subespacio vectorial) en  $\mathbb{R}^n$  que contiene a  $X$ . Es decir,  $convX$  ( $affX$ ,  $spanX$ ) es la intersección de todos los conjuntos convexos (variedades afines, espacios vectoriales) en  $\mathbb{R}^n$  que contienen a  $X$ .

Dado un conjunto  $X \subseteq \mathbb{R}^n$  no vacío se define su *dimensión* como  $dimX = dim affX$ .

De acuerdo con el Teorema 2.3 de [11],  $convX$  es el conjunto de todas las combinaciones convexas de elementos de  $X$ . Similarmente,  $affX$  ( $spanX$ ) es el conjunto de todas las combinaciones afines (combinaciones lineales) de elementos de  $X$ .

**Definición 2.6.** Sea  $X \subset \mathbb{R}^n$  un conjunto convexo. Se define el *interior relativo* de  $X$  como el conjunto

$$riX = \{x \in affX : \exists \epsilon > 0, (x + \epsilon B) \cap (affX) \subseteq X\}.$$

**Ejemplo 2.7.** Consideremos el segmento

$$X = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : 1 \leq x_1 \leq 5, x_2 = 4 \right\},$$

el cual está representado en la Figura 2.3. El interior de este conjunto es  $\emptyset$  pero el interior

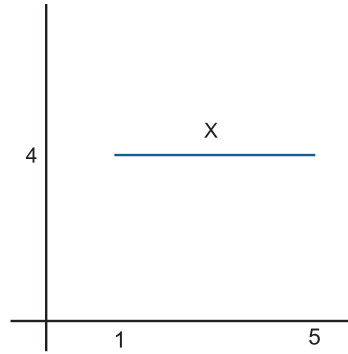


Figura 2.3: El interior de  $X$  es vacío pero el interior relativo de  $X$  es no vacío.

relativo de  $X$  es

$$ri X = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : 1 < x_1 < 5, x_2 = 4 \right\}.$$

El interior relativo de un conjunto convexo  $X$  coincide con el interior topológico usual cuando  $aff X = \mathbb{R}^n$ . Las combinaciones convexas entre puntos en la frontera y el interior relativo de un conjunto convexo pertenecen al interior relativo del conjunto convexo, esta propiedad es conocida como *lema de Accesibilidad*.

**Teorema 2.8** (Teorema 6.1 de [11]). *Sea  $X \subseteq \mathbb{R}^n$  un conjunto convexo. Si  $x_1 \in ri X$  y  $x_2 \in cl X$  entonces  $\lambda x_1 + (1 - \lambda)x_2 \in int X$  para cada  $0 \leq \lambda < 1$ .*

Como consecuencia del Teorema 2.8 tenemos que el interior de un conjunto convexo arbitrario y la cerradura de un conjunto convexo con interior no vacío son conjuntos convexos.

Tal como lo enuncia el Teorema 2.1 de [11], la intersección de una colección arbitraria de conjuntos convexos es también un conjunto convexo.

**Definición 2.9.** Un conjunto  $C \subseteq \mathbb{R}^n$  es un *cono* si para todo  $x \in C$  y  $\lambda > 0$  se cumple  $\lambda x \in C$ . Sea  $X \subset \mathbb{R}^n$ , el *cono convexo generado por  $X$*  está definido por

$$cone X = conv\{\lambda x : x \in X, \lambda \geq 0\}.$$

El *cono dual* de  $X$  está dado por

$$X^* = \{p \in \mathbb{R}^n : x'p \leq 0 \forall x \in X\}.$$

El *cono polar* se define como

$$X^o = \{p \in \mathbb{R}^n : x'p \geq 0 \forall x \in X\}.$$

Geoméricamente, el conjunto  $\text{cone}X$  es la envoltura convexa del cono que consta de todos los rayos que contienen elementos de  $X$ ,  $X^*$  consta de todos los vectores en  $\mathbb{R}^n$  que forman un ángulo mayor o igual que 90 grados con todos los elementos de  $X$  y  $X^o$  está formado por los vectores que forman un ángulo menor o igual que 90 grados con todos los elementos de  $X$ . La Figura 2.4 ilustra estos conjuntos.

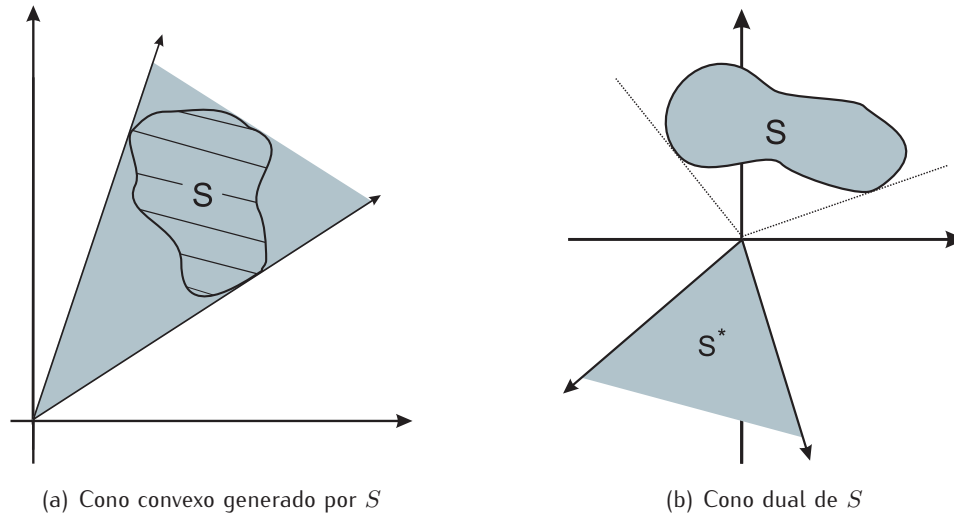


Figura 2.4: Conos convexo y dual.

**Definición 2.10.** Se dice que el conjunto  $X \subseteq \mathbb{R}^n$  es un *poliedro* si es el conjunto solución de un sistema finito de desigualdades lineales.

Como poliedros especiales tenemos a los polítopos. Un *polítopo* se define como la envoltura convexa de un conjunto finito de puntos. El Teorema siguiente enuncia que un polítopo, así como un cono convexo generado por un conjunto finito de puntos, son poliedros.

**Teorema 2.11** (Teorema 3.5 de [7]). *Sea  $X \subseteq \mathbb{R}^n$  un conjunto finito no vacío, entonces*

1. *El polítopo  $\text{conv}X$  es un poliedro acotado.*
2. *El cono finitamente generado  $\text{cone}X$  es el conjunto solución de un sistema de desigualdades homogéneo, es decir, es un cono poliédrico.*

Como consecuencia inmediata de este teorema tenemos que un cono convexo finitamente generado es cerrado.

**Definición 2.12.** Sea  $X \subseteq \mathbb{R}^n$  y  $\bar{x} \in X$ . Se dice que  $\mathbf{d} \in \mathbb{R}^n$  es una *dirección factible* de  $X$  en  $\bar{x}$  si existe  $\epsilon > 0$  tal que  $\bar{x} + \lambda \mathbf{d} \in X$ , para cada  $0 \leq \lambda < \epsilon$ . El conjunto de direcciones factibles del conjunto  $X$  en el punto  $\bar{x}$  se denota por  $D(X, \bar{x})$ .

**Proposición 2.13.** Si  $X \subseteq \mathbb{R}^n$  es un conjunto convexo entonces para todo  $\bar{x} \in X$ ,  $D(X, \bar{x})$  es un cono convexo.

*Demostración.* Veamos que  $D(X, \bar{x})$  es un cono. Sea  $\alpha > 0$  arbitrario, si  $\mathbf{d} \in D(X, \bar{x})$  entonces existe  $\epsilon > 0$  tal que

$$\bar{x} + \delta \mathbf{d} \in X, \text{ para cada } 0 \leq \delta < \epsilon,$$

de lo anterior se sigue que

$$\bar{x} + \frac{\delta}{\alpha}(\alpha \mathbf{d}) \in X, \text{ para cada } 0 \leq \delta < \epsilon,$$

es decir,  $\alpha \mathbf{d} \in D(X, \bar{x})$ . Por lo tanto  $D(X, \bar{x})$  es un cono.

Veamos ahora que  $D(X, \bar{x})$  es un conjunto convexo. Sean  $\mathbf{d}_1, \mathbf{d}_2 \in D(X, \bar{x})$  y  $\lambda \in (0, 1)$  arbitrarios. Como  $\mathbf{d}_1 \in D(X, \bar{x})$  entonces existe  $\epsilon_1 > 0$  tal que

$$\bar{x} + \delta \mathbf{d}_1 \in X, \text{ para cada } 0 \leq \delta < \epsilon_1.$$

Como  $\mathbf{d}_2 \in D(X, \bar{x})$  entonces existe  $\epsilon_2 > 0$  tal que

$$\bar{x} + \delta \mathbf{d}_2 \in X, \text{ para cada } 0 \leq \delta < \epsilon_2.$$

Sea  $\epsilon = \min\{\epsilon_1, \epsilon_2\}$ . Como  $X$  es convexo entonces

$$\lambda(\bar{x} + \delta \mathbf{d}_1) + (1 - \lambda)(\bar{x} + \delta \mathbf{d}_2) \in X, \text{ para cada } 0 \leq \delta < \epsilon.$$

Por otra parte

$$\lambda(\bar{x} + \delta \mathbf{d}_1) + (1 - \lambda)(\bar{x} + \delta \mathbf{d}_2) = \bar{x} + \delta(\lambda \mathbf{d}_1 + (1 - \lambda) \mathbf{d}_2) \in X, \text{ para cada } 0 \leq \delta < \epsilon,$$

se sigue que

$$\lambda \mathbf{d}_1 + (1 - \lambda) \mathbf{d}_2 \in D(X, \bar{x}).$$

Por lo tanto  $D(X, \bar{x})$  es convexo. □

Si un conjunto convexo es no acotado, entonces existe al menos una dirección en la que el conjunto es no acotado. Tales direcciones son llamadas de *recesión*, cuya definición es la siguiente:

**Definición 2.14.** Un vector  $\mathbf{d} \in \mathbb{R}^n$  es llamado *dirección de recesión* de  $X$  si  $\forall \mathbf{x} \in X$ ,  $\mathbf{x} + \lambda \mathbf{d} \in X$  para cada  $\lambda \geq 0$ . El conjunto de direcciones de recesión de  $X$  está denotado y definido por

$$\text{rec}X = \{\mathbf{d} \in \mathbb{R}^n : \mathbf{d} \text{ es dirección de recesión de } X\}.$$

De forma análoga a la demostración de la Proposición 2.13 se verifica que  $\text{rec}X$  es un cono convexo, el cual es llamado *cono de recesión* de  $X$ .

Enseguida tenemos la definición de punto extremo y dirección extrema de un conjunto convexo.

**Definición 2.15.** Si  $X \subseteq \mathbb{R}^n$  es convexo, entonces un punto  $\mathbf{x} \in X$  es un *punto extremo* de  $X$  si  $\mathbf{x} = \lambda \mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2$  con  $\mathbf{x}_1, \mathbf{x}_2 \in X$  y  $\lambda \in (0, 1)$  implica  $\mathbf{x}_1 = \mathbf{x}_2$ . El *conjunto de puntos extremos* de  $X$  será denotado por  $\text{ext}X$ . De forma similar,  $\mathbf{d} \in \text{rec}X$  es una *dirección extrema* de  $X$  si  $\mathbf{d} = \lambda \mathbf{d}_1 + (1 - \lambda)\mathbf{d}_2$  con  $\mathbf{d}_1$  y  $\mathbf{d}_2$  direcciones de recesión de  $X$  y cierto  $\lambda \in (0, 1)$  implica  $\mathbf{d}_1 = \mathbf{d}_2$ .

**Ejemplo 2.16.** Consideremos el conjunto siguiente

$$X = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : x_1 \geq 0, x_2 \geq 0, 2x_1 + 5 \geq x_2, \frac{1}{2}x_1 - 2 \leq x_2 \right\}$$

el cual está representado en la Figura 2.5(a). Los puntos extremos de  $F$  son

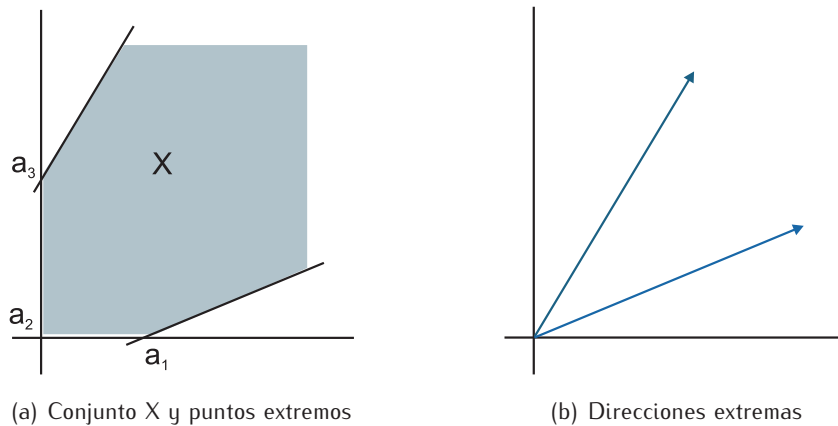


Figura 2.5: Poliedro con puntos extremos y direcciones extremas.

$$\mathbf{a}_1 = \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \mathbf{a}_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \mathbf{a}_3 = \begin{pmatrix} 0 \\ 5 \end{pmatrix},$$

y las direcciones extremas de  $F$  son

$$\mathbf{d}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \mathbf{d}_2 = \begin{pmatrix} \frac{1}{2} \\ 1 \end{pmatrix}.$$

Además de la representación como conjunto solución de un sistema de desigualdades, llamada *representación externa*, los poliedros pueden representarse como una combinación convexa de sus puntos extremos y sus direcciones extremas, la cual es llamada *representación interna*. El Teorema de representación de conjuntos convexos (Teorema 19.1 de [11]) asegura que si  $X \subseteq \mathbb{R}^n$  es convexo con puntos extremos  $\mathbf{a}_1, \dots, \mathbf{a}_k$  y direcciones extremas  $\mathbf{d}_1, \dots, \mathbf{d}_l$ , entonces para cualquier  $\bar{\mathbf{x}} \in X$  existen escalares  $\lambda_1, \dots, \lambda_k$  y  $\mu_1, \dots, \mu_l$  tales que

$$\bar{\mathbf{x}} = \sum_{i=1}^k \lambda_i \mathbf{a}_i + \sum_{j=1}^l \mu_j \mathbf{d}_j, \quad (2.2)$$

donde  $\sum_{i=1}^k \lambda_i = 1$ ,  $\lambda_i \geq 0$  para  $i = 1, \dots, k$  y  $\mu_j \geq 0$  para  $j = 1, \dots, l$ .

El Corolario 19.1.1 de [11] indica que un poliedro tiene un número finito de puntos extremos y direcciones extremas. El *Teorema de Weyl* proporciona una caracterización de los puntos extremos de un poliedro en términos de su representación externa:

**Teorema 2.17** (Teorema 5.2 de [7]). *Sea el poliedro  $F = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}'_i \mathbf{x} \geq b_i, i = 1, \dots, m\}$ . Si  $\dim \text{span}\{\mathbf{a}_i : i = 1, 2, \dots, m\} \geq 2$  entonces  $\bar{\mathbf{x}} \in F$  es un punto extremo de  $F$  si y solo si  $\dim \text{span}\{\mathbf{a}_i : \mathbf{a}'_i \bar{\mathbf{x}} = b_i\} = n$ .*

*Demostración.* Sea  $\bar{\mathbf{x}} \in F$  tal que  $\dim \text{span}\{\mathbf{a}_i : \mathbf{a}'_i \bar{\mathbf{x}} = b_i\} < n$ . Como

$$\dim \text{span}\{\mathbf{a}_i : \mathbf{a}'_i \bar{\mathbf{x}} = b_i\} + \dim \text{span}\{\mathbf{a}_i : \mathbf{a}'_i \bar{\mathbf{x}} = b_i\}^\perp = n,$$

donde  $\text{span}\{\mathbf{a}_i : \mathbf{a}'_i \bar{\mathbf{x}} = b_i\}^\perp$  es el complemento ortogonal de  $\text{span}\{\mathbf{a}_i : \mathbf{a}'_i \bar{\mathbf{x}} = b_i\}$ , entonces  $\dim \text{span}\{\mathbf{a}_i : \mathbf{a}'_i \bar{\mathbf{x}} = b_i\}^\perp \geq 1$ . Por lo tanto, existe  $\mathbf{y} \in \mathbb{R}^n$  no nulo tal que

$$\mathbf{a}'_i \mathbf{y} = 0, \text{ para cada } \mathbf{a}_i \text{ tal que } \mathbf{a}'_i \bar{\mathbf{x}} = b_i. \quad (2.3)$$

Veamos que  $\mathbf{y} \in D(F, \bar{\mathbf{x}})$ . De (2.3) tenemos que, para cada  $\mathbf{a}_i$  tal que  $\mathbf{a}'_i \bar{\mathbf{x}} = b_i$ ,

$$\mathbf{a}'_i(\bar{\mathbf{x}} + \epsilon \mathbf{y}) = \mathbf{a}'_i \bar{\mathbf{x}} + \epsilon \mathbf{a}'_i \mathbf{y} = \mathbf{a}'_i \bar{\mathbf{x}} = b_i, \text{ para todo } \epsilon > 0.$$

Por otra parte, si  $\mathbf{a}_i$  es tal que  $\mathbf{a}'_i \bar{\mathbf{x}} > b_i$ , consideremos los dos casos siguientes:

(a)  $\mathbf{a}'_i \mathbf{y} \geq 0$ : En este caso  $\mathbf{a}'_i(\bar{\mathbf{x}} + \epsilon \mathbf{y}) = \mathbf{a}'_i \bar{\mathbf{x}} + \epsilon \mathbf{a}'_i \mathbf{y} = b_i + \epsilon \mathbf{a}'_i \mathbf{y} > b_i$ , para todo  $\epsilon > 0$ .



- (b)  $\mathbf{a}'_i \mathbf{y} < 0$ : En este caso basta definir  $\epsilon_i = \frac{\mathbf{a}'_i \bar{\mathbf{x}} - b_i}{|\mathbf{a}'_i \mathbf{y}|} > 0$  para obtener  $\mathbf{a}'_i(\bar{\mathbf{x}} + \delta \mathbf{y}) > b_i$ , para cada  $0 \leq \delta \leq \epsilon_i$ .

De lo anterior concluimos que  $\mathbf{y} \in D(F, \bar{\mathbf{x}})$ . De forma análoga podemos verificar que  $-\mathbf{y} \in D(F, \bar{\mathbf{x}})$ . Existe  $\delta > 0$  tal que  $\bar{\mathbf{x}} \pm \delta \mathbf{y} \in F$ , entonces

$$\bar{\mathbf{x}} = \frac{1}{2}(\bar{\mathbf{x}} + \delta \mathbf{y}) + \frac{1}{2}(\bar{\mathbf{x}} - \delta \mathbf{y}),$$

de aquí,  $\bar{\mathbf{x}}$  no es un punto extremo de  $F$ .

Supongamos ahora que  $\bar{\mathbf{x}} \in F$  es tal que  $\dim \text{span}\{\mathbf{a}_i : \mathbf{a}'_i \bar{\mathbf{x}} = b_i\} = n$  y que  $\bar{\mathbf{x}}$  no es punto extremo de  $F$ . Entonces existen  $\mathbf{x}_1, \mathbf{x}_2 \in F$  distintos tales que  $\bar{\mathbf{x}} = \lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2$ , para cierto  $\lambda \in (0, 1)$ .

Sea  $\mathbf{y} = \mathbf{x}_2 - \mathbf{x}_1$ . Entonces  $\bar{\mathbf{x}} + \lambda \mathbf{y} = \lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2 + \lambda(\mathbf{x}_2 - \mathbf{x}_1) = \mathbf{x}_2 \in F$  y por lo tanto  $\mathbf{y} \in D(F, \bar{\mathbf{x}})$ . Por otra parte, si  $0 < \epsilon < 1 - \lambda$  entonces

$$\bar{\mathbf{x}} - \epsilon \mathbf{y} = \lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2 + \epsilon \mathbf{x}_1 - \epsilon \mathbf{x}_2 = (\lambda + \epsilon) \mathbf{x}_1 + (1 - (\lambda + \epsilon)) \mathbf{x}_2 \in F,$$

por ser una combinación convexa de  $\mathbf{x}_1$  y  $\mathbf{x}_2$ .

Existe  $\mathbf{a}_k \in \{\mathbf{a}_i : \mathbf{a}'_i \bar{\mathbf{x}} = b_i\}$  tal que  $\mathbf{a}'_k \mathbf{y} \neq b_i$ , puesto que en caso contrario se tendría

$$\mathbf{a}'_i(\bar{\mathbf{x}} - \mathbf{y}) = b_i - b_i = 0, \quad \text{para todo } i = 1, \dots, m,$$

y por lo tanto

$$\text{span}\{\mathbf{a}_i : \mathbf{a}'_i \bar{\mathbf{x}} = b_i\} \subseteq \text{span}\{\bar{\mathbf{x}} - \mathbf{y}\}^\perp,$$

de donde

$$2 \leq \dim \text{span}\{\mathbf{a}_i : \mathbf{a}'_i \bar{\mathbf{x}} = b_i\} \leq \text{span}\{\bar{\mathbf{x}} - \mathbf{y}\}^\perp = 1,$$

lo cual es una contradicción. Consideremos los casos siguientes

- (a)  $\mathbf{a}'_k(\bar{\mathbf{x}} + \lambda \mathbf{y}) > b_k$ : En este caso  $\mathbf{a}'_k \mathbf{y} > \frac{1}{\lambda}(b_k - \mathbf{a}'_k \bar{\mathbf{x}}) = 0$ . Luego

$$\mathbf{a}'_k(\bar{\mathbf{x}} - \epsilon \mathbf{y}) = \mathbf{a}'_k \bar{\mathbf{x}} - \epsilon \mathbf{a}'_k \mathbf{y} = b_k - \epsilon \mathbf{a}'_k \mathbf{y} < b_k,$$

de lo anterior obtenemos la contradicción  $\bar{\mathbf{x}} - \epsilon \mathbf{y} \notin F$ .

- (b)  $\mathbf{a}'_k(\bar{\mathbf{x}} + \lambda \mathbf{y}) < b_k$ : En este caso es evidente la contradicción  $\bar{\mathbf{x}} + \lambda \mathbf{y} \notin F$ .

□

Si  $C \subseteq \mathbb{R}^n$  es un cono convexo, entonces  $C$  solamente puede tener como punto extremo al vector nulo, como está enunciado a continuación.

**Proposición 2.18.** Si  $C \subseteq \mathbb{R}^n$  un cono convexo, entonces  $C$  no tiene puntos extremos no nulos.

*Demostración.* Supongamos que  $x \in C$  es un punto extremo de  $C$  diferente de cero. Entonces para cualesquiera  $x_1, x_2 \in C$  y para todo  $\lambda \in (0, 1)$ , la igualdad  $x = \lambda x_1 + (1 - \lambda)x_2$  implica que  $x_1 = x_2$ . Como  $C$  es un cono, entonces

$$x_1 := \frac{1}{2}x = x - \frac{1}{2}x \in C,$$

y

$$x_2 := \frac{3}{2}x = x + \frac{1}{2}x \in C.$$

Claramente  $x_1 \neq x_2$ . Notemos que  $x = \frac{1}{2}x_1 + (1 - \frac{1}{2})x_2$  lo cual contradice que  $x$  sea punto extremo de  $C$ .  $\square$

Así, cada punto de un cono convexo puede representarse como una combinación cónica de sus direcciones extremas, en otras palabras, un cono convexo coincide con la envoltura convexa del cono generado por sus direcciones extremas, como lo veremos en el Corolario siguiente.

**Corolario 2.19.** Sea  $C$  un cono convexo y  $d_1, \dots, d_r$  las direcciones extremas de  $C$ . Entonces para cada  $x \in C$  existe un conjunto  $\{d_1, \dots, d_k\}$  de direcciones extremas de  $C$  y escalares positivos  $\alpha_i, i = 1, \dots, k$  tales que  $x = \sum_{i=1}^k \alpha_i d_i$ .

*Demostración.* De acuerdo con la representación de conjuntos convexos cada punto de  $C$  puede ser representado como una combinación convexa de puntos extremos de  $C$  más una combinación cónica de direcciones extremas de  $C$ . Según la Proposición 2.18,  $C$  no tiene puntos extremos no nulos. Así para cada  $x \in C$  existen escalares  $\alpha_i \geq 0, i = 1, \dots, r$  tales que

$$x = \sum_{i=1}^r \alpha_i d_i. \quad (2.4)$$

Ignorando los términos tales que  $\alpha_i = 0$  en la expresión 2.4 se obtiene el resultado.  $\square$

El siguiente resultado proporciona una representación para los elementos del interior de un cono convexo. Incluimos la demostración hecha en [4].

**Teorema 2.20** (Teorema 2.5 de [4]). Sea  $\emptyset \neq C \subseteq \mathbb{R}^n$ . Las siguientes afirmaciones son equivalentes:

(i)  $x \in \text{int cone}C$ ;

(ii) existen puntos  $x_i \in C$  y escalares  $\lambda_i > 0, i = 1, 2, \dots, r$ , tales que  $\text{span}\{x_1, \dots, x_r\} = \mathbb{R}^n$  y  $x = \sum_{i=1}^r \lambda_i x_i$ .

*Demostración.* (i)  $\Rightarrow$  (ii). Sea  $\mathbf{x} \in \text{int cone}C$  entonces existen vectores  $\mathbf{z}_i \in \text{cone}C$ ,  $i = 1, 2, \dots, r$  y escalares positivos  $\beta_1, \dots, \beta_r$  tales que

$$\text{span}\{\mathbf{z}_1, \dots, \mathbf{z}_r\} = \mathbb{R}^n \text{ y } \mathbf{x} = \sum_{i=1}^r \beta_i \mathbf{z}_i.$$

De acuerdo con el Corolario 2.19 para cada  $i = 1, \dots, r$  existen escalares  $\gamma_{ij} > 0$  y vectores  $\mathbf{d}_{ij} \in C$ ,  $j = 1, 2, \dots, s_i$ , tales que

$$\mathbf{z}_i = \sum_{j=1}^{s_i} \gamma_{ij} \mathbf{d}_{ij}.$$

Por lo tanto

$$\mathbf{x} = \sum_{i=1}^r \beta_i \sum_{j=1}^{s_i} \gamma_{ij} \mathbf{d}_{ij} = \sum_{i=1}^r \sum_{j=1}^{s_i} \beta_i \gamma_{ij} \mathbf{d}_{ij},$$

renombrando los términos en la expresión anterior obtenemos (ii).

(ii)  $\Rightarrow$  (i). Sea el conjunto  $\{\mathbf{x}_1, \dots, \mathbf{x}_r\}$  tal que  $\text{span}\{\mathbf{x}_1, \dots, \mathbf{x}_r\} = \mathbb{R}^n$ . Consideremos el mapeo  $\varphi : \mathbb{R}^r \rightarrow \mathbb{R}^n$  dado por  $\varphi(\alpha_1, \dots, \alpha_r) = \sum_{i=1}^r \alpha_i \mathbf{x}_i$ , el cual es lineal y continuo. En particular,  $\{(\alpha_1, \dots, \alpha_r) : \alpha_i > 0, i = 1, \dots, r\}$  es un abierto en  $\mathbb{R}^r$  por lo que su imagen  $\{\sum_{i=1}^r \alpha_i \mathbf{x}_i \in \mathbb{R}^n : \alpha_i > 0, i = 1, \dots, r\} \subset \text{cone}C$  es un abierto en  $\mathbb{R}^n$ , de lo cual concluimos que  $\mathbf{x} \in \text{int cone}C$ .  $\square$

## 2.2. Separación y soporte de conjuntos convexos

En esta sección presentamos las propiedades de separación de conjuntos convexos y de soporte en un punto en la frontera de un conjunto convexo.

**Definición 2.21.** Sean  $X_1$  y  $X_2$  subconjuntos de  $\mathbb{R}^n$  no vacíos. Se dice que el hiperplano  $H = \{\mathbf{x} : \mathbf{p}'\mathbf{x} = \alpha\}$  separa  $X_1$  y  $X_2$  si  $X_1$  está contenido en uno de los semiespacios definidos por  $H$  mientras que  $X_2$  está contenido en el otro semiespacio, es decir, si se tiene alguna de las siguientes condiciones:

- a)  $\mathbf{p}'\mathbf{x} \geq \alpha$  para todo  $\mathbf{x} \in X_1$  y  $\mathbf{p}'\mathbf{x} \leq \alpha$  para todo  $\mathbf{x} \in X_2$ ,
- b)  $\mathbf{p}'\mathbf{x} \leq \alpha$  para todo  $\mathbf{x} \in X_1$  y  $\mathbf{p}'\mathbf{x} \geq \alpha$  para todo  $\mathbf{x} \in X_2$ .

Se dice que la separación es *estricta* si las desigualdades anteriores son estrictas, la separación es *propia* si  $X_1$  y  $X_2$  no están ambos contenidos en  $H$ , la separación es fuerte si existe  $\epsilon > 0$  tal que  $X_1 + \epsilon B$  está contenido en uno de los semiespacios mientras que  $X_2 + \epsilon B$  está contenido en el otro.

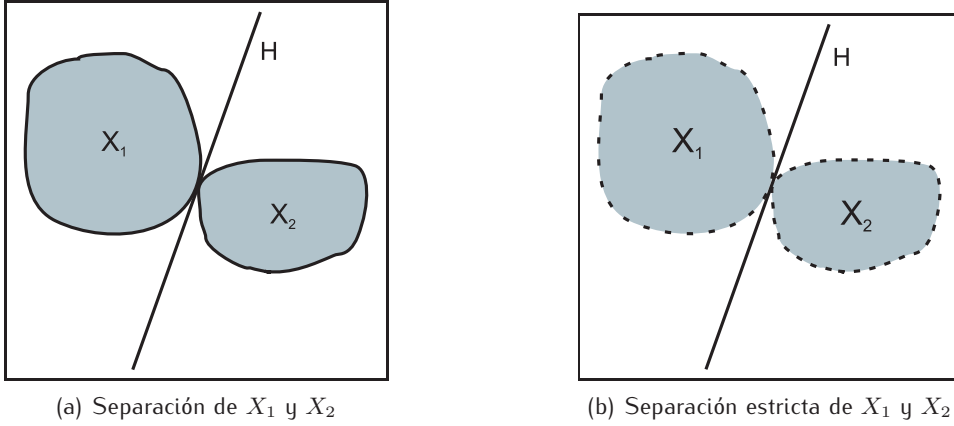


Figura 2.6: Separación de conjuntos convexos.

En [11] se puede consultar una exposición detallada de condiciones que garantizan la separación de conjuntos convexos, así como demostraciones constructivas de algunos de estos resultados. En este trabajo sólo mencionamos algunos resultados a los que haremos referencia posteriormente.

**Teorema 2.22** (Teorema 11.3 de [11]). *Sean  $X_1$  y  $X_2$  dos subconjuntos convexos no vacíos de  $\mathbb{R}^n$ . Para que exista un hiperplano que separa  $X_1$  y  $X_2$  propiamente es necesario y suficiente que  $riX_1$  no se intersecte con  $riX_2$ .*

El lema de Farkas para conos, enunciado a continuación, establece que si  $C$  es un cono convexo cerrado, entonces  $C^{oo}$  es igual a  $C$ .

**Lema 2.23** (Lema 3.1 de [7]). *Si  $C$  es un cono convexo cerrado entonces  $C^{oo} = C$ .*

*Demostración.* Sea  $\bar{x} \in C^{oo}$ , es decir,  $\bar{x}'\mathbf{p} \geq 0$  para cada  $\mathbf{p} \in C^0$ . Supongamos que  $\bar{x} \notin C$ . Entonces, por el Teorema 2.22, existen  $\mathbf{p} \in \mathbb{R}^n$  no nulo y  $\alpha \in \mathbb{R}$  tales que  $\mathbf{p}'\mathbf{x} \leq \alpha$  para todo  $\mathbf{x} \in C$  y  $\mathbf{p}'\bar{x} > \alpha$ .

Como  $\mathbf{0} \in C$ , entonces

$$\mathbf{p}'\bar{x} > \alpha \geq \mathbf{p}'\mathbf{0} = 0. \quad (2.5)$$

Por otra parte, para cualesquiera  $\mathbf{x} \in C$ , tenemos que  $\lambda\mathbf{x} \in C$ , para todo  $\lambda > 0$ . Luego  $\mathbf{p}'(\lambda\mathbf{x}) \leq \alpha$ , de donde  $\mathbf{p}'\mathbf{x} \leq \frac{\alpha}{\lambda}$ . Tomando el límite cuando  $\lambda$  tiende a infinito,

$$\mathbf{p}'\mathbf{x} \leq 0, \text{ para todo } \mathbf{x} \in C. \quad (2.6)$$

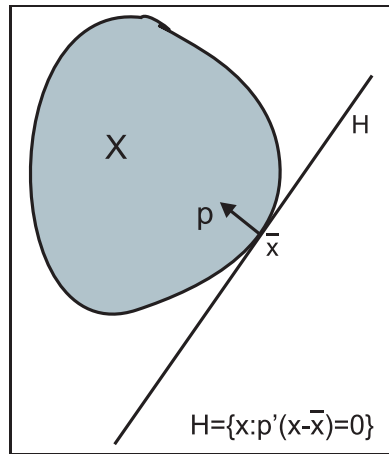
La desigualdad (2.6) implica que  $\mathbf{p} \in C^0$ , notemos que por la desigualdad (2.5) tenemos que  $\mathbf{p}'\bar{x} > 0$ , lo cual contradice que  $\bar{x} \in C^{oo}$ . Por lo tanto,  $\bar{x} \in C$ . Así  $C^{oo} \subseteq C$ .

Sea  $\bar{x} \in C$ , entonces para todo  $z \in C^0$  tenemos  $z'\bar{x} \leq 0$ , lo cual implica  $\bar{x} \in C^{00}$ . Por lo tanto  $C \subseteq C^{00}$ .

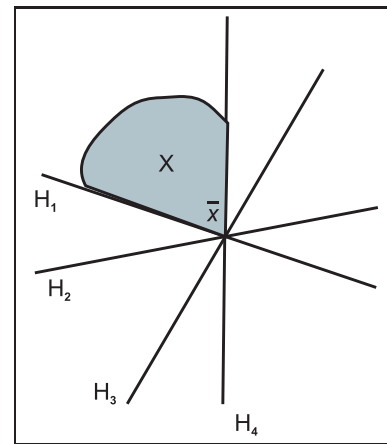
□

La separación de un conjunto convexo y un punto en la frontera del mismo se denomina soporte, el cual es una generalización de la definición usual de tangencia.

**Definición 2.24.** Sea  $X \subseteq \mathbb{R}^n$  un conjunto convexo. Se dice que el hiperplano  $H = \{x \in \mathbb{R}^n : p'x = \alpha\}$  es un *hiperplano de soporte* para el conjunto  $X$  en el punto  $\bar{x} \in X$  si  $\bar{x} \in H$  y  $X$  está contenido en alguno de los semiespacios cerrados generados por  $H$ .



(a) Existe un único hiperplano de soporte de  $X$  en el punto  $\bar{x}$



(b) Existe una infinidad de hiperplanos de soporte de  $X$  en  $\bar{x}$

Figura 2.7: Hiperplanos de soporte en la frontera de un conjunto convexo.

Podemos notar en la Figura 2.7(a) que la existencia de un hiperplano de soporte de  $X$  en  $\bar{x}$  implica  $\bar{x} \in \partial X$ . A continuación veremos la argumentación de este hecho dada en [7]. Supongamos que el hiperplano  $H$  soporta al conjunto  $X$  en el punto  $\bar{x} \in X$ . Tenemos que para todo  $x \in X$ ,  $p'x \leq p'\bar{x}$ . Sea  $y_k := \bar{x} - \frac{p}{k}$ , entonces

$$p'y_k = p' \left( \bar{x} - \frac{p}{k} \right) = p'\bar{x} - \frac{\|p\|^2}{k} < p'\bar{x},$$

lo anterior implica que  $y_k \notin X$ , luego  $\{y_k\} \in \mathbb{R}^n \setminus X$  es una sucesión que converge a  $\bar{x}$ . De lo anterior se deduce que  $\bar{x} \in \partial(\mathbb{R}^n \setminus X)$ . Como  $\bar{x} \in X$ , se sigue que  $\bar{x} \in \partial X$ .

**Teorema 2.25** (Teorema 11.6 de [11]). *Sea  $X$  un conjunto convexo en  $\mathbb{R}^n$  y  $C$  un subconjunto convexo no vacío de  $X$ . Para que exista un hiperplano de soporte para  $X$  que contiene a  $C$  tal que no contiene a  $X$  es necesario y suficiente que  $C$  sea disjunto de  $riX$ .*

El teorema anterior garantiza que para cualquier punto  $\bar{x}$  en la frontera de un conjunto convexo cerrado  $X$ , tal que  $\bar{x} \notin riX$  existe un hiperplano que soporta al conjunto  $X$ , no necesariamente único (ver la Figura 2.7(b)).

### 2.3. Conjunto factible de un problema de programación lineal

Dado un problema lineal, la primera cuestión que desearíamos aclarar es si el problema es consistente, es decir, si el conjunto solución de su sistema de desigualdades es no vacío; y en caso de ser consistente, podríamos preguntarnos si en las desigualdades involucradas existe redundancia, es decir, si existen desigualdades que pueden ser eliminadas sin modificar el conjunto solución. Como podemos observar, el estudio de la consistencia de un problema, es en realidad el estudio de la consistencia del sistema de desigualdades asociado al problema.

En el Teorema 2.27 de esta sección mencionamos una caracterización de los sistemas de desigualdades que son consistentes y en el Teorema 2.28 veremos una caracterización de las desigualdades que resultan redundantes en un sistema consistente. Posteriormente mencionamos algunas propiedades de los conjuntos poliédricos.

Dado un sistema de desigualdades lineales de la forma  $\{\mathbf{a}_i' \mathbf{x} \geq b_i, i \in I\}$ , con conjunto solución  $F$ , se define su *cono característico* como sigue:

$$K = \text{cone} \left\{ \begin{pmatrix} \mathbf{a}_i \\ b_i \end{pmatrix}, i \in I; \begin{pmatrix} \mathbf{0} \\ -1 \end{pmatrix} \right\}.$$

Entonces el cono polar de  $K$  está dado por

$$K^0 = \left\{ \begin{pmatrix} \mathbf{x} \\ x_{n+1} \end{pmatrix} \in \mathbb{R}^{n+1} : \mathbf{a}_i' \mathbf{x} + b_i x_{n+1} \geq 0, -x_{n+1} \geq 0 \right\}.$$

**Lema 2.26** (Lema 3.2 de [7]). *Consideremos el sistema de desigualdades  $\{\mathbf{a}_i' \mathbf{x} \geq b_i, i \in I\}$  y sean  $F$  su conjunto factible y  $K$  su cono característico.*

(i)  $F = \emptyset$  si y sólo si  $x_{n+1} = 0$ , para todo  $\begin{pmatrix} \mathbf{x} \\ x_{n+1} \end{pmatrix} \in K^0$ .

(ii)  $\bar{x} \in F$  si y sólo si  $\begin{pmatrix} \bar{x} \\ -1 \end{pmatrix} \in K^0$ .

*Demostración.* (i) Sea  $\begin{pmatrix} \bar{x} \\ \bar{x}_{n+1} \end{pmatrix} \in K^0$  tal que  $\bar{x}_{n+1} \neq 0$  entonces  $\bar{x}_{n+1} < 0$ . Dividiendo por  $\bar{x}_{n+1}$  ambos lados de la desigualdad  $\mathbf{a}'_i \bar{x} + b_i \bar{x}_{n+1} \geq 0$ ,  $i \in I$ , concluimos que  $-\frac{\bar{x}}{\bar{x}_{n+1}} \in F$ . Por lo tanto  $F \neq \emptyset$ .

Recíprocamente, si existe  $\bar{x} \in F$  entonces  $\mathbf{a}'_i \bar{x} + b_i(-1) \geq 0$ , para todo  $i \in I$ . Por lo tanto,  $\begin{pmatrix} \bar{x} \\ -1 \end{pmatrix} \in K^0$  y  $x_{n+1} = -1 \neq 0$ .

(ii)  $\mathbf{a}'_i \bar{x} \geq b_i$ , para cada  $i \in I$  si y sólo si  $\begin{pmatrix} \bar{x} \\ -1 \end{pmatrix}' \begin{pmatrix} \mathbf{a}_i \\ b_i \end{pmatrix} \geq 0$ , para cada  $i \in I$ , si y solo si  $\begin{pmatrix} \bar{x} \\ -1 \end{pmatrix} \in K^0$ .

□

El siguiente resultado conocido como *Teorema de existencia de Zhu*, caracteriza a los sistemas de desigualdades lineales consistentes. Los dos resultados siguientes son válidos en general aún cuando el conjunto de índices  $T$  sea infinito.

**Teorema 2.27** (Teorema 3.7 de [7]). *El sistema de desigualdades  $\{\mathbf{a}'_t x \geq b_t, t \in T\}$  es consistente si y sólo si  $\begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} \notin K$ .*

*Demostración.* Supongamos que el sistema es consistente y sea  $\bar{x} \in F$ . De acuerdo con el Lema 2.26 (ii),  $\begin{pmatrix} \bar{x} \\ -1 \end{pmatrix} \in K^0$ , de aquí  $\begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} \notin K^{00}$  puesto que  $\begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix}' \begin{pmatrix} \bar{x} \\ -1 \end{pmatrix} = -1 < 0$ .

Segun el Lema 2.23,  $K^{00} = K$ , por lo tanto  $\begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} \notin K$ .

Recíprocamente, si el sistema es inconsistente, entonces, según el Lema 2.26 (i),  $K^0$  está contenido en el hiperplano

$$H = \left\{ \begin{pmatrix} \bar{x} \\ x_{n+1} \end{pmatrix} \in \mathbb{R}^{n+1} : x_{n+1} = 0 \right\}.$$

Por otra parte, si  $\begin{pmatrix} \bar{x} \\ x_{n+1} \end{pmatrix} \in H$  entonces  $\begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix}' \begin{pmatrix} \bar{x} \\ x_{n+1} \end{pmatrix} = 0$ . Luego  $\begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} \in K^{00} = K$ .

□

Se dice que la desigualdad  $\mathbf{a}'\mathbf{x} \geq b$  es consecuencia del sistema  $\{\mathbf{a}'_i\mathbf{x} \geq b_i, i \in I\}$  si cada solución de este sistema satisface también la desigualdad. En otras palabras una desigualdad es consecuencia de cierto sistema si al incluirla en él resulta redundante. Usando el cono característico del sistema se puede dar una caracterización de las desigualdades redundantes:

**Teorema 2.28** (Teorema 3.6 de [7]). *Sea  $\{\mathbf{a}'_i\mathbf{x} \geq b_i, i \in I\}$  un sistema consistente y sea  $K$  su cono característico. Una desigualdad  $\mathbf{a}'\mathbf{x} \geq b$  es consecuencia de dicho sistema si y sólo si  $\begin{pmatrix} \mathbf{a} \\ b \end{pmatrix} \in K$ .*

*Demostración.* Supongamos que  $\mathbf{a}'\mathbf{x} \geq b$  es consecuencia del sistema  $\{\mathbf{a}'_i\mathbf{x} \geq b_i, i \in I\}$ . Puesto que  $K^{00} = K$  entonces basta probar que  $\begin{pmatrix} \mathbf{a} \\ b \end{pmatrix} \in K^{00}$  para demostrar la condición necesaria.

Es decir, se desea probar que  $\begin{pmatrix} \bar{\mathbf{x}} \\ \bar{x}_{n+1} \end{pmatrix} \in K^0$  implica  $\mathbf{a}'\bar{\mathbf{x}} + b\bar{x}_{n+1} \geq 0$ . Como  $\bar{x}_{n+1} \leq 0$ , pueden darse dos casos:

Caso 1:  $\bar{x}_{n+1} < 0$ .

Por ser  $K^0$  un cono, y como  $\gamma = -\frac{1}{\bar{x}_{n+1}} > 0$  se tiene

$$\gamma \begin{pmatrix} \bar{\mathbf{x}} \\ \bar{x}_{n+1} \end{pmatrix} = \begin{pmatrix} \gamma\bar{\mathbf{x}} \\ -1 \end{pmatrix} \in K^0,$$

esto implica que, para todo  $i \in I$ ,  $\mathbf{a}'_i(\gamma\bar{\mathbf{x}}) \geq b_i$ . Como la desigualdad  $\mathbf{a}'\mathbf{x} \geq b$  es consecuencia del sistema  $\{\mathbf{a}'_i\mathbf{x} \geq b_i, i \in I\}$ , entonces  $\mathbf{a}'\mathbf{x} \geq b$  o, de forma equivalente,  $\mathbf{a}'\bar{\mathbf{x}} + b\bar{x}_{n+1} \geq 0$ .

Caso 2:  $\bar{x}_{n+1} = 0$ .

En este caso  $\begin{pmatrix} \bar{\mathbf{x}} \\ 0 \end{pmatrix} \in K^0$ . Tomando  $\bar{\mathbf{z}} \in F$  arbitrario, y como  $\begin{pmatrix} \bar{\mathbf{z}} \\ -1 \end{pmatrix} \in K^0$ , se tiene

$$(1-\lambda) \begin{pmatrix} \bar{\mathbf{x}} \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} \bar{\mathbf{z}} \\ -1 \end{pmatrix} = \begin{pmatrix} (1-\lambda)\bar{\mathbf{x}} + \lambda\bar{\mathbf{z}} \\ -\lambda \end{pmatrix} \in K^0$$

para todo  $\lambda \in (0, 1)$ . Dicho punto está considerado en el caso 1, por lo cual tenemos que  $\mathbf{a}'[(1-\lambda)\bar{\mathbf{x}} + \lambda\bar{\mathbf{z}}] - \lambda b \geq 0$ . Tomando el límite cuando  $\lambda \rightarrow 0$  se obtiene  $\mathbf{a}'\bar{\mathbf{x}} \geq 0$ .

Recíprocamente, si  $\begin{pmatrix} \mathbf{a} \\ b \end{pmatrix} \in K = K^{00}$ , entonces, de acuerdo con el Lema 3.2 (ii) de [7], para todo  $\bar{\mathbf{x}} \in F$  se cumple

$$\begin{pmatrix} \mathbf{a} \\ b \end{pmatrix}' \begin{pmatrix} \bar{\mathbf{x}} \\ -1 \end{pmatrix} \geq 0.$$

lo cual implica  $\mathbf{a}'\bar{\mathbf{x}} \geq b$ , para todo  $\bar{\mathbf{x}} \in F$ . Por lo tanto  $\mathbf{a}'\bar{\mathbf{x}} \geq b$  es consecuencia del sistema  $\{\mathbf{a}'_i\mathbf{x} \geq b_i, i \in I\}$ .  $\square$



Existe una correspondencia natural entre desigualdades lineales y semiespacios cerrados. Consideremos una desigualdad lineal de la forma siguiente:

$$\mathbf{a}'\mathbf{x} \geq b, \quad (2.7)$$

donde  $\mathbf{a} \in \mathbb{R}^n$ ,  $b \in \mathbb{R}$  son fijos y  $\mathbf{x} \in \mathbb{R}^n$ . Si  $\mathbf{a} = \mathbf{0}$  y  $b_i \leq 0$  entonces el conjunto solución de la desigualdad (2.7) es  $\mathbb{R}^n$ , si  $\mathbf{a} = \mathbf{0}$  y  $b_i > 0$  entonces el conjunto solución de la desigualdad (2.7) es  $\emptyset$ . Si  $\mathbf{a} \neq \mathbf{0}$ , entonces el conjunto solución de la desigualdad (2.7) es el semiespacio positivo  $H^+ = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}'\mathbf{x} \geq b\}$  determinado por el hiperplano  $H = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}'\mathbf{x} = b\}$ . Geométricamente, para  $n = 2$  el conjunto solución debe tener la forma representada en la Figura 2.8.

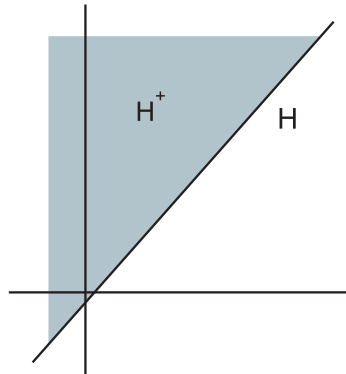


Figura 2.8: Semiespacio positivo determinado por el hiperplano  $H$ .

El conjunto factible de un problema lineal semi-infinito es la intersección de los conjuntos solución de las desigualdades de su sistema de desigualdades asociado, cada uno de los cuales es un convexo cerrado y por lo tanto el conjunto factible es un conjunto convexo y cerrado. En el caso ordinario el conjunto factible es la intersección finita de semiespacios cerrados, es decir, un poliedro. Si el conjunto factible es acotado entonces es un polítopo (ver la Figura 2.9).

De acuerdo con el Corolario 9.6.1 de [8], en la clase de problemas lineales semi-infinitos que satisfacen  $\dim \text{span}\{\mathbf{a}_t, t \in T\} = n$  se tiene que si un problema es resoluble entonces existe un punto extremo que es solución óptima.

En programación lineal semi-infinita es posible que el conjunto de puntos extremos del conjunto factible sea infinito, tal como lo ilustra el Ejemplo 2.29. En programación lineal ordinaria un punto extremo es la intersección de los hiperplanos que son determinados por restricciones cuyo conjunto de gradientes tiene dimensión  $n$ , según la caracterización presentada en el Teorema 2.17. Así, el conjunto de puntos extremos de un poliedro es finito, pues

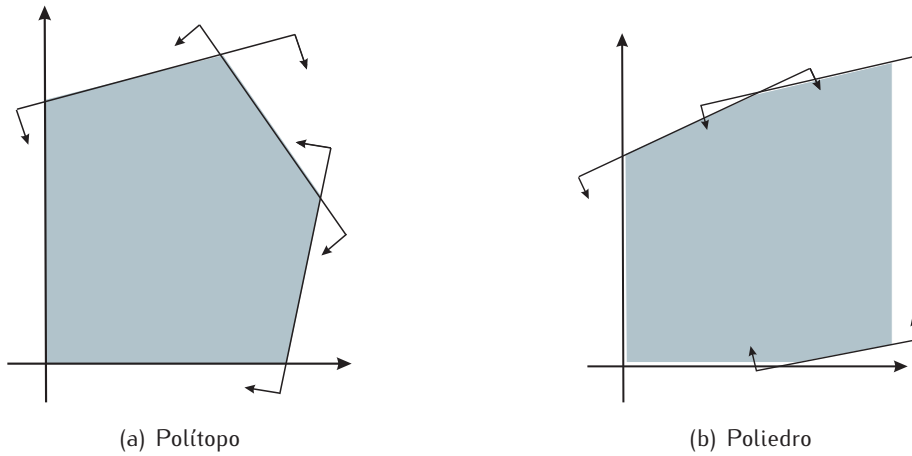


Figura 2.9: Conjuntos poliédricos.

está acotado por  $\binom{m}{n}$ . Lo anterior es de gran importancia en el desarrollo de métodos de solución, puesto que si se cuenta con una caracterización de una solución óptima entonces es posible diseñar métodos iterativos que evalúen la optimalidad de los puntos extremos. El método simplex es un algoritmo que realiza este procedimiento evaluando la condición de optimalidad de Karush-Kuhn-Tucker, y, dado que existe un número finito de puntos extremos, se asegura que el algoritmo terminará en un número finito de iteraciones.

**Ejemplo 2.29.** Consideremos el problema  $\sigma$  dado por

$$\begin{aligned} \text{Min} \quad & x_1 + 3x_2 \\ \text{s.a} \quad & x_1 \geq 0 \\ & -x_1 \geq -1 \\ & x_2 - 2tx_1 \geq -t^2, \quad \text{para cada } t \in [0, 1]. \end{aligned}$$

Cada restricción  $x_2 - 2tx_1 \geq -t^2$  se puede reescribir como

$$x_2 - t^2 \geq 2t(x_1 - t), \quad \text{para cada } t \in [0, 1].$$

Es decir, cada una de estas restricciones representa la tangente a la curva  $x_2 = x_1^2$  en el punto  $t \in [0, 1]$  (ver Figura 2.10) por lo que el conjunto factible de este problema es

$$F = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : 0 \leq x_1 \leq 1, x_2 \geq x_1^2 \right\}.$$

El conjunto de puntos extremos de  $F$  es

$$\text{ext}F = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : x_1 \in [0, 1], x_2 = x_1^2 \right\}.$$

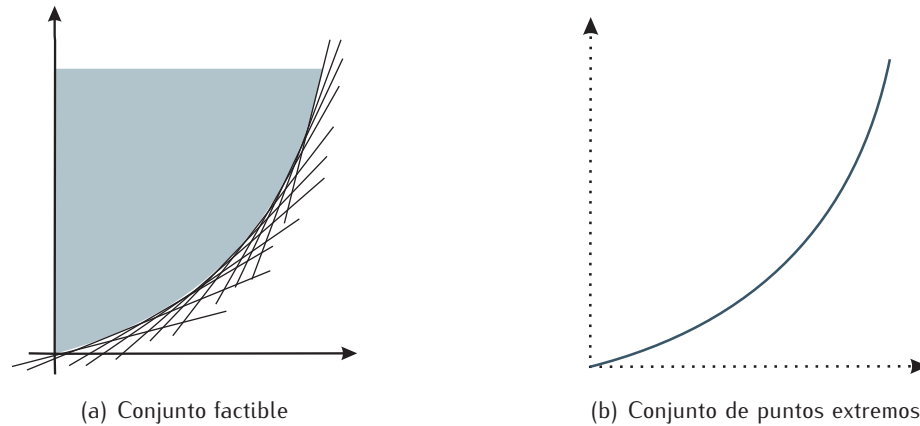


Figura 2.10: Conjunto convexo con infinitos puntos extremos.

**Definición 2.30.** Sea  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  una función. Se dice que un vector  $\mathbf{d} \in \mathbb{R}^n$  es una *dirección de descenso* en  $\bar{\mathbf{x}} \in \mathbb{R}^n$  si existe  $\delta > 0$  tal que  $f(\bar{\mathbf{x}} + \lambda \mathbf{d}) < f(\bar{\mathbf{x}})$ , para cada  $\lambda \in (0, \delta)$ .

Como veremos en el Teorema 3.1, si  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es una función diferenciable entonces las direcciones de descenso en el punto  $\bar{\mathbf{x}}$  son aquéllas que forman un ángulo agudo con el vector  $-\nabla f(\bar{\mathbf{x}})$ . Claramente la función objetivo de un problema lineal  $f(\mathbf{x}) = \mathbf{c}'\mathbf{x}$  es diferenciable y además  $\nabla f(\mathbf{x}) = \mathbf{c}$ . Intuitivamente, podemos decir que un problema lineal puede ser resuelto geoméricamente eligiendo un punto factible y avanzando desde este punto en dirección del vector  $-\mathbf{c}$  mientras no se pierda la factibilidad<sup>1</sup>.

Las curvas de nivel de la dirección objetivo son hiperplanos de la forma  $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{c}'\mathbf{x} = z\}$ , donde  $z \in \mathbb{R}$ . Si el conjunto factible es acotado, entonces el problema estará resuelto cuando se haya encontrado el hiperplano  $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{c}'\mathbf{x} = V\}$ . La intersección de este hiperplano con el conjunto factible es el conjunto de soluciones  $F^*$ . Por ejemplo, en el caso en que  $n = 2$  dicha intersección puede ser solo un punto, un segmento o un rayo.

En las Figuras 2.11, 2.12 y 2.13 suponemos que se resuelve un problema de minimización, por lo cual, la dirección sobre la cual el valor de la función objetivo disminuye está dada por el vector  $-\mathbf{c}$ . Las líneas punteadas representan las curvas de nivel de la función objetivo.

Si el conjunto factible es un poliedro no acotado, entonces tiene al menos una dirección de recesión. En este caso puede suceder que el problema tenga solución no acotada si  $-\mathbf{c} \in \text{rec}X^*$  (como en la Figura 2.13), es decir, para cada constante  $M \in \mathbb{R}$  existe  $\mathbf{x} \in F$  tal que  $\mathbf{c}'\mathbf{x} < M$ .

<sup>1</sup>Los algoritmos basados en esta idea se llaman métodos de gradiente descendente, y son muy populares en optimización numérica.

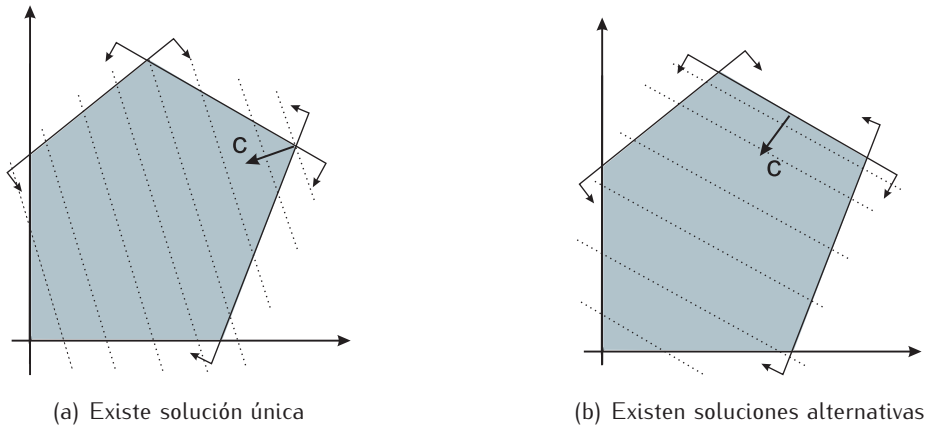


Figura 2.11: Soluciones cuando el conjunto factible es acotado.

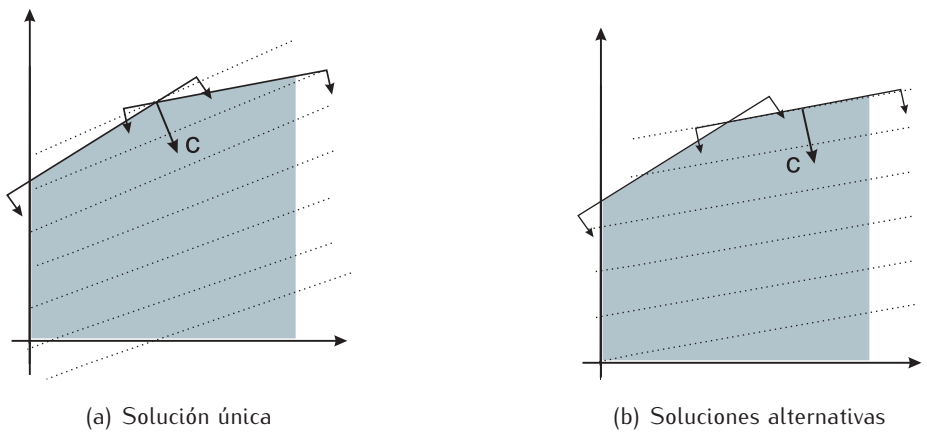


Figura 2.12: Soluciones cuando el conjunto factible es no acotado.

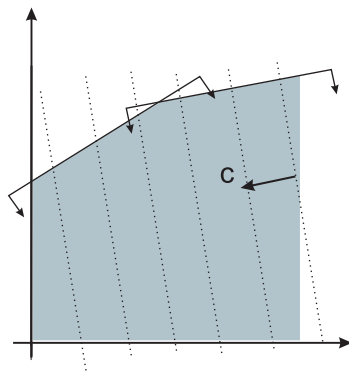


Figura 2.13: Solución no acotada.

## 2.4. Funciones convexas

En problemas de optimización en los que existe un número finito de restricciones de la forma  $g_i(\mathbf{x}) \geq 0$ , la concavidad de las restricciones  $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$  nos permitirá deducir la convexidad del conjunto factible  $F$ . Además, la convexidad de la función objetivo, para un problema de minimización, garantiza la suficiencia de la condición necesaria de primer orden para optimalidad local además de que optimalidad local sobre un conjunto convexo es equivalente a optimalidad global, tal como lo enuncia la Proposición 3.14. Dado lo anterior, son importantes los criterios para evaluar la convexidad de una función dada. En general, verificar si una función es convexa no es una tarea sencilla. Para funciones diferenciables, existen condiciones necesarias y suficientes de convexidad, como veremos en el Teorema 2.34. Además, si la función es dos veces diferenciable, entonces el análisis de su matriz Hessiana permite evaluar la convexidad de la función, como lo veremos en el Teorema 2.35.

**Definición 2.31.** Sea  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  una función y  $S \subseteq \mathbb{R}^n$  un conjunto convexo no vacío. Se dice que  $f$  es *convexa* en  $S$  si para cualesquiera  $\mathbf{x}_1, \mathbf{x}_2 \in S$

$$f(\lambda \mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2) \leq \lambda f(\mathbf{x}_1) + (1 - \lambda)f(\mathbf{x}_2),$$

para todo  $\lambda \in (0, 1)$ . Si la desigualdad anterior se satisface de forma estricta entonces se dice que  $f$  es *estrictamente convexa* en  $S$ . Se dice que  $f$  es *cóncava* en  $S$  si  $-f$  es convexa en  $S$  y se dice que es *estrictamente cóncava* en  $S$  si  $-f$  es estrictamente convexa en  $S$ .

Algunos ejemplos de estos tipos de funciones:

- (i) La función  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  dada por  $f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = x_1^2 + x_2^2$  es estrictamente convexa.
- (ii) La función  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  dada por  $f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \log x_1 + \log x_2$  es estrictamente cóncava.
- (iii) La función  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  dada por  $f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$ , donde  $\mathbf{A}$  es una matriz con entradas reales de  $m \times n$  y  $\mathbf{b} \in \mathbb{R}^m$ , es convexa y cóncava.
- (iv) Una función de la forma  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  dada por  $f(\mathbf{x}) = \mathbf{x}'\mathbf{H}\mathbf{x} + \mathbf{a}'\mathbf{x} + b$ , donde  $\mathbf{H}$  es una matriz real de  $n \times n$  simétrica,  $\mathbf{a} \in \mathbb{R}^n$  y  $b \in \mathbb{R}$ , se denomina *función cuadrática*. Como veremos en el Teorema 2.35, una función cuadrática es convexa si la matriz  $\mathbf{H}$  es definida positiva.

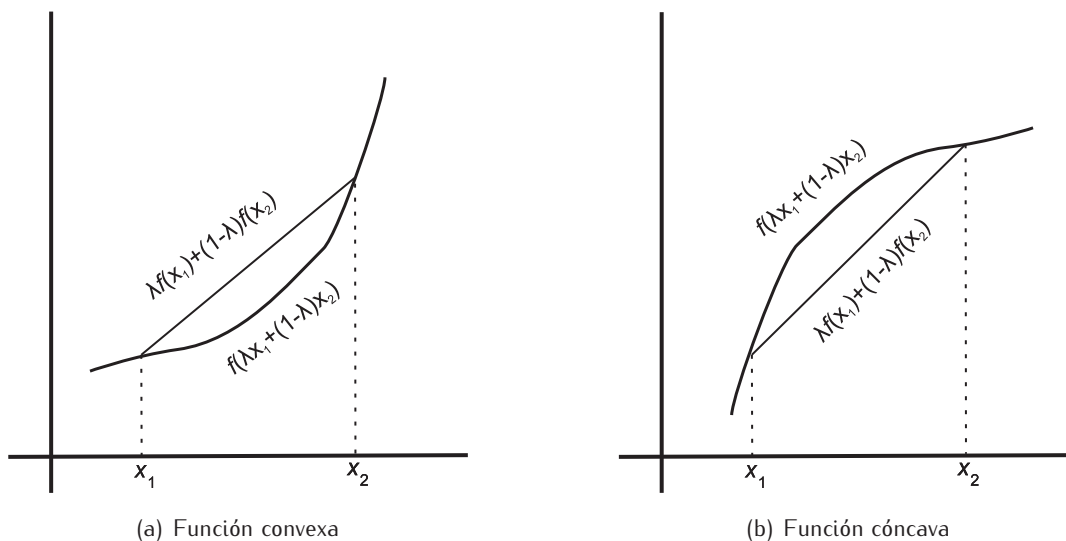


Figura 2.14: Funciones convexa y cóncava.

A continuación se presentan algunas propiedades de las funciones convexas.

**Definición 2.32.** Sea  $S \subseteq \mathbb{R}^n$  no vacío y  $f : S \rightarrow \mathbb{R}$ . El *epígrafo* de  $f$ , denotado por  $epif$ , se define como el subconjunto de  $\mathbb{R}^{n+1}$  siguiente

$$epif = \{(\mathbf{x}, y) : \mathbf{x} \in S, y \in \mathbb{R}, y \geq f(\mathbf{x})\}.$$

El *hipógrafo* de  $f$ , denotado  $hipf$ , se define como el subconjunto de  $\mathbb{R}^{n+1}$  siguiente:

$$hipf = \{(\mathbf{x}, y) : \mathbf{x} \in S, y \in \mathbb{R}, y \leq f(\mathbf{x})\}.$$

**Teorema 2.33.** Sea  $S \subseteq \mathbb{R}^n$  convexo. Entonces se cumple:

- (i) Si la función  $f : S \rightarrow \mathbb{R}$  es convexa entonces para cada  $\alpha \in \mathbb{R}$  el conjunto de nivel inferior  $S_\alpha = \{\mathbf{x} \in S : f(\mathbf{x}) \leq \alpha\}$  es convexo (Lema 3.1.2 de [2]).
- (ii) La función  $f : S \rightarrow \mathbb{R}$  es convexa en  $S$  si y sólo si  $epif$  es convexo (Teorema 3.2.2 de [2]).

*Demostración.* (i) Sean  $\mathbf{x}_1, \mathbf{x}_2 \in S_\alpha$ , entonces  $f(\mathbf{x}_1) \leq \alpha$  y  $f(\mathbf{x}_2) \leq \alpha$ . De la convexidad de

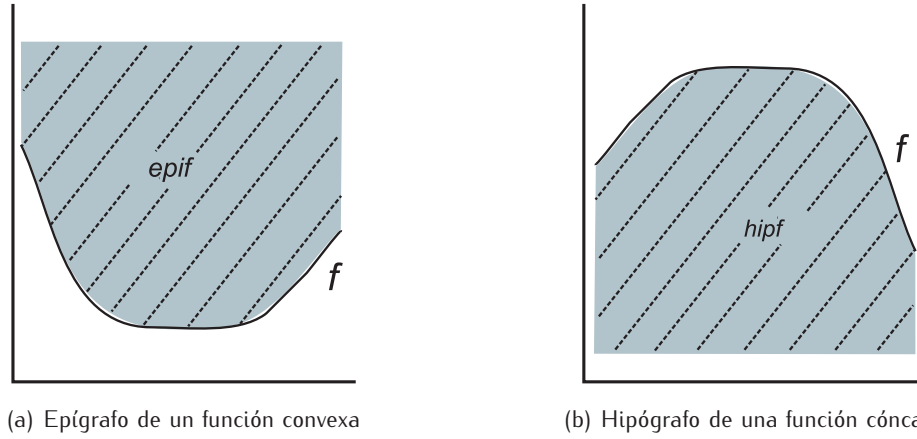


Figura 2.15: El epígrafo de una función convexa y el hipógrafo de una función cóncava son conjuntos convexos.

$f$  se tiene que para todo  $\lambda \in (0, 1)$ ,

$$\begin{aligned} f(\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2) &\leq \lambda f(\mathbf{x}_1) + (1 - \lambda) f(\mathbf{x}_2) \\ &\leq \lambda \alpha + (1 - \lambda) \alpha \\ &= \alpha, \end{aligned}$$

esto implica que, para todo  $\lambda \in (0, 1)$ ,  $\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2 \in S_\alpha$  y por lo tanto  $S_\alpha$  es convexo.

(ii) Supongamos que  $f$  es convexa. Sean  $(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2) \in epif$  cualesquiera y  $\lambda \in (0, 1)$ ,

$$\begin{aligned} f(\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2) &\leq \lambda f(\mathbf{x}_1) + (1 - \lambda) f(\mathbf{x}_2) \\ &\leq \lambda y_1 + (1 - \lambda) y_2, \end{aligned}$$

lo anterior implica  $(\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2, \lambda y_1 + (1 - \lambda) y_2) = \lambda(\mathbf{x}_1, y_1) + (1 - \lambda)(\mathbf{x}_2, y_2) \in epif$ , y por lo tanto  $epif$  es convexo.

Supongamos ahora que  $epif$  es convexo. Obviamente para cada  $\mathbf{x}_1, \mathbf{x}_2 \in S$ ,  $(\mathbf{x}_1, f(\mathbf{x}_1))$  y  $(\mathbf{x}_2, f(\mathbf{x}_2))$  son elementos de  $epif$ . Por la convexidad de  $epif$ , para cada  $\lambda \in (0, 1)$ , se satisface:

$$\begin{aligned} \lambda(\mathbf{x}_1, f(\mathbf{x}_1)) + (1 - \lambda)(\mathbf{x}_2, f(\mathbf{x}_2)) &\in epif, \\ (\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2, \lambda f(\mathbf{x}_1) + (1 - \lambda) f(\mathbf{x}_2)) &\in epif, \end{aligned}$$

es decir

$$f(\lambda \mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2) \leq \lambda f(\mathbf{x}_1) + (1 - \lambda)f(\mathbf{x}_2),$$

de donde,  $f$  es convexa.

□

Del teorema anterior se deducen propiedades análogas para funciones cóncavas.

(i) Si la función  $f : S \rightarrow \mathbb{R}$  es cóncava entonces el *conjunto de nivel superior* definido como sigue

$$S^\alpha = \{\mathbf{x} \in S : f(\mathbf{x}) \geq \alpha\},$$

es convexo para cada  $\alpha \in \mathbb{R}$ .

(ii) Una función  $f : S \rightarrow \mathbb{R}$  es cóncava en  $S$  si y sólo si *hipf* es convexo.

El teorema siguiente proporciona una caracterización geométrica de las funciones diferenciables que son convexas:

**Teorema 2.34** (Teorema 3.3.3 de [2]). *Sea  $S$  un subconjunto de  $\mathbb{R}^n$  abierto, convexo y no vacío; y sea  $f : S \rightarrow \mathbb{R}$  una función diferenciable en  $S$ . Entonces  $f$  es convexa si y sólo si para cada  $\bar{\mathbf{x}} \in S$  se tiene*

$$f(\mathbf{x}) \geq f(\bar{\mathbf{x}}) + \nabla f(\bar{\mathbf{x}})'(\mathbf{x} - \bar{\mathbf{x}}), \quad \text{para todo } \mathbf{x} \in S.$$

*Similarmente,  $f$  es estrictamente convexa si y sólo si para cada  $\bar{\mathbf{x}} \in S$  se tiene*

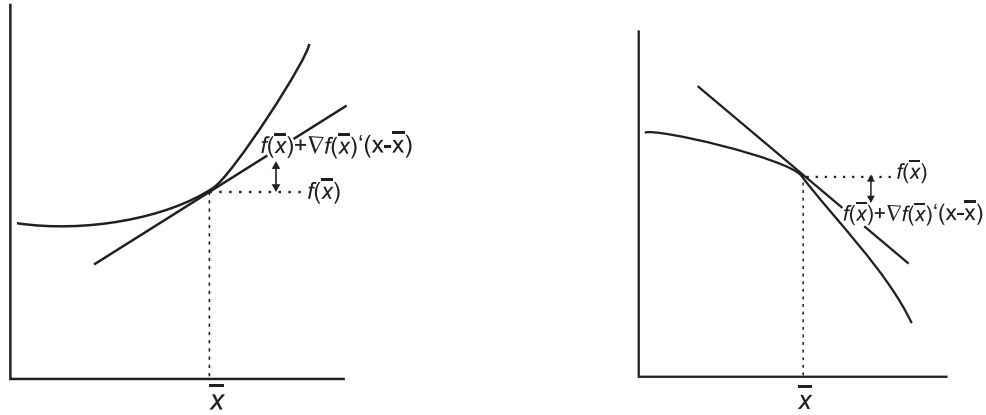
$$f(\mathbf{x}) > f(\bar{\mathbf{x}}) + \nabla f(\bar{\mathbf{x}})'(\mathbf{x} - \bar{\mathbf{x}}), \quad \text{para todo } \mathbf{x} \neq \bar{\mathbf{x}} \text{ en } S.$$

Como se menciona en [2], el teorema anterior permite derivar métodos de aproximación, ya sea usando la aproximación de primer orden para  $f(\mathbf{x})$  dada por la función afín  $f(\bar{\mathbf{x}}) + \nabla f(\bar{\mathbf{x}})'(\mathbf{x} - \bar{\mathbf{x}})$ , o la aproximación de primer orden para las restricciones. Estas aproximaciones lineales proporcionan cotas superiores o inferiores de la solución del problema original, según  $f$  sea convexa o cóncava.

Si una función  $f$  es dos veces diferenciable entonces existe un criterio para evaluar su convexidad verificando algunas propiedades de su matriz Hessiana:

**Teorema 2.35** (Teorema 3.3.7 de [2]). *Sea  $S$  un conjunto abierto y convexo en  $\mathbb{R}^n$ . Una función dos veces diferenciable  $f : S \rightarrow \mathbb{R}$  es convexa en  $S$  si y sólo si la matriz Hessiana  $\mathbf{H}(\bar{\mathbf{x}})$  es positiva semidefinida para todo  $\bar{\mathbf{x}} \in S$ .*





(a) La aproximación de primer orden es una cota inferior para una función convexa

(b) La aproximación de primer orden es una cota superior para una función cóncava

Figura 2.16: Teorema 2.34.

*Demostración.* Supongamos que  $f$  es convexa. Se debe probar que  $\mathbf{x}'\mathbf{H}(\bar{\mathbf{x}})\mathbf{x} \geq 0$  para todo  $\mathbf{x} \in \mathbb{R}^n$ . Sea  $\mathbf{x} \in \mathbb{R}^n$ , como  $S$  es abierto entonces  $\bar{\mathbf{x}} + \lambda\mathbf{x} \in S$  para  $\lambda$  tal que  $|\lambda| \neq 0$  es suficientemente pequeño. Según el teorema 2.34 tenemos

$$f(\bar{\mathbf{x}} + \lambda\mathbf{x}) \geq f(\bar{\mathbf{x}}) + \lambda\nabla f(\bar{\mathbf{x}})'\mathbf{x}, \tag{2.8}$$

y de la segunda expansión en serie de Taylor de  $f$  en el punto  $\bar{\mathbf{x}}$ , al evaluar en  $\bar{\mathbf{x}} + \lambda\mathbf{x}$  obtenemos

$$f(\bar{\mathbf{x}} + \lambda\mathbf{x}) = f(\bar{\mathbf{x}}) + \lambda\nabla f(\bar{\mathbf{x}})'\mathbf{x} + 1/2\lambda^2\mathbf{x}'\mathbf{H}(\bar{\mathbf{x}})\mathbf{x} + \lambda^2\|\mathbf{x}\|^2\alpha(\bar{\mathbf{x}}; \lambda\mathbf{x}). \tag{2.9}$$

Multiplicando (2.8) por  $-1$  y sumando con (2.9) se obtiene:

$$1/2\lambda^2\mathbf{x}'\mathbf{H}(\bar{\mathbf{x}})\mathbf{x} + \lambda^2\|\mathbf{x}\|^2\alpha(\bar{\mathbf{x}}; \lambda\mathbf{x}) \geq 0.$$

Dividiendo esta última ecuación por  $1/2\lambda^2 > 0$  y tomando el límite cuando  $\lambda \rightarrow 0$  se sigue que

$$\mathbf{x}'\mathbf{H}(\bar{\mathbf{x}})\mathbf{x} \geq 0.$$

Supongamos ahora que la matriz Hessiana es positiva semidefinida en cada punto de  $S$ . Consideremos dos puntos cualesquiera  $\mathbf{x}$  y  $\bar{\mathbf{x}}$  en  $S$ . De acuerdo con el Teorema del valor medio tenemos

$$f(\mathbf{x}) = f(\bar{\mathbf{x}}) + \nabla f(\bar{\mathbf{x}})'(\mathbf{x} - \bar{\mathbf{x}}) + \frac{1}{2}(\mathbf{x} - \bar{\mathbf{x}})'\mathbf{H}(\hat{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}}), \tag{2.10}$$

donde  $\hat{x} = \lambda\bar{x} + (1 - \lambda)x$  para cierto  $\lambda \in (0, 1)$ . Dado que  $S$  es convexo,  $\hat{x} \in S$  luego  $\mathbf{H}(\hat{x})$  es positiva semidefinida, por consiguiente  $(x - \hat{x})'\mathbf{H}(\hat{x})(x - \hat{x}) \geq 0$ . De (2.10) concluimos que

$$f(x) \geq f(\bar{x}) + \nabla f(\bar{x})'(x - \bar{x}).$$

La desigualdad anterior es válida para cada  $x, \bar{x} \in S$ , luego  $f$  es convexa según el Teorema 2.34.  $\square$

A continuación veremos un Teorema en el que se establecen condiciones que permiten verificar si una matriz simétrica es positiva semidefinida o positiva definida.

**Teorema 2.36** (Teorema 3.3.12 de [2]). *Sea  $\mathbf{H}$  una matriz simétrica con  $n$  filas.*

1. Si  $h_{ii} \leq 0$  para algún  $i \in \{1, 2, \dots, n\}$ , entonces  $\mathbf{H}$  no es positiva definida; y si  $h_{ii} < 0$  para algún  $i \in \{1, 2, \dots, n\}$ , entonces  $\mathbf{H}$  no es positiva semidefinida.
2. Si  $h_{ii} = 0$  para algún  $i \in \{1, 2, \dots, n\}$  y existe  $j \in \{1, 2, \dots, n\}$  tal que  $h_{ij} \neq 0$  entonces  $\mathbf{H}$  no es positiva semidefinida.
3. Si  $n = 1$  entonces  $\mathbf{H}$  es positiva semidefinida (positiva definida) si y sólo si  $h_{11} \geq 0$  ( $h_{11} > 0$ ). Si  $n \geq 2$  sea la matriz  $\mathbf{H}$  en forma particionada:

$$\mathbf{H} = \begin{bmatrix} h_{11} & \mathbf{q}' \\ \mathbf{q} & \mathbf{G} \end{bmatrix},$$

donde, por el inciso anterior, si  $h_{11} = 0$  entonces podemos asumir  $\mathbf{q} = \mathbf{0}$ , puesto que en caso contrario  $\mathbf{H}$  no es positiva semidefinida. En otro caso asumimos  $h_{11} > 0$  y aplicando operaciones elementales de Gauss-Jordan,  $\mathbf{H}$  se puede transformar en la matriz siguiente:

$$\mathbf{H}_{nueva} = \begin{bmatrix} h_{11} & \mathbf{q}' \\ \mathbf{0} & \mathbf{G}_{nueva} \end{bmatrix}.$$

Entonces  $\mathbf{G}_{nueva}$  es una matriz simétrica de  $(n-1) \times (n-1)$  y  $\mathbf{H}$  es positiva semidefinida si y sólo si  $\mathbf{G}_{nueva}$  es positiva semidefinida. Mas aún, si  $h_{11} > 0$ , entonces  $\mathbf{H}$  es positiva definida si y sólo si  $\mathbf{G}_{nueva}$  es positiva definida.

Enseguida veremos un ejemplo en el cual verificamos si una función es semidefinida positiva analizando, por medio del Teorema anterior, si su matriz Hessiana es semidefinida positiva.

**Ejemplo 2.37.** Sea  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$  la función dada por

$$f(\mathbf{x}) = (x_1 \ x_2 \ x_3) \begin{bmatrix} 2 & 0 & 6 \\ 0 & 0 & 0 \\ 6 & 0 & 20 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + (1 \ 0 \ -2) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Para determinar si la función  $f$  es convexa debemos verificar si su matriz Hessiana

$$H = \begin{bmatrix} 2 & 0 & 6 \\ 0 & 0 & 0 \\ 6 & 0 & 20 \end{bmatrix}$$

es positiva semidefinida. Aplicando la reducción de Gauss-Jordan a la matriz  $H$  obtenemos

$$H_1 = \left[ \begin{array}{c|cc} 2 & 0 & 6 \\ \hline 0 & 0 & 0 \\ 0 & 0 & 2 \end{array} \right]$$

Notamos que la matriz  $G_1 = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$  es positiva semidefinida pues la matriz  $G_2 = [2]$  lo es, por lo tanto  $H$  es positiva semidefinida y  $f$  es convexa.



## Capítulo 3

# Criterios de optimalidad y unicidad

Una de las tareas principales en la teoría de optimización es la determinación de condiciones que nos permitan identificar la optimalidad de una solución factible. En la sección 3.1 veremos una caracterización de las direcciones de descenso para problemas de optimización en los que las funciones involucradas son diferenciables. La condición necesaria de primer orden para optimalidad local, que presentamos en la sección 3.2, involucra el llamado cono de tangentes. La tarea de calcular el cono de tangentes de forma explícita es difícil en general, sin embargo, con alguna condición adicional sobre las restricciones, el teorema de Karush-Kuhn-Tucker (KKT) simplifica la verificación de la condición de primer orden al sustituir el cálculo del cono de tangentes utilizando el cono de restricciones activas, esto es de lo que tratan las secciones 3.3 y 3.4.

El Teorema de KKT es una condición necesaria cuando existe diferenciables de las funciones que definen un problema además de una cualificación de restricciones que permita sustituir el cono de tangentes por el llamado cono activo. Con algunas hipótesis de convexidad sobre las funciones involucradas en el problema, esta condición resulta ser también suficiente, tal como lo demostraremos posteriormente en el Teorema 3.15. En programación lineal semi-infinita se ha generalizado el Teorema de KKT para ciertas clases de problemas, de las cuales mencionaremos los problemas tipo Farkas-Minkowski. En el Teorema 3.22 una caracterización para optimalidad en el caso semi-infinito en general que involucra el cono de direcciones factibles.

En la última sección de este capítulo mencionamos criterios para unicidad de la solución.

### 3.1. Direcciones de descenso

En lo que sigue haremos referencia a un problema de optimización que tiene la forma siguiente:

$$(P) \quad \begin{array}{l} \text{Minimizar } f(\mathbf{x}) \\ \text{sujeto a } \quad g_i(\mathbf{x}) \geq 0, \quad \text{para cada } i \in I, \end{array}$$

donde  $I = \{1, 2, \dots, m\}$  es un conjunto de índices y  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$  son funciones.  $F$  representará al conjunto factible del problema (P).

Se dice que  $\bar{\mathbf{x}} \in F$  es un mínimo local para el problema (P) si existe una vecindad de  $\bar{\mathbf{x}}$ ,  $N(\bar{\mathbf{x}})$ , tal que  $f(\bar{\mathbf{x}}) \leq f(\mathbf{x})$  para todo  $\mathbf{x} \in N(\bar{\mathbf{x}}) \cap F$ , y se dice que  $\bar{\mathbf{x}}$  es un mínimo global si  $f(\bar{\mathbf{x}}) \leq f(\mathbf{x})$  para todo  $\mathbf{x} \in F$ .

De acuerdo con la Definición 2.30,  $\mathbf{d}$  es una *dirección de descenso* para  $f$  en  $\bar{\mathbf{x}}$  si existe  $\delta > 0$  tal que  $f(\bar{\mathbf{x}} + \lambda \mathbf{d}) < f(\bar{\mathbf{x}})$ , para cada  $\lambda \in (0, \delta)$ . El siguiente resultado caracteriza las direcciones de descenso para una función diferenciable.

**Teorema 3.1** (Teorema 4.1.2 de [2]). *Supongamos que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es una función diferenciable en  $\bar{\mathbf{x}}$ . Si el vector  $\mathbf{d} \in \mathbb{R}^n$  satisface  $\nabla f(\bar{\mathbf{x}})' \mathbf{d} < 0$ , entonces  $\mathbf{d}$  es una dirección de descenso de  $f$  en  $\bar{\mathbf{x}}$ . Recíprocamente, si  $\mathbf{d} \in \mathbb{R}^n$  es una dirección de descenso en  $\bar{\mathbf{x}}$  entonces  $\nabla f(\bar{\mathbf{x}})' \mathbf{d} < 0$ .*

*Demostración.* Sea  $\lambda > 0$ . Según la hipótesis de diferenciability de  $f$  podemos escribir

$$f(\bar{\mathbf{x}} + \lambda \mathbf{d}) = f(\bar{\mathbf{x}}) + \lambda \nabla f(\bar{\mathbf{x}})' \mathbf{d} + \lambda \|\mathbf{d}\| \alpha(\bar{\mathbf{x}}; \lambda \mathbf{d}),$$

donde  $\alpha(\bar{\mathbf{x}}; \lambda \mathbf{d}) \rightarrow 0$  cuando  $\lambda \rightarrow 0$ ; entonces, si  $\lambda \neq 0$

$$\frac{f(\bar{\mathbf{x}} + \lambda \mathbf{d}) - f(\bar{\mathbf{x}})}{\lambda} = \nabla f(\bar{\mathbf{x}})' \mathbf{d} + \|\mathbf{d}\| \alpha(\bar{\mathbf{x}}; \lambda \mathbf{d}),$$

tomando límite cuando  $\lambda \rightarrow 0$  obtenemos

$$\lim_{\lambda \rightarrow 0} \frac{f(\bar{\mathbf{x}} + \lambda \mathbf{d}) - f(\bar{\mathbf{x}})}{\lambda} = \nabla f(\bar{\mathbf{x}})' \mathbf{d} < 0.$$

Así, existe  $\delta > 0$  tal que  $f(\bar{\mathbf{x}} + \lambda \mathbf{d}) - f(\bar{\mathbf{x}}) < 0$  para todo  $\lambda \in (0, \delta)$ .

Supongamos ahora que  $\mathbf{d}$  es una dirección de descenso en  $\bar{\mathbf{x}}$ , es decir, existe  $\delta > 0$  tal que, para todo  $\lambda \in (0, \delta)$ ,  $f(\bar{\mathbf{x}} + \lambda \mathbf{d}) < f(\bar{\mathbf{x}})$ , entonces

$$\nabla f(\bar{\mathbf{x}})' \mathbf{d} = \lim_{\lambda \rightarrow 0} \frac{f(\bar{\mathbf{x}} + \lambda \mathbf{d}) - f(\bar{\mathbf{x}})}{\lambda} < 0.$$

□

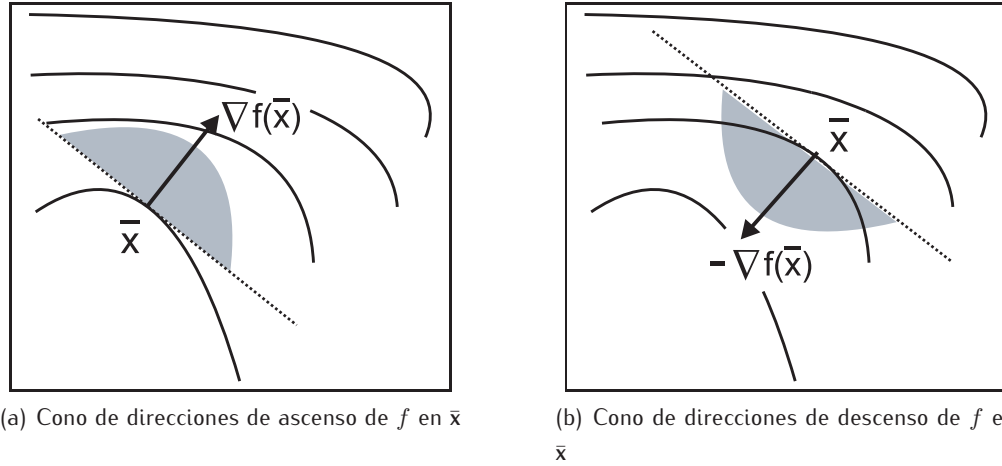


Figura 3.1: En la figura se muestran las curvas de nivel de  $f$ . Los conos están sombreados.

Como consecuencia del resultado anterior tenemos que, si  $\mathbf{d} \in \mathbb{R}^n$  satisface  $\nabla f(\bar{x})' \mathbf{d} > 0$  entonces  $\mathbf{d}$  es una dirección de ascenso en  $\bar{x}$ . El conjunto de direcciones de descenso, así como el conjunto de direcciones de ascenso de una función diferenciable en un punto son conos convexos, como veremos a continuación.

**Corolario 3.2.** Sea  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  una función diferenciable. Para todo  $\bar{x} \in \mathbb{R}^n$  el conjunto de direcciones de descenso de  $f$  en  $\bar{x}$  es un cono convexo.

*Demostración.* Sea  $\bar{x} \in \mathbb{R}^n$ . De acuerdo con el teorema anterior, el conjunto de direcciones de descenso de  $f$  en el punto  $\bar{x}$  está caracterizado como  $C = \{\mathbf{d} \in \mathbb{R}^n : \nabla f(\bar{x})' \mathbf{d} < 0\}$ . Veamos que  $C$  es un cono. Sea  $\mathbf{d} \in C$ , entonces  $\nabla f(\bar{x})' \mathbf{d} < 0$ . De lo anterior tenemos que para cualesquiera  $\lambda > 0$ ,  $\nabla f(\bar{x})' (\lambda \mathbf{d}) = \lambda \nabla f(\bar{x})' \mathbf{d} < 0$ , es decir  $\lambda \mathbf{d} \in C$ , de donde  $C$  es un cono.

Veamos que  $C$  es convexo. Sean  $\mathbf{d}_1, \mathbf{d}_2 \in C$  y sea  $\lambda \in (0, 1)$ , entonces

$$\nabla f(\bar{x})' (\lambda \mathbf{d}_1 + (1 - \lambda) \mathbf{d}_2) = \lambda \nabla f(\bar{x})' \mathbf{d}_1 + (1 - \lambda) \nabla f(\bar{x})' \mathbf{d}_2 < 0,$$

lo cual implica  $\lambda \mathbf{d}_1 + (1 - \lambda) \mathbf{d}_2 \in C$ . Por lo tanto,  $C$  es un cono convexo. □

Como  $\mathbf{d}$  y  $\lambda \mathbf{d}$ ,  $\lambda > 0$ , definen la misma dirección, podemos tomar como direcciones representativas del cono a las direcciones de norma euclídeana menor o igual que 1. Sobre este conjunto se verifica que la dirección de mayor descenso es  $-\frac{\nabla f(\bar{x})}{\|\nabla f(\bar{x})\|}$ , como veremos en el Corolario siguiente:

**Corolario 3.3.** Sea  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  una función diferenciable en  $\bar{x}$ . Para cualquier vector  $\mathbf{d} \in \mathbb{R}^n$  dirección de descenso para  $f$  en  $\bar{x}$  tal que  $\|\mathbf{d}\| \leq 1$  existe  $\epsilon > 0$  tal que, para todo  $\lambda \in (0, \epsilon)$ , se satisface  $f(\bar{x} + \lambda\bar{\mathbf{d}}) \leq f(\bar{x} + \lambda\mathbf{d})$ , donde  $\bar{\mathbf{d}} = -\frac{\nabla f(\bar{x})}{\|\nabla f(\bar{x})\|}$ .

*Demostración.* Sea  $\mathbf{d} \in \mathbb{R}^n$  una dirección de descenso para  $f$  en  $\bar{x}$  tal que  $\|\mathbf{d}\| \leq 1$ . Como  $\|\mathbf{d}\| \leq 1 = \|\bar{\mathbf{d}}\|$ , entonces  $\|\nabla f(\bar{x})\|\|\mathbf{d}\| \leq \|\nabla f(\bar{x})\|\|\bar{\mathbf{d}}\|$ . Por otra parte

$$\|\nabla f(\bar{x})\|\|\bar{\mathbf{d}}\| = |\nabla f(\bar{x})'\bar{\mathbf{d}}| = \|\nabla f(\bar{x})\|,$$

de lo anterior

$$|\nabla f(\bar{x})'\mathbf{d}| \leq |\nabla f(\bar{x})'\bar{\mathbf{d}}|. \quad (3.1)$$

Como  $\mathbf{d}$  y  $\bar{\mathbf{d}}$  son direcciones de descenso de  $f$  en  $\bar{x}$ , entonces  $\nabla f(\bar{x})'\mathbf{d} < 0$  y  $\nabla f(\bar{x})'\bar{\mathbf{d}} < 0$ , así (3.1) implica

$$\nabla f(\bar{x})'\bar{\mathbf{d}} \leq \nabla f(\bar{x})'\mathbf{d}, \quad (3.2)$$

de donde se sigue

$$\lim_{\lambda \rightarrow 0} \frac{f(\bar{x} + \lambda\bar{\mathbf{d}}) - f(\bar{x})}{\lambda} = \nabla f(\bar{x})'\bar{\mathbf{d}} \leq \nabla f(\bar{x})'\mathbf{d} = \lim_{\lambda \rightarrow 0} \frac{f(\bar{x} + \lambda\mathbf{d}) - f(\bar{x})}{\lambda},$$

por lo tanto, existe  $\epsilon > 0$  tal que, para todo  $\lambda \in (0, \epsilon)$ , se satisface  $f(\bar{x} + \lambda\bar{\mathbf{d}}) \leq f(\bar{x} + \lambda\mathbf{d})$ .  $\square$

## 3.2. Condición necesaria de primer orden

**Definición 3.4.** Sea  $S \subseteq \mathbb{R}^n$  no vacío y sea  $\bar{x} \in \text{cl}S$ . El *cono de tangentes* de  $S$  en el punto  $\bar{x}$ , denotado por  $T(\bar{x})$  está definido como el conjunto de todas las direcciones  $\mathbf{d}$  tales que  $\mathbf{d} = \lim_{k \rightarrow \infty} \lambda_k(\mathbf{x}_k - \bar{x})$  donde  $\lambda_k > 0$ ,  $\mathbf{x}_k \in S$  para todo  $k$ , y  $\mathbf{x}_k \rightarrow \bar{x}$ . A cada elemento del cono de tangentes se le denomina *dirección tangencial*.

En general, siempre se cumple que cualquier dirección factible en un punto de un conjunto convexo también es una dirección tangencial.

**Proposición 3.5.** Si  $X \subseteq \mathbb{R}^n$  es un conjunto convexo entonces para todo  $\bar{x} \in X$ ,  $D(X, \bar{x}) \subseteq T(\bar{x})$ .

*Demostración.* Si  $\mathbf{d} \in D(X, \bar{x})$  entonces existe  $\epsilon > 0$  tal que  $\bar{x} + \epsilon\mathbf{d} \in X$ . Definiendo  $\mathbf{x}_k = \bar{x} + \frac{1}{k}\mathbf{d}$ , para cada  $k \in \mathbb{N}$ , obtenemos que  $\mathbf{x}_k \rightarrow \bar{x}$  y además

$$\mathbf{d} = \lim_{k \rightarrow \infty} k(\mathbf{x}_k - \bar{x}).$$

Por lo tanto  $\mathbf{d} \in T(\bar{x})$ .  $\square$



Si el conjunto factible es un poliedro, como ocurre en programación lineal ordinaria, entonces toda dirección tangencial es una dirección factible en cualquier punto del conjunto factible (Proposición 3.4 de [7]).

La condición necesaria de primer orden, enunciada a continuación, establece que no existen direcciones tangenciales que sean direcciones de descenso en un óptimo local.

**Teorema 3.6** (Proposición 3.2 de [7]). *Si  $\bar{x} \in F$  es un mínimo local de  $(P)$  entonces  $\nabla f(\bar{x})' \mathbf{d} \geq 0$ , para cada  $\mathbf{d} \in T(\bar{x})$ . En particular, si  $\bar{x} \in \text{int}F$ , entonces  $\nabla f(\bar{x}) = \mathbf{0}$ .*

*Demostración.* Si  $\mathbf{d} \in T(\bar{x})$  entonces existe una sucesión  $\{\mathbf{x}_k\} \in F$  tal que  $\mathbf{x}_k \rightarrow \bar{x}$  y existen escalares  $\lambda_k \in \mathbb{R}$  tales que  $\mathbf{d} = \lim_{k \rightarrow \infty} \lambda_k (\mathbf{x}_k - \bar{x})$ . Por la diferenciabilidad de  $f$ , para cada  $k \in \mathbb{N}$  podemos escribir

$$f(\mathbf{x}_k) = f(\bar{x}) + \nabla f(\bar{x})'(\mathbf{x}_k - \bar{x}) + \|\mathbf{x}_k - \bar{x}\| \alpha(\bar{x}; \mathbf{x}_k - \bar{x}),$$

donde  $\lim_{\mathbf{x}_k \rightarrow \bar{x}} \alpha(\bar{x}; \mathbf{x}_k - \bar{x}) = 0$ . Debido a la optimalidad local de  $\bar{x}$ , y como  $\mathbf{x}_k \rightarrow \bar{x}$ , entonces  $f(\bar{x}) \leq f(\mathbf{x}_k)$ , para  $k$  suficientemente grande. Por otra parte

$$\nabla f(\bar{x})'(\mathbf{x}_k - \bar{x}) + \|\mathbf{x}_k - \bar{x}\| \alpha(\bar{x}; \mathbf{x}_k - \bar{x}) = f(\mathbf{x}_k) - f(\bar{x}),$$

tomando el límite cuando  $k \rightarrow \infty$ ,

$$\lim_{k \rightarrow \infty} (\nabla f(\bar{x})'(\mathbf{x}_k - \bar{x}) + \|\mathbf{x}_k - \bar{x}\| \alpha(\bar{x}; \mathbf{x}_k - \bar{x})) \geq 0.$$

De aquí,

$$\lim_{k \rightarrow \infty} \nabla f(\bar{x})'(\mathbf{x}_k - \bar{x}) \geq 0.$$

Multiplicando ambos lados de la desigualdad por  $\lambda_k$

$$\lim_{k \rightarrow \infty} \nabla f(\bar{x})' \lambda_k (\mathbf{x}_k - \bar{x}) \geq 0.$$

Esto es,

$$\nabla f(\bar{x})' \mathbf{d} \geq 0.$$

Si  $\bar{x} \in \text{int}F$  entonces  $D(F, \bar{x}) = \mathbb{R}^n$  y de acuerdo con la Proposición 3.5  $D(F, \bar{x}) \subset T(\bar{x})$ , por lo que  $T(\bar{x}) = \mathbb{R}^n$ , luego  $\mathbf{d} = -\nabla f(\bar{x}) \in T(\bar{x})$ . Como  $\nabla f(\bar{x})' \mathbf{d} \geq 0$  entonces

$$0 \leq \nabla f(\bar{x})' \mathbf{d} = -\nabla f(\bar{x})' \nabla f(\bar{x}) = -\|\nabla f(\bar{x})\|^2,$$

lo anterior es válido sólo si  $\nabla f(\bar{x}) = \mathbf{0}$ . □

Según la condición necesaria de primer orden, los posibles candidatos a soluciones óptimas locales son los puntos  $\bar{x} \in F$  que satisfacen  $\nabla f(\bar{x})' \mathbf{d} \geq 0$ , para cada  $\mathbf{d} \in T(\bar{x})$ . Sin embargo, la verificación de esta condición no es fácil, debido a que se debe conocer  $T(\bar{x})$ , el cual es difícil de determinar de forma explícita. El teorema de Karush-Kuhn-Tucker salva esta dificultad al sustituir el cálculo del cono de tangentes por el cono generado por los gradientes de las restricciones activas, si se satisface una condición sobre las restricciones que garanticen la igualdad de estos conjuntos.

### 3.3. Restricciones activas

Si  $\bar{x} \in F$ , las restricciones que satisfacen  $g_i(\bar{x}) = 0$  se llaman *restricciones activas* en  $\bar{x}$ . El conjunto de índices de las restricciones activas en  $\bar{x}$  se denota con  $I(\bar{x})$ , es decir,  $I(\bar{x}) = \{i \in I : g_i(\bar{x}) = 0\}$ .

**Definición 3.7.** El *cono activo* en  $\bar{x}$  es el cono convexo generado por los gradientes de las restricciones activas en  $\bar{x}$ , es decir

$$A(\bar{x}) := \text{cone}\{\nabla g_i(\bar{x}) : i \in I(\bar{x})\}.$$

Para ver la relación que existe entre el cono formado por los gradientes de las restricciones activas y el cono de direcciones tangenciales en un punto consideremos el siguiente ejemplo:

**Ejemplo 3.8.** Consideremos el problema siguiente

$$\begin{aligned} & \text{Minimizar} && x_1 + x_2 \\ & \text{sujeto a} && -x_2 + 1 \geq 0 \\ & && x_2 - x_1^2 \geq 0. \end{aligned}$$

Tenemos dos restricciones dadas por  $g_1 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = -x_2 + 1$  y  $g_2 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_2 - x_1^2$ .

Consideremos los puntos factibles  $\bar{x}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$  y  $\bar{x}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ . Los conos tangenciales en estos puntos son

$$T(\bar{x}_1) = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : x_2 \geq 0 \right\}$$

y

$$T(\bar{x}_2) = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : x_2 \leq 0 \right\}$$

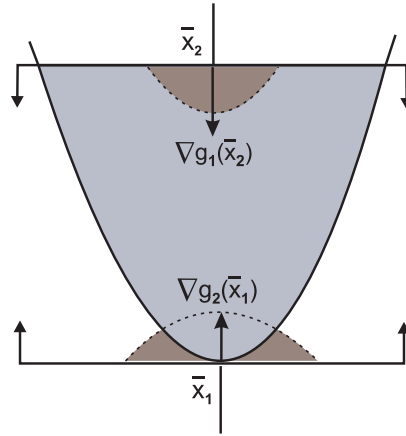


Figura 3.2: Ejemplo 3.8.

los cuales están representados en color más obscuro en la Figura 3.2. El gradiente de la restricción activa en  $\bar{x}_1$  es  $\nabla g_2(\bar{x}_1) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$  y el gradiente de la restricción activa en  $\bar{x}_2$  es

$$\nabla g_1(\bar{x}_2) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

Notemos que  $x \in T(\bar{x}_1)$  si y solo si  $\nabla g_2(\bar{x}_1)'x \geq 0$ , de la misma forma  $x \in T(\bar{x}_2)$  si y solo si  $\nabla g_1(\bar{x}_2)'x \geq 0$ .

El cono polar de  $A(\bar{x})$  es el conjunto  $A(\bar{x})^0 = \{d \in \mathbb{R}^n : \nabla g_i(\bar{x})'d \geq 0, \forall i \in I(\bar{x})\}$ . Así, en los dos puntos factibles del ejemplo anterior el cono tangencial coincide con el polar del cono activo. Aunque esta igualdad no es válida en general, siempre se tiene una contención, tal como se muestra a continuación.

**Proposición 3.9** (Proposición 3.3 de [7]). *Para cualesquiera  $\bar{x} \in F$  se cumple  $T(\bar{x}) \subseteq A(\bar{x})^0$ .*

*Demostración.* Si  $I(\bar{x}) = \emptyset$  entonces nada hay que probar pues  $T(\bar{x}) = A(\bar{x})^0 = \mathbb{R}^n$ . Supongamos  $I(\bar{x}) \neq \emptyset$ . Sea  $d \in T(\bar{x})$ , existe una sucesión  $\{x_k\} \subseteq F$  tal que  $x_k \rightarrow \bar{x}$  y escalares positivos  $\lambda_1, \lambda_2, \dots$ , tales que  $d = \lim_{k \rightarrow \infty} \lambda_k(x_k - \bar{x})$ . Sea  $g_i \in I(\bar{x})$ , según la diferenciabilidad de  $g$  tenemos

$$g_i(x_k) = g_i(\bar{x}) + \nabla g_i(\bar{x})'(x_k - \bar{x}) + \|x_k - \bar{x}\|\alpha(\bar{x}; x_k - \bar{x}),$$

tomando límite cuando  $k \rightarrow \infty$  en la igualdad anterior, y como  $g_i(\bar{x}) = 0$ , se tiene

$$\lim_{k \rightarrow \infty} (\nabla g_i(\bar{x})'(x_k - \bar{x}) + \|x_k - \bar{x}\|\alpha(\bar{x}; x_k - \bar{x})) = \lim_{k \rightarrow \infty} g_i(x_k),$$

pero  $\mathbf{x}_k \in F$  implica  $g_i(\mathbf{x}_k) \geq 0$  y, por consiguiente,  $\lim_{k \rightarrow \infty} g_i(\mathbf{x}_k) \geq 0$ . Además  $\lim_{k \rightarrow \infty} \|\mathbf{x}_k - \bar{\mathbf{x}}\| \alpha(\bar{\mathbf{x}}; \mathbf{x}_k - \bar{\mathbf{x}}) = 0$ . Luego

$$\lim_{k \rightarrow \infty} \nabla g_i(\bar{\mathbf{x}})'(\mathbf{x}_k - \bar{\mathbf{x}}) \geq 0,$$

multiplicando por  $\lambda_k$

$$\nabla g_i(\bar{\mathbf{x}})' \mathbf{d} = \lim_{k \rightarrow \infty} \nabla g_i(\bar{\mathbf{x}})' \lambda_k (\mathbf{x}_k - \bar{\mathbf{x}}) \geq 0.$$

□

Para reemplazar en la condición necesaria de primer orden  $T(\bar{\mathbf{x}})$  por  $A(\bar{\mathbf{x}})^\circ$ , se debe satisfacer la otra contención  $T(\bar{\mathbf{x}}) \supseteq A(\bar{\mathbf{x}})^\circ$ . El siguiente ejemplo muestra que esta contención no siempre ocurre:

**Ejemplo 3.10.** Consideremos el problema siguiente:

$$\begin{aligned} & \text{Minimizar} && x_1 + x_2 \\ & \text{sujeto a} && x_1^2 + (x_2 - 1)^2 \leq 1 \\ & && x_1^2 + (x_2 + 1)^2 \leq 1. \end{aligned}$$

Las restricciones son  $g_1 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = -x_1^2 - (x_2 - 1)^2 + 1$  y  $g_2 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = -x_1^2 - (x_2 + 1)^2 + 1$  (Ver

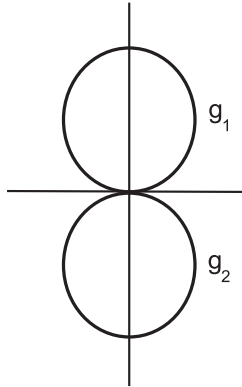


Figura 3.3: Ejemplo donde  $T(\bar{\mathbf{x}}) \not\supseteq A(\bar{\mathbf{x}})^\circ$ .

la Figura 3.3). El conjunto factible consta solamente del punto  $\bar{\mathbf{x}} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ , de aquí  $T(\bar{\mathbf{x}}) = \{\mathbf{0}\}$  puesto que  $\mathbf{d} \in T(\bar{\mathbf{x}})$  si y solo si

$$\mathbf{d} = \lim_{k \rightarrow \infty} \lambda_k (\mathbf{x}_k - \bar{\mathbf{x}}), \quad (3.3)$$

donde  $\{\mathbf{x}_k\}$  es una sucesión en  $F$  que converge a  $\bar{\mathbf{x}}$  y  $\lambda_k > 0$ . Como  $F = \{\bar{\mathbf{x}}\}$ , es claro que la única expresión de la forma (3.3) es el vector  $\mathbf{0}$ .

Por otra parte, las dos restricciones son activas en  $\bar{\mathbf{x}}$  con gradientes  $\nabla g_1(\bar{\mathbf{x}}) = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$  y  $\nabla g_2(\bar{\mathbf{x}}) = \begin{pmatrix} 0 \\ -2 \end{pmatrix}$ , respectivamente. Luego el cono activo en  $\bar{\mathbf{x}}$  es

$$A(\bar{\mathbf{x}}) = \text{cone} \left\{ \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \end{pmatrix} \right\} = \left\{ \lambda \begin{pmatrix} 0 \\ 1 \end{pmatrix} : \lambda \in \mathbb{R} \right\},$$

y su cono polar positivo es

$$A(\bar{\mathbf{x}})^\circ = \left\{ \lambda \begin{pmatrix} 1 \\ 0 \end{pmatrix} : \lambda \in \mathbb{R} \right\}.$$

Claramente  $T(\bar{\mathbf{x}}) \subsetneq A(\bar{\mathbf{x}})^\circ$ .

Una *cualificación de restricciones* es una condición sobre las restricciones de tal forma que se cumpla la igualdad  $T(\bar{\mathbf{x}}) = A(\bar{\mathbf{x}})^\circ$ , llamada *cualificación de restricciones de Abadie*. En la siguiente Proposición mencionamos dos propiedades que son cualificaciones de restricciones, las cuales se pueden consultar en la sección 5.2 de [2].

**Proposición 3.11.** *Cada una de las siguientes condiciones es una cualificación de restricciones.*

- (i) *Cualificación de regularidad: El conjunto  $\{\nabla g_i(\bar{\mathbf{x}}), i \in I(\bar{\mathbf{x}})\}$  es linealmente independiente.*
- (ii) *Cualificación de Slater: Las funciones  $\{g_i, i \in I\}$  son cóncavas y además  $\text{int}F \neq \emptyset$ .*

El resultado que sigue establece que en cualquier punto de un poliedro el cono de tangentes coincide con el cono de direcciones factibles y el cono de tangentes es igual al cono polar del cono activo. Incluimos la demostración tomada de [7].

**Proposición 3.12** (Proposición 3.4 de [7]). *Si  $F = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}'_i \mathbf{x} \geq b_i, i \in I\}$ , entonces para cada  $\bar{\mathbf{x}} \in F$  se tiene  $T(\bar{\mathbf{x}}) = D(F, \bar{\mathbf{x}})$  y  $T(\bar{\mathbf{x}}) = A(\bar{\mathbf{x}})^\circ$ .*

*Demostración.* Sea  $\bar{\mathbf{x}} \in F$ . Consideremos los dos casos siguientes

- (i)  $I(\bar{\mathbf{x}}) = \emptyset$ . En este caso  $\mathbf{a}'_i \bar{\mathbf{x}} > b_i$ , para cada  $i \in I$ . Para cada  $i = 1, \dots, m$ , tenemos que  $\bar{\mathbf{x}}$  pertenece al semiespacio abierto

$$\{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}'_i \mathbf{x} > b_i\},$$

por lo tanto existe  $\delta_i > 0$  tal que

$$\bar{x} + \delta B \subseteq \{x \in \mathbb{R}^n : a'_i x > b_i\}.$$

Sea  $\delta = \min\{\delta_i, i \in I\}$  entonces  $\bar{x} + \delta B \in F$ , es decir  $\bar{x} \in \text{int } F$ . Se sigue que  $D(\bar{x}) = T(\bar{x}) = \mathbb{R}^n$ . Por otra parte,  $A(\bar{x}) = \text{cone } \emptyset = \{0\}$ , luego  $A(\bar{x})^0 = \mathbb{R}^n$  y por lo tanto se cumple  $T(\bar{x}) = A(\bar{x})$ .

(ii)  $I(\bar{x}) \neq \emptyset$ . Para demostrar  $A(\bar{x})^0 \subseteq T(\bar{x})$  basta verificar que  $A(\bar{x}) \subseteq D(F, \bar{x})$ , puesto que  $D(F, \bar{x}) \subseteq T(\bar{x})$ . Sea  $d \in A(\bar{x})^0$  arbitrario, entonces  $a'_i d \geq 0$  para cada  $i \in I(\bar{x})$ . Si para todo  $i = 1, \dots, m$  se satisface  $a'_i d \geq 0$  entonces  $a'_i(\bar{x} + \lambda d) \geq 0$ , para cualquier  $\lambda > 0$  y por lo tanto  $d \in D(F, \bar{x})$ . Si  $i \notin I(\bar{x})$  entonces  $a'_i \bar{x} > b_i$ , luego

$$\lim_{\lambda \rightarrow 0^+} a'_i(\bar{x} + \lambda d) > b_i,$$

de donde, existe  $\lambda_i > 0$  tal que  $a'_i(\bar{x} + \lambda_i d) > b_i$ . Haciendo  $\epsilon = \min\{\lambda_i, i \notin I(\bar{x})\}$  obtenemos  $\bar{x} + \epsilon d \in F$  lo cual implica que  $d \in D(F, \bar{x})$ .

□

### 3.4. Teorema de Karush-Kuhn-Tucker

En las secciones previas hemos considerado el problema de optimización siguiente:

$$(P) \quad \begin{array}{l} \text{Minimizar} \quad f(x) \\ \text{sujeto a} \quad g_i(x) \geq 0, \quad \text{para cada } i \in I, \end{array}$$

donde  $I = \{1, \dots, m\}$  es un conjunto de índices y  $f, g_i : \mathbb{R}^n \rightarrow \mathbb{R}$  son funciones. Hemos visto que para evaluar la condición necesaria de primer orden para optimalidad local en un punto factible del problema (P) se requiere de conocer el cono de tangentes. Para los problemas que satisfacen una cualificación de restricciones es posible sustituir el cono de tangentes por el cono polar del cono activo el cual se puede determinar de forma directa. A continuación veremos el Teorema de Karush-Kuhn-Tucker.

**Teorema 3.13** (Teorema de Karush-Kuhn-Tucker). *Si  $\bar{x} \in F$  es un mínimo local de (P) y se satisface una cualificación de restricciones entonces  $\nabla f(\bar{x}) \in A(\bar{x})$ . Por lo tanto, existen escalares no negativos  $\lambda_i, i \in I(\bar{x})$  (llamados multiplicadores de KKT), tales que*

$$\nabla f(\bar{x}) = \sum_{i \in I(\bar{x})} \lambda_i \nabla g_i(\bar{x}),$$

donde la suma es 0 cuando  $I(\bar{x}) = \emptyset$ .

*Demostración.* Dado que  $\bar{x}$  es un mínimo local para  $(P)$ , entonces la condición necesaria de primer orden garantiza  $\nabla f(\bar{x})'d \geq 0$ , para cada  $d \in T(\bar{x})$ ; es decir,  $\nabla f(\bar{x}) \in T(\bar{x})^0$ . Como se satisface una cualificación de restricciones entonces  $T(\bar{x}) = A(\bar{x})^0$ . Así  $T(\bar{x})^0 = A(\bar{x})^{00}$ . Dado que  $A(\bar{x})$  es un cono convexo cerrado, al aplicar el lema de Farkas para conos obtenemos  $A(\bar{x})^{00} = A(\bar{x})$ . Esto implica  $T(\bar{x})^0 = A(\bar{x})^{00} = A(\bar{x})$ . Por lo tanto  $\nabla f(\bar{x}) \in A(\bar{x})$ . □

Los puntos que satisfacen la condición necesaria anterior son llamados *puntos de KKT*. El teorema de Karush-Kuhn-Tucker no es una condición suficiente en general, sin embargo lo es bajo algunas hipótesis de convexidad.

A continuación presentamos algunas propiedades de las funciones convexas:

**Proposición 3.14.** *Consideremos el problema  $(P)$ . Entonces:*

- (i) *Si para todo  $i \in I$ ,  $g_i$  es cóncava, entonces el conjunto factible  $F$  es convexo.*
- (ii) *Si  $f$  es convexa y el conjunto factible es convexo, entonces cualquier mínimo local del problema  $(P)$  es también un mínimo global.*

*Demostración.* (i) El conjunto factible está dado por  $F = \bigcap_{i \in I} F_i$ , donde

$$F_i = \{x \in \mathbb{R}^n : g_i(x) \geq 0\} = \{x \in \mathbb{R}^n : -g_i(x) \leq 0\},$$

el cual es un conjunto convexo según la Proposición 2.33. Así  $F$  es convexo, puesto que es la intersección de conjuntos convexas.

- (ii) Sea  $\bar{x} \in F$  es un mínimo local del problema  $(P)$ . Supongamos que existe  $x^* \in F$  tal que  $f(x^*) < f(\bar{x})$ . Como  $f$  es convexa, para cada  $\lambda \in (0, 1)$  se satisface

$$\begin{aligned} f(\lambda \bar{x} + (1 - \lambda)x^*) &\leq \lambda f(\bar{x}) + (1 - \lambda)f(x^*) \\ &< f(\bar{x}), \end{aligned}$$

lo cual, para  $\lambda$  suficientemente pequeño, contradice la optimalidad local de  $\bar{x}$ . Así, no existe  $x^* \in F$  con  $f(x^*) < f(\bar{x})$ . Concluimos que  $\bar{x}$  es un mínimo global. □

Asumiendo la convexidad de la función  $f$  y la concavidad de las restricciones  $g_i$  que son activas en  $\bar{x}$  probaremos la suficiencia del Teorema de KKT. Las hipótesis de convexidad de las funciones involucradas que asumimos no son necesarias pero en la mayoría de las aplicaciones prácticas de ciertas áreas se cumplen. La demostración realizada en la sección 4.2 de [2] solamente requiere algunos conceptos de convexidad relajados, como pseudoconvexidad y cuasiconvexidad (Ver la referencia [2]).

**Teorema 3.15** (Suficiencia del Teorema de KKT). *Consideremos el problema (P). Si  $\mathbf{x}^*$  es un punto de KKT,  $f$  es convexa y diferenciable en  $\mathbf{x}^*$ , y para cada  $i \in I(\mathbf{x}^*)$  las funciones  $g_i$  son cóncavas y diferenciables en  $\mathbf{x}^*$  entonces  $\mathbf{x}^*$  es una solución óptima global del problema (P).*

*Demostración.* Si  $I(\mathbf{x}^*) \neq \emptyset$ , definamos  $S = \{\mathbf{x} \in \mathbb{R}^n : g_i(\mathbf{x}) \geq 0 \text{ para cada } i \in I(\mathbf{x}^*)\}$ . Claramente  $F \subseteq S$  y además  $S$  es convexo según el primer punto de la Proposición 3.14.

Para cada  $\mathbf{x} \in S$  se satisface  $g_i(\mathbf{x}^*) = 0 \leq g_i(\mathbf{x})$  ya que las restricciones definidas por las funciones  $g_i$  son activas en  $\mathbf{x}^*$ . De la concavidad de las funciones  $g_i$  se tiene

$$g_i(\mathbf{x}) \leq g_i(\mathbf{x}^*) + \nabla g_i(\mathbf{x}^*)'(\mathbf{x} - \mathbf{x}^*),$$

de aquí

$$\nabla g_i(\mathbf{x}^*)'(\mathbf{x} - \mathbf{x}^*) \geq g_i(\mathbf{x}) - g_i(\mathbf{x}^*) \geq 0.$$

De acuerdo con lo anterior,  $\nabla g_i(\mathbf{x}^*)'(\mathbf{x} - \mathbf{x}^*) \geq 0$ , para cada  $i \in I(\mathbf{x}^*)$ . Al multiplicar cada una de estas desigualdades por el multiplicador de KKT  $\lambda_i$  correspondiente obtenemos  $\lambda_i \nabla g_i(\mathbf{x}^*)'(\mathbf{x} - \mathbf{x}^*) \geq 0$ , y sumando sobre  $I(\mathbf{x}^*)$

$$\sum_{i \in I(\mathbf{x}^*)} \lambda_i \nabla g_i(\mathbf{x}^*)'(\mathbf{x} - \mathbf{x}^*) \geq 0. \quad (3.4)$$

Por hipótesis  $\nabla f(\mathbf{x}^*) = \sum_{i \in I(\mathbf{x}^*)} \lambda_i \nabla g_i(\mathbf{x}^*)$ , luego, según (3.4)

$$\nabla f(\mathbf{x}^*)'(\mathbf{x} - \mathbf{x}^*) = \sum_{i \in I(\mathbf{x}^*)} \lambda_i \nabla g_i(\mathbf{x}^*)'(\mathbf{x} - \mathbf{x}^*) \geq 0.$$

Por otra parte, como  $f$  es convexa,  $f(\mathbf{x}) \geq f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)'(\mathbf{x} - \mathbf{x}^*)$ ; y por lo tanto

$$0 \leq \nabla f(\mathbf{x}^*)'(\mathbf{x} - \mathbf{x}^*) \leq f(\mathbf{x}) - f(\mathbf{x}^*), \quad \text{para cada } \mathbf{x} \in S.$$

De lo anterior se concluye  $f(\mathbf{x}^*) \leq f(\mathbf{x})$ , para cada  $\mathbf{x} \in S$  de donde se sigue  $f(\mathbf{x}^*) \leq f(\mathbf{x})$ , para cada  $\mathbf{x} \in F$ .

Si  $I(\mathbf{x}^*) = \emptyset$ , entonces, por el teorema de KKT, tenemos  $\nabla f(\mathbf{x}^*) = \mathbf{0}$ . Luego

$$f(\mathbf{x}) \geq f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)'(\mathbf{x} - \mathbf{x}^*) = f(\mathbf{x}^*),$$

por lo tanto  $f(\mathbf{x}^*) \leq f(\mathbf{x})$ , para cada  $\mathbf{x} \in F$ . □

Según la Proposición 3.12, en programación lineal ordinaria siempre se satisface la igualdad  $T(\bar{\mathbf{x}}) = A(\bar{\mathbf{x}})^0$ . Además las funciones objetivo y restricciones son funciones convexas y cóncavas (puesto que cualquier función lineal lo es), y por lo tanto el teorema de Karush-Kuhn-Tucker es una condición necesaria y suficiente de optimalidad.



El teorema de KKT está enunciado para problemas en los que las restricciones son de la forma  $g_i(\mathbf{x}) \geq 0$ . En el caso de problemas de programación lineal podemos considerar cada restricción  $\mathbf{a}'_i \mathbf{x} \geq b_i$  como  $g_i(\mathbf{x}) = \mathbf{a}'_i \mathbf{x} - b_i \geq 0$ . Así,  $\nabla g_i(\mathbf{x}) = \mathbf{a}_i$ . En este caso el teorema anterior queda enunciado como sigue.

**Teorema 3.16** (Teorema de Karush-Kuhn-Tucker para programación lineal ordinaria). *Sea  $\sigma = (\mathbf{c}, \mathbf{A}, \mathbf{b}) \in \theta$ .  $\mathbf{x}^* \in F^*$  si y sólo si  $\mathbf{c} \in A(\mathbf{x}^*)$ , es decir, existen escalares no negativos  $\lambda_i, i \in I(\mathbf{x}^*)$  tales que*

$$\mathbf{c} = \sum_{i \in I(\mathbf{x}^*)} \lambda_i \mathbf{a}_i$$

donde la suma es  $\mathbf{0}$  cuando  $I(\mathbf{x}^*) = \emptyset$ .

En el caso finito el cono de restricciones activas, así como cualquier cono finitamente generado, es cerrado. Sin embargo en el caso semi-infinito esto no es siempre válido, como lo ilustra el ejemplo siguiente:

**Ejemplo 3.17.** *Consideremos el problema  $\sigma$  dado por:*

$$\begin{aligned} & \text{Minimizar } x_1 + x_2 \\ & \text{sujeto a } x_1 + \frac{1}{n}x_2 \geq 0, \quad \text{para cada } n \in \mathbb{N}. \end{aligned}$$

Claramente  $\mathbf{x}^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \in F$ . Además el conjunto de gradientes de las restricciones activas en  $\mathbf{x}^*$  es  $\left\{ \begin{pmatrix} 1 \\ 1/n \end{pmatrix} \in \mathbb{R}^2 : n \in \mathbb{N} \right\}$ . El cono activo en este punto es

$$A(\mathbf{x}^*) = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : x_1 \geq x_2, x_2 > 0 \right\} \cup \{\mathbf{0}\},$$

el cual no es cerrado como lo ilustra la Figura 3.4.

La cerradura de ciertos conos es una hipótesis fundamental para extender los resultados de programación lineal ordinaria a ciertas clases de problemas lineales semi-infinitos. Según la definición dada en [4], se dice que un parámetro  $\sigma \in \Pi$  es *Farkas-Minkowski* si

$$\text{cone} \left[ \left\{ \begin{pmatrix} \mathbf{a}_t \\ b_t \end{pmatrix}, t \in T \right\} \cup \left\{ \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} \right\} \right]$$

es un cono cerrado. En [8] se ha demostrado que un sistema es Farkas-Minkowski si y solo si cualquier desigualdad que sea consecuencia de dicho sistema es también consecuencia de un subsistema finito.

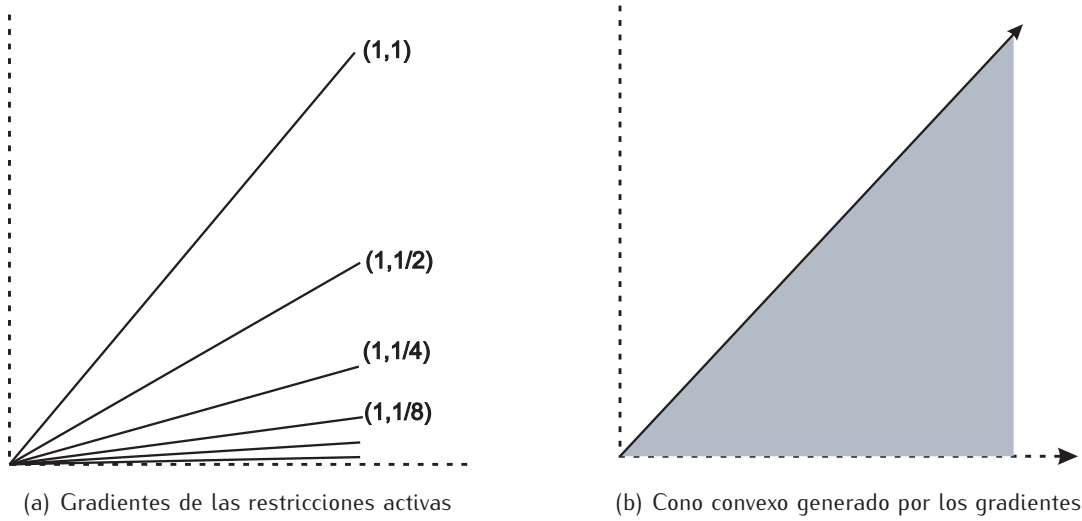


Figura 3.4: Ejemplo de cono convexo no cerrado.

La clase de problemas lineales semi-infinitos del tipo Farkas-Minkowski ha sido estudiada ampliamente por varios autores. En particular [4] extiende a esta clase el Teorema de Karush-Kuhn-Tucker usual.

**Teorema 3.18** (Teorema 2.4 de [4]). *Sea  $\sigma \in \Pi$  un parámetro Farkas-Minkowski y sea  $\mathbf{x}^* \in F$ . Entonces,  $\mathbf{x}^* \in F^*$  si y solo si  $\mathbf{c} \in A(\mathbf{x}^*)$ .*

Otra propiedad del caso finito es que en cualquier punto en la frontera del conjunto factible existe al menos una restricción activa<sup>1</sup>, lo cual no sucede en general en el caso semi-infinito, tal como podemos observar en el ejemplo siguiente:

**Ejemplo 3.19.** *Consideremos el problema siguiente*

$$\begin{aligned} & \text{Minimizar} && -x_1 \\ & \text{sujeto a} && -x_1 \geq -1 - 1/n, \quad n \in \mathbb{N} \\ & && x_1 \geq 0 \\ & && x_2 \geq 0. \end{aligned}$$

El conjunto factible de este problema es  $F = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : 0 \leq x_1 < 1, x_2 \geq 0 \right\}$ . Además

$\mathbf{x}^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \in \partial F \cap F^*$ , pero  $I(\mathbf{x}^*) = \emptyset$ , puesto que  $\mathbf{x}^*$  satisface todas las restricciones de

<sup>1</sup>puesto que el conjunto factible es la intersección finita de semiespacios cerrados.

forma estricta. La Figura 3.5 corresponde a este ejemplo.

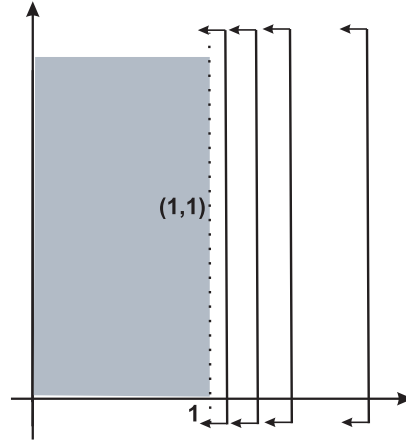


Figura 3.5: Cono activo vacío.

En este ejemplo  $x^* \in F^*$  y  $I(x^*) = \emptyset$ .

El ejemplo anterior muestra que el Karush-Kuhn-Tucker no es válido para el caso semi-infinito usando el concepto de restricciones activas. Los autores de [9] introdujeron el concepto de restricciones activas modificadas, definido como sigue.

**Definición 3.20.** Sea  $\sigma \in \Pi$  tal que  $F \neq \emptyset$ . Dado  $\bar{x} \in F$  el conjunto de restricciones activas modificadas está definido por

$$D(\bar{x}) := \left\{ \mathbf{a} \in \mathbb{R}^n : \begin{pmatrix} \mathbf{a} \\ \mathbf{a}'\bar{x} \end{pmatrix} \in clD \right\},$$

donde  $D = \left\{ \begin{pmatrix} \mathbf{a}_t \\ b_t \end{pmatrix}, t \in T \right\}$ . Es decir,  $\mathbf{a} \in D(\bar{x})$  si y solo si existe una sucesión  $\left\{ \begin{pmatrix} \mathbf{a}_{t_s} \\ b_{t_s} \end{pmatrix} \right\}_{s \in \mathbb{N}}$  en  $D$  tal que

$$\mathbf{a} = \lim_{s \rightarrow \infty} \mathbf{a}_{t_s} \text{ y } \mathbf{a}'\bar{x} = \lim_{s \rightarrow \infty} b_{t_s}.$$

Para cada  $\bar{x} \in F$ ,  $D(\bar{x})$  es un conjunto cerrado que contiene al conjunto de gradientes de las restricciones activas en  $\bar{x}$ . Además en el caso continuo estos dos conjuntos coinciden, es decir,  $D(\bar{x}) = \{\mathbf{a}_t \mid \mathbf{a}'_t \bar{x} = b_t\}$ . Usando el concepto anterior, los autores de [9] extendieron el teorema de Karush-Kuhn-Tucker a la clase de problemas de programación lineal semi-infinita tales que el mapeo  $a : T \rightarrow \mathbb{R}^n$  es acotado:

$$\Lambda := \left\{ \sigma \in \Pi : \sup_{t \in T} \|\mathbf{a}_t\| < \infty \right\}.$$

A continuación mencionamos tal extensión del Teorema de KKT usando el concepto de restricciones activas modificadas.

**Teorema 3.21** (Teorema 3.2 de [9]). *Sea  $\sigma \in \Pi$  tal que  $F \neq \emptyset$ .*

(a) *Sea  $\sigma \in \Lambda$  que satisface la condición fuerte de Slater. Si  $\mathbf{x}^* \in F^*$ , entonces*

$$\mathbf{c} \in \text{cone} \{ \mathbf{a} \in \mathbb{R}^n : \mathbf{a} \in D(\mathbf{x}^*) \}. \quad (3.5)$$

(b) *Si (3.5) se satisface para algún  $\mathbf{x}^* \in F$ , entonces  $\mathbf{x}^* \in F^*$ .*

Se dice que  $\sigma \in \Pi$  satisface la condición fuerte de Slater si existe  $\mathbf{x} \in F$  y  $\epsilon > 0$  tales que, para todo  $t \in T$ ,  $\mathbf{a}_t' \mathbf{x} \geq b_t + \epsilon$ .

Además de la generalización de la condición de KKT para programación lineal semi-infinita mencionada previamente, existe una condición de optimalidad la cual enuncia que en una solución óptima no existen direcciones factibles que sean de descenso:

**Teorema 3.22** (Teorema 7.1 de [8]). *Sea  $\sigma \in \Pi$ .  $\mathbf{x}^* \in F^*$  si y sólo si  $\mathbf{c} \in D(F, \mathbf{x}^*)^0$ .*

*Demostración.* Veamos que  $\mathbf{d} \in D(F, \mathbf{x}^*)$  si y sólo si existe  $\mathbf{x} \in F$  y  $\alpha > 0$  tales que  $\mathbf{d} = \alpha(\mathbf{x} - \mathbf{x}^*)$ . Sea  $\mathbf{d} \in D(F, \mathbf{x}^*)$  entonces existe  $\epsilon > 0$  tal que  $\mathbf{x}^* + \lambda \mathbf{d} \in F$ , para cada  $0 \leq \lambda < \epsilon$ . Sea  $\mathbf{x} := \mathbf{x}^* + \epsilon \mathbf{d}$  entonces  $\mathbf{x} \in F$ . Además  $\mathbf{x} - \mathbf{x}^* = \mathbf{x}^* + \epsilon \mathbf{d} - \mathbf{x}^* = \epsilon \mathbf{d}$  y por lo tanto  $\mathbf{d} = \frac{1}{\epsilon}(\mathbf{x} - \mathbf{x}^*)$ . Recíprocamente, sean  $\mathbf{x} \in F$  y  $\alpha > 0$ . Sea  $\mathbf{d} = \alpha(\mathbf{x} - \mathbf{x}^*)$ , entonces  $\mathbf{x}^* + \frac{1}{\alpha} \mathbf{d} = \mathbf{x} \in F$ , es decir  $\mathbf{d} \in D(F, \mathbf{x}^*)$ .

Tenemos que  $\mathbf{x}^* \in F^*$  si y sólo si

$$\mathbf{c}' \mathbf{x}^* \leq \mathbf{c}' \mathbf{x}, \text{ para todo } \mathbf{x} \in F,$$

si y sólo si

$$\mathbf{c}' \alpha(\mathbf{x} - \mathbf{x}^*) = \alpha(\mathbf{c}' \mathbf{x} - \mathbf{c}' \mathbf{x}^*) \geq 0, \text{ para todo } \mathbf{x} \in F \text{ y } \alpha > 0,$$

si y sólo si

$$\mathbf{c}' \mathbf{d} \geq 0, \text{ para cada } \mathbf{d} \in D(F, \mathbf{x}^*),$$

si y sólo si  $\mathbf{c} \in D(F, \mathbf{x}^*)^0$ . □

Para verificar esta caracterización se requiere conocer el cono  $D(F, \mathbf{x}^*)$  el cual es difícil de determinar en forma explícita. Para reemplazar este cono por otro cono que involucra los datos del problema se usa una cualificación de restricciones. Una cualificación de restricciones es la llamada condición local Farkas-Minkowski, definida a continuación.

**Definición 3.23.** Se dice que el sistema semi-infinito  $\{a'_t x \geq b_t, t \in T\}$  consistente es *localmente Farkas-Minkowski* si cualquier desigualdad que sea consecuencia de dicho sistema que determine un hiperplano de soporte para su conjunto solución  $F$  es también consecuencia de un subsistema finito. Un parámetro  $\sigma \in \Pi$  tal que  $F \neq \emptyset$  es localmente Farkas-Minkowski si su correspondiente sistema de desigualdades lo es.

La propiedad localmente Farkas-Minkowski es una cualificación de restricciones como lo muestra el siguiente resultado:

**Teorema 3.24** (Teorema 5.7 de [8]). *Un parámetro  $\sigma \in \Pi$  tal que  $F \neq \emptyset$  es localmente Farkas-Minkowski si y solo si  $A(x) = D(F, x)^0$ , para cada  $x \in F$ .*

### 3.5. unicidad

En esta sección estudiamos el concepto de unicidad en programación lineal. En seguida veremos un ejemplo en el cual se satisface el Teorema de KKT en un punto de la frontera del conjunto factible y por lo tanto es solución óptima pero no es la única solución óptima.

**Ejemplo 3.25.** *Consideremos el ejemplo siguiente:*

$$\text{Maximizar } x_1 + x_2,$$

*sujeto a las restricciones*

$$x_1 + x_2 \leq 6 \tag{3.6}$$

$$2x_1 + x_2 \leq 8 \tag{3.7}$$

$$x_1 \geq 0$$

$$x_2 \geq 0.$$

*El conjunto factible de este ejemplo está representado en la Figura 3.6. Notemos que la restricción (3.7) es activa en cada punto del segmento cuyos extremos son los puntos  $A = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$  y  $B = \begin{pmatrix} 6 \\ 0 \end{pmatrix}$ . El gradiente de la restricción (3.7) en cada punto de este segmento es  $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$  el cual coincide con el gradiente de la función objetivo. Por lo tanto el conjunto óptimo de este problema es  $F^* = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : 2x_1 + x_2 = 8, 2 \leq x_1 \leq 6, 0 \leq x_2 \leq 4 \right\}$ . Notemos que en cada punto óptimo el gradiente de la restricción activa pertenece a la frontera del cono activo.*

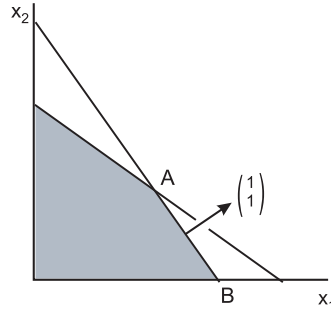


Figura 3.6: La solución no es única.

En general, dado un problema  $\sigma$  de programación lineal ordinaria y un punto factible  $\bar{x}$  de este problema, la condición  $c \in \partial A(\bar{x})$  garantiza la optimalidad de  $\bar{x}$ , pero no la unicidad. En efecto, como  $c \in \partial A(\bar{x})$ , entonces tomando el mayor conjunto linealmente independiente en  $\{\mathbf{a}_i : i \in I(\bar{x})\}$ , digamos  $\{\mathbf{a}_1, \dots, \mathbf{a}_p\}$ , tenemos  $c = \sum_{i=1}^p \lambda_i \mathbf{a}_i$ , donde, si  $p = n$ , no todos los escalares  $\lambda_i$  son positivos<sup>2</sup>. Distingamos dos casos:

1. Si  $p < n$ , entonces cualquier otro punto factible  $x^*$  tal que  $I(\bar{x}) = I(x^*)$  es también una solución óptima, puesto que

$$c' \bar{x} = c' x^* = \sum_{i=1}^p \lambda_i b_i.$$

2. Si  $p = n$ , entonces  $\{\mathbf{a}_1, \dots, \mathbf{a}_p\}$  es una base de  $\mathbb{R}^n$ , y por lo tanto existen escalares  $\alpha_1, \dots, \alpha_n$  tales que  $\bar{x} = \sum_{i=1}^p \alpha_i \mathbf{a}_i$ . Por otra parte, supongamos que en la representación  $c = \sum_{i=1}^p \lambda_i \mathbf{a}_i$  se tiene  $\lambda_k = 0$ . Cualquier punto factible de la forma

$$x^* = \sum_{i=1}^{k-1} \alpha_i \mathbf{a}_i + \alpha_k^1 \mathbf{a}_k + \sum_{i=k+1}^n \alpha_i \mathbf{a}_i,$$

con  $\alpha_k^1 \neq \alpha_k$ , es también una solución óptima, puesto que

$$c' \bar{x} = c' x^* = \sum_{i=1}^{k-1} \lambda_i b_i + \sum_{i=k+1}^n \lambda_i b_i.$$

Para garantizar la unicidad de la solución en programación convexa, se podría fortalecer la hipótesis de convexidad de la función objetivo a convexidad estricta (ver Teorema 3.4.2 de [2]). Es claro que las funciones lineales, aunque son convexas, no son estrictamente convexas;

<sup>2</sup>Si  $p = n$  y todos los escalares son positivos entonces, de acuerdo con el Teorema 2.20,  $\bar{x} \in \text{int}A(\bar{x})$ .

por lo que los criterios de unicidad para programación convexa no son aplicables en el caso lineal.

A continuación presentamos una caracterización de unicidad de la solución en programación lineal ordinaria.

**Teorema 3.26** (Teorema 4.5 de [7]). *Sea  $\sigma = (c, A, b) \in \theta$ .  $x^* \in F$  es la única solución óptima de  $\sigma$  si y sólo si  $c \in \text{int}A(x^*)$ .*

*Demostración.* Supongamos que  $x^* \in F$  es la única solución óptima de  $\sigma$ . Entonces  $c \in A(x^*)$ , según el Teorema 3.16. Si  $c \in \partial A(x^*)$ , como  $A(x^*)$  es un cono convexo cerrado, entonces según el Teorema 2.25, existe un hiperplano de soporte de  $A(x^*)$  en  $c$ ; es decir, existe un vector  $d \neq 0$  tal que

$$d'c = 0 \text{ y } d'x \geq 0, \text{ para todo } x \in A(x^*),$$

de donde  $d \in A(x^*)^0$ . De acuerdo con la Proposición 3.12,  $A(x^*)^0 = D(F, x^*)$  y por lo tanto  $d$  es una dirección factible de  $F$  en  $x^*$ , es decir, existe  $\epsilon > 0$  tal que  $x^* + \lambda d \in F$  para todo  $\lambda \in [0, \epsilon)$ . Por otra parte

$$c'(x^* + \lambda d) = c'x^* + \lambda c'd = c'x^*,$$

lo cual implica que  $x^*$  no es la única solución óptima. Por lo tanto  $c \notin \partial A(x^*)$ , por consiguiente  $c \in \text{int}A(x^*)$ .

Supongamos ahora que  $c \in \text{int}A(x^*)$ . Entonces  $x^*$  es una solución óptima, según el Teorema 3.16. Supongamos que existe otra solución óptima  $\hat{x} \neq x^*$ . Sea  $d := \hat{x} - x^* \neq 0$ . Tenemos que  $d$  es una dirección factible de  $F$  en  $x^*$  puesto que  $x^* + d = \hat{x} \in F$ , luego  $d \in D(F, x^*) = T(x^*) = A(x^*)^0$ , es decir

$$d'x \geq 0, \text{ para cada } x \in A(x^*). \tag{3.8}$$

Por otra parte,

$$d'c = d'(\hat{x} - x^*) = V - V = 0. \tag{3.9}$$

De (3.8) y (3.9) deducimos que el hiperplano  $\{x \in \mathbb{R}^n : d'x = 0\}$  soporta a  $A(x^*)$  en  $c$ , lo cual implica que  $c \in \partial A(x^*)$ . Por lo tanto  $x^*$  es la única solución óptima de  $\sigma$ .  $\square$





## Capítulo 4

# Resultados principales

En este capítulo revisaremos algunas propiedades que caracterizan la estabilidad de los sistemas de desigualdades lineales consistentes y problemas lineales resolubles. En la sección 4.1 daremos una introducción al estudio de los criterios de estabilidad que se han establecido en la literatura para el caso semi-infinito. En el Teorema 4.8 trasladaremos algunos de estos criterios de estabilidad a problemas de programación lineal ordinaria consistentes. En la sección 4.2 demostraremos de forma constructiva que el conjunto de problemas lineales con solución única es denso en el conjunto de problemas lineales resolubles.

### 4.1. Estabilidad

La teoría de estabilidad en programación lineal semi-infinita trata del estudio del comportamiento de las propiedades de un problema cuando se perturban sus datos. Dichas perturbaciones pueden surgir de forma natural al usar aproximaciones, debidas por ejemplo, a la estimación o redondeo de ciertas cantidades involucradas en el problema.

Se dice que un problema que tiene una propiedad especial, como consistencia o resolubilidad, es estable respecto a esta propiedad si los problemas suficientemente “próximos” conservan tal propiedad, es decir, al hacer perturbaciones pequeñas en los datos de un problema nominal, el problema nuevo obtenido también tiene la propiedad del problema nominal.

En lo que sigue presentamos conceptos que usaremos para establecer criterios de estabilidad. Para esto presentamos a continuación las definiciones de multifunción y de continuidad de una multifunción.

**Definición 4.1.** Sean  $X, Y$  dos conjuntos no vacíos. Una *multifunción* de  $X$  en  $Y$  es una correspondencia  $\psi$  entre  $X$  y  $Y$  tal que a cada  $x \in X$  le asigna el subconjunto  $\psi(x)$  de  $Y$ .

Una multifunción  $\psi$  entre  $X$  y  $Y$  será denotada por  $\psi : X \rightrightarrows Y$ .

**Definición 4.2.** Sea  $\psi : X \rightrightarrows Y$  una multifunción entre los espacios topológicos  $X$  y  $Y$ . Se dice que  $\psi$  es *semicontinua inferiormente* en  $x \in X$ , si para cada conjunto abierto  $W \subseteq Y$  con  $W \cap \psi(x) \neq \emptyset$  existe un abierto  $V \subseteq X$  que contiene al punto  $x$  tal que  $W \cap \psi(z) \neq \emptyset$ , para todo  $z \in V$ . Se dice que  $\psi$  es *semicontinua superiormente* en  $x \in X$ , si para cada conjunto abierto  $W \subseteq Y$  tal que  $\psi(x) \subset W$  existe un abierto  $V \subseteq X$  que contiene a  $x$ , tal que  $\psi(z) \subseteq W$ , para cada  $z \in V$ . Si  $\psi$  es continua tanto inferiormente como superiormente en  $x$ , se dice que es continua en  $x$ .

Para estudiar la estabilidad de la consistencia y resolubilidad de los problemas estudiaremos la semicontinuidad inferior y superior de la multifunción conjunto factible y la multifunción conjunto óptimo que definimos a continuación.

**Definición 4.3.** Se define el *mapeo conjunto factible* como la multifunción  $\mathcal{F} : \Pi \rightrightarrows \mathbb{R}^n$  tal que a cada  $\sigma \in \Pi$  le asigna el conjunto factible  $F$  correspondiente al problema definido por  $\sigma$ . Se define la *multifunción conjunto óptimo* como la multifunción  $\mathcal{F}^* : \Pi \rightrightarrows \mathbb{R}^n$  tal que a cada parámetro  $\sigma \in \Pi$  le asigna el conjunto óptimo  $F^*$  correspondiente al problema definido por  $\sigma$ .

Enseguida tenemos la definición de la llamada condición de Slater la cual usaremos como criterio para establecer la estabilidad de los problemas consistentes.

**Definición 4.4.** Sea  $\sigma \in \Pi$ . Se dice que  $\sigma$  satisface la *condición de Slater* si existe  $\bar{x} \in \mathbb{R}^n$  tal que  $\mathbf{a}'_t \bar{x} > b_t$ , para todo  $t \in T$ , y en tal caso se dice que  $\bar{x}$  es un *punto de Slater*. Se dice que  $\sigma$  satisface la *condición fuerte de Slater* si existe  $\epsilon > 0$  y un punto  $\bar{x} \in \mathbb{R}^n$  llamado *punto fuerte de Slater*, tal que  $\mathbf{a}'_t \bar{x} > b_t + \epsilon$ , para todo  $t \in T$ .

Es claro que la condición fuerte de Slater implica la condición de Slater. Para problemas continuos estas condiciones son equivalentes, como lo veremos en el siguiente resultado.

**Proposición 4.5.** Si  $\sigma \in \Pi$  continuo, entonces  $\sigma$  satisface la condición de Slater si y sólo si satisface la condición fuerte de Slater.

*Demostración.* Veamos que la condición de Slater implica la condición fuerte de Slater. Supongamos que  $\sigma$  satisface la condición de Slater, entonces existe  $\bar{x}$  tal que  $\mathbf{a}'_t \bar{x} > b_t$  para cada  $t \in T$ . Como  $\sigma$  es continuo entonces la función  $g : T \rightarrow \mathbb{R}$  definida por  $g(t) := \mathbf{a}'_t \bar{x} - b_t$  es continua sobre el compacto  $T$  y está acotada inferiormente por 0 puesto que  $\bar{x}$  es punto de Slater, por lo tanto existe  $t_0 \in T$  que es mínimo de  $g$ , es decir

$$0 < g(t_0) \leq g(t), \text{ para cada } t \in T. \quad (4.1)$$

Sea  $\epsilon = g(t_0) = \mathbf{a}'_{t_0} \bar{\mathbf{x}} - b_{t_0}$ , de (4.1) se sigue que

$$\mathbf{a}'_t \bar{\mathbf{x}} \geq b_t + \epsilon, \text{ para todo } t \in T,$$

de lo anterior concluimos que  $\bar{\mathbf{x}}$  es punto fuerte de Slater para  $\sigma$ .

□

Para estudiar la estabilidad de las propiedades de problemas lineales consideramos los subconjuntos de  $\Pi$  siguientes; el conjunto de problemas que tienen soluciones factibles

$$\Pi_c := \{\sigma \in \Pi : F \neq \emptyset\},$$

el conjunto de problemas con solución óptima

$$\Pi_s := \{\sigma \in \Pi : F^* \neq \emptyset\},$$

y el conjunto de problemas con solución óptima única

$$\Pi_u := \{\sigma \in \Pi_s : F^* = \{\mathbf{x}^*\}\}.$$

Cada uno de los conjuntos anteriores es un espacio topológico equipado con la topología relativa.

En el Capítulo 10 de [8] se han estudiado las propiedades de semicontinuidad de la multifunción conjunto óptimo, además de la función valor óptimo  $\vartheta : \Pi \rightarrow [-\infty, +\infty]$ , tal que a cada parámetro  $\sigma \in \Pi$  le asigna su valor óptimo  $\vartheta(\sigma)$ .

El capítulo 6 de [8] se estudia la estabilidad de los sistemas de ecuaciones lineales. Entre otros resultados enunciado y probados en [8] nosotros solo utilizaremos las caracterizaciones de  $\text{int}\Pi_c$  que mencionamos a continuación.

**Teorema 4.6** (Teorema 6.1 de [8]). *Sea  $\sigma \in \Pi_c$ . Las siguientes afirmaciones son equivalentes:*

- (a)  $\mathcal{F}$  es semicontinuo inferiormente en  $\sigma$ ;
- (b)  $\sigma \in \text{int}\Pi_c$ ;
- (c)  $\sigma$  satisface la condición fuerte de Slater.

Con respecto a la estabilidad de problemas resolubles, el siguiente resultado caracteriza los problemas que pertenecen al conjunto  $\text{int}\Pi_c$  como los que son estables en cuanto a consistencia y tienen conjunto óptimo acotado y no vacío, esta última condición es equivalente a  $\mathbf{c} \in \text{int cone}\{a_t, t \in T\}$ .

**Teorema 4.7** (Proposición 1 (iv) de [6]).  $\sigma \in \text{int}\Pi_s$  si y solo si  $\sigma \in (\text{int}\Pi_c) \cap \Pi_s$  y  $c \in \text{int cone}\{a_t, t \in T\}$ .

Para trasladar el estudio de estabilidad a programación lineal ordinaria consideramos el conjunto  $\theta$  de los problemas de programación lineal ordinaria que tienen  $n$  variables y  $m$  restricciones. Consideramos los subconjuntos de  $\theta$  denotados con  $\theta_c$ ,  $\theta_s$  y  $\theta_u$  definidos como el conjunto de problemas consistentes, el conjunto de problemas resolubles y el conjunto de problemas con solución única, respectivamente.

En lo que sigue caracterizamos el interior de  $\theta_c$  usando la condición de Slater y la semicontinuidad inferior de la multifunción  $\mathcal{F}$ , estableciendo propiedades análogas a las propiedades establecidas para programación lineal semi-infinita antes mencionadas.

**Teorema 4.8.** Sea  $\sigma \in \theta_c$ . Las siguientes afirmaciones son equivalentes:

- (i)  $\sigma$  satisface la condición de Slater;
- (ii)  $\text{int}F \neq \emptyset$ ;
- (iii)  $\sigma \in \text{int}\theta_c$ ;
- (iv)  $\mathcal{F}$  es semicontinuo inferiormente en  $\sigma$ .

*Demostración.* Primero verificamos (i)  $\Leftrightarrow$  (ii). Supongamos que  $\sigma$  satisface la condición de Slater y sea  $\bar{x}$  un punto de Slater. Para cada  $i \in \{1, 2, \dots, m\}$ ,  $\bar{x}$  es elemento del semiespacio abierto  $\{x \in \mathbb{R}^n : a'_i x > b_i\}$ , por lo tanto existe  $r_i > 0$  tal que  $\bar{x} + r_i B \subseteq \{x \in \mathbb{R}^n : a'_i x > b_i\}$ . Sea  $r = \min\{r_i, i \in I\}$ , entonces  $\bar{x} + rB \subseteq F$ , de aquí  $\bar{x} \in \text{int}F$ .

Recíprocamente, si  $\text{int}F \neq \emptyset$  entonces existe  $\bar{x} \in \text{int}F$ , es decir, existe  $\epsilon > 0$  tal que  $\bar{x} + \epsilon B \subseteq F$ . Supongamos que existe  $i \in \{1, \dots, m\}$  tal que  $a'_i \bar{x} = b_i$ . Entonces el hiperplano  $H = \{x \in \mathbb{R}^n : a'_i x = b_i\} = \{x \in \mathbb{R}^n : a'_i(x - \bar{x}) = 0\}$  soporta al conjunto  $F$  en  $\bar{x}$ , es decir

$$a'_i(x - \bar{x}) \geq 0, \text{ para todo } x \in F. \quad (4.2)$$

Sea  $d = -a_i$ ,  $d$  es una dirección factible en  $\bar{x}$  puesto que para cada  $0 < \lambda < \frac{\epsilon}{\|a_i\|}$

$$\|(\bar{x} + \lambda d) - \bar{x}\| = \|\lambda d\| = \lambda \|a_i\| < \frac{\epsilon}{\|a_i\|} \|a_i\| = \epsilon,$$

es decir,  $\bar{x} + \lambda d \in \bar{x} + \epsilon B \subseteq F$ . Por otra parte

$$a'_i(\bar{x} + \lambda d - \bar{x}) = \lambda a'_i d = -\lambda a'_i a_i = -\lambda \|a_i\|^2 < 0,$$

lo cual se contradice con (4.2). Por lo tanto para todo  $i = 1, \dots, m$  tenemos  $a'_i \bar{x} > b_i$ , es decir,  $\bar{x}$  es un punto de Slater.

Veamos que (i)  $\Leftrightarrow$  (iii). Supongamos que  $\sigma$  satisface la condición de Slater y sea  $\mathbf{x}_0 \in F$  un punto de Slater, entonces  $\mathbf{a}'_i \mathbf{x}_0 > b_i$ , para cada  $i \in I$ . Sean  $\delta = \min\{\mathbf{a}'_i \mathbf{x}_0 - b_i, i \in I\} > 0$  y  $\epsilon = \min\{\frac{\delta}{2}, \frac{\delta}{2n\|\mathbf{x}_0\|_\infty}\} > 0$ . Sea  $\sigma_1 = (\mathbf{c}^1, \mathbf{A}^1, \mathbf{b}^1) \in \sigma + \epsilon B$  arbitrario. Entonces

$$d(\sigma, \sigma_1) = \|\mathbf{A} - \mathbf{A}^1\| + \|\mathbf{b} - \mathbf{b}^1\|_\infty + \|\mathbf{c} - \mathbf{c}^1\|_\infty < \epsilon,$$

lo cual implica que, para todo  $i \in I$ :

$$\|\mathbf{a}_i - \mathbf{a}_i^1\|_\infty < \epsilon, \quad |b_i - b_i^1| < \epsilon,$$

donde  $\mathbf{a}_i$  y  $\mathbf{a}_i^1$  representan la fila  $i$  de la matriz  $\mathbf{A}$  y de la matriz  $\mathbf{A}^1$ , respectivamente.

Por otra parte, si  $\mathbf{x} \in \mathbb{R}^n$  entonces  $\|\mathbf{x}\|_2 \leq \sqrt{n}\|\mathbf{x}\|_\infty$ , luego

$$\begin{aligned} |\mathbf{a}_i^1 \mathbf{x}_0 - b_i^1 - (\mathbf{a}_i \mathbf{x}_0 - b_i)| &= |(\mathbf{a}_i^1 - \mathbf{a}_i) \mathbf{x}_0 + (b_i - b_i^1)| \\ &\leq |(\mathbf{a}_i^1 - \mathbf{a}_i) \mathbf{x}_0| + |b_i - b_i^1| \\ &\leq \|(\mathbf{a}_i^1 - \mathbf{a}_i)\|_2 \|\mathbf{x}_0\|_2 + |b_i - b_i^1| \\ &\leq n \|(\mathbf{a}_i^1 - \mathbf{a}_i)\|_\infty \|\mathbf{x}_0\|_\infty + |b_i - b_i^1| \\ &< n\epsilon \|\mathbf{x}_0\|_\infty + \epsilon \\ &\leq \frac{\delta}{2n\|\mathbf{x}_0\|_\infty} n\|\mathbf{x}_0\|_\infty + \frac{\delta}{2} \\ &= \delta/2 + \delta/2 = \delta, \end{aligned}$$

de aquí

$$\mathbf{a}_i^1 \mathbf{x}_0 - b_i^1 > \mathbf{a}_i \mathbf{x}_0 - b_i - \delta, \quad \text{para todo } i \in I, \quad (4.3)$$

Según la definición de  $\delta$  tenemos que  $\mathbf{a}_i \mathbf{x}_0 - b_i - \delta \geq 0$ , para cada  $i \in I$ ; por lo tanto, (4.3) implica

$$\mathbf{a}_i^1 \mathbf{x}_0 - b_i^1 > 0, \quad \text{para cada } i \in I,$$

es decir,  $\mathbf{x}_0 \in \text{int}F_1$ . Así,  $\sigma_1 \in \theta_c$ , para cualesquiera  $\sigma_1 \in \sigma + \epsilon B$ . Concluimos que  $\sigma \in \text{int}\theta_c$ .

Sea  $\sigma = (\mathbf{c}, \mathbf{A}, \mathbf{b}) \in \text{int}\theta_c$ , entonces existe  $\epsilon > 0$  tal que  $\sigma + \epsilon B \subseteq \theta_c$ . Definamos el problema

$$\sigma_1 = (\mathbf{c}, \mathbf{A}, \mathbf{b}^1),$$

que se obtiene al hacer una perturbación al problema  $\sigma$  en el vector  $\mathbf{b}$ , de la forma siguiente:

$$\mathbf{b}^1 := \mathbf{b} + \bar{\delta},$$

donde  $\bar{\delta} = (\epsilon/2, \dots, \epsilon/2)'$ . Tenemos que

$$d(\sigma, \sigma_1) = \|\mathbf{c} - \mathbf{c}\|_\infty + \|\mathbf{A} - \mathbf{A}\| + \|\mathbf{b} - \mathbf{b}^1\|_\infty = \|\mathbf{b} - \mathbf{b} - \bar{\delta}\|_\infty = \|\bar{\delta}\|_\infty = \epsilon/2 < \epsilon,$$

y por tanto  $\sigma_1 \in \theta_c$ . Luego existe  $\mathbf{x}_1 \in F_1$  tal que

$$\mathbf{a}'_i \mathbf{x}_1 \geq b_i^1 = b_i + \epsilon/2, \quad \text{para cada } i \in I. \quad (4.4)$$

Como  $\epsilon/2 > 0$ , de (4.4) obtenemos

$$\mathbf{a}'_i \mathbf{x}_1 > b_i, \quad \text{para cada } i \in I.$$

La desigualdad anterior implica que  $\mathbf{x}_1$  es un punto de Slater para  $\sigma$ , y por lo tanto  $\sigma$  satisface la condición de Slater.

Veamos que (iii)  $\Leftrightarrow$  (iv). Sea  $\sigma \in \text{int}\theta_c$  y sea  $W \subseteq \mathbb{R}^n$  abierto tal que  $F \cap W \neq \emptyset$ . Sea  $\bar{\mathbf{x}} \in F \cap W$ . Si  $\bar{\mathbf{x}} \notin \text{int}F$ , entonces tomamos  $\bar{\mathbf{y}} \in \text{int}F$  (tenemos que  $\text{int}F \neq \emptyset$  puesto que  $\sigma \in \text{int}\theta_c$ ). Como  $W$  es abierto entonces  $W = \text{int}W$ , luego  $\bar{\mathbf{x}} \in \text{int}W$ . De acuerdo con el Teorema 2.8, existe  $\lambda > 0$  suficientemente pequeño tal que

$$\bar{\mathbf{x}} + \lambda(\bar{\mathbf{y}} - \bar{\mathbf{x}}) \in W \cap \text{int}F.$$

De lo anterior podemos asumir  $\bar{\mathbf{x}} \in \text{int}F \cap W$ . Tenemos ahora que  $\bar{\mathbf{x}} \in W$  es un punto de Slater para  $\sigma$ , entonces  $\bar{\mathbf{x}} \in F_1$  para cualesquiera  $\sigma_1 \in \sigma + \epsilon B$ , donde  $\epsilon = \frac{1}{2} \min\{\frac{\delta}{2}, \frac{\delta}{2n\|\mathbf{x}_0\|_\infty}\}$ . Definiendo  $V := \sigma + \epsilon B \subseteq \theta_c$  concluimos que  $\bar{\mathbf{x}} \in F_1 \cap W$  para cualquiera  $\sigma_1 \in V$ , y por lo tanto  $\mathcal{F}$  es semicontinuo inferiormente en  $\sigma$ .

Supongamos ahora la semicontinuidad inferior de  $\mathcal{F}$  en  $\sigma$ . Tomando  $W = \mathbb{R}^n$  tenemos  $W \cap F \neq \emptyset$  y por lo tanto existe una vecindad de  $\sigma$ ,  $V$ , tal que  $F_1 = F_1 \cap \mathbb{R}^n \neq \emptyset$  para cualquier  $\sigma_1 \in V$ , de donde concluimos  $\sigma \in \text{int}\theta_c$ . □

Enseguida veremos un resultado que usaremos en la demostración del Teorema 4.10.

**Lema 4.9.** *Sea  $\{\mathbf{a}_1, \dots, \mathbf{a}_n\} \subseteq \mathbb{R}^n$  un conjunto linealmente independiente. Si  $\mathbf{a} \in \text{int cone}\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ , entonces existe  $\epsilon > 0$  tal que  $\mathbf{b} \in \text{int cone}\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$  para cualquier conjunto  $\{\mathbf{b}, \mathbf{b}_i, i = 1, \dots, n\} \subseteq \mathbb{R}^n$  tal que  $\|\mathbf{a} - \mathbf{b}\|_\infty < \epsilon$  y  $\max\{\|\mathbf{a}_i - \mathbf{b}_i\|_\infty, i = 1, \dots, n\} < \epsilon$*

*Demostración.* Supongamos que para todo  $\epsilon > 0$  existen  $\mathbf{b}, \mathbf{b}_i \in \mathbb{R}^n$  tales que  $\|\mathbf{a} - \mathbf{b}\|_\infty < \epsilon$ ,  $\max\{\|\mathbf{a}_i - \mathbf{b}_i\|_\infty, i = 1, \dots, n\} < \epsilon$  y  $\mathbf{b} \notin \text{int cone}\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ . Entonces tomando sucesivamente  $\epsilon = 1/n$  podemos construir sucesiones  $\{\mathbf{b}_r\}_{r \in \mathbb{N}}$  y  $\{\mathbf{b}_{ir}\}_{r \in \mathbb{N}}$  tales que  $\mathbf{b}_r \rightarrow \mathbf{a}$ ,  $\mathbf{b}_{ir} \rightarrow \mathbf{a}_i$ ,  $i = 1, \dots, n$  y  $\mathbf{b}_r \notin \text{int cone}\{\mathbf{b}_{1r}, \dots, \mathbf{b}_{nr}\}$ . Si  $\mathbf{a} \in \text{int cone}\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ , entonces existe  $\delta > 0$  tal que  $\mathbf{a} + \delta B \in \text{cone}\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ . Existe  $N \in \mathbb{N}$  suficientemente grande tal que  $\mathbf{a} + \delta B \in \text{cone}\{\mathbf{b}_{1r}, \dots, \mathbf{b}_{nr}\}$  y  $\|\mathbf{a} - \mathbf{b}_r\| < \delta$  para  $r \geq N$ , lo cual no puede suceder. Por lo tanto  $\mathbf{a} \notin \text{int cone}\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ . □

El Teorema 4.7 caracteriza el interior del conjunto de problemas lineales resolubles como el conjunto de los problemas que pertenecen al interior del conjunto de problemas consistentes y tales que su conjunto óptimo es acotado y no vacío. En particular, tenemos que los problemas lineales que pertenecen al interior del conjunto de problemas lineales consistentes y tienen una única solución óptima pertenecen al interior del conjunto de problemas resolubles, tal como veremos en el resultado siguiente.

**Teorema 4.10.**  $\theta_u \cap \text{int}\theta_c \subset \text{int}\theta_s$ .

*Demostración.* Sea  $\sigma \in \theta_u \cap \text{int}\theta_c$ . Sea  $\bar{x}$  la única solución óptima de  $\sigma$ . De acuerdo con el Teorema 3.26 tenemos  $\mathbf{c} \in \text{int}A(\bar{x})$ . El Teorema 2.20 garantiza que existe un conjunto linealmente independiente  $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$  tal que  $\mathbf{c} \in \text{int cone}\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ . Según el Lema anterior, existe  $\epsilon_1 > 0$  tal que para cualesquiera  $\mathbf{c}^1, \mathbf{a}_i^1 \in \mathbb{R}^n$  tales que  $\|\mathbf{a} - \mathbf{c}^1\|_\infty < \epsilon_1$  y  $\max\{\|\mathbf{a}_i - \mathbf{a}_i^1\|_\infty, i = 1, \dots, n\} < \epsilon_1$  implica  $\mathbf{c}^1 + \epsilon_1 B \in \text{int cone}\{\mathbf{a}_1^1, \dots, \mathbf{a}_n^1\}$ . Por otra parte, como  $\theta \in \text{int}\theta_c$ , entonces, según el Teorema 4.8, existe  $\epsilon_2 > 0$  tal que  $\sigma_1 = (\mathbf{c}^1, \mathbf{A}^1, \mathbf{b}^1) \in \theta_c$  para cualesquiera  $\sigma_1 \in \sigma + \epsilon_2 B$ . Definiendo  $\epsilon = \min\{\epsilon_1, \epsilon_2\}$  concluimos que  $\sigma_1 \in \text{int}\theta_c$  y  $\mathbf{c}^1 \in \text{int cone}\{\mathbf{a}_1^1, \dots, \mathbf{a}_n^1\}$ , es decir,  $\sigma_1 \in \theta_s$  para cualesquiera  $\sigma_1 \in \sigma + \epsilon B$ . Por lo tanto  $\sigma \in \text{int}\theta_s$ . □

## 4.2. Resultado principal

En esta sección demostraremos el resultado principal de esta tesis. A continuación ilustramos la demostración partiendo de un ejemplo sencillo que tiene solución al que aproximaremos con un problema que tiene una única solución óptima.

**Ejemplo 4.11.** Consideremos el problema siguiente

$$\begin{aligned} & \text{Minimizar } x_1 \\ & \text{sujeto a } \quad x_1 \geq 1 \\ & \quad \quad \quad -x_1 \geq -1. \end{aligned}$$

Este problema está definido por  $\sigma = (\mathbf{c}, \mathbf{A}, \mathbf{b})$ , donde

$$\mathbf{c} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \mathbf{A} = \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

El conjunto factible de este problema es  $F = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : x_1 = 1 \right\}$ , el cual está representado en la Figura 4.1. Como  $F^* = F$ , este problema no tiene solución única, además  $\sigma \notin \text{int}\theta_c$ , puesto que  $\text{int}F = \emptyset$ .

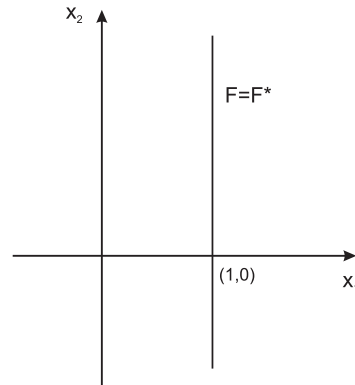


Figura 4.1: Problema inicial, cuyo conjunto solución es infinito.

1. Primero a partir de los datos de  $\sigma$  construimos un problema  $\sigma_1$  tal que  $d(\sigma, \sigma_1) < \epsilon/3$  y  $\sigma_1 \in \text{int}\theta_c$ . Hacemos esto con un desplazamiento en el vector  $\mathbf{b}$ . El problema  $\sigma_1$  resultante es el que sigue:

$$\begin{aligned} & \text{Minimizar } x_1 \\ & \text{sujeto a } \quad x_1 \geq 1 - \epsilon/3 \\ & \quad \quad -x_1 \geq -1 - \epsilon/3. \end{aligned}$$

Este problema está definido por los datos  $\sigma_1 = (\mathbf{c}, \mathbf{A}, \mathbf{b}^1)$ , donde  $\mathbf{b}^1 = \mathbf{b} - \begin{pmatrix} \epsilon/3 \\ \epsilon/3 \end{pmatrix}$ . El

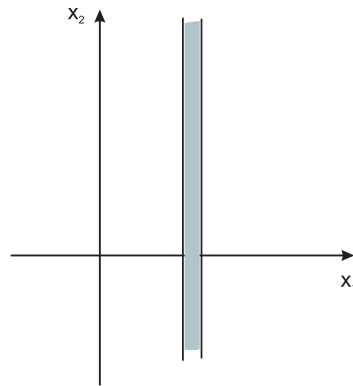


Figura 4.2: Se satisface la condición de Slater (paso 1).

conjunto factible de este problema es

$$F_1 = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : 1 - \epsilon/3 \leq x_1 \leq 1 \right\},$$



el cual está ilustrado en la Figura 4.2. El conjunto óptimo de este problema es

$$F^* = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : x_1 = 1 - \epsilon/3 \right\}.$$

Como  $\text{int}F_1 \neq \emptyset$  entonces  $\sigma^1 \in \text{int}\theta_c$  lo cual garantiza que al hacer pequeñas modificaciones en sus datos no se pierda la factibilidad.

2. A partir del problema  $\sigma_1$  construimos ahora un problema  $\sigma_2 = (c, A^1, b^1)$  modificando los gradientes de las restricciones de  $\sigma_1$  de tal forma que el rango de la matriz  $A^1$  sea igual a 2. Notemos que para esto es necesario que existan al menos 2 restricciones.

$$\begin{aligned} &\text{Minimizar } x_1 \\ &\text{sujeto a } \quad x_1 + \frac{\epsilon}{3}x_2 \geq 1 - \epsilon/3 \\ &\quad \quad \quad -x_1 \geq -1 - \epsilon/3. \end{aligned}$$

Los datos de este problema son  $(c, A^1, b^1)$ , donde  $A^1 = \begin{pmatrix} 1 & \epsilon/3 \\ -1 & 0 \end{pmatrix}$ . Notemos que

$$d(\sigma_1, \sigma_2) = \|c - c\|_\infty + \|A^1 - A\| + \|b^1 - b^2\|_\infty = \epsilon/3.$$

El conjunto factible de este problema es

$$F_2 = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : x_2 \geq -2, 1 - \epsilon/3(1 - x_2) \leq x_1 \leq 1 + \epsilon/3 \right\},$$

representado en la Figura 4.3. Tenemos que  $x^* = \begin{pmatrix} 1 + \epsilon/3 \\ -2 \end{pmatrix} \in F_2$  y además  $\dim A_2(x^*) = 2$ , puesto que  $A_2(x^*)$  es el cono convexo generado por el conjunto linealmente independiente  $\left\{ \begin{pmatrix} 1 \\ \epsilon/3 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right\}$ . Así  $\text{int}A(x^*) \neq \emptyset$ .

3. Modificamos el vector  $c$  en el problema  $\sigma_2$  para obtener el problema  $\hat{\sigma}$  definido por  $\hat{\sigma} = (\hat{c}, A^1, b^1)$  donde  $\hat{c} = \begin{pmatrix} 0 \\ \epsilon/3 \end{pmatrix}$ .

$$\begin{aligned} &\text{Minimizar } \frac{\epsilon}{3}x_2 \\ &\text{sujeto a } \quad x_1 + \frac{\epsilon}{3}x_2 \geq 1 - \epsilon/3 \\ &\quad \quad \quad -x_1 \geq -1 - \epsilon/3. \end{aligned}$$

Tenemos  $\hat{c} \in \text{int}A(x^*) = \text{int cone} \left\{ \begin{pmatrix} 1 \\ \epsilon/3 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right\}$  y por lo tanto  $x^*$  es la única solución óptima de  $\hat{\sigma}$ .

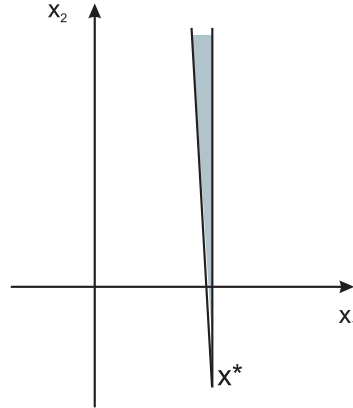


Figura 4.3: Cono activo en  $x^*$  no vacío (paso 2).

En el resultado siguiente demostramos que dado un conjunto de  $n$  vectores en  $\mathbb{R}^n$  podemos aproximarlo tanto como deseemos por un conjunto de  $n$  vectores linealmente independientes en  $\mathbb{R}^n$ .

**Lema 4.12.** *Dado un conjunto  $\{\mathbf{a}_1, \dots, \mathbf{a}_n\} \subseteq \mathbb{R}^n$  y un escalar  $\epsilon > 0$  arbitrarios, existe un conjunto  $\{\mathbf{b}_1, \dots, \mathbf{b}_n\} \subseteq \mathbb{R}^n$  linealmente independiente tal que  $\max\{\|\mathbf{a}_i - \mathbf{b}_i\|_\infty, i = 1, \dots, n\} < \epsilon$ .*

*Demostración.* Sea  $\{\mathbf{b}_1, \dots, \mathbf{b}_k\}$  el mayor subconjunto linealmente independiente de  $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ . Si  $k = n$ , nada hay que probar. Supongamos que  $1 \leq k < n$ . Sea  $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  la base canónica de  $\mathbb{R}^n$ . Inicializando  $j := k$ , realizamos lo siguiente:

1. Sea  $l = \min\{i : \mathbf{e}_i \notin \text{span}\{\mathbf{b}_1, \dots, \mathbf{b}_j\}, i = 1, \dots, n\}$ . El mínimo del conjunto  $\{i : \mathbf{e}_i \notin \text{span}\{\mathbf{b}_1, \dots, \mathbf{b}_j\}\}$  existe puesto que  $j < n$ .
2. Consideremos los dos casos siguientes

(a)  $\mathbf{a}_{j+1} \in \text{span}\{\mathbf{b}_1, \dots, \mathbf{b}_j\}$ . En este caso sea  $\mathbf{b}_{j+1} := \mathbf{a}_{j+1} + \frac{\epsilon}{2}\mathbf{e}_l$ . Entonces

$$\|\mathbf{a}_{j+1} - \mathbf{b}_{j+1}\|_\infty = \left\| \mathbf{a}_{j+1} - \mathbf{a}_{j+1} - \frac{\epsilon}{2}\mathbf{e}_l \right\|_\infty = \epsilon/2 < \epsilon,$$

además  $\mathbf{b}_{j+1} \notin \text{span}\{\mathbf{b}_1, \dots, \mathbf{b}_j\}$ , pues de lo contrario se tendría  $\frac{1}{2}\mathbf{e}_l = \mathbf{b}_{j+1} - \mathbf{a}_{j+1} \in \text{span}\{\mathbf{b}_1, \dots, \mathbf{b}_j\}$  lo cual contradice la elección de  $l$ .

(b)  $\mathbf{a}_{j+1} \notin \text{span}\{\mathbf{b}_1, \dots, \mathbf{b}_j\}$ . En este caso sea  $\mathbf{b}_{j+1} := \mathbf{a}_{j+1}$ .

En cualquier caso  $\|\mathbf{b}_{j+1} - \mathbf{a}_{j+1}\| < \epsilon$  y  $\{\mathbf{b}_1, \dots, \mathbf{b}_{j+1}\}$  es linealmente independiente.

3. Si  $j+1 < n$  hacer  $j = j+1$  y regresar al primer paso. Si  $j+1 = n$ , entonces el conjunto construido  $\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$  es linealmente independiente y para cada  $i \in \{1, \dots, n\}$  se satisface  $\|\mathbf{a}_i - \mathbf{b}_i\|_\infty < \epsilon$ .

Realizando a lo más  $n-1$  veces el procedimiento anterior podemos construir el conjunto deseado.

□

Enseguida enunciamos y demostramos el resultado principal de este trabajo.

**Teorema 4.13.** *El conjunto de problemas de programación lineal ordinaria que tienen solución única  $\theta_u$  es un subconjunto denso del conjunto de problemas de programación lineal ordinaria resolubles  $\theta_s$ .*

*Demostración.* Sea  $\sigma = (\mathbf{c}, \mathbf{A}, \mathbf{b}) \in \theta_s$ , demostraremos que para todo  $\epsilon > 0$  existe  $\hat{\sigma} \in \theta_u$  tal que  $d(\sigma, \hat{\sigma}) < \epsilon$ .

Sea  $\epsilon > 0$  y sea  $\bar{\mathbf{x}} \in F^*$ . Para cada  $i \in I$  sea  $b_i^1 = b_i - \epsilon/3$ . El problema  $\sigma_1 = (\mathbf{c}, \mathbf{A}, \mathbf{b}^1)$  satisface la condición de Slater puesto que  $\mathbf{a}_i^1 \bar{\mathbf{x}} \geq b_i > b_i - \epsilon/3 = b_i^1$ , para cada  $i \in I$ ; por consiguiente, según el Teorema 4.8,  $\sigma_1 \in \text{int}\theta_c$ . Además

$$d(\sigma, \sigma_1) = \|\mathbf{c} - \mathbf{c}\|_\infty + \|\mathbf{A} - \mathbf{A}\| + \|\mathbf{b} - \mathbf{b}^1\|_\infty = \|\mathbf{b} - \mathbf{b}^1\|_\infty = \epsilon/3.$$

Veamos que  $I(\bar{\mathbf{x}}) \neq \emptyset$  para el problema  $\sigma$ . Supongamos que  $I(\bar{\mathbf{x}}) = \emptyset$  entonces  $\mathbf{a}_i^1 \bar{\mathbf{x}} > b_i$ , para cada  $i \in I$ , es decir,  $\bar{\mathbf{x}}$  es un punto de Slater para  $\sigma$ . De acuerdo con el Teorema 4.8,  $\bar{\mathbf{x}} \in \text{int}F$  y según el Teorema 3.1, existe una dirección de descenso en  $\bar{\mathbf{x}}$  lo cual contradice la optimalidad de  $\bar{\mathbf{x}}$  para  $\sigma$ ; por lo tanto,  $I(\bar{\mathbf{x}}) \neq \emptyset$  y por consiguiente  $A(\bar{\mathbf{x}}) \neq \emptyset$ . El Teorema 3.16 garantiza que  $\mathbf{c} \in A(\bar{\mathbf{x}})$ . Sea  $\{\mathbf{a}_1^2, \mathbf{a}_2^2, \dots, \mathbf{a}_k^2\}$  el conjunto de direcciones extremas de  $A(\bar{\mathbf{x}})$ , entonces  $\mathbf{c} \in \text{cone}\{\mathbf{a}_1^2, \mathbf{a}_2^2, \dots, \mathbf{a}_k^2\} = A(\bar{\mathbf{x}})$ . Sea  $\beta = \min\{\frac{\epsilon}{3}, \frac{\delta}{2}, \frac{\delta}{2n\|\bar{\mathbf{x}}\|_\infty}\}$ , donde  $\delta = \min\{\mathbf{a}_i^1 \bar{\mathbf{x}} - b_i, i \in I\} > 0$ . Como  $m \geq n$  podemos tomar  $n-k$  vectores distintos del conjunto  $\{\mathbf{a}_i : i \in I\} \setminus \{\mathbf{a}_1^2, \dots, \mathbf{a}_k^2\}$  y aplicando el procedimiento descrito en el Lema 4.12 con el escalar  $\beta$  podemos completar el conjunto linealmente independiente  $\{\mathbf{a}_1^2, \dots, \mathbf{a}_k^2, \mathbf{a}_{k+1}^2, \dots, \mathbf{a}_n^2\}$ . Entonces el parámetro  $\sigma_2 = (\mathbf{c}, \mathbf{A}^2, \mathbf{b}^1)$ , donde la matriz  $\mathbf{A}^2$  se obtiene de la matriz  $\mathbf{A}$  al reemplazar cada fila  $\mathbf{a}_i$  por  $\mathbf{a}_i^2$ , satisface lo siguiente

(a)  $d(\sigma_1, \sigma_2) = \|\mathbf{c} - \mathbf{c}\|_\infty + \|\mathbf{A} - \mathbf{A}^2\| + \|\mathbf{b}^1 - \mathbf{b}^1\|_\infty = \|\mathbf{A} - \mathbf{A}^2\| < \beta \leq \epsilon/3.$

(b) Según el Teorema 4.8 tenemos que  $\sigma_2 \in \text{int}\theta_c$ , puesto que

$$d(\sigma_1, \sigma_2) < \beta \leq \min\left\{\frac{\delta}{2}, \frac{\delta}{2n\|\bar{\mathbf{x}}\|_\infty}\right\}.$$

La independencia lineal del conjunto  $\{\mathbf{a}_1^2, \dots, \mathbf{a}_n^2\}$  implica la existencia de  $\mathbf{x}^*$ , solución única del sistema  $\{\mathbf{a}_i^{2'} \mathbf{x} = b_i, i = 1, \dots, n\}$ , tal que  $\dim A_2(\mathbf{x}^*) = n$  y por lo tanto,  $\text{int} A_2(\mathbf{x}^*) \neq \emptyset$ . Tenemos que

$$\mathbf{c} \in \text{cone}\{\mathbf{a}_1^2, \dots, \mathbf{a}_k^2\} \subseteq \text{cone}\{\mathbf{a}_1^2, \dots, \mathbf{a}_n^2\} = A(\mathbf{x}^*).$$

De acuerdo con el Corolario 2.19 existe un subconjunto de  $\{\mathbf{a}_1^2, \dots, \mathbf{a}_k^2\}$ , digamos  $\{\mathbf{a}_1^2, \dots, \mathbf{a}_p^2\}$  y escalares positivos  $\lambda_1, \dots, \lambda_p$  tales que

$$\mathbf{c} = \sum_{i=1}^p \lambda_i \mathbf{a}_i^2.$$

Como  $\dim A_2(\mathbf{x}^*) = n$  podemos tomar  $n-p$  vectores linealmente independientes  $\{\mathbf{a}_{p+1}, \dots, \mathbf{a}_n\}$  de  $A_2(\mathbf{x}^*)$ . Sea  $\mu = \epsilon / \left(6 \sum_{i=p+1}^n \|\mathbf{a}_i^2\|_\infty\right)$  y sea  $\mathbf{c}^1 := \mathbf{c} + \mu (\mathbf{a}_{p+1}^2 + \dots + \mathbf{a}_n^2)$ , el Teorema 2.20 garantiza que  $\mathbf{c}^1 \in \text{int} A_2(\mathbf{x}^*)$  y además

$$\|\mathbf{c}^1 - \mathbf{c}\|_\infty = \left\| \mathbf{c} + \mu \sum_{i=p+1}^n \mathbf{a}_i^2 - \mathbf{c} \right\|_\infty = \mu \left\| \sum_{i=p+1}^n \mathbf{a}_i^2 \right\|_\infty \leq \mu \sum_{i=p+1}^n \|\mathbf{a}_i^2\|_\infty = \epsilon/6 < \epsilon/3.$$

Definamos por último el problema  $\hat{\sigma} := (\mathbf{c}^1, \mathbf{A}^2, \mathbf{b}^1)$ . De acuerdo con lo anterior  $\mathbf{c}^1 \in \text{int} A_{\hat{\sigma}}(\mathbf{x}^*)$ , puesto que  $A_2(\mathbf{x}^*) = A_{\hat{\sigma}}(\mathbf{x}^*)$ . El Teorema 3.26 nos permite concluir que  $\mathbf{x}^*$  es solución única de  $\hat{\sigma}$ . Mas aún,

$$d(\hat{\sigma}, \sigma_2) = \|\mathbf{c}^1 - \mathbf{c}\|_\infty + \|\mathbf{A}^2 - \mathbf{A}^2\| + \|\mathbf{b}^1 - \mathbf{b}^1\|_\infty = \|\mathbf{c}^1 - \mathbf{c}\|_\infty < \epsilon/3,$$

se sigue que

$$d(\sigma, \hat{\sigma}) \leq d(\sigma, \sigma_1) + d(\sigma_1, \sigma_2) + d(\sigma_2, \hat{\sigma}) < 3\epsilon/3 = \epsilon.$$

□

En la demostración anterior, aproximamos un problema lineal que tiene al menos una solución óptima con un problema lineal que tiene una única solución óptima y es estable respecto a la propiedad de consistencia.

## Capítulo 5

# Conclusiones

Para demostrar que el conjunto de problemas con una única solución óptima es denso en el conjunto de problemas resolubles, estudiamos la teoría de análisis convexo que sustenta las caracterizaciones de optimalidad y unicidad en programación lineal y las caracterizaciones de estabilidad respecto a las propiedades de consistencia y resolubilidad.

El análisis de estabilidad en el conjunto de problemas de programación lineal ordinaria  $\theta$  considerado en este trabajo nos proporcionó información sobre su estructura e hizo posible conocer algunas relaciones existentes entre subconjuntos distintos de problemas. Así, en el Teorema 4.8 caracterizamos en varios sentidos los problemas que son estables respecto a la propiedad de consistencia. En la demostración del resultado principal, partiendo de un problema con al menos una solución óptima construimos un problema próximo el cual tiene una única solución óptima, además el problema nuevo es estable respecto a consistencia y por lo tanto, como consecuencia del Teorema 4.10, es también estable respecto a la propiedad de resolubilidad.

Muchas propiedades y resultados establecidos en programación lineal ordinaria se pueden extender al conjunto de problemas de programación lineal semi-infinita que son continuos, es decir cuando  $T$  es un conjunto Hausdorff compacto y las funciones  $a : T \rightarrow \mathbb{R}^n$  y  $b : T \rightarrow \mathbb{R}$  son continuas. En la clase de problemas lineales continuos, si un problema tiene soluciones óptimas entonces existe un punto extremo que es solución óptima y en cada punto en la frontera del conjunto factible existe al menos una restricción activa, estas propiedades no se cumplen en programación lineal semi-infinita en general, tal como lo ilustra el Ejemplo 3.19. Algunos autores han generalizado varios resultados al conjunto de problemas acotados donde el conjunto  $T$  no tiene estructura topológica. Un resultado análogo al resultado principal tratado en esta tesis se demostró primero en 1998 en el artículo [9] para la clase de problemas

de programación lineal semi-infinita continuos, posteriormente se generalizó en 2003 en [6] para problemas de programación lineal semi-infinita acotados.

Recientemente se ha introducido el concepto de acotación local [5]. Dado  $\gamma > 0$ , se dice que la desigualdad  $\mathbf{a}'_i \mathbf{x} \geq b_i$  es  $\gamma$ -activa en  $\bar{\mathbf{x}}$  si existe  $\mathbf{y} \in B(\bar{\mathbf{x}}, \gamma)$  tal que  $\mathbf{a}'_t \mathbf{y} = b_t$ . Dado un problema de programación lineal semi-infinita  $\sigma$  y un punto  $\bar{\mathbf{x}} \in F$ , se denota con  $W(\bar{\mathbf{x}}, \gamma)$  al conjunto de los gradientes de las restricciones  $\gamma$ -activas en  $\bar{\mathbf{x}}$ . Se dice que el problema de programación lineal semi-infinita  $\sigma$  es acotado localmente en  $\bar{\mathbf{x}} \in F$  si existe  $\gamma > 0$  tal que el conjunto  $W(\bar{\mathbf{x}}, \gamma)$ . Es un problema abierto generalizar este resultado a la clase de problemas acotados localmente.

# Topología

En este apéndice presentamos los conceptos y la notación de topología de conjuntos que usamos en la tesis. Esta teoría se puede consultar en [10].

**Definición .1.** Sea  $X$  un conjunto. Una *topología*  $\tau$  en el conjunto  $X$  es una colección de subconjuntos de  $X$  que tiene las siguientes propiedades:

- (i)  $\emptyset$  y  $X$  son elementos de  $\tau$ .
- (ii) La unión arbitraria de elementos de  $\tau$  es un elemento de  $\tau$ .
- (iii) La intersección finita de elementos de  $\tau$  es un elemento de  $\tau$ .

Si  $\tau$  es una topología en el espacio  $X$ , se dice que  $X$  es un *espacio topológico* y se denota con  $(X, \tau)$ , o simplemente  $X$ , si no hay confusión sobre la topología  $\tau$ . A los subconjuntos de  $X$  que pertenecen a  $\tau$  se les llama *abiertos*.

**Definición .2.** Sea  $(X, \tau)$  un espacio topológico y sea  $x \in X$ . Se dice que  $V \subseteq X$  es una *vecindad* de  $x$  si existe  $A \in \tau$  tal que  $x \in A \subseteq V$ .

**Definición .3.** Una *base* para una topología en el conjunto  $X$  es una colección  $\mathcal{B}$  de subconjuntos de  $X$  tal que

- (i) Para cada  $x \in X$ , existe al menos un elemento  $B \in \mathcal{B}$  que contiene a  $x$ .
- (ii) Si  $B_1$  y  $B_2$  son dos elementos de  $\mathcal{B}$  y  $x \in B_1 \cap B_2$ , entonces existe  $B_3 \in \mathcal{B}$  tal que  $x \in B_3 \subseteq B_1 \cap B_2$ .

Si  $\mathcal{B}$  es una base para una topología en  $X$ , la *topología  $\tau$  generada por  $\mathcal{B}$*  es la colección de todos los subconjuntos  $U$  de  $X$  tales que para cada  $x \in U$  existe  $B \in \mathcal{B}$  que cumple  $x \in B \subseteq U$ .

**Definición .4.** Sea  $(X, \tau)$  un espacio topológico y sea  $x \in X$ . Una colección de vecindades de  $x$ ,  $\mathcal{B}$ , es una *base local* de  $x$  si para cualesquiera vecindad  $V$  de  $x$  existe  $A \in \mathcal{B}$  tal que  $A \subseteq V$ . Se dice que el espacio  $(X, \tau)$  satisface el *primer axioma de numerabilidad* si cada punto  $x \in X$  tiene una base local numerable.

**Definición .5.** Sea  $(X, \tau)$  un espacio topológico. Se dice que  $X$  es *de Hausdorff* si para todo  $x, y \in X$  con  $x \neq y$  existen dos abiertos  $U, V \in \tau$  tales que  $x \in U, y \in V$  y  $U \cap V = \emptyset$ .

Los espacios topológicos que satisfacen el primer axioma de numerabilidad permiten caracterizar los conceptos convergencia, continuidad, compacidad por medio de sucesiones. La propiedad Hausdorff garantiza la unicidad del límite de cualquier sucesión convergente. Estas propiedades no se satisfacen para cualquier espacio topológico en general.

**Definición .6.** Sea  $X$  un conjunto. Una métrica es una función  $d : X \times X \rightarrow \mathbb{R}$  que cumple las siguientes condiciones

$$(i) \quad d(x, y) \geq 0 \text{ y } d(x, y) = 0 \text{ si y solo si } x = y,$$

$$(ii) \quad d(x, y) = d(y, x),$$

$$(iii) \quad d(x, y) \leq d(x, z) + d(z, y).$$

Si existe una métrica definida en el conjunto  $X$  entonces se dice que  $X$  es un espacio métrico.

**Definición .7.** Sea  $(X, d)$  un espacio métrico. La *bola unitaria*  $B$  en  $X$  está definida por  $B = \{x \in X : d(x, 0) < 1\}$ . Sea  $a \in X$  y  $r \geq 0$ , se define la bola abierta de radio  $r$  con centro en  $a$  como el conjunto  $B(a, r) = \{x \in X : d(x, a) < r\}$ . Si el espacio métrico  $X$  tiene estructura de espacio vectorial entonces la bola abierta con centro en  $a$  y radio  $r$  es el conjunto  $a + rB$ .

Una métrica sobre un espacio induce de forma natural una topología sobre  $X$ , la cual tiene como base al conjunto de todas las bolas abiertas.

Todo espacio métrico  $(X, d)$  equipado con la topología inducida por la métrica  $d$  satisface el primer axioma de numerabilidad y la propiedad Hausdorff, puesto que para cada  $x \in X$ , la colección de conjuntos  $\{y \in X : d(x, y) < \frac{1}{n} : n \in \mathbb{N}\}$  es una base local numerable, además dados  $x, y \in X$  tales que  $x \neq y$  los abiertos  $U := \{z \in X : d(z, x) < \frac{d(x, y)}{2}\}$  y  $V := \{z \in X : d(z, y) < \frac{d(x, y)}{2}\}$  son tales que  $x \in U, y \in V$  y  $U \cap V = \emptyset$ .

**Definición .8.** Sea  $(X, d)$  un espacio métrico y  $Y \subseteq X$ . El *interior* de  $Y$  está definido como sigue:

$$\text{int}Y = \{x \in X : \text{existe } \epsilon > 0 \text{ tal que } B(x, \epsilon) \subseteq Y\},$$



la *frontera* de  $Y$  está dada por:

$$\partial Y = \{x \in X : \text{para todo } \epsilon > 0 \ B(x, \epsilon) \cap Y \neq \emptyset \text{ y } B(x, \epsilon) \cap X \setminus Y \neq \emptyset\}.$$

La *cerradura* o clausura de  $Y$  está dada por

$$clY = \{x \in X : \text{para todo } \epsilon > 0, Y \cap B(x, \epsilon) \neq \emptyset\}.$$

**Definición .9.** Sea  $(X, d)$  un espacio métrico. Se dice que una sucesión  $\{x_n\}$  en  $X$  *converge* a  $p \in X$  si para cada  $\epsilon > 0$  existe  $N \in \mathbb{N}$  tal que  $n \geq N$  implica  $d(x_n, p) < \epsilon$ .

**Definición .10.** Sea  $(X, d)$  un espacio métrico. Un subconjunto  $Y \subseteq X$  es *denso* en  $X$  si  $clY = X$ .

La propiedad de densidad del conjunto  $Y$  en el conjunto  $X$  se puede enunciar también de las siguientes dos formas equivalentes:

- (i)  $Y$  es denso en  $X$  si para todo  $x \in X$  y cualquier  $\epsilon > 0$  existe  $y \in Y$  tal que  $d(x, y) < \epsilon$ .
- (ii)  $Y$  es denso en  $X$  si para todo  $x \in X$  existe una sucesión en  $Y$  que converge a  $x$ .



# Diferenciabilidad

En este apéndice mencionamos las definiciones de función diferenciable y función dos veces diferenciable las cuales se pueden consultar en la referencia [2].

**Definición .11.** Sea  $X \subseteq \mathbb{R}^n$  no vacío,  $\bar{x} \in \text{int}X$  y  $f : X \rightarrow \mathbb{R}$  una función. Se dice que  $f$  es *diferenciable* en  $\bar{x}$  si existe un vector  $\nabla f(\bar{x}) \in \mathbb{R}^n$  llamado *gradiente* de  $f$  en  $\bar{x}$  y una función real  $\alpha$  que satisface  $\alpha(\bar{x}; \mathbf{x}) \rightarrow 0$  cuando  $\mathbf{x} \rightarrow \bar{x}$  tales que

$$f(\mathbf{x}) = f(\bar{x}) + \nabla f(\bar{x})'(\mathbf{x} - \bar{x}) + \|\mathbf{x} - \bar{x}\|\alpha(\bar{x}; \mathbf{x}), \quad \forall \mathbf{x} \in X. \quad (1)$$

El vector  $\nabla f(\bar{x})$  está dado por

$$\nabla f(\bar{x})' = \left( \frac{\partial f(\bar{x})}{\partial x_1}, \frac{\partial f(\bar{x})}{\partial x_2}, \dots, \frac{\partial f(\bar{x})}{\partial x_n} \right).$$

Se dice que  $f$  es *dos veces diferenciable* en  $\bar{x}$  si es diferenciable en  $\bar{x}$  y existe una matriz  $\mathbf{H}(\bar{x})$  llamada *matriz Hessiana* de  $f$  en  $\bar{x}$  y una función  $\alpha$  que satisface  $\alpha(\bar{x}; \mathbf{x}) \rightarrow 0$  cuando  $\mathbf{x} \rightarrow \bar{x}$  tales que para todo  $x \in X$  se satisface

$$f(\mathbf{x}) = f(\bar{x}) + \nabla f(\bar{x})'(\mathbf{x} - \bar{x}) + 1/2(\mathbf{x} - \bar{x})'\mathbf{H}(\bar{x})(\mathbf{x} - \bar{x}) + \|\mathbf{x} - \bar{x}\|^2\alpha(\bar{x}; \mathbf{x}) \quad (2)$$

Si  $\mathbf{H}(\bar{x})$  existe entonces se satisface  $(\mathbf{H}(\bar{x}))_{ij} = \frac{\partial^2 f(\bar{x})}{\partial x_i \partial x_j}$ . La expresión 1 es el *polinomio de Taylor de primer grado* de la función  $f$  en  $\bar{x}$ , mientras que la expresión 2 es el *polinomio de Taylor de segundo grado* de  $f$  en  $\bar{x}$ .



# Bibliografía

- [1] Bazaraa M. S., Jarvis J. J. "Programación Lineal y Flujo en Redes". Limusa, México 1981.
- [2] Bazaraa M. S., Sherali H. D., Shetty C. M. "Nonlinear Programming, Theory and Algorithms". Tercera Edición. John Wiley, 2006.
- [3] Cánovas, M. J. y Parra, J. *Estabilidad en programación semi-infinita lineal: un enfoque cuantitativo*. Boletín de la SEIO Vol 22, número 3 (2006).
- [4] Fisher, T. *Strong Unicity and Alternation for Linear Optimization*, Journal of Optimization Theory and Applications Vol 69 (1991), 251-267.
- [5] Goberna M. A., Lancho A., Todorov, M. I., Vera de Serio V. N. *Locally upper bounded linear semi-infinite systems with applications*, por aparecer en AMO.
- [6] Goberna M. A., López M. A., Todorov, M. I. *A Generic Result in Linear Semi-Infinite Optimization*, Applied Mathematics and Optimization 48 (2003), 181-193.
- [7] Goberna M.A., Puente R., Jornet V. "Optimización Lineal, Teoría, Métodos y Modelos". McGraw-Hill, Madrid, 2004.
- [8] Goberna M.A., López M.A. "Linear Semi-Infinite Optimization". John Wiley, 1998.
- [9] Helbig, S., Todorov, M. I. *Unicity results for general linear semi-infite optimization problems using a new concept of active constraints*, Applied Mathematics and Optimization 38 (1998), 21-43.
- [10] Munkres, J. "Topology". Segunda Edición. Prentice Hall, 2000.
- [11] Rockafellar, R. "Convex Analysis". Princeton University Press, 1970.
- [12] Rodríguez A. M. "Nuevos resultados sobre sistemas lineales y conjuntos convexos". Tesis de doctorado, Universidad de Alicante (2001).