

**UN ALGORITMO PRIMAL PARA EL PROBLEMA CONTINUO
DE PROGRAMACION SEMI-INFINITA LINEAL**

Memoria presentada por Dña. María Teresa León Mendoza para optar al grado de Doctora en Ciencias Matemáticas.

Realizada bajo la dirección de Dña. Maria Enriqueta Vercher González,
Profesora Titular de Estadística e Investigación Operativa de la Facultad de
Matematicas de la Universitat de Valencia.



UMI Number: U607796

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U607796

Published by ProQuest LLC 2014. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

UNIVERSITAT DE VALÈNCIA
BIBLIOTECA CIÈNCIES

Registre 1434

DATA 10-6-92

SIGNATURA T.D. 128
BIBLIOTECA

Nº LIBIS:

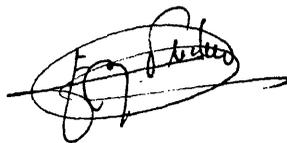
29 ans.

i18957936
b16767172

DÑA. ENRIQUETA VERCHER GONZALEZ, PROFESORA TITULAR DE ESTADISTICA E INVESTIGACION OPERATIVA DE LA FACULTAD DE MATEMATICAS DE LA UNIVERSITAT DE VALENCIA.

CERTIFICA: Que la Memoria titulada "UN ALGORITMO PRIMAL PARA EL PROBLEMA CONTINUO DE PROGRAMACION SEMI-INFINITA LINEAL" ha sido realizada bajo su dirección en el Departamento de Estadística e Investigación Operativa por MARIA TERESA LEON MENDOZA y constituye su Tesis para optar al Grado de Doctor en Ciencias Matemáticas.

Y para que conste, en cumplimiento de la legislación vigente, presenta ante la Facultad de Matemáticas de la Universitat de València la referida Tesis, firmando el presente certificado en Valencia, a 8 de noviembre de mil novecientos noventa y uno.

A handwritten signature in black ink, enclosed within a hand-drawn oval. The signature is cursive and appears to read 'Enriqueta Vercher'.

Fdo: Maria Enriqueta Vercher González.

A mi madre.

Quiero dar las gracias, en primer lugar, a Enriqueta, que no sólo ha sido la directora del trabajo, sino que ha puesto tanto empeño e ilusión como yo misma en que salieran bien las cosas.

También a Paco, Angel, a los dos Vicentes, Juana, Luis, Mario, los dos Enriques, Ramón, Tama, Susa, Guillermo, Antonio, Gudo, Jose, Mari Carmen, Maite, Jose Manuel, María, Maribel y Carmen, mis compañeros del Departamento por conseguir que venir a trabajar a la Facultad sea un placer.

No puedo dejar de citar a mi padre que me inculcó el "gusanillo" de las matemáticas y siempre me ha estimulado a seguir adelante.

Y finalmente, a mi hermana y mis amigos por su interés en mi trabajo y su cariño.

INDICE

CAPITULO I: Introducción a la Programación Semi-infinita Lineal

Introducción	1
1.1 Planteamiento del problema y algunos resultados sobre dualidad	3
1.2 Fundamentos de Programación Lineal infinito dimensional	8
1.3 Una extensión del Simplex para programas continuos semi-infinitos	13
1.4 Aplicaciones de la Programación Semi-infinita	19

CAPITULO II: Condiciones de optimalidad para Programación Semi-infinita Lineal

Introducción	23
2.1 Condiciones de optimalidad	25
2.2 Un test de optimalidad	37
2.3 Cualificaciones de restricciones locales en PSI Lineal	41
2.4 Cualificaciones de restricciones globales en PSI Lineal	46

CAPITULO III: Un algoritmo de purificación

Introducción	53
3.1 Condiciones de recesión	55
3.2 Algoritmo de purificación	65
3.3 Análisis de la convergencia	71
3.4 Resultados numéricos	74

CAPITULO IV: Reglas de descenso para el problema de Programación Semi-infinita Lineal

Introducción	79
4.1 Reglas de descenso desde puntos extremos no degenerados	81
4.2 Reglas de descenso desde puntos extremos degenerados	93
4.3 La solución posible básica inicial	103

APENDICE

A.1 Ejemplos del Capítulo II	110
A.2 Ejemplos del Capítulo III	113
A.3 Ejemplos del Capítulo IV	115
A.4 Ejemplo y gráficas	118

BIBLIOGRAFIA	119
--------------	-----

CAPITULO I

INTRODUCCION A LA PROGRAMACION

SEMI-INFINITA LINEAL

CAPITULO I. Introducción a la Programación Semi-infinita Lineal.

La programación semi-infinita (PSI) puede considerarse como aquella parte de la Programación Matemática que estudia los problemas en los que, siendo finito el número de variables, el número de restricciones es infinito.

Muchos problemas de optimización en Investigación Operativa, Ingeniería, Economía, etc pueden ser formulados como programas semi-infinitos.

El primer trabajo en el que se habla explícitamente de la Programación Semi-infinita (en particular de la PSI lineal) es de Charnes, Cooper y Kortanek (62), quienes utilizaban resultados de Haar (24) sobre sistemas semi-infinitos de inecuaciones lineales para probar un teorema de dualidad fuerte. Aunque ya en los trabajos de John (48), Krein (51) y Rogosinski (58, 62) aparecen referencias implícitas a problemas semi-infinitos.

En los últimos 30 años la programación semi-infinita ha tenido un gran desarrollo, tanto desde el punto de vista teórico como en el computacional (este último a partir de los años 70) y el de las aplicaciones (véanse por ejemplo Hettich (79), Fiacco y Kortanek (83), Anderson y Philpott (85), Polak (87) y Anderson y Nash (87)).

En principio, uno puede plantearse la cuestión de si la formulación semi-infinita es relevante, o si una "inspirada discretización" del problema semi-infinito conduciría a una buena solución óptima del problema original. Realmente si un problema discreto fuera siempre más fácil de tratar no valdría la pena considerar el modelo semi-infinito, pero hay dos aspectos que hacen interesante, en la práctica, este planteamiento:

1) Atendiendo a criterios de fácil manipulación y bajo requerimiento de almacenamiento es preferible un modelo con restricciones dependiente de un parámetro s , formulado a través de una desigualdad del tipo $f(x_1, x_2, \dots, x_n; s) \leq a(s)$, a otro modelo compuesto por un gran número de restricciones no relacionadas entre sí.

2) En el punto óptimo del problema semi-infinito hay únicamente un número finito de restricciones activas. En el caso, bastante usual, de que el conjunto de índices S sea un subconjunto de \mathbb{R}^p , ello determina una alta eficiencia de los llamados "métodos continuos".



El objetivo de esta memoria es desarrollar un método tipo Simplex para resolver el problema continuo de programación semi-infinita lineal. En el Capítulo II hacemos un estudio de las condiciones de optimalidad necesarias para caracterizar las soluciones óptimas. Hemos tratado de conseguir que, tanto el test de optimalidad que presentamos como la cualificación de restricciones que se maneja, sean lo más operativos posible.

A continuación, en el Capítulo III, se introduce un algoritmo de purificación que nos permite ó bien pasar de una solución posible a una solución posible básica, al menos tan buena como la inicial, ó bien detectar la no acotación del problema. Se ha demostrado también que el método converge, a lo sumo, en n etapas.

En el Capítulo IV abordamos el problema de cómo generar direcciones de descenso y tratamos de forma diferente el caso de puntos extremos degenerados y el de no degenerados. Dado que las reglas que presentamos pueden, a veces, conducirnos a soluciones posibles no básicas necesitamos utilizar el algoritmo de purificación del Capítulo III para llegar de nuevo a una solución básica. También tratamos la cuestión de la solución posible básica inicial, utilizando un esquema de dos fases para encontrarla.

En cada capítulo se dan numerosos ejemplos para complementar e ilustrar los desarrollos teóricos que van apareciendo. Además, se han resuelto completamente todos los problemas planteados, mediante la aplicación del método que hemos desarrollado.

En cuanto a este primer capítulo, en la Sección 1.1 introducimos el problema de programación semi-infinita lineal haciendo énfasis en el de tipo continuo, así como algunos resultados fundamentales de Teoría de la dualidad que necesitaremos a lo largo de la memoria. En la Sección 1.2 resumimos los principales aspectos del estudio de Nash (85) sobre la estructura y fundamentos teóricos del programa lineal infinito-dimensional. Este trabajo ha sido decisivo para el desarrollo de los métodos tipo Simplex para el problema primal, y en particular para el nuestro. En la Sección 1.3, tratamos de la aplicación que Anderson y Nash (87) hacen de esta teoría sobre el problema continuo de PSIL, aunque sin incidir en las reglas de descenso que proponen estos autores, pues lo haremos brevemente en el Capítulo IV. Y, por último, en la Sección 1.4 comentamos algunas aplicaciones de la programación semi-infinita.

Sección 1.1 Planteamiento del problema y algunos resultados sobre dualidad.

El problema semi-infinito lineal puede formularse como sigue:

$$(P) \quad \text{Min } c^T x \\ \text{s.a. } a^T(s)x \geq b(s) \quad s \in S$$

donde $c, x, a(s) \in \mathbb{R}^n, \forall s \in S$ y $b(s) \in \mathbb{R}$.

En esta memoria trabajaremos con el llamado problema continuo de PSIL (ver p.e. Anderson y Lewis (89)) y supondremos que las funciones $a_i(s) i=1, \dots, n$ y $b(s)$ son funciones analíticas sobre S , que es un conjunto compacto de \mathbb{R}^p . En el Capítulo II asumiremos simplemente la continuidad de esas funciones.

Aunque existen otros planteamientos del problema dual, vamos a tratar con el problema dual standard del PSIL (Charnes, Cooper y Kortanek (63)):

$$(D) \quad \text{Max } \sum_{s \in S} \lambda(s)b(s) = \Psi(\lambda) \\ \text{s.a. } \sum_{s \in S} \lambda(s)a(s) = c \\ \lambda \in \mathbb{R}_+^{(S)}$$

Siendo $\mathbb{R}_+^{(S)}$ el cono de las sucesiones finitas positivas generalizadas incluido en $\mathbb{R}^{(S)}$, donde $\mathbb{R}^{(S)} = \{ \alpha: S \rightarrow \mathbb{R} \mid \alpha(s) = 0 \quad \forall s \text{ excepto para un subconjunto finito de } S \}$.

La teoría de la dualidad tiene una gran importancia en programación semi-infinita. Las raíces de la teoría de la dualidad en PSI se encuentran en la teoría de sistemas de desigualdades lineales (Duffin (56), Charnes, Cooper y Kortanek (63, 65)), en la teoría de aproximación uniforme de funciones (Duffin y Karlovitz (68)) y en la teoría clásica de momentos (Duffin (67), Gustafson (70), Gustafson y Kortanek (73)). Los problemas de dimensión infinita que aparecían en la práctica fueron los que llevaron a tratar de extender los resultados de la teoría finito-dimensional. Los primeros intentos se referían, más o menos en igual proporción, a la teoría de la dualidad y al desarrollo de extensiones directas del método simplex, más tarde los trabajos se

concentraron en el estudio de la dualidad. En su libro, Anderson y Nash (87) presentan de forma clara y precisa los resultados obtenidos en ambas vertientes.

Recordemos algunos resultados de dualidad para PSI que nos serán de utilidad posteriormente.

Denotaremos por $\Lambda = \{\lambda \in \mathbb{R}_+^{(S)} : \sum \lambda(s)a(s) = c\}$ y $F = \{x \in C \mid a(s)^T x \geq b(s), s \in S\}$ los conjuntos factibles del problema primal y dual, respectivamente. Siendo, a su vez, $V(P) = \inf \{c^T x : x \in F\}$ y $V(D) = \sup \{\Psi(\lambda) : \lambda \in \Lambda\}$ los valores óptimos de los problemas primal y dual.

Un primer resultado de dualidad para PSI, que además nos permite clasificar el par dual (P)-(D), es que la dualidad débil se satisface. Por tanto, para el par dual (P)-(D) se tiene que un problema de optimización tipo (P) puede estar en uno y sólo uno de los siguientes estados (Glashoff y Gustafson (83)):

I.C.: inconsistente, cuando $F = \emptyset$. Por definición tomamos $V(P) = +\infty$.

B: acotado $F \neq \emptyset$ y $V(P)$ finito.

UB: no acotado $F \neq \emptyset$ y $V(P) = -\infty$

Análogamente para un problema del tipo (D):

IC: inconsistente $\Lambda = \emptyset$, tomamos $V(D) = -\infty$

B: acotado $\Lambda \neq \emptyset$ y $V(D)$ finito.

UB: no acotado $\Lambda \neq \emptyset$ y $V(D) = +\infty$

Entonces, la afirmación del lema de dualidad débil puede representarse mediante el siguiente diagrama de estados:

		<u>Primal</u>		
		<u>IC</u>	<u>B</u>	<u>UB</u>
<u>Dual</u>	IC	1	2	4
	B	3	5	x
	UB	6	x	x

El símbolo x señala los casos que no pueden ocurrir. Los casos 1, 2, 3, 4, 5 y 6, que no están excluidos por el teorema de dualidad, pueden aparecer en la práctica (nótese que 2 y 3 no pueden darse en la programación lineal finita). Si $v(P) - v(D) > 0$ diremos que hay un fallo de dualidad.

Hay ejemplos que ponen de manifiesto el hecho de que el dual (D) de un problema de optimización depende no sólo de la función objetivo y del conjunto factible, sino de la formulación de (P), es decir de la forma en que el conjunto de vectores posibles de (P) es descrito a través de desigualdades lineales. Además, Charnes Cooper y Kortanek (62) demostraron que siempre puede encontrarse un sistema de desigualdades lineales $\{a^*(s)^T x \geq b^*(s), \text{ para } s \in S\}$ equivalente al original de forma que no exista fallo de dualidad. El problema es que no hay métodos sencillos que permitan encontrarlo fácilmente, y entonces lo que suele buscarse son condiciones que nos aseguren que no hay fallo de dualidad para un problema de optimización lineal dado.

Para poder dar algunas propiedades del par dual, (P)-(D) nos interesa introducir algunos conceptos y notación.

Definición 1.1

Decimos que (P) es *regular* si tanto (P) como (D) son resolubles (se alcanza el valor óptimo) y $v(P) = v(D)$.

La regularidad de (P) es una propiedad muy útil para cualquier aproximación computacional. Convendrá, por tanto, estudiar bajo qué condiciones ambos problemas tienen solución y se cumple $v(P)=v(D)$. Otro motivo por el que interesa estudiar las propiedades del problema dual es que nos permite comprobar si una solución posible del problema primal es o no óptima. De hecho, si planteamos el sistema no lineal:

$$\sum_{s \in S} \lambda(s)a(s) = c \quad (1)$$

$$a(s)^T x = b(s), \text{ si } \lambda(s) > 0 \quad (2)$$

$$\lambda \in R_+^{(S)}$$

y encontramos una solución $(x^0, \lambda^0(s))$, de forma que x^0 sea factible primal, tendremos resueltos (P) y (D). Nótese que multiplicando cada ecuación de (2) por su correspondiente $\lambda^0(s)$ y sumando, se obtiene la igualdad dual : $c^T x^0 = \sum_{s \in S} \lambda^0(s)b(s)$. Luego x^0 y $\lambda^0(s)$ son soluciones óptimas de (P) y (D), respectivamente. En este caso los valores $\lambda^0(s_i)$, $i=1, \dots, m$ son los multiplicadores de Lagrange asociados a la solución óptima x^0 .



Definición 1.2

Sea A un subconjunto de \mathbb{R}^n . Llamaremos envoltura cónica convexa de A , con la notación $CC(A)$, a la envoltura cónica de la envoltura convexa de A . Es fácil comprobar que:

$$CC(A) = \{z: z = \sum_{i=1}^q \lambda_i a_i \quad \lambda_i \geq 0 \quad i=1, \dots, q \quad a_i \in A \quad i=1, \dots, q, q \geq 1\}$$

Las restricciones del problema primal se expresan en términos del conjunto de vectores $A_S = \{a(s): s \in S\} \subset \mathbb{R}^n$. Ahora bien, puede verse que $\{s_1, \dots, s_q, \lambda_1, \dots, \lambda_q\}$ es solución posible del dual si, y solo si, el vector c puede escribirse como combinación lineal no negativa de los vectores $a(s_1), \dots, a(s_q)$ con coeficientes $\lambda_1, \dots, \lambda_q$. Entonces, (D) tiene soluciones posibles si y solo si $c \in CC(A_S)$. Para abreviar denotaremos por M_n a $CC(A_S)$, que es el cono de momentos del problema de optimización.

Otros resultados interesantes son los siguientes:

Teorema 1.1

Supongamos que $v(D)$ es finito y $c \in \text{int } M_n$. Entonces $v(D) = v(P)$ y (P) tiene solución óptima.

Definición 1.3

Un sistema consistente que posea la propiedad de que toda relación consecuente lo es de un subsistema finito se dice que es un sistema de Farkas-Minkowski.

Nota. La condición $c \in \text{int } M_n$ es suficiente para que un sistema consistente sea de Farkas-Minkowski. Goberna, López y Pastor (81) han caracterizado tales sistemas, que juegan un papel importante en el establecimiento de resultados sobre dualidad y optimización.

Teorema 1.2

El sistema consistente $\{a(s)^T x \geq b(s), s \in S\}$ es de Farkas-Minkowski. si y solo si $CC(\bar{A}_S)$ es un cono cerrado. Siendo: $\bar{A}_S = \{(a(s)^T, b(s))^T : s \in S\}$.

Teorema 1.3

Supongamos que el sistema de restricciones de (P) es de Farkas-Minkowski. Entonces:

- (i) $v(P) = -\infty$ sí y solo sí (D) es inconsistente.
- (ii) si $v(P)$ es finito, entonces $v(P) = v(D)$ y (D) tiene solución óptima.

La dualidad en Programación semi-infinita se estudió para tratar de establecer nuevos métodos numéricos. Una ventaja que proporciona estudiar el problema dual es que permite encontrar cotas inferiores para el valor de (P), pero lo más importante es que las restricciones de (D) están formuladas como igualdades, a diferencia de las restricciones de (P). Esta fue la base de todo un tipo de métodos que resuelven el problema dual en lugar de abordar el problema (P) directamente. Notemos que para transformar las restricciones del problema primal en igualdades (tal como se hace en el caso finito) hay que introducir infinitas variables de "holgura", o una función de holgura, y esta idea no se había considerado hasta Anderson (85). Con lo que no se había intentado una extensión del método Simplex que trabajase directamente con el problema primal, sin hacer previas discretizaciones.

Sección 1.2 Fundamentos de Programación Lineal infinito-dimensional.

En esta sección nos dedicaremos, fundamentalmente, a resumir el estudio de Nash (85) sobre la estructura y fundamentos geométricos del programa lineal infinito-dimensional.

En los últimos años ha habido una cierta tendencia hacia la integración de algunas áreas de la teoría de la optimización, aparentemente distintas, a través del estudio de los problemas planteados en conjuntos abstractos, capaces de abarcar tanto los problemas finitos como los infinitos dimensionales. En particular, será interesante dilucidar en qué medida las propiedades de los problemas lineales dependen del conjunto en que se plantea el problema.

Desde luego, no siempre pueden trasladarse los conceptos de la PL finita al caso infinito-dimensional. Muchas de las técnicas finitas dependen de las propiedades de n , pero, por otro lado, se necesita muy poca estructura algebraica para plantear un problema lineal. De hecho, la formulación de un programa lineal sólo requiere la estructura de espacio vectorial y que se defina la noción de positividad. Veamos a continuación, cual es la situación en el caso infinito:

Sean X y Z espacios vectoriales reales, sea P un cono convexo. Supongamos que X está parcialmente ordenado por la relación \leq , definida por: $x \leq y$ si $y - x \in P$, siendo $x, y \in X$. Si denotamos por θ el vector nulo de cualquier espacio vectorial, entonces $x \in P$ si y solamente si $x \geq \theta$, y diremos que x es positivo.

Sean X^* y Z^* los duales algebraicos de X y Z respectivamente, i.e. el conjunto de todas las funciones lineales sobre esos espacios. Si $x \in X$ y $c^* \in X^*$ denotamos la imagen de x mediante c^* por $\langle c^*, x \rangle$. Puede definirse en X^* un cono positivo $P^* = \{c^* \in X^* : \langle c^*, x \rangle \geq 0, \text{ para todo } x \in P\}$, que también tiene un orden parcial asociado. Podemos definir el programa lineal como sigue:

Sea A una aplicación de X en Z , b un elemento de Z y $c^* \in X^*$, entonces el programa lineal primal es:

$$\begin{aligned} \text{EP:} \quad & \text{Min } \langle c^*, x \rangle \\ \text{s.a.} \quad & Ax = b \\ & x \in X \quad x \geq \theta \end{aligned}$$

Se trata de un problema con restricciones en forma de igualdad, para considerar un problema con desigualdades necesitaríamos introducir un orden en Z y reemplazar $Ax=b$ por $Ax \geq b$. De todas formas, es fácil reducir un problema con desigualdades a uno del tipo EP, introduciendo variables de holgura. Podemos asociar a EP un problema dual:

Sea A^* la aplicación adjunta de A , esta aplicación de Z^* en X^* verifica que $\langle A^*z^*, x \rangle = \langle z^*, Ax \rangle, \forall x \in X, z^* \in Z^*$. Entonces asociamos a EP un problema dual algebraico:

$$\begin{aligned} \text{AD: Max } & \langle z^*, b \rangle \\ \text{s.a. } & A^* z^* \leq c^* \\ & z^* \in Z^* \end{aligned}$$

Es fácil, entonces, establecer el siguiente resultado:

Teorema 1.4

Si x es posible para EP y z^* posible para AD, entonces: $\langle c^*, X \rangle \geq \langle Z^*, b \rangle$. Se da la igualdad dual sí, y sóloamente sí, x y z^* son soluciones óptimas a EP y AD, respectivamente, y entonces, $\langle A^*z^* - c^*, x \rangle = 0$.

El algoritmo del Simplex en programación lineal finita explota esta relacion construyendo en cada etapa una solución del problema primal y una solución complementaria x^* para las restricciones de igualdad del problema dual. Si esta solución dual es posible (o sea, además satisface las restricciones de positividad, i.e. $c^* - A^*z^* = x^* \geq \theta$), entonces ambas son óptimas. Si la solución dual no es posible, entonces es un gradiente reducido para el problema primal, en el sentido de que para x^1 y x^2 soluciones factibles de EP se tiene que $\langle c^*, (x^1 - x^2) \rangle = \langle x^*, (x^1 - x^2) \rangle$, e indica las direcciones de perturbación que mejoraran el valor de la función objetivo.

Las propiedades de las soluciones básicas son esenciales para estas construcciones, así pues, vamos a examinarlas. Cuando se trata de definir solución posible básica para un EP uno se encuentra con algunas dificultades. Si decimos que una solución x es básica si es la única solución a las restricciones, entre el conjunto de funciones cuyo soporte está contenido en el soporte de x , podríamos esperar que las

soluciones básicas incluyan todos los puntos extremos de la región factible, pero esto no ocurre siempre así (como pone de manifiesto un ejemplo debido a Perold (81)). Lo que necesitamos es dar una definición algebraica de solución posible básica.

Definición 1.4

Para cualquier $x \in P$ definimos la solidificación de x como $B(x) = \{\xi \in X \mid x + \lambda \xi \geq \theta, x - \lambda \xi \geq \theta \text{ para algún } \lambda > 0\}$.

Definición 1.5

Si representamos por $N(A)$ el espacio nulo de A , entonces x es básica sí $B(x) \cap N(A) = \{\theta\}$.

El resultado siguiente sirve para demostrar que las soluciones básicas son siempre puntos extremos, mientras que la propiedad de soporte mínimo de las soluciones básicas en PL finita, se convierte en que su solidificación es minimal.

Teorema 1.5

Las siguientes proposiciones son equivalentes:

- (1) x es una solución posible básica.
- (2) x es un punto extremo del conjunto factible.
- (3) x es una solución factible y si ξ es otra solución factible con $B(\xi) \subset B(x) \rightarrow \xi = x$.

Lo que este teorema deja claro es que la propiedad crucial de las soluciones básicas, que cada solución define un subespacio $B(x)$ sobre el cual se resuelven con unicidad las restricciones, es consecuencia del hecho de que ellas son puntos extremos del conjunto factible.

Definición 1.6

Una solución posible básica x es degenerada si $D(x)$ está contenido estrictamente en X , o equivalentemente, si la imagen mediante A de $B(x)$ es más pequeña que AX , donde $D(x) = B(x) \oplus N(A)$.

Podemos apreciar la analogía con la degeneración en PL finita: cuando la base no tiene bastantes elementos o la matriz de la base no es de rango completo.

Una cuestión fundamental que ha de resolverse antes de plantearse un método tipo simplex es la existencia de soluciones básicas óptimas. El problema EP tendrá una solución básica óptima cuando c^* alcance su mínimo sobre un punto extremo del conjunto factible. A partir de condiciones simples y no restrictivas sobre la relación de orden en X , es posible garantizar la existencia de una solución básica óptima con tal de que exista alguna solución óptima. En particular, para los problemas que nos interesan que son el primal continuo y su dual standard podemos asegurar esta propiedad.

La definición de costes reducidos puede extenderse también para el problema general de programación lineal, aunque ahora se hablará de funcional de coste reducido x^* que se define como $x^* = (I - A_B^{-1}A)c^*$, siendo A_B la restricción de A a $B(x)$, que es invertible.

Nótese que x^* está definido sobre $D(x)$ (no necesariamente sobre todo X). Al igual que en PL finita, una solución básica x es óptima si, y solo si, x^* es un funcional positivo sobre $D(x)$. Entonces, el coste reducido proporciona un test de optimalidad de una solución posible básica. Si la solución actual no es óptima, cualquier perturbación posible ξ para la cual $\langle x^*, \xi_N \rangle < 0$ mejorará el valor de la función objetivo; por tanto podremos movernos hacia una mejor solución encontrando un $\xi \in N(x)$ tal que $\langle x^*, \xi \rangle < 0$, añadiendo algún múltiplo positivo de ξ a x , y manteniendo la factibilidad con la suma de algún elemento apropiado de $B(x)$, siendo $N(x)$ un complemento directo de $B(x)$ en $D(x)$.

Lo que conduce a que nos preguntemos si podremos encontrar una perturbación de esta clase que mejore la solución posible básica, y además, si podremos caracterizar estas perturbaciones para que nos ayuden en las operaciones de pivote en el algoritmo del simplex.

Pasamos a discutir, a continuación la operación de pivote. Veamos, en primer lugar, cual es el conjunto de todas las perturbaciones factibles de una solución básica dada x . Sea $K(x)$ el cono de las direcciones factibles en x , definido por:

$$K(x) = \{ \xi \in X : x + \lambda \xi \text{ es factible para algún } \lambda > 0 \}$$

Entonces, puede verse que, si x es una solución posible básica, tenemos que $\xi \in K(x)$ si, y solo si, $\xi = p_N(y)$, donde $y \in D(x) \cap P$, y $p_N(\cdot)$ es la proyección de $D(x)$ sobre $N(A)$. Así, pues, el problema de construir una operación de pivote está en determinar

los elementos de $D(x) \cap P$ para los que se cumple la relación anterior, y elegir entre ellos una dirección que nos conduzca a otra solución básica.

En el simplex finito la operación de pivote corresponde exactamente a este tipo de perturbación. Entre todas las aristas del conjunto factible que surgen de la solución básica actual se elige una de las que produce una mejora y la nueva solución está al final de la arista. Para poder extender esta operación al caso infinito necesitamos que $K(x)$ tenga rayos extremos, que podamos caracterizarlos y que nos conduzcan a una mejor solución.

En el caso finito el orden establecido y la dimensionalidad hacen que esta operación sea fácil. Se forma $N(x)$ como el complemento directo de orden de $B(x)$, es decir $\xi \in D(x)$ es positivo si y solo si ξ_B y ξ_N son positivos. Los problemas para extender esta construcción al caso infinito son: la posibilidad de que no existan rayos extremos en $K(x)$ y que no exista un complemento directo de orden $B(x)$. Incluso en el caso finito la degeneración de x puede significar que solo es posible obtener una fluctuación trivial en una dirección dada. En el caso infinito esta dificultad es importante y habrá que estudiar muy bien el problema de la degeneración.

En la práctica la operación de pivote puede construirse de este modo solo en ciertas clases de problemas, como aquellos que se expresan en espacios de sucesiones. En cambio, cuando son los correspondientes a espacios de funciones continuas, esta clase de construcción es casi siempre degenerada y debe pensarse en otras operaciones de pivote.

Sección 1.3 Una extensión del simplex para programas continuos semi-infinitos.

Anderson y Nash (87) aplican la teoría desarrollada en la sección anterior para espacios infinito- dimensionales, a problemas específicos de programación semi-infinita. Hay dos grandes clases de problemas semi-infinitos lineales: aquellos cuyo conjunto de índices es numerable y aquellos cuyo conjunto de índices es un subconjunto cualquiera, generalmente compacto de \mathbb{R}^P . Se les llama problemas numerables y continuos, respectivamente. Estos autores han desarrollado algoritmos tipo simplex para ambas clases de problemas, en cada caso discutiendo dos algoritmos diferentes, uno para el problema primal y otro para el dual.

Vamos a ver las versiones de las definiciones y los resultados de la sección anterior para el problema primal continuo de PSIL.

$$\begin{aligned} \text{(PSI1) Min } & c^T x \\ \text{s.a. } & a^T(s) x \geq b(s) \quad s \in [0,1] \\ & x \in \mathbb{R}^n \end{aligned}$$

siendo a, b funciones analíticas de $[0,1]$ en \mathbb{R}^n y \mathbb{R} , respectivamente. Añadiendo una función de holgura $z(s)$ tendremos:

$$\begin{aligned} \text{(PSI2) Min } & c^T x \\ \text{s.a. } & a^T(s) x - z(s) = b(s) \quad s \in [0,1] \\ & x \in \mathbb{R}^n \quad z(s) \geq 0, \quad s \in [0,1] \end{aligned}$$

Supongamos que $(x; z)$ es posible para (PSI2), siendo z la función de holgura asociada a x . Definimos $\text{constr}(x) = \{s: a^T(s) x = b(s)\}$ como el conjunto de puntos donde z se anula, i.e. el conjunto de restricciones activas. Como z es analítica, entonces o bien tiene un número finito de ceros o es idénticamente nula. Supondremos que $b(s)$ no está en el espacio generado por $a_i(s)$ $i=1,2,\dots,n$, así que no puede ser idénticamente nula. Entonces $\text{constr}(x)$ es un conjunto finito que denotaremos por $\text{constr}(x) = \{s_1, \dots, s_m\}$. Nos interesa también conocer el orden de multiplicidad de los

ceros de $z(\cdot)$.

Sea el entero $d(i)$, $i=1,2,\dots,n$, definido de forma que:

$$\begin{aligned} z^{(k)}(s_i) &= 0 & k &= 0,1,\dots, d(i) \\ z^{(k)}(s_i) &\neq 0 & k &= d(i) + 1 & i &= 1,2,\dots,m \end{aligned}$$

Cuando $s_i \neq 0$, $s_i \neq 1$, es fácil comprobar que $d(i)$ debe ser un número impar.

Lema 1.6

Sea $v \in C[0,1]$ analítica. Entonces $v \in B(z)$ si, y solamente si, $v^{(j)}(s_i) = 0$, $j=0,1,\dots,d(i)$ $i=1,2,\dots,m$.

Usando este lema podemos caracterizar las soluciones básicas del (PSI2). Para el problema (PSI2) tenemos:

$$B(x,z) = \{ (u,v) : u \in \mathbb{R}^n, v \in C[0,1], v \in B(z) \}$$

$$N(A) = \{ (u,v) : u \in \mathbb{R}^n, v(s) = a^T(s) u, s \in [0,1] \}$$

Sabemos que $(x; z)$ es una solución básica si $B(x; z) \cap N(A) = \{0\}$, entonces x es básica si $a^T(s) u$ en $B(z)$ implica $u=0$.

Teorema 1.7

Si $(x; z)$ es posible para (PSI2), entonces $(x; z)$ es básica si, y solamente si, las filas de

\tilde{A} son linealmente independientes. Siendo \tilde{A} la matriz cuyas columnas son $a^{(j)}(s_i)$ $j=0,1,\dots,d(i)$ $i=1,2,\dots,m$, i.e. $\tilde{A} = [a(s_1), a'(s_1), \dots, a^{(d(1))}(s_1), a(s_2), \dots, a(s_m), a'(s_m), \dots, a^{(d(m))}(s_m)]$.

Teorema 1.8

Una solución posible básica $(x; z)$ para el problema (PSI2) es no degenerada cuando la matriz \tilde{A} es invertible.

Por lo tanto, si $(x; z)$ es una solución básica no degenerada $x = (\tilde{A}^{-1})^T \bar{b}$.

Consideremos el funcional de coste reducido x^* . Para $(u,v) \in \mathbb{R}^n \times A([0,1])$, tenemos que $\langle x^*, (u,v) \rangle = c^T u - c^T (\tilde{A}^{-1})^T q$, donde q es un vector de elementos

$$\frac{\partial^j}{\partial t^j} (a(s_i)^T u - v(s_i)) \quad j=0,1, \dots, d(i) \quad i=1,2, \dots, m. \text{ Entonces: } \langle x^*, (u,v) \rangle = c^T (\tilde{A}^{-1})^T \tilde{v}.$$

Se puede comprobar que para que x^* sea no negativo debe cumplirse que los

elementos de $\bar{A}^{-1}c$ deben ser todos cero, excepto aquellos que corresponden a $v(s_i)$ en \bar{v} que han de ser no negativos.

Reordenamos las columnas de \bar{A} para simplificar la notación, primero $a(s_i)$ $i=1, \dots, m$, seguidas por $a^{(j)}(s_i)$ $j=1, \dots, d(i)$ $i=1, 2, \dots, m$. Además tomamos los elementos de \bar{b} ordenados de la misma forma. Definimos los vectores $\lambda \in \mathbb{R}^m$ y $\rho \in \mathbb{R}^{n-m}$ mediante $\begin{bmatrix} \lambda \\ \rho \end{bmatrix} = \bar{A}^{-1}c$. Entonces se tiene el siguiente resultado:

Teorema 1.9

Una solución básica no degenerada del (PSI2) es óptima si, y solamente si, $\lambda \in \mathbb{R}^m$ es no negativo y $\rho \in \mathbb{R}^{n-m}$ es cero.

Basándose en este teorema y con unas reglas de pivote que comentaremos en el Capítulo IV, Anderson y Lewis (89) han diseñado un algoritmo tipo Simplex para el problema primal. El problema del criterio de optimalidad que acabamos de introducir es que sólo es aplicable sobre puntos extremos no degenerados. En el Capítulo II daremos un criterio que puede utilizarse también para el caso degenerado.

A continuación vamos a resumir la operación de pivote del algoritmo tipo simplex planteado por Goberna y Jornet (88) para el problema dual. Los autores hacen también un estudio teórico que les permite caracterizar los puntos y las direcciones extremas del problema dual standard.

Existen otros algoritmos de este tipo para el problema dual (Anderson y Nash (87)), pero que no emplean la formulación del dual standard, por lo que hemos preferido comentar estos resultados que se ajustan más a nuestro planteamiento del problema. Este procedimiento sigue bastante fielmente el espíritu del simplex finito ya que pasa de un punto extremo a otro adyacente a través de una arista del conjunto factible.

Definición 1.7

Dado T un subconjunto de S , se dice que es un conjunto básico si $\{a(s), s \in T\}$ es una base de \mathbb{R}^n y existe algún $\lambda \in \Lambda$ tal que $\text{supp } \lambda$ está contenido en T .

Obviamente, este λ es único, así como lo es la solución del sistema

$\{a(s)^T x = b(s), s \in T\}$, que denotaremos por x . Para un conjunto básico T , con puntos asociados λ y x , pueden ocurrir dos casos:

Caso 1: $\inf\{a(s)^T x - b(s) : s \in S\} = 0$. De lo contrario,

Caso 2: sea $u \in S - T$ tal que $a(u)^T x - b(u) < 0$, es decir que x no sea factible para el problema (P). Si $T = \{s_1, \dots, s_n\}$, existirán $\alpha_1, \dots, \alpha_n$, únicos tales que $a(u) = \alpha_1 a(s_1) + \dots + \alpha_n a(s_n)$. Definimos $\rho = \delta - \gamma \in R^{(S)}$ siendo:

$$\delta(s) = \begin{cases} 1 & \text{si } s = u \\ 0 & \text{si } s \neq u \end{cases} \quad \text{y} \quad \gamma(s) = \begin{cases} \alpha_i & s = s_i \quad i = 1, \dots, n \\ 0 & s \notin S \end{cases}$$

De nuevo pueden aparecer dos situaciones:

Caso 2.1 : $\alpha_i \leq 0 \quad i = 1, \dots, n$. De lo contrario,

Caso 2.2 : sea $\mu = \min\{\lambda(s_i)/\alpha_i, \alpha_i > 0\}$ y sea $\lambda(s_j)/\alpha_j = \mu \geq 0$. Consideremos, entonces el nuevo conjunto T' dado por: $T' = (T - \{s_j\}) \cup \{u\} \subset S$

Teorema 1.10

Se cumplen las siguientes proposiciones:

(i) En el Caso 1, x y λ son soluciones óptimas de (P) y (D), respectivamente y $c^T x = \Psi(\lambda)$.

(ii) En el Caso 2.1, el problema (P) es inconsistente y el problema (D) es no acotado. Además si $a(u)$ pertenece a la variedad lineal generada por $\{a(s), s \in \text{supp } \lambda\}$, entonces $\{\lambda + \mu \rho, \mu \geq 0\}$ es una arista infinita de Λ .

(iii) En el Caso 2.2, T' es un conjunto básico y su punto asociado de Λ , es $\lambda^1 = \lambda + \mu \rho$. Además si $a(u)$ pertenece a la variedad lineal generada por $\{a(s), s \in \text{supp } \lambda\}$ entonces $\Psi(\lambda^1) < \Psi(\lambda)$ y λ^1, λ son puntos extremos adyacentes.

Nota. En la etapa r -ésima del algoritmo ó bien se detiene (Casos 1 y 2.1) ó bien conseguimos un nuevo conjunto básico $T_{r+1} \neq T_r$. Pero si λ^r es un punto extremo degenerado es posible que $\lambda^{r+1} = \lambda^r$ y no se producirán mejoras en el valor de la función objetivo, aunque también es posible que partiendo de un punto extremo degenerado lleguemos al óptimo. Nótese, además, que un conjunto básico existe si, y sólo si, $\text{rango } \{a(s), s \in S\} = n$.

No se tiene ningún resultado sobre la convergencia finita del método Simplex primal de Anderson y Nash (87), ni del simplex dual de Goberna y Jornet (88).

Actualmente ninguno de los algoritmos está implementado y se encuentran, más bien, en un estadio teórico, aunque se han aplicado sobre diferentes ejemplos.

Además del esfuerzo realizado por Anderson y Nash (87) para presentar posibles implementaciones del simplex para problemas semi-infinitos primales ó duales, numerables ó continuos, han aparecido otros trabajos con un propósito unificador análogo. Uno de los más recientes es el de Rudolph (91), según él hay dos posibles interpretaciones geométricas para la regla de descenso del método Simplex. Ambas interpretaciones son admisibles debido a la correspondencia entre los puntos extremos del conjunto posible y las soluciones básicas de las ecuaciones $Ax=b$:

(i) En la primera interpretación, que él atribuye a Kantorovic, el algoritmo hace que pasemos de un punto extremo del poliedro descrito por las restricciones a otro punto extremo del mismo, mejorando (o al menos sin empeorar) el valor de la función objetivo.

(ii) Para la segunda interpretación, debida a Dantzig, lo que hace el método Simplex es mover un simplex (subcono) en el correspondiente cono de los momentos en la geometría del espacio de columnas de la matriz A que describe las restricciones.

Las extensiones al caso semi-infinito del método Simplex de las que se habla en esta memoria se basan en la primera de estas interpretaciones. Sin embargo, Rudolph opina que la segunda es más natural y esperanzadora con respecto a la eficiencia de la extensión del Simplex.

En su trabajo presenta un algoritmo Simplex semi-infinito para problemas lineales sobre espacios localmente convexos, definiendo de forma general lo que entiende por problema semi-infinito primal (psi) y por problema semi-infinito dual (dsi). Según su esquema, el que para nosotros es el problema primal continuo sería un dsi. Al particularizar su esquema algorítmico a nuestro problema obtenemos algunos resultados contradictorios que nos impiden aplicarlo. Rudolph (91) afirma que su procedimiento es convergente y que una regla de pivote más debil puede generar una sucesión que converja a una solución posible no óptima.

Desde luego, existen otros tipos de procedimientos para resolver los problemas de programación semi-infinita. Siguiendo a Hettich (79) puede decirse que hay 3 formas de encontrar soluciones aproximadas a estos problemas: técnicas de discretización, algoritmos de cambio y métodos contínuos.

Los métodos de discretización generan un subconjunto finito S' de S y encuentran x' que minimiza la función objetivo sobre las restricciones correspondientes a S' . Si el tamaño de la partición escogida es suficientemente pequeño, entonces es probable que x' sea una buena solución óptima aproximada. Los métodos de discretización se aplican a problemas con un conjunto S con una estructura apropiada y a menudo fallan si S es no acotado. A nivel de resultados numéricos, cabe citar los trabajos de Gustafson (70, 71), de Hettich (83, 86) y de Hettich y Zencke (82). Además conviene señalar que existen algunos problemas, como el llamado "problema de satisfabilidad" (ver Polak y He (91)), que sólo conceptualmente son semi-infinitos, y es imprescindible discretizar para poder resolverlos.

Los métodos de tipo semi-continuo (o métodos de cambio) generan una sucesión creciente de subconjuntos finitos de S , $\{S^k\}_{k \geq 1}$ tales que $S^{k+1} \supseteq S^k$ para todo k , y encuentran x^k que minimiza la función objetivo sobre las restricciones correspondientes a S^k , y donde $\{S^k\}_{k \geq 1}$ se construye de modo que cualquier punto de acumulación de la sucesión $\{x^k\}_{k \geq 1}$ es una solución óptima del programa semi-infinito. Para resolver problemas semi-infinitos lineales se han desarrollado, p.e. los algoritmos de Roleff (79) y el algoritmo de las tres fases de Glashoff y Gustafson (83).

En los métodos continuos la aproximación al óptimo se basa en la posibilidad de que bajo determinadas condiciones el problema de programación semi-infinita puede reducirse localmente a un problema finito equivalente. La desventaja de estos métodos es que las hipótesis para su aplicación son, a menudo, bastante restrictivas y casi siempre su convergencia es local.

Algunos procedimientos globalmente convergentes para resolver programas semi-infinitos son: el método exacto de la función de penalización de Conn y Gould (87) y el método lagrangiano proyectado de Coope y Watson (85). En Tanaka et al (88) puede encontrarse un estudio comparativo de varios algoritmos para programación semi-infinita no lineal.

Sección 1.4 Aplicaciones de la Programación Semi-infinita.

El interés de la programación semi-infinita se justifica también por sus múltiples aplicaciones y su conexión con otras teorías matemáticas. Existen numerosos temas en ciencias físicas y sociales que involucran conceptos y modelización con conjuntos infinitos de desigualdades y un número finito de variables. Cabe señalar algunas de ellas: diseño en Ingeniería, desigualdades variacionales y problemas de punto de silla, control parabólico no lineal, diseño de regresión experimental y teoría de momentos, programación lineal continua, programación geométrica, teoría de la decisión secuencial, teoría de conjuntos difusos, etc. Veamos algunas.

La Teoría de la aproximación constituye una de las más importantes aplicaciones de la programación semi-infinita. Es sabido que muchos problemas de aproximación pueden ser planteados como programas semi-infinitos. Veamos, en particular, los problemas de aproximación de Chebyshev y los de aproximación L_1 .

La formulación del problema de aproximación uniforme que se da a continuación puede encontrarse en Anderson y Nash (85). Sea f una función continua definida sobre $[a,b]$ y u_1, \dots, u_n un conjunto de funciones definidas también sobre $[a,b]$ y continuas. Supongamos que deseamos aproximar f mediante una combinación lineal de las u_i , entonces nuestro objetivo será minimizar $\max |f(s)-p(s)|$, siendo $p(s) = \alpha_1 u_1(s) + \alpha_2 u_2(s) + \dots + \alpha_n u_n(s)$. Este problema recibe el nombre de problema de aproximación uniforme o minimax, y puede ser formulado como un problema semi-infinito lineal: se introduce un escalar r , que es una variable, para medir el error máximo y queda

(P_a) Minimizar r

$$\begin{aligned} \text{s.a.} \quad & \sum_{i=1, n} \alpha_i u_i(s) + r \geq f(s) \quad s \in [a, b] \\ & -\sum_{i=1, n} \alpha_i u_i(s) + r \geq -f(s) \quad s \in [a, b]. \end{aligned}$$

Un método muy conocido y eficaz para resolver el problema de aproximación uniforme es el algoritmo de cambio de Remez, que puede aplicarse cuando las funciones $u_i(\cdot)$ $i=1, \dots, n$ satisfacen la condición de Haar (en particular si $u_i(\cdot) = s^{i-1}$, $s \in [a,b]$ se verifica esta condición). Anderson y Nash (85) demuestran que ese

procedimiento puede verse como una extensión del simplex para el dual de (P_a) .

También pueden encontrarse otros métodos para resolver problemas de aproximación uniforme en los trabajos de Watson (75) y Hoffman y Klostermair (76).

El problema de aproximación L_1 puede formularse también como un problema de PSI lineal. En Glashoff y Gustafson (83) puede encontrarse un estudio muy detallado de este problema, que puede enunciarse como sigue: dada una función f continuamente diferenciable y un conjunto de funciones u_1, \dots, u_n que constituyen un sistema de Chebyshev de orden 2 sobre $[a, b]$, se trata de aproximar la función f mediante una combinación lineal $Q = y^T u$ de las funciones u_i $i=1, 2, \dots, n$, de forma que sea mínima $\|Q-f\|$. Esto puede escribirse como el siguiente problema de optimización:

$$\begin{aligned} &\text{minimizar} \quad \int |Q(s)-f(s)| w(s) ds \\ \text{s.a.} \quad &\sum_{i=1, n} x_i u_i(s) \geq f(s) \quad s \in S \end{aligned}$$

Cuando $Q(s) \geq f(s)$ en $[a, b]$, $w(s)$ es una función de peso positiva fija, continua sobre $[a, b]$. Puede formularse equivalentemente como:

$$\begin{aligned} &\text{minimizar} \quad c^T x \\ \text{s.a.} \quad &\sum_{i=1, n} x_i u_i(s) \geq f(s) \quad s \in S \\ &\text{donde } c_i = \int_{s \in S} u_i(s) w(s) ds \end{aligned}$$

Hay numerosos problemas test de este tipo en la bibliografía de la programación semi-infinita lineal sobre los que se prueba la eficacia de diferentes métodos (ver p.e. Ferris y Philpott (89)).

Es muy frecuente el caso en que $S=[0, 1]$, $u_i(s)=s^{i-1}$ y $w(s)=1$. Se tiene entonces que $c_i=1/i$. En el Capítulo IV aplicamos el algoritmo tipo Simplex, que hemos desarrollado, para resolver algunos de estos problemas de aproximación.

Tal como se ha dicho anteriormente, hay una fuerte conexión entre la teoría de la dualidad para PSIL y los problemas de momentos de la teoría de la probabilidad. En los primeros trabajos sobre dualidad en PSI (Charnes, Cooper y Kortanek (63)) ya aparece clara esta conexión, asimismo, en los trabajos de Gustafson (70) y Gustafson,

Kortanek y Rom (70) se tratan clases particulares de problemas de momentos como programas semi-infinitos lineales y se establecen métodos numéricos para la resolución de los mismos. Las conexiones entre ambas teorías han sido posteriormente discutidas por Glashoff (79) y Kemperman (83). Recientemente Kortanek (91) ha utilizado la teoría de la dualidad en PSI para tratar un problema de Inferencia Estadística : el problema de la máxima verosimilitud multinomial.

La PSI ha sido también utilizada en el desarrollo de algoritmos de planos de corte para el problema de programación geométrica, reformulado por Charnes, Cooper y Kortanek (71) como un problema de PSIL (Gustafson y Kortanek (73), Gribik (79)). Gribik y Lee (78) demuestran que pueden obtenerse resultados similares si se aplica el algoritmo central de planos de corte de Elzinga y Moore (75) al programa de programación geométrica original, que si se aplica el algoritmo de Gribik (79) a la formulación semi-infinita. Si en ambos casos para generar el corte se utiliza la restricción más violada, es más profundo el corte producido en el programa semi-infinito, y además el algoritmo de Elzinga y Moore necesita más iteraciones para obtener una solución óptima (para un porcentaje de optimalidad dado).

Aunque los resultados teóricos obtenidos no implican que los algoritmos basados en la programación semi-infinita sean mejores, la experiencia computacional con dieciseis problemas test parece indicar que es así. Recientemente Rajgopal y Bricker (90) han vuelto a dedicarse a este tema y han desarrollado dos algoritmos para programación geométrica que trabajan con programas semi-infinitos lineales.

El problema de transporte semi-infinito ha sido estudiado por Lowe y Hurter (76) y Todd (78) como un problema de optimización infinito-dimensional al que llamaron "problema de transporte generalizado con un área de mercado". Pero no es más que un problema de transporte con un número finito de fuentes y un número de destinatarios (mercado) infinito. Un ejemplo del mismo aparece cuando se trata de efectuar la distribución de un producto desde un conjunto discreto de fuentes a un conjunto de usuarios distribuidos sobre un área .

La Teoría de la dualidad de este problema ha sido discutida por Todd (78) y por Kortanek y Yamasaki (82). Aunque estos últimos autores establecen condiciones de optimalidad para el par de programas duales que extienden los resultados clásicos del problema de transporte finito. Las restricciones impuestas por Todd para desarrollar la

dualidad no se cumplen, de hecho, en el problema clásico de transporte. Por último, Anderson y Nash (87) caracterizan las soluciones básicas del problema de transporte semi-infinito primal y dual y plantean la posibilidad de resolverlo mediante un algoritmo para el dual.

En Polak (87) se pueden encontrar varios ejemplos de aplicación de la PSI para modelizar y resolver problemas de diseño en Ingeniería: diseño de estructuras resistentes a los terremotos, diseño de un sistema de control MIMO (multi-input, multi-output) y diseño de un amplificador de banda ancha. Una de las primeras aplicaciones de la PSI, que prueba la existencia de solución óptima de un problema de control de polución ambiental fue desarrollado por Gorr, Gustafson y Kortanek (72).

Otros trabajos son los que se siguen desarrollando para demostrar que determinados problemas convenientemente reformulados pueden resolverse como un PSIL, proponiendo los correspondientes algoritmos para hacerlo de esta manera :

(a) En el trabajo de Tichatschke, Hettich y Still (89) se presentan resultados de dualidad entre programas lineales generalizados e inexactos y se describe un tipo especial de programas semi-infinitos en conexión con los programas anteriormente mencionados. Para resolver programas inexactos se plantea un programa auxiliar que es resoluble explícitamente, pero, este problema es una reformulación del problema semi-infinito reducido. Por tanto, todos los métodos numéricos para resolver problemas de PSIL pueden usarse para el tratamiento numérico de los problemas lineales inexactos y generalizados.

(b) Anderson, Lewis y Wu (89) consideran un tipo de programa lineal infinito-dimensional, planteado sobre un espacio de medidas, llamado problema de capacidad. Este problema se relaciona con el de determinar la capacidad electrostática de un conductor, y aparece también en ciertos tipos de juegos bipersonales de suma nula.

Se discute la teoría de la dualidad para este problema y se dan condiciones bajo las cuales la solución óptima es una medida con soporte finito. Cuando se restringe el problema al caso de medidas con soporte finito pueden caracterizarse los puntos extremos de la región posible, así es posible aplicar algún método simplex semi-infinito para resolver el problema de capacidad.

CAPITULO II.

CONDICIONES DE OPTIMALIDAD PARA PROGRAMACION SEMI-INFINTA LINEAL

CAPITULO II. Condiciones de Optimalidad para PSIL.

Introducción.

Cuando nos enfrentamos a un problema de optimización nos encontramos con dos cuestiones a resolver : una es caracterizar las soluciones óptimas del problema y la otra diseñar un algoritmo para hallarlas. En este capítulo hemos abordado la primera cuestión , es decir, la caracterización de las soluciones óptimas del problema continuo de PSIL, en los capítulos siguientes presentaremos el algoritmo diseñado para la resolución del problema.

Así, pues, en este capítulo estableceremos una caracterización de las soluciones óptimas del problema continuo de Programación Semi-Infinita Lineal :

$$(PSI1) \quad \text{Min } c^T x \\ \text{s.a. } a^T(s)x \geq b(s) \quad s \in S$$

donde el conjunto S es un compacto de \mathbb{R}^p . Supondremos que las componentes a_1, \dots, a_n de la función a y la función b son continuas en S .

Anderson y Lewis (89) han estudiado este problema para el caso en que a_1, \dots, a_n y b son funciones analíticas en S , y dan una caracterización de los puntos extremos y una condición de optimalidad aplicable a puntos extremos no degenerados, que tengan un número finito de restricciones activas. La condición de optimalidad que han introducido lleva implícita la cualificación de restricciones de independencia lineal de los gradientes activos.

Estos autores señalan la dificultad de establecer un test sencillo de optimalidad en el caso degenerado, pues el coste reducido ya no proporciona información, suficiente tal y como sucede en el caso no degenerado. Sin embargo, sabemos que muchas soluciones posibles básicas en PSIL son degeneradas (Nash (85)). Además, en algunos problemas semi-infinitos, debido a la estructura local de la región factible, los puntos extremos degenerados aparecen agrupados en dominios afines, lo que dificulta encontrar reglas de pivote que permitan evitarlos. Así que, desde luego, necesitamos disponer de criterios de optimalidad válidos también para esta clase de puntos.

La teoría de Kuhn y Tucker ha dado muy buenos resultados en el estudio de la optimalidad en Programación Matemática, aunque es bien sabido que las

correspondientes condiciones necesarias exigen introducir una condición de regularidad sobre las restricciones, salvo en el caso finito lineal. En la Sección 2.1, tras comprobar que los candidatos a óptimos son puntos posibles con alguna restricción activa (no necesariamente puntos de la frontera), hemos establecido una condición necesaria de optimalidad de tipo Kuhn-Tucker para el problema de PSIL. Para ello hemos utilizado una condición de regularidad, extensión de la cualificación de restricciones de Mangasarian-Fromovitz, que juega el papel de una cualificación de restricciones local. Además, hemos demostrado que la caracterización establecida por Anderson y Lewis (89) es un caso particular de esta, cuando se tiene un punto extremo no degenerado.

La caracterización que hemos establecido nos ha permitido desarrollar un sencillo test de optimalidad basado en la resolución de un problema lineal finito, que se presenta en la Sección 2.2. Este test es aplicable a todo punto posible con un número finito de restricciones activas que satisfaga la cualificación de Mangasarian-Fromovitz.

En la Sección 2.3 hemos llevado a cabo un estudio comparativo de la cualificación de Mangasarian-Fromowitz con algunas de las cualificaciones de restricciones locales, que han aparecido en PSI lineal y convexa, siguiendo la línea de Bazaraa y Shetty (76) para Programación no Lineal. También incluimos numerosos ejemplos que clarifican las relaciones entre esta y las demás cualificaciones.

Finalmente, en la Sección 2.4 presentamos un resultado muy interesante que revela el hecho de que si se verifica la cualificación de Mangasarian-Fromovitz en alguna solución posible con un número finito o infinito de restricciones activas, entonces se cumple la cualificación de restricciones de Slater, que es de tipo global. Por tanto, no es necesario volver a comprobar ninguna cualificación de restricciones para los demás puntos posibles.

Si bien esta situación no coincide con la que se presenta en PL finita, nos encontramos con que para una amplia clase de problemas de Programación Semi-infinita lineal no es necesario comprobar ninguna otra cualificación de restricciones. Por ejemplo, para aquellos que tengan al menos un punto extremo no degenerado, ya que según comprobaremos, en este tipo de instancias se satisface la cualificación de Mangasarian-Fromovitz.

Sección 2.1 Condiciones de Optimalidad.

Podemos escribir el problema anterior en forma standard introduciendo la función de holgura $z(s)=a^T(s)x-b(s)$, así el problema (PSI1) puede ser también formulado :

$$\begin{aligned} \text{(PSI2) Min } & c^T x \\ \text{s.a. } & a^T(s)x - z(s) = b(s) \quad s \in S \\ & x \in \mathbb{R}^n \quad z(s) \geq 0 \quad s \in S \end{aligned}$$

Veremos en primer lugar que, al igual que sucede en PL finita, son candidatos a óptimo aquellos puntos posibles para (PSI2) que tienen alguna restricción activa, como se comprueba en el lema siguiente:

Lema 2.1

Sea $(x; z)$ un punto posible para (PSI2). Si $\text{constr}(x)=\emptyset$, entonces $x \in \text{int}(F)$.

Demostración.

Veamos que para cualquier dirección $d \in \mathbb{R}^n$ existe $\bar{x} \in F$ tal que $\bar{x}=x+\lambda d$ con $\lambda > 0$. Dado que $\text{constr}(x)=\emptyset$, entonces $z(s) > 0 \quad s \in S$. S es un compacto, por tanto existe $\varepsilon > 0$ de manera que $z(s) > \varepsilon, s \in S$. Distinguimos dos casos:

- (1) Si $a^T(s)d \geq 0 \quad \forall s \in S$, entonces $\bar{x}=x+\lambda d \in F$ para cualquier $\lambda > 0$.
- (2) Si para d existe algún s tal que $a^T(s)d < 0$, definimos $\varphi = \max\{-a^T(s)d : s \in S\}$, $\varphi > 0$.
Sea $\gamma = \inf\{-z(s)/a^T(s)d : a^T(s)d < 0\} \geq 0$, veamos que $\gamma > 0$.
Para cualquier s tal que $a^T(s)d < 0$ se tiene $-z(s)/a^T(s)d > -\varepsilon/a^T(s)d > \varepsilon/\varphi$, por lo tanto $\gamma \geq \varepsilon/\varphi > 0$. Y por construcción $\bar{x}=x+\gamma d \in F$. ♦

Corolario 2.1.1

Sea $x \in F$ tal que $\text{constr}(x)=\emptyset$, entonces x no es un mínimo .

Demostración.

Según el lema anterior $x \in \text{int}(F)$ y para cualquier dirección d existe $\bar{x} \in F$ tal que $\bar{x}=x+\lambda d$ con $\lambda > 0$, consideremos $d=-c$, d es dirección de mejora ya que $c^T d = -\|c\|^2 < 0$, entonces x no podría ser un óptimo. ♦

Sin embargo, el recíproco del Lema 2.1 no es cierto, como se pone de manifiesto en el siguiente ejemplo:



Ejemplo 2.1

Consideremos el problema siguiente:

$$(P_1) \text{ Min } x_1 + x_2 \\ \text{s.a. } sx_1 + s^2x_2 \geq 0 \quad s \in [0,1]$$

Es fácil comprobar que: $F = \{x \in \mathbb{R}^2: x_1 \geq 0, x_1 + x_2 \geq 0\}$. Sea $\bar{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \in \text{int}(F)$, tenemos que $\text{constr}(\bar{x}) = \{0\} \neq \emptyset$ siendo $a(0) = 0_n$.

Ahora bien, este es el único caso en que se da esta circunstancia. Es decir, si $x \in \text{int}(F)$, su $\text{constr}(x) \neq \emptyset$ y $s \in \text{constr}(x)$, entonces necesariamente $a(s) = 0_n$. En otro caso, $d = -a(s)$ sería una dirección de \mathbb{R}^n tal que $x + \lambda d \notin F$ para ningún valor positivo de λ , lo que contradiría la suposición de que $x \in \text{int}(F)$.

Por otra parte, podríamos preguntarnos qué sucede si hay una solución posible x tal que 0_n es uno de los gradientes activos. Tenemos el siguiente resultado que nos asegura que todos los demás puntos posibles tendrán también el cero como gradiente activo.

Lema 2.2

Si existe un $\bar{s} \in S$ tal que $a(\bar{s}) = 0_n$, entonces ó $\bar{s} \in \text{constr}(x)$, para todo $x \in F$, ó $\bar{s} \notin \text{constr}(x)$, para ningún $x \in F$.

Demostración.

Tomamos una solución posible cualquiera x , su función de holgura será $z(s) = a^T(s)x - b(s)$, $z(s) \geq 0$, para todo $s \in S$. En particular se tiene que $z(\bar{s}) = -b(\bar{s}) \geq 0$. Así que, hay dos posibilidades:

- (i) $b(\bar{s}) = 0 = z(\bar{s})$, y $\bar{s} \in \text{constr}(x)$, o
- (ii) $b(\bar{s}) < 0$ y $z(\bar{s}) > 0$, con lo que $\bar{s} \notin \text{constr}(x)$. ♦

Por tanto, en un problema en el que detectemos una solución posible x , tal que 0_n es un gradiente activo, para todas las demás soluciones posibles ocurrirá lo mismo. Es decir tenemos una restricción redundante, que si es posible es conveniente eliminar.

Desde luego, también existen problemas cuyas soluciones óptimas son tales que el conjunto de gradientes activos se reduce a 0_n , i.e. $A(x) = \{0_n\}$:

Ejemplo 2.2 (en Glashoff y Gustafson (83))

Consideremos el siguiente problema:

$$(P_2) \text{ Min } x_1 \\ \text{s.a. } sx_1 + s^2x_2 \geq s^2 \quad s \in [0,1]$$

Sea $\bar{x} = \begin{bmatrix} 0 \\ 2 \end{bmatrix} \in F$, $z(s) = s^2 \quad s \in [0,1]$ y $\text{constr}(\bar{x}) = \{0\}$. Se tiene que $A(\bar{x}) = \{0_n\}$, veamos que \bar{x} es un óptimo. Notemos que $c^T \bar{x} = 0$ y a continuación comprobaremos que $v(P_2) = 0$ (en realidad, que $v(P_2) \geq 0$).

Para cada $x \in F$, tenemos que $x_1 \geq 0$, pues:

$$sx_1 + s^2x_2 \geq s^2 \quad s \in [0,1] \Leftrightarrow s(x_1 + sx_2) \geq s^2 \quad s \in [0,1] \Leftrightarrow x_1 + sx_2 \geq s \quad s \in (0,1] \Rightarrow x_1 \geq 0$$

por tanto $v(P_2) = 0$ y \bar{x} es óptimo para (P_2) .

Así pues, los candidatos a óptimo serán soluciones posibles con un número finito o infinito de restricciones activas. Vamos a introducir a continuación, una condición necesaria de optimalidad para puntos posibles con un número finito de restricciones activas, basada en el conjunto de gradientes activos y la siguiente condición:

Definición 2.1

Sea $(x; z)$ un punto posible para el (PSI2) tal que $A(x) = \{a(s), s \in \text{constr}(x)\} \neq \emptyset$. Diremos que x satisface la cualificación de restricciones de Mangasarian- Fromovitz (M-F) si existe $d \in \mathbb{R}^n$ tal que $a(s)^T d > 0$ para todo $s \in \text{constr}(x)$.

Nota . Esta formulación, que manejamos a nivel teórico, no es operativa. Por ello es útil comprobar la equivalencia de la cualificación M-F con la condición $0_n \notin \text{co}(A(x))$.

Distinguimos dos casos:

(a) Si $\text{constr}(x)$ es finito ambas condiciones son equivalentes :

Si $(x; z)$ verifica la condición M-F, es decir, existe $d \in \mathbb{R}^n$ tal que $a^T(s_i)d > 0$ $i=1, \dots, m$, siendo $\text{constr}(x) = \{s_1, \dots, s_m\}$, entonces por el Teorema de Gordan (vease p.e. Mangasarian (69)) el sistema (1) es inconsistente:

$$(1) \quad \begin{aligned} 0_n &= \sum_{i=1}^m \alpha_i a(s_i) \\ \alpha_i &\geq 0 \quad i=1, \dots, m \end{aligned}$$

o, lo que es lo mismo, $0_n \notin \text{co}(A(x))$.

(b) Si el cardinal de $\text{constr}(x)$ es infinito la demostración se hace aplicando el Teorema de Gordan generalizado, necesitando comprobar previamente que el conjunto $\text{co}\{A(x)\}$ es cerrado:

(1) El conjunto de índices activos $\text{constr}(x)$ es cerrado por ser la antiimagen mediante una función continua de un cerrado: $\text{constr}(x) = \{s \in S : z(s) = 0\} = z^{-1}(\{0\})$. Además se trata de un subconjunto de S , compacto, así pues $\text{constr}(x)$ es compacto.

La función a es continua, por tanto, $a(\{\text{constr}(x)\}) = A(x)$ es un compacto y $\text{co}\{A(x)\}$ es un compacto.

(2) Puesto que el conjunto $\text{co}\{A(x)\}$ es cerrado, es posible aplicar el Teorema de Gordan generalizado (López y Vercher (83)) y se tiene, de nuevo, que el cumplimiento de la cualificación M-F es equivalente a que $0_n \notin \text{co}(A(x))$.

Para un punto posible x , la condición $0_n \notin \text{co}(A(x))$ caracteriza las denominadas soluciones normales para problemas de optimización con infinitas restricciones (ver Hestenes (75)).

Teorema 2.3

Sea $(x; z)$ un punto posible para el (PSI2) cuyo $\text{constr}(x) = \{s_1, s_2, \dots, s_m\}$. Supongamos que verifica la cualificación M-F. Entonces $(x; z)$ es óptimo para el (PSI2) si, y sólo si, existen $\lambda_i \geq 0, i=1, \dots, m$ tales que $c = \sum_{i=1}^m \lambda_i a(s_i)$.

Demostración.

(\leftarrow) Aplicamos el Teorema de Holgura Complementaria para probar esta implicación. Sea x una solución posible para (PSI1) cuya función de holgura es $z(s) = a^T(s)x - b(s)$. Si suponemos que $c = \sum_{i=1}^m \lambda_i a(s_i)$, entonces $\{s_1, s_2, \dots, s_m, \lambda_1, \lambda_2, \dots, \lambda_m\}$ es una solución posible para el (PSI1*) y puesto que $\lambda_i (a^T(s_i)x - b(s_i)) = \lambda_i z(s_i) = 0 \quad i=1, \dots, m$, del Teorema de Holgura Complementaria (Glashoff-Gustafson (83)) se sigue que x es solución óptima para (PSI1) y $\{s_1, s_2, \dots, s_m, \lambda_1, \lambda_2, \dots, \lambda_m\}$ solución óptima para (PSI1*). Además $(x; z)$ es solución óptima para el (PSI2).

(\rightarrow) Lo demostraremos por reducci3n al absurdo. Supongamos que no existen $\lambda_1, \lambda_2, \dots, \lambda_m \geq 0$ tales que $c = \sum_{i=1}^m \lambda_i a(s_i)$, entonces el sistema (I) es inconsistente:

$$(I) \quad \begin{aligned} c &= \sum_{i=1}^m \lambda_i a(s_i) \\ \lambda_i &\geq 0 \quad i=1, \dots, m \end{aligned}$$

Aplicando el Lema de Farkas se tiene que existe $u \in \mathbb{R}^n$ tal que $a^T(s_i)u \geq 0 \quad i=1, \dots, m$ y $c^T u < 0$. Por otra parte $(x; z)$ verifica M-F, as3i pues, existe $d \in \mathbb{R}^n$ tal que $a^T(s_i)d > 0 \quad i=1, \dots, m$.

Sea $y = u + \gamma d$, $\gamma > 0$, entonces $c^T y = c^T u + \gamma c^T d$ y eligiendo γ suficientemente peque1o podemos asegurar que $c^T y < 0$. Por tanto y es una direcci3n de mejora. Adem3s $a^T(s_i)y = a^T(s_i)u + \gamma a^T(s_i)d > 0 \quad i=1, \dots, m$. [1]

Vamos a comprobar que podemos movernos desde x en la direcci3n y . Definimos $\mu = \inf\{-z(s)/a^T(s)y : a^T(s)y < 0, s \in S\}$, $\mu \geq 0$, veamos que $\mu > 0$. Supongamos que $\mu = 0$, tomamos $K \in \mathbb{N}$ suficientemente grande y para cada $j \geq K$ definimos $\epsilon_j = 1/j$. Dado que $\mu = 0$, para ϵ_j existir3a $t_j \in S$ tal que $a^T(t_j)y < 0$, siendo $z(t_j) < 1/j$. As3i $\{t_j\}_{j \geq K}$ es una sucesi3n en un compacto, entonces existir3a una subsucesi3n $\{t_{j_k}\}_{k \geq 1}$ convergente a $t^0 \in S$, es decir, $\lim t_{j_k} = t^0$. Por continuidad de $z(s)$ y $a(s)$ tenemos que:

- (1) $z(t^0) = \lim z(t_{j_k}) = 0$, por tanto $t^0 \in \text{constr}(x)$.
- (2) $a^T(t^0)y = \lim a^T(t_{j_k})y \leq 0$, pero esto contradice [1].

Luego, necesariamente $\mu > 0$.

Construimos $x_1 = x + \mu y$, se comprueba f3cilmente que $x_1 \in F$ y como $c^T x_1 < c^T x$, tendr3amos una soluci3n posible mejor que la soluci3n 3ptima lo que es una contradicci3n. ♦

La cualificaci3n de restricciones de Mangasarian-Fromovitz est3 definida tanto para puntos con un n3mero finito de restricciones activas como para aquellos que tienen un n3mero infinito de ellas. En este 3ltimo caso se pierde, en gran parte, la cualidad que la hace m3s atractiva : su operatividad, ya que no es sencillo comprobar si el elemento neutro de \mathbb{R}^n pertenece a la envoltura convexa de un conjunto infinito de vectores.

Notemos que comprobar la cualificaci3n M-F en un punto (x, z) cuyo $\text{constr}(x) = \{s_1, \dots, s_m\}$, se reduce a resolver el siguiente PL:

$$\begin{aligned} \text{Encontrar } v = \max \sum_{i=1}^m \alpha_i \\ \text{s.a. } \sum_{i=1}^m \alpha_i a(s_i) = 0 \\ 0 \leq \alpha_i \leq 1 \quad i=1, \dots, m \end{aligned}$$

y verificar si $v=0$, pues sólo en este caso se cumple la cualificación M-F.

La versión del Teorema 2.3 para puntos con infinitas restricciones activas se presenta en el siguiente resultado :

Teorema 2.4

Sea $(x; z)$ un punto posible para el (PSI2) cuyo $\text{constr}(x) \neq \emptyset$, y supongamos que verifica la cualificación M-F. Entonces $(x; z)$ es óptimo para el (PSI2) si, y sólo si, existen un subconjunto $\{s_1, s_2, \dots, s_m\}$ de $\text{constr}(x)$ y unos escalares $\lambda_i \geq 0$, $i=1, \dots, m$ tales que $c = \sum_{i=1}^m \lambda_i a(s_i)$ (equivalentemente, si y sólo si, $c \in K\{A(x)\}$).

Demostración.

La prueba de la condición suficiente es completamente análoga a la del Teorema 2.3. En cuanto a la condición necesaria:

Sabemos que $K\{A(x)\}$ es un cerrado (Corolario 9.6.1 en Rockafellar (70)), ya que $\text{co}(A(x))$ es un compacto que no contiene a 0_n , pues por hipótesis se cumple M-F en $(x; z)$.

Supongamos que $c \notin K\{A(x)\} = \text{cl } K\{A(x)\}$, aplicando el Teorema de Farkas generalizado (Goberna et al (84)) se tiene que existe $u \in \mathbb{R}^n$ tal que $a(s)^T u \geq 0$ para todo $s \in \text{constr}(x)$ y $c^T u < 0$. Por verificarse M-F, sabemos que existe $d \in \mathbb{R}^n$ tal que $a(s)^T d > 0$ para todo $s \in \text{constr}(x)$. Consideremos $y = u + \gamma d$, para $\gamma > 0$ suficientemente pequeño tenemos que $c^T y < 0$ y $a(s)^T y > 0$, para todo $s \in \text{constr}(x)$.

El resto de la demostración coincide con la correspondiente a un número finito de restricciones activas: podemos movernos desde x en la dirección y hasta una solución posible mejor, lo que es absurdo. \blacklozenge

Nótese que la cualificación de M-F permite asegurar que, también en el caso de infinitas restricciones activas, el conjunto $K\{A(x)\}$ es cerrado; aunque no es posible afirmar directamente que $M_{n+1} = K\left\{\begin{pmatrix} a(s) \\ b(s) \end{pmatrix} \mid s \in S\right\}$ sea cerrado, cualificación de restricciones utilizada por Glashoff y Gustafson (83).

En el ejemplo siguiente, hemos comprobado que se verifica M-F, sobre una solución posible no óptima, pero en general no será fácil hacerlo.

Ejemplo 2.3

$$\text{Sea (P}_3\text{) } \begin{aligned} &\text{Min } x_1 + 1/2 x_2 + 1/3 x_3 \\ &\text{s.a. } s_1 x_1 + s_2 x_2 + x_3 \geq s_1^2 + s_2 \quad s \in [0,1] \times [0,1] \end{aligned}$$

Consideremos la solución posible $x=(1, 1, 0)^T$ cuyos $z(s_1, s_2)=s_1 - s_1^2$ y $\text{constr}(x)=\left\{ \begin{pmatrix} 0 \\ s_2 \end{pmatrix}, \begin{pmatrix} 1 \\ s_2 \end{pmatrix} : s_2 \in [0,1] \right\}$. Dado que la última componente de cualquier vector $a(s)$ es igual a 1, puede comprobarse fácilmente que se verifica la cualificación M-F, puesto que $0_3 \notin \text{co}\{A(x)\}$, donde $A(x)=\left\{ \begin{pmatrix} 0 \\ s_2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ s_2 \\ 1 \end{pmatrix} : s_2 \in [0,1] \right\}$. Asimismo, se comprueba que $c=\begin{pmatrix} 1 \\ 1/2 \\ 1/3 \end{pmatrix} \notin K\{A(x)\}$ y, por tanto, x no es solución óptima, según se ha demostrado en el Teorema 2.4.

Interpretación geométrica. La cualificación de M-F, que juega el papel de una cualificación de restricciones local para los problemas de programación semi-infinita lineal, permite asegurar que podemos movernos partiendo de una solución x a lo largo de una dirección d . Para ello se necesita que $a^T(s)d > 0$ para todo $s \in \text{constr}(x)$, si estas desigualdades no se satisfacen estrictamente tal movimiento puede no ser posible. Veámoslo:

Ejemplo 2.4

Consideremos el problema :

$$\text{(P}_4\text{) } \begin{aligned} &\text{Min } x_1 + 1/2 x_2 \\ &\text{s.a. } x_1 + s x_2 \geq 1/(1+s^2) \quad s \in [0,1] \end{aligned}$$

(1) Tomamos $x^0=\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ que es una solución posible para (P₄), cuya función de holgura

$$\text{es } z^0(s)=1-1/(1+s^2) \geq 0 \text{ para } s \in [0,1], \text{ constr}(x^0)=\{0\} \text{ y } a(0)=\begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Sea $u = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$, se tiene que $a^T(s_i)u \geq 0 \quad \forall s_i \in \text{constr}(x)$ y $c^T u < 0$. Si nos planteamos hallar

$\mu = \inf\{-z(s)/a^T(s)u : a^T(s)u < 0, s \in S\}$ queda:

$$\mu = \inf\{-(1-1/(1+s^2))/(-s) : -s < 0, s \in S\} = \inf\{1-1/(1+s^2)/s : s \in (0,1]\} = 0$$

$$\left(\lim_{s \rightarrow 0} 1-1/(1+s^2)/s = \lim_{s \rightarrow 0} 2s/(1+s^2)^2 = 0\right).$$

En este caso no podemos movernos desde x^0 en la dirección factible de descenso u , pero dado que en x^0 se satisface la cualificación M-F, sabemos que existe una dirección que podemos sumar a u para movernos sin perder la factibilidad y mejorando el valor de la función objetivo, ya que x^0 no es solución óptima del problema: $c \neq \lambda a(0)$. La obtención de esta dirección de mejora se estudiará con detalle en el Capítulo IV.

(2) De la interpretación geométrica de este problema de aproximación L_1 puede obtenerse el punto posible:

$$x = \begin{pmatrix} 1.103553 \\ -0.603553 \end{pmatrix} \text{ cuyo } \text{constr}(x) = \{0.41421, 1\} \quad d(0.41421) = 1, d(1) = 0.$$

La matriz \tilde{A} no es invertible, luego x es una solución posible básica degenerada y con el criterio establecido en el Teorema 1.9 no podemos asegurar si es o no óptima. A continuación comprobaremos que se trata de un óptimo mediante el Teorema 2.3.

Planteamos el problema lineal que nos permite decidir si un punto satisface o no la cualificación M-F, y obtenemos que es así, ya que $v=0$. Ahora, si existen unos multiplicadores no negativos de manera que c sea combinación de $a(s_1)$ y $a(s_2)$ tenemos un óptimo y, lo que es más importante, si no existen tales multiplicadores no lo será. Pero

$$c = \begin{bmatrix} 1 \\ 1/2 \end{bmatrix} = \lambda_1 \begin{bmatrix} 1 \\ .41421 \end{bmatrix} + \lambda_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Leftrightarrow \lambda_1 = .8535482, \lambda_2 = .1464518$$

luego, efectivamente x es una solución óptima de (P_4) .

Es, pues, imprescindible disponer de criterios de optimalidad para puntos extremos degenerados pues, a diferencia de lo que ocurre en Programación Lineal, en Programación Semi-Infinita Lineal muchas soluciones posibles básicas son degeneradas y no puede hacerse mucho por evitarlas, pues ciertos tipos de degeneración son típicos de la Programación infinito dimensional (Nash (85)).

Frente a la condición de optimalidad anterior (Teoremas 2.3 y 2.4), que puede comprobarse sobre cualquier punto posible $(x; z)$ con un número finito o no de

restricciones activas, la que desarrollan Anderson y Lewis (89) sólo puede ser aplicada a puntos extremos no degenerados con un número finito de restricciones activas. Su caracterización aparece dentro de una extensión del método Simplex para programación semi-infinita lineal y, sólo necesitan un criterio de optimalidad para puntos extremos.

Vamos a comprobar que el criterio de optimalidad del Teorema 2.3, coincide con el demostrado por Anderson y Lewis (89) cuando se aplica a puntos extremos no degenerados de un problema (PSI2) y S es un compacto de R . Cuando además a_1, \dots, a_n y b son funciones analíticas, se tiene que la función de holgura de cualquier punto posible ó tiene un número finito de ceros ó es idénticamente nula. Por lo cual, siempre es posible comprobar si un punto dado verifica la cualificación M-F.

Teorema 2.5

Sea $(x; z)$ un punto extremo no degenerado del (SIP2), tal que $\text{constr}(x) \neq \emptyset$. Bajo las hipótesis del Teorema 1.9, las condiciones de optimalidad del Teorema 1.9 y del Teorema 2.3 son equivalentes.

Demostración.

Sea $(x; z)$ un punto extremo no degenerado con $\text{constr}(x) = \{s_1, s_2, \dots, s_m\}$ entonces \tilde{A} es invertible siendo \tilde{A} la matriz $n \times n$:

$$\tilde{A} = (a(s_1), \dots, a(s_m), a'(s_1), a''(s_1), \dots, a^{d(1)}(s_1), a'(s_2), a''(s_2), \dots, a^{d(2)}(s_2), \dots, a^{d(m)}(s_m)).$$

Entonces los vectores de $A(x)$ son linealmente independientes y $(x; z)$ verifica M-F.

Según el criterio del Teorema 1.9, x es óptima si, y sólo si,

$$\tilde{A}^{-1}c = \begin{bmatrix} \lambda \\ \rho \end{bmatrix} \text{ con } \lambda \geq 0_m \text{ y } \rho = 0_{n-m} \Leftrightarrow c = \tilde{A} \begin{bmatrix} \lambda \\ \rho \end{bmatrix} \text{ con } \lambda \geq 0_m \text{ y } \rho = 0_{n-m}$$

$$\Leftrightarrow c = \tilde{A} \begin{bmatrix} \lambda \\ 0_{n-m} \end{bmatrix} \text{ con } \lambda \geq 0_m \Leftrightarrow c = \sum_{i=1}^m \lambda_i a(s_i) \text{ con } \lambda \geq 0_m \Leftrightarrow (x; z) \text{ es óptima según el}$$

Teorema 2.3. ♦

Veamos algunas condiciones que garantizan que un punto posible verifica la cualificación M-F.

Teorema 2.6

Sea $(x; z)$ una solución posible para el (PSI2), tal que $\text{constr}(x) \neq \emptyset$. Las siguientes condiciones son suficientes para que se verifique la cualificación M-F:

- (1) Los vectores de $A(x)$ son linealmente independientes.
- (2) $(x; z)$ es un punto extremo no degenerado.
- (3) $b(s) > 0, \forall s \in S$ (o $b(s) < 0 \forall s \in S$).
- (4) $\{a(s): s \in S\}$ tiene la propiedad del signo opuesto.

Demostración.

(1) Trivial.

(2) Sea $(x; z)$ un punto extremo no degenerado de (PSI2). Sea $C[S]$ la clase de las funciones continuas en S , se cumplirá que (ver Nash, (85)):

$X = \mathbb{R}^n \times C[S] = B((x; z)) \oplus N(A)$, donde $N(A) = \{(d; h) \in \mathbb{R}^n \times C[S]: h(s) = a^T(s)d \text{ para todo } s \in S\}$, y $B((x; z))$ es la solidificación de $(x; z)$.

Para $(d; h) \in B((x; z))$ y de la definición de $B((x; z))$ se sigue que: $(z + \lambda h)(s) \geq 0$ y $(z - \lambda h)(s) \geq 0$ para todo $s \in S$. Para $s \in \text{constr}(x)$ tenemos que $z(s) = 0$, entonces $\lambda h(s) \geq 0$ y $-\lambda h(s) \geq 0$, para $\lambda \geq 0$. En consecuencia, $h(s) = 0$ para $s \in \text{constr}(x)$.

Sean $p \in \mathbb{R}^n$ y M una constante positiva tales que $(p; M) \in X$. Entonces existirá $(d; h) \in B((x; z))$ tal que $(p - d; M - h) \in N(A)$, con lo que $a^T(s)(p - d) = M - h(s)$, para todo $s \in S$. En particular, para $s \in \text{constr}(x)$: $a^T(s)(p - d) = M > 0$. Luego existe una dirección $u \in \mathbb{R}^n$ tal que $a^T(s)u > 0$, para $s \in \text{constr}(x)$, por tanto se verifica la cualificación M-F.

(3) Supongamos que $b(s) > 0, \forall s \in S$. Para $s \in \text{constr}(x)$ se tiene que $a^T(s)x = b(s)$. Supongamos que $0_n \in \text{co}(A(x))$, sabemos por el Teorema de Caratheodory que existiran $\{s_1, s_2, \dots, s_{n+1}\} \in \text{constr}(x)$ e $y \geq 0_{n+1}, y \neq 0_{n+1}$ tales que se tiene que $\sum_{i=1}^{n+1} y_i a(s_i) = 0_n$, entonces $\sum_{i=1}^{n+1} y_i a^T(s_i)x = 0_n^T x = 0$. Pero, por otra parte: $\sum_{i=1}^{n+1} y_i a^T(s_i)x = \sum_{i=1}^{n+1} y_i b(s_i) > 0$, y llegaríamos a una contradicción, luego debe cumplirse la cualificación M-F.

(4) El conjunto $\{a(s): s \in S\}$ tiene la propiedad del signo opuesto si $\alpha \in \mathbb{R}^{(S)} - \{0\}$ y $\sum_{s \in S} \alpha(s)a(s) = 0_n$ implican que existen $s, t \in S$ tales que $\alpha(s)\alpha(t) < 0$ (ver Charnes, Cooper y Kortanek, (63)). En tal caso, desde luego, cualquier punto posible verificará la cualificación M-F. ♦

Notas. En el caso de que el punto extremo no degenerado tenga un número finito de restricciones activas, la demostración del apartado (2) se sigue inmediatamente de su

caracterización, pues los gradientes activos son linealmente independientes. Además, dado que el Teorema 1.9 ha sido demostrado para soluciones posibles que son puntos extremos no degenerados, es inmediato que esta es una condición de regularidad que ha sido exigida implícitamente para caracterizar las soluciones óptimas.

La versión local de (3) es (3'): $b(s) > 0, \forall s \in \text{constr}(x)$ (o $b(s) < 0, \forall s \in \text{constr}(x)$). Tanto (3) como (3') son condiciones fáciles de comprobar.

La propiedad del signo opuesto (OSP) ha sido utilizada en programación semi-infinita lineal para estudiar algunas propiedades del problema dual (PSI1*). En particular, la condición (4) también implica que el conjunto factible dual no tiene direcciones de recesión y está generado por sus puntos extremos (Kortanek y Strojwas, (85)). La condición de optimalidad establecida en los Teoremas 2.3 y 2.4 es del tipo Kuhn -Tucker, extendida al caso semi-infinito lineal. Se sabe que las versiones semi-infinitas de las teorías de Fritz-John y de Kuhn -Tucker fallan para caracterizar la optimalidad, incluso en el caso lineal, a menos que se exijan ciertas condiciones de regularidad sobre las restricciones que son las conocidas como cualificaciones de restricciones (vease p.e. Gehner (74) y Goberna et al (84)).

De entre las cualificaciones de restricciones utilizadas en programación semi-infinita cabe destacar la de Slater. En particular, la condición de Slater y una extensión de la cualificación de Abadie han sido utilizadas en Krabs (79) para caracterizar las soluciones óptimas de un problema de optimización no lineal infinito-dimensional sobre un espacio vectorial normado. En Glashoff y Gustafson (83) se trabaja con la cualificación de Slater y otras condiciones de clausura para establecer resultados de optimalidad en PSIL.

Otros autores han establecido caracterizaciones de optimalidad que no necesitan cualificaciones de restricciones. Veamos una de las más sencillas, en el sentido de que al menos puede comprobarse. Ben-Tal, Kerzner y Zlobec (80) dan una condición de óptimo para un tipo muy amplio de problemas que incluye al que estudiamos y que no exige una cualificación de restricciones, aunque sí que se satisfaga la llamada "uniform mean value property". El (PSI1) verifica esta propiedad, así pues, sus soluciones óptimas pueden caracterizarse utilizando estos resultados. Aunque, debido a su complejidad, parece más práctico emplear cualificaciones de restricciones. Adaptada a nuestro problema y con nuestra notación dicha caracterización quedaría como sigue :

Lema 2.7 (Ben-Tal, Kerzner y Zlobec (80))

Sea $x \in F$, dicha solución es óptima si, y sólo si, $\forall \alpha > 0$ y $\forall \Omega \subseteq \text{constr}(x)$ el sistema (III) es inconsistente. Siendo

$$\begin{aligned} d^T c &< 0 \\ d^T a(s) &> 0 \quad \forall s \in \Omega \\ \text{(III)} \quad d^T a(s) &= 0 \quad \forall s \in \text{constr}(x) - \Omega \\ \frac{d^T a(s)}{z(s)} &\geq -\frac{1}{\alpha} \quad \forall s \in S - \text{constr}(x) \end{aligned}$$

Consideremos de nuevo el Ejemplo 2.2. Veamos que al óptimo que habíamos encontrado no podemos asociarle un punto de Kuhn-Tucker.

Sea $\bar{x} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$, con $\text{constr}(\bar{x}) = \{0\}$, $A(\bar{x}) = \{0_n\}$ y $c = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Se tiene claramente que $c = \lambda a(0)$ con $\lambda \geq 0$ es inconsistente. En realidad c no puede escribirse como combinación lineal positiva de los vectores $\{a(s) : s \in S\}$ puesto que, para $\alpha(s) \geq 0$,

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \sum \alpha(s) a(s) \Leftrightarrow \sum \alpha(s) s = 1 \text{ y } \sum \alpha(s) s^2 = 0 \text{ lo que es imposible.}$$

En cambio, si podemos aplicar la caracterización anterior, pues se tiene que:

$\bar{x} = \begin{bmatrix} 0 \\ 2 \end{bmatrix} \in F$, sea $\alpha > 0$, $\text{constr}(\bar{x}) = \{0\} \Rightarrow \Omega = \emptyset$ u $\Omega = \{0\}$. Tomamos $\alpha > 0$, entonces:

i) $\Omega = \emptyset$. El sistema (III) en este caso es:

$$\begin{aligned} d_1 &< 0 \\ 0 &= 0 \end{aligned} \quad \Leftrightarrow \quad \begin{aligned} d_1 &< 0 \\ d_1 + s d_2 &\geq -\frac{1}{\alpha} s \quad \forall s \in (0, 1] \end{aligned} \text{ inconsistente.}$$

$$\frac{s d_1 + s^2 d_2}{s^2} \geq -\frac{1}{\alpha} \quad \forall s \in (0, 1]$$

ii) $\Omega = \{0\}$. El sistema (III) es: $d_1 < 0$
 $0 > 0$ inconsistente.

En consecuencia, dicho punto es un óptimo de (P_2) .

Notemos que este criterio es poco operativo y nos proponemos utilizarlo sólo si fallan las cualificaciones de restricciones.

Sección 2.2 Un test de optimalidad.

En esta sección desarrollamos un test de optimalidad para puntos posibles con un número finito de restricciones activas que cumplan la cualificación M-F. La resolución de un problema lineal finito permite determinar si uno de estos puntos es o no óptimo del problema (PSI1).

Goberna, López y Pastor (81) han establecido condiciones de tipo Kuhn-Tucker, para el problema diferenciable de programación semi-infinita convexa (ver también Hettich-Jongen (78)), basadas en los siguientes elementos que pueden asociarse a cualquier punto posible, $x^* \in F$: los conjuntos: $P = \{y \in \mathbb{R}^n : c^T y \leq 0, a^T(s)y \geq 0, \text{ para todo } s \in \text{constr}(x^*)\}$ y $P_0 = \{y \in P : c^T y = 0\}$.

A continuación demostramos el correspondiente criterio de optimalidad para el problema (PSI1), utilizando la cualificación de restricciones M-F.

Teorema 2.8

Sea $x^* \in F$ con $\text{constr}(x^*) \neq \emptyset$. Si $P = P_0$, entonces x^* es un óptimo para el problema (PSI1). Si x^* satisface la cualificación M-F, el recíproco también es cierto.

Demostración.

Supongamos que x^* no fuese óptimo, entonces existiría $\bar{x} \in F$ tal que $c^T \bar{x} < c^T x^*$. Tomaríamos $d = \bar{x} - x^*$, entonces $c^T d = c^T \bar{x} - c^T x^* < 0$. Además, como $\bar{x} \in F$, para cada $s \in \text{constr}(x^*)$ tendríamos que $a^T(s)d = a^T(s)\bar{x} - a^T(s)x^* = a^T(s)\bar{x} - b(s) \geq 0$. Por tanto $d \in P/P_0$, con lo que llegaríamos a una contradicción.

Recíprocamente, si x^* es un óptimo para el (PSI1) y satisface M-F siendo $\text{constr}(x^*) \neq \emptyset$, por el Teorema 2.4 sabemos que existirán $\lambda_i \geq 0, i=1, \dots, q$ y $s_i \in \text{constr}(x^*) i=1, \dots, q$ tales que $c = \sum_{i=1}^q \lambda_i a(s_i)$. Así pues, $P = P_0$, ya que si $P \neq P_0$, existiría $d \in \mathbb{R}^n$ tal que $c^T d < 0$ y $a^T(s_i)d \geq 0$, para $i=1, \dots, q$ y, entonces $0 > c^T d = \sum_{i=1}^q \lambda_i a^T(s_i)d \geq 0$. ♦

El resultado anterior sugiere el siguiente Test de Optimalidad. Para un punto posible $(x^*; z^*)$ cuyo $\text{constr}(x^*) = \{s_1, \dots, s_m\}$, resolvemos el siguiente problema de programación lineal:

Min $c^T d$

s.a. $a^T(s_i)d \geq 0 \quad i=1, \dots, m$

$-1 \leq d_j \leq 1 \quad j=1, \dots, n$

Sea d^* una solución óptima y sea $\theta^* = c^T d^*$. Hay dos posibilidades:

(i) Si $\theta^* = 0$, $P = P_0$ y, por tanto x^* es una solución óptima.

(ii) Si $\theta^* < 0$, $P \neq P_0$ y, por el lema de Farkas, no existe ningún $\lambda \geq 0_m$ tal que

$c = \sum_{i=1}^m \lambda_i a(s_i)$. En este caso no podemos afirmar que x^* no sea un óptimo para el (PSI1), a menos que se verifique la cualificación de restricciones de M-F o alguna otra condición de regularidad.

Nota. Si $P = \{0_n\}$ la solución óptima es única:

Sea x' otra solución óptima, entonces $c^T x' = c^T x^*$. Tomamos $d = x' - x^*$, entonces $c^T d = 0$ y $a(s)^T d = a(s)^T x' - a(s)^T x^* = a(s)^T x' - b(s) \geq 0$ para $s \in \text{constr}(x^*)$. Por tanto, $d \in P$ y por hipótesis $P = \{0_n\}$, luego $x' = x^*$.

Supongamos que el problema (PSI1) satisface alguna cualificación de restricciones, de modo que podemos asegurar, en el caso (ii), que x^* no es una solución óptima. Entonces el test puede darnos más información calculando algunos parámetros relacionados con el vector óptimo d^* :

(1) Si $\beta = \inf \{a^T(s)d^*, s \in S\} \geq 0$, entonces d^* es una dirección de recesión y el problema (PSI1) es no acotado.

(2) Si $\beta < 0$, hallamos $\mu^* = \inf \{-z(s)/a^T(s)d^* : a^T(s)d^* < 0, s \in S\}$. Entonces:

(a) si $\mu^* > 0$, tenemos que $x^* + \mu^* d^*$ es un punto posible mejor que x^* .

(b) si $\mu^* = 0$, d^* no puede utilizarse como dirección de descenso, aunque según veremos en el Capítulo IV, puede transformarse adecuadamente para generar una dirección factible de descenso.

A continuación veremos, sobre diferentes ejemplos, las distintas situaciones que pueden presentarse al aplicar el test.

Ejemplo 2.5

Volviendo al problema (P_2) , sabemos que $\bar{x} = (0, 2)^T$ es una solución óptima con $v(P_2) = 0$, la función de holgura es $z(s) = s^2$, el $\text{constr}(\bar{x}) = \{0\}$ y $A(\bar{x}) = \{0_2\}$, con lo que $P = \{d \in \mathbb{R}^2 : d_1 \leq 0\} \neq \{d \in P : d_1 = 0\} = P_0$ siendo $\theta^* = -1$. La cualificación de M-F no se verifica y \bar{x} no tiene asociado ningún punto de Kuhn-Tucker. De hecho, el correspondiente problema dual no tiene soluciones posibles.

Veamos un ejemplo en el que primero comprobamos que el punto posible disponible cumple la cualificación M-F y después, mediante la aplicación del test de optimalidad, llegamos a la conclusión de que tal punto no es óptimo.

Ejemplo 2.6

$$(P_6) \quad \text{Inf } \sum_{i=1,..,7} (1/i) x_i$$

$$\text{s.a. } \sum_{i=1,..,7} s^{i-1} x_i \geq 1/(1+s^2) \quad s \in [0, 1].$$

Después de dos iteraciones, el algoritmo de purificación (León y Vercher (91)) encuentra un punto extremo degenerado:

Iteración	x	constr(x), d(i)
1	(1,0,-1,1/4,1/2,-1/4,0)	{0,1}, d(0)=2, d(1)=2
2	(1,0,-1,0,5/4,-1,1/4)	{0,1}, d(0)=3, d(1)=3

Para x^2 se cumple trivialmente M-F. Puede observarse que $b(s) > 0$ o bien que los vectores $a(0)$ y $a(1)$ son linealmente independientes.

Si aplicamos el test de optimalidad a x^2 , se resuelve el PL siguiente:

$$\text{Min } \sum_{i=1,..,7} (1/i) d_i$$

$$\text{s.a. } d_1 \geq 0$$

$$d_1 + d_2 + \dots + d_7 \geq 0$$

$$-1 \leq d_i \leq 1 \quad i=1,..,7$$

y se llega a $d^* = (0, -1, -1, -1, 1, 1, 1)$ con $\theta^* = -0.573809 < 0$. Luego, podemos asegurar que x^2 no es una solución óptima de (P_6) . Sin embargo, la dirección d^* no permite pasar a una solución mejor, puesto que si intentamos llevar a cabo un paso de descenso a lo largo de ella tenemos que $\mu^* = 0$.

En el siguiente ejemplo, sin embargo, la dirección generada tras aplicar el test de optimalidad sí nos permite pivotar a partir de la solución actual.

Ejemplo 2.7

Consideremos el problema:

$$(P_7) \min \quad \{x_1 + (1/2)x_2\}$$
$$\text{s.a.} \quad sx_1 + s^2x_2 \geq s^3 - 1 \quad s \in [-1, 1]$$

Sea $x^\circ = (1, -1)^T$ una solución posible para (P_7) y $z^\circ(s) = 1 + s - s^2 - s^3$ su función de holgura, siendo $\text{constr}(x^\circ) = \{-1, 1\}$, $d(1) = 0$, $d(-1) = 1$ y $A(x^\circ)$ una matriz no invertible. Así pues, x° es un punto extremo degenerado y satisface la cualificación M-F. El test de optimalidad da como resultado $\theta^* = -1/2$ y $d^* = (-1, 1)^T$. Hallamos $\mu^* = \inf\{(1+s^2)/s : s \in (0, 1)\} = 4$, por tanto tenemos una solución posible mejor $x^1 = (-3, 3)^T$ tal que $\text{constr}(x^1) = \{1\}$, y $d(1) = 2$. Se puede comprobar que este nuevo punto no es óptimo y que es también un punto extremo degenerado.

Por último, veamos el caso en que el test detecta la no acotación de un problema.

Ejemplo 2.8

Consideremos el problema:

$$(P_8) \min \quad \{-x_1 + x_2\}$$
$$\text{s.a.} \quad x_1 + sx_2 \geq 6s^2 - 4s^3 - 3s^4 + 2s^5 \quad s \in [-3/2, 3/2]$$

Sea $x^\circ = (3, -2)^T$ una solución posible para (P_8) y $z^\circ(s)$ su función de holgura, con $\text{constr}(x^\circ) = \{-1, 1, 3/2\}$, $d(-1) = 1 = d(1)$ y $d(3/2) = 0$. Entonces x° es un punto extremo degenerado, y es fácil comprobar que verifica la cualificación M-F. Aplicamos el test de optimalidad y encontramos $\theta^* = -5/3$ y $d^* = (1, -2/3)^T$. Como $\beta = \inf \{a^T(s)d^*, s \in S\} \geq 0$, tenemos que d^* es una dirección de recesión y el problema (P_8) no está acotado.

Los ejemplos anteriores muestran la importancia del test cuando se aplica para puntos extremos degenerados, porque permite detectar la no acotación del problema ó encontrar, en unos casos directamente y en otros tras ser modificada, direcciones de mejora cuando la solución actual no es óptima.

Sección 2.3 Cualificaciones de restricciones locales en Programación Semi-Infinita Lineal.

Con la intención de situar la cualificación de restricciones de Mangasarian-Fromovitz en el contexto de las cualificaciones de restricciones que permiten determinar condiciones de optimalidad de tipo Kuhn-Tucker, hemos estudiado las generalizaciones de las distintas cualificaciones de restricciones que han sido introducidas para problemas finitos no lineales (ver Bazaraa-Shetty (76)) y hemos revisado algunas de las que ya han sido establecidas tanto para Programación Semi-Infinita Lineal como para PSI Convexa.

Podemos distinguir dos tipos de condiciones de optimalidad para PSI Convexa según consideremos el caso diferenciable o el caso no diferenciable. Aunque en nuestro problema todas las funciones involucradas son diferenciables no conviene perder de vista el origen de las condiciones que hemos estudiado, puesto que han sido establecidas para problemas más generales y con hipótesis más fuertes, algunas de las cuales se cumplen trivialmente en Programación Semi-Infinita Lineal.

Dado que los problemas (PSI1) y (PSI2) son equivalentes para referirnos a cualquiera de ellos utilizaremos la notación (PSIL).

En primer lugar hemos comprobado que la extensión de la cualificación de M-F es equivalente a la generalización al caso semi-infinito lineal de la cualificación de restricciones de Cottle establecida en programación no lineal.

Definición 2.2

Sea x un punto posible cuyo $\text{constr}(x) \neq \emptyset$. Diremos que satisface la cualificación de restricciones de Cottle, si y sólo si, $C \subseteq \text{cl } G_0$ donde $C = \{d: a^T(s)d \geq 0 \quad s \in \text{constr}(x)\}$ y $G_0 = \{d: a^T(s)d > 0 \quad s \in \text{constr}(x)\}$.

Teorema 2.9

Para el problema (PSIL) la cualificación de restricciones de Mangasarian-Fromovitz es equivalente a la cualificación de restricciones de Cottle.

Demostración.

Sea $x \in F$ con $\text{constr}(x) \neq \emptyset$. Supongamos que x satisface Cottle. Si x no cumpliera la cualificación M-F, entonces el sistema $\{a^T(s)d > 0, s \in \text{constr}(x)\}$ sería inconsistente y, por tanto, $G_0 = \emptyset$. En ese caso $\text{cl } G_0 = \emptyset$ y por hipótesis $C = \emptyset$. Pero $0_n \in C$, con lo que llegaríamos a una contradicción. Así pues, x satisface la cualificación M-F.

Recíprocamente, si x satisface M-F, tendríamos que $G_0 \neq \emptyset$. Ahora bien, $G_0 = \bigcap \{C_s : s \in \text{constr}(x)\} \neq \emptyset$, siendo $C_s = \{d : a^T(s)d > 0\}$, con $\text{int } C_s = C_s$. Entonces, por el Teorema 6.5 de Rockafellar (70), $\text{cl } G_0 = \text{cl } [\bigcap \{C_s : s \in \text{constr}(x)\}] = \bigcap \{\text{cl } C_s : s \in \text{constr}(x)\} = \{d : a^T(s)d \geq 0 : s \in \text{constr}(x)\} = C$ y, por tanto, en el punto x se satisface la cualificación de restricciones de Cottle. ♦

Nota. La condición $G_0 \neq \emptyset$, para aquellos $x \in F$ cuyo $\text{constr}(x) \neq \emptyset$, ha sido también utilizada como condición de regularidad en Krabs (79) para problemas semi-infinitos no lineales.

Otra cualificación local de restricciones estudiada para Programación Semi-infinita Convexa no diferenciable es la condición de punto regular Lagrangiano (Definición 3.3. en López y Vercher (83)), que puede considerarse como una extensión de la cualificación de Guignard.

Recordemos la noción de cono de tangentes a F en un punto $\bar{x} \in \text{cl } F$:

$$T(F, \bar{x}) = \{z : z = \lim_{k \rightarrow \infty} \lambda_k(x_k - \bar{x}) \text{ donde } \lambda_k > 0, x_k \in F \text{ y } \lim_{k \rightarrow \infty} x_k = \bar{x}\} = \text{cl } K(F, \bar{x}).$$

Siendo $K(F, \bar{x})$ el cono convexo generado por $(F - \bar{x})$.

En general, para el problema de PSI convexa se demuestra en López y Vercher (83) que si $\bar{x} \in F$, entonces $T(F, \bar{x}) \subseteq B(\bar{x})^*$, donde $B(\bar{x}) = -A(\bar{x}) = \{-a(s) : s \in \text{constr}(\bar{x})\}$ y $B(\bar{x})^* = \{d \in \mathbb{R}^n : a(s)^T d \geq 0, s \in \text{constr}(\bar{x})\}$ es el polar de $B(\bar{x})$. A continuación vemos que para nuestro problema puede darse otra demostración bastante sencilla de este mismo resultado:

Lema 2.10

Si $\bar{x} \in F$, entonces $T(F, \bar{x}) \subseteq B(\bar{x})^*$.

Demostración.

Sea $z \in T(F, \bar{x})$, se trata de comprobar que $a(s)^T z \geq 0$ para todo $s \in \text{constr}(\bar{x})$. Por definición, sabemos que $z = \lim_{k \rightarrow \infty} \lambda_k (x_k - \bar{x})$ donde $\lambda_k > 0$, $x_k \in F$ y $\lim_{k \rightarrow \infty} x_k = \bar{x}$.

Para cada $s \in \text{constr}(\bar{x})$ tendremos que $a(s)^T z = a(s)^T \lim_{k \rightarrow \infty} \lambda_k (x_k - \bar{x}) = \lim_{k \rightarrow \infty} \lambda_k a(s)^T (x_k - \bar{x}) = \lim_{k \rightarrow \infty} \lambda_k (a(s)^T x_k - b(s)) \geq 0$. ♦

Definición 2.3

Para el problema (PSIL) diremos que $\bar{x} \in F$ es un punto regular Lagrangiano si:

- (i) $T(F, \bar{x}) = \mathbb{R}^n$, cuando $\text{constr}(\bar{x}) = \emptyset$
- (ii) (a) $B(\bar{x})^* \subseteq T(F, \bar{x})$ y
(b) $K\{B(\bar{x})\}$ es cerrado, cuando $\text{constr}(\bar{x}) \neq \emptyset$.

Si $\text{constr}(\bar{x}) = \emptyset$ el punto posible \bar{x} es trivialmente regular Lagrangiano, aunque, según hemos visto en el Lema 2.1, no será candidato a óptimo. Si el conjunto de índices activos es finito no es necesario exigir (ii)-(b), ya que $K\{B(\bar{x})\}$ es un cono finitamente generado. Por otra parte, si $z(s)$ es idénticamente nula sólo necesitamos exigir (ii)-b:

Lema 2.11

Sea $x \in F$ con función de holgura $z(s)$ idénticamente nula. Entonces x es un punto regular Lagrangiano si, y sólo si, el cono de los momentos $M_n = K\{a(s) : s \in S\}$ es cerrado.

Demostración.

Sea $d \in B(x)^*$, sabemos que para todo $s \in S$ $a(s)^T d \geq 0$ y, por tanto, $x + \lambda d \in F$, para cada $\lambda > 0$. Tomamos $\lambda = 1$, entonces $x_1 = x + d \in F$ y $d = x_1 - x \in \text{cl}K(F - x) = T(F, x)$. Por tanto, $B(x)^* \subseteq T(F, x)$. En cuanto a (ii)-b, es evidente que $K\{B(x)\}$ es cerrado si, y sólo si, el cono de los momentos M_n es cerrado. ♦

Nota. La condición (ii)-(a) no se cumple trivialmente si la función de holgura no es idénticamente nula. Consideremos el problema del Ejemplo 2.1:

$$(P_1) \text{ Min } x_1 + x_2 \\ \text{s.a. } s x_1 + s^2 x_2 \geq 0 \quad s \in [0, 1]$$

Para el punto $\bar{x} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, tenemos que $B(\bar{x})^* = \mathbb{R}^2$ y $T(F, \bar{x}) = \{x \in \mathbb{R}^2: x_1 \geq 0\}$, luego \bar{x} no es un punto regular Lagrangiano.

Este mismo problema, considerando $x_1 = 0$, nos proporciona un ejemplo en el que $z_1(s)$ es idénticamente nula y $M_n = \{(s, s^2)^T: s \in [0, 1]\}$ no es cerrado, por lo que el punto x_1 no es tampoco regular Lagrangiano, aunque sí se trata de un óptimo como puede comprobarse fácilmente.

Vamos a relacionar la cualificación de restricciones de Mangasarian-Fromovitz con la regularidad Lagrangiana. Por un lado, cuando se trabaja con puntos posibles x cuyo $\text{constr}(x)$ es finito, la ventaja que presenta M-F, frente a la condición de punto regular Lagrangiano, es que se puede comprobar con más facilidad aunque, por contra, la cualificación M-F es más fuerte en el sentido de que todo punto x que satisfaga M-F será regular Lagrangiano. El recíproco no es cierto.

En cuanto al caso de que tengamos puntos con infinitas restricciones activas, aunque no es tan evidente la ventaja de comprobarla, es posible que en muchos casos sea más sencilla la cualificación M-F que la condición de regularidad Lagrangiana.

Teorema 2.12

Sea $\bar{x} \in F$ un punto que satisface la condición M-F, entonces \bar{x} es un punto regular Lagrangiano.

Demostración.

Sabemos que existe $d \in \mathbb{R}^n$ tal que $a^T(s)d > 0$ para $s \in \text{constr}(\bar{x})$. Supongamos que \bar{x} no fuera regular Lagrangiano, entonces existiría $y \in B(\bar{x})^* \setminus T(F, \bar{x})$.

Como $y \notin T(F, \bar{x}) = \text{cl } K(F - \bar{x})$, existirá un entorno de y : $U(y) \subseteq \mathbb{R}^n \setminus \text{cl } K(F - \bar{x})$.

Consideremos $v = y + \gamma d$, escogiendo $\gamma > 0$ suficientemente pequeño se tiene que $v \in U(y)$, por tanto $v \notin \text{cl } K(F - \bar{x})$.

Sin embargo, notemos que $a^T(s)v = a^T(s)y + \gamma a^T(s)d > 0$ para $s \in \text{constr}(\bar{x})$. Sea $\mu = \inf \left\{ \frac{-z(s)}{a^T(s)v} : a^T(s)v < 0 \right\}$ análogamente a la demostración del Teorema 2.3 se comprueba que $\mu > 0$. Sea $x_1 = \bar{x} + \mu v$, $x_1 \in F$ por construcción. Además $v = \frac{1}{\mu}(x_1 - \bar{x})$ y entonces $v \in \text{cl } K(F - \bar{x})$, con lo que llegamos a una contradicción.

Tanto en el caso de cardinal de constr (\bar{x}) finito, en el que es evidentemente cierto, como en el infinito en que lo hemos probado, podemos afirmar que $K\{B(\bar{x})\}$ es cerrado. Por tanto, \bar{x} es regular Lagrangiano. ♦

Veamos que el recíproco no se cumple.

Ejemplo 2.9

Consideremos el siguiente problema

(P₉) Min x_1

$$\text{s.a. } (s-1)x_1 + s(s-2)x_2 \geq 0 \quad s \in [0,2]$$

Tenemos que $F = \left\{ \begin{bmatrix} 0 \\ \delta \end{bmatrix} : \delta \leq 0 \right\}$ y todas las soluciones posibles son óptimas.

Sea $\bar{x} = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \in F$ que no es punto extremo, con $\text{constr}(\bar{x}) = \{0,2\}$ y $A(\bar{x}) = \left\{ \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$.

Dado que $0_2 = \frac{1}{2} \begin{bmatrix} -1 \\ 0 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, \bar{x} no verifica M-F.

Sin embargo \bar{x} si es regular Lagrangiano puesto que: $B(\bar{x})^* = \{d \in \mathbb{R}^2 : a^T(s_i)d \geq 0$

$$s_i \in \text{constr}(\bar{x})\} = \{d \in \mathbb{R}^2 : d_1 = 0\} = \left\{ \begin{bmatrix} 0 \\ \delta \end{bmatrix} : \delta \in \mathbb{R} \right\}.$$

Veamos que $B(\bar{x})^* \subseteq T(F, \bar{x})$.

(i) Si $\delta < 0$

$$\begin{bmatrix} 0 \\ \delta \end{bmatrix} = \lim_{k \rightarrow \infty} -k\delta \left(\begin{bmatrix} 0 \\ -1-1/k \end{bmatrix} - \begin{bmatrix} 0 \\ -1 \end{bmatrix} \right), \text{ es decir } \lambda_k = -k\delta > 0 \text{ y } x^k = \begin{bmatrix} 0 \\ -1-1/k \end{bmatrix} \in F \text{ para } k \geq 1$$

(ii) Si $\delta > 0$

$$\begin{bmatrix} 0 \\ \delta \end{bmatrix} = \lim_{k \rightarrow \infty} k\delta \left(\begin{bmatrix} 0 \\ -1+1/k \end{bmatrix} - \begin{bmatrix} 0 \\ -1 \end{bmatrix} \right), \lambda_k = k\delta > 0 \text{ y } x^k = \begin{bmatrix} 0 \\ -1+1/k \end{bmatrix} \in F \text{ para } k \geq 2$$

(iii) Si $\delta = 0$, tomamos $x^k = \bar{x}$, $\lambda_k = 1 \quad \forall k$.

En cada caso, tenemos que $\begin{bmatrix} 0 \\ \delta \end{bmatrix} = \lim_{k \rightarrow \infty} \lambda_k (x^k - \bar{x})$.

Sección 2.4 Cualificaciones de restricciones globales para Programación Semi-Infinita Lineal.

Para el problema de PSI convexa se ha estudiado ya la relación entre la condición de regularidad lagrangiana y la condición de Slater. En López y Vercher (83) se demuestra que la condición de Slater implica que todo punto posible es regular Lagrangiano, y que ambas condiciones garantizan que cada solución óptima tiene un punto de Kuhn-Tucker asociado.

Ahora veremos un resultado muy interesante acerca de la relación entre M-F y la cualificación de restricciones de Slater, la más utilizada de las condiciones de regularidad en programación semi-infinita.

Definición 2.4

Para el problema continuo (PSIL) la condición de Slater puede enunciarse: existe $\bar{x} \in F$ tal que $a^T(s)\bar{x} > b(s)$, para todo $s \in S$.

Notemos que tanto la cualificación M-F como la de punto regular Lagrangiano son cualificaciones de restricciones de tipo local, mientras que la condición de Slater es global.

La condición de Slater es una de las cualificaciones de restricciones más utilizadas en PNL, y es bien conocido que esta condición es más fuerte que la cualificación de Cottle. En particular, en Krabs (79) puede encontrarse una demostración de este resultado para un problema de optimización en espacios vectoriales normados con infinitas restricciones (Lema III 3.10). A continuación presentamos una prueba alternativa para nuestro caso.

Teorema 2.13

Si el problema (PSIL) satisface la condición de Slater, entonces la cualificación M-F se verifica para todo punto posible con alguna restricción activa.

Demostración.

Sea $x \in F$ con $\text{constr}(x) \neq \emptyset$. Por satisfacerse la condición de Slater, existe $\bar{x} \in F$ tal que $a^T(s)\bar{x} > b(s)$, $\forall s \in S$. Sea $d = \bar{x} - x$, entonces $a^T(s)d = a^T(s)\bar{x} - a^T(s)x = a^T(s)\bar{x} - b(s) > 0$ para todo $s \in \text{constr}(x)$, por tanto se satisface F-M. ♦

Otra demostración de este resultado puede obtenerse a partir de los trabajos de Kurcyusz (76) y Zowe (78) en los que se establece la equivalencia entre dos condiciones de regularidad tipo Slater, que son cualificaciones de restricciones válidas para el problema de Programación Matemática en espacios de Banach. En particular, una de esas cualificaciones podía ser explicitada como una condición sobre la parte activa de las restricciones y garantizaba la existencia de multiplicadores Lagrangianos en el caso de que el punto candidato fuese óptimo.

Hasta este momento tenemos la siguiente cadena de implicaciones:

Independencia lineal

⇓

Slater \Rightarrow M-F \equiv Cottle \Rightarrow regularidad Lagrangiana.

Además, la condición de Slater no implica que los vectores de $A(x)$ sean linealmente independientes, para aquellos puntos $x \in F$ tales que $A(x) \neq \emptyset$ y $\text{card}(\text{constr}(x)) \leq n$.

Ejemplo 2.10

(P₁₀) Min $x_1 + 1/2x_2$
 s.a. $x_1 + s^2x_2 \geq 0 \quad s \in [-1,1]$

Comprobamos que se satisface la cualificación de Slater tomando $\bar{x} = (1, 0)^T$, pues $z(\bar{x}) = 1 > 0 \quad \forall s \in [-1,1]$.

Consideramos $x = (1, -1)^T$ cuya función de holgura es $z(s) = 1 - s^2$ y $\text{constr}(x) = \{-1, 1\}$. Tenemos que $A(x) = \{a(1), a(-1)\}$, que son vectores linealmente dependientes.

Si pensamos en la cualificación M-F como en la cualificación de Cottle, el Teorema 2.13 responde a lo que cabía esperar en el contexto de las implicaciones conocidas entre cualificaciones de restricciones en PNL. Ahora bien, en el siguiente teorema se establece un resultado nuevo: "la cualificación de restricciones de Mangasarian- Fromovitz implica la de Slater", que no ocurre, por ejemplo, en programación no lineal finita.

Teorema 2.14

Si existe algún punto posible que verifique M-F, entonces se cumple la condición de Slater.

Demostración.

Supongamos que la cualificación de Slater no se cumple. Para cada $x \in F$ existirá $s(x) \in S$ tal que $z(s)=0$, es decir, todo punto posible tiene al menos una restricción activa. Por hipótesis hay un punto \bar{x} posible de forma que existe $d \in \mathbb{R}^n$ tal que $a^T(s)d > 0 \forall s \in \text{constr}(\bar{x})$. Para cada $\lambda \geq 0$ definimos $x^\lambda = \bar{x} + \lambda d$ y $z^\lambda(s) = \bar{z}(s) + \lambda a^T(s)d$.

Distinguimos dos casos:

(1) Si $a^T(s)d \geq 0 \forall s \in S$, entonces $\forall \lambda \geq 0 \ x^\lambda \in F$ y $z^\lambda(s) > 0 \forall s \in S$.

(2) Si existe $s_0 \in S$ tal que $a^T(s_0)d < 0$, sean $\rho = \max\{-a^T(s)d : s \in S\}$, $\rho > 0$ y $\mu = \inf\{\frac{-z(s)}{a^T(s)d} : a^T(s)d < 0\}$. Razonando análogamente al Teorema 2.3 tenemos que $\mu > 0$.

Sea $\lambda = \mu/2 > 0$, vamos a comprobar que $z^\lambda(s) > 0$ para todo $s \in S$:

(a) Es obvio para $s \in \text{constr}(\bar{x})$ y para $s \in S - \text{constr}(\bar{x})$ tales que $a^T(s)d \geq 0$,

(b) Para $s \in S - \text{constr}(\bar{x})$ tales que $a^T(s)d < 0$, tenemos que $z^\lambda(s) = \bar{z}(s) + (\mu/2)a^T(s)d > \bar{z}(s) + \lambda a^T(s)d \geq 0$, luego $z^\lambda(s) > 0 \forall s \in S$.

En cada caso hemos encontrado algún $\lambda > 0$ tal que $x^\lambda \in F$ y $z^\lambda(s) > 0$ para todo $s \in S$, así pues llegamos a una contradicción por lo que podemos concluir que se verifica Slater. ♦

Corolario 2.14.1

Si un punto posible \bar{x} satisface la cualificación M-F, entonces todos los demás puntos posibles con un número finito (infinito) de restricciones activas la satisfacen también.

Demostración.

Inmediata a partir de los Teoremas 2.13 y 2.14. ♦

Por lo tanto, acabamos de demostrar para el problema continuo de PSIL que la condición de Slater es equivalente a que exista un punto posible \bar{x} tal que $0_n \notin \text{co}(A(\bar{x}))$, propiedad sencilla de verificar en el caso de un punto posible con finitas restricciones activas.

Como ya hemos visto en la Sección 2.1 hay una familia de problemas cuyo sistema de restricciones está formulado de manera que 0_n es un gradiente activo para todas las soluciones posibles del problema. En estos casos no se verificará la condición de Slater, por lo que tendríamos que utilizar otras cualificaciones más débiles: la

cualificación local de punto regular Lagrangiano o la cualificación de restricciones de Farkas-Minkowski, que veremos a continuación.

Hasta ahora hemos estudiado las extensiones al caso semi-infinito lineal de las cualificaciones de restricciones de Mangasarian- Fromovitz, Cottle, Guignard y Slater, establecidas inicialmente para problemas de Programación No Lineal. A continuación consideramos una condición específica de PSI que no tendría sentido en problemas finitos, la cualificación de restricciones de Farkas-Minkowski (ver Goberna et al (81)).

Definición 2.5

Diremos que un problema (PSI1) satisface la cualificación de restricciones de Farkas-Minkowski si el sistema $\{a^T(s)x \geq b(s) : s \in S\}$ es un sistema de Farkas-Minkowski.

López y Vercher (83) demuestran para el problema semi-infinito convexo, cuando el conjunto F de soluciones posibles es acotado, que la condición de Slater implica que se verifica la propiedad de Farkas-Minkowski y que la cualificación de Farkas-Minkowski permite asociar a cada solución óptima un punto de Kuhn-Tucker. En el siguiente teorema presentamos una demostración alternativa de este resultado para nuestro problema.

Teorema 2.15

Si el problema (PSIL) verifica la condición de Slater, entonces también se satisface la cualificación de Farkas-Minkowski.

Demostración.

Consideremos $\alpha: S \rightarrow (0, +)$ definida $\alpha(s)=1$ para cada $s \in S$. El conjunto $\left\{ \begin{pmatrix} a(s) \\ b(s) \end{pmatrix} : s \in S \right\}$ es compacto, ya que es la imagen continua de un conjunto compacto. Además por verificarse la condición de Slater existe un punto interior algebraico. El sistema $\left\{ a(s)^T x \geq b(s) : s \in S \right\}$ es canónicamente cerrado, y por tanto es un sistema de Farkas-Minkowski (ver corolario 3.1.1 en Goberna et al (81)). Por lo tanto el problema (PSIL) satisface la cualificación de Farkas- Minkowski. ♦

En general el recíproco no es cierto.

Ejemplo 2.11

Consideremos de nuevo el problema del Ejemplo 2.9.

(P₉) Min x_1

$$\text{s.a. } (s-1)x_1 + s(s-2)x_2 \geq 0 \quad s \in [0,2]$$

Sabemos que $F = \left\{ \begin{bmatrix} 0 \\ \delta \end{bmatrix} : \delta \leq 0 \right\}$ y que todas las soluciones posibles son óptimas, aunque

solo la solución $\tilde{x} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ es un punto extremo. Consideremos $x(\delta) = \begin{bmatrix} 0 \\ \delta \end{bmatrix}$ con $\delta < 0$, entonces $z(\delta) = \delta s(s-2) \geq 0$. No hay soluciones posibles sin restricciones activas y no se cumple Slater.

Sin embargo, el sistema $\{(s-1)x_1 + s(s-2)x_2 \geq 0 \quad s \in [0,2]\}$ es de Farkas-

Minkowski, pues el cono $K = K \left\{ \begin{bmatrix} a(s) \\ 0 \end{bmatrix} : s \in [0,2] \right\} = \left\{ \begin{bmatrix} \alpha \\ \beta \\ 0 \end{bmatrix} : \beta \leq 0 \right\}$ es cerrado

(Corolario 3.1.1 en Goberna et al (81)).

En cuanto a la relación que existe entre las cualificaciones de restricciones de regularidad Lagrangiana y de Farkas-Minkowski, tenemos que el hecho de que no se verifique la cualificación de restricciones de Farkas-Minkowski no significa que no haya puntos regulares Lagrangianos.

Ejemplo 2.12

(P₁₂) Min $x_1 + x_2$

$$\text{s.a. } sx_1 + s^2x_2 \geq s^3 \quad s \in [0,1]$$

El sistema que hemos de estudiar es $\{sx_1 + s^2x_2 \geq s^3 : s \in [0,1]\}$. Una relación consecuente de este sistema es: $x_1 \geq 0$. Sin embargo vamos a ver que no es relación consecuente para ningún subsistema finito:

$$s_1x_1 + s_1^2x_2 \geq s_1^3$$

$$s_2x_1 + s_2^2x_2 \geq s_2^3$$

.....

$$s_kx_1 + s_k^2x_2 \geq s_k^3$$

Podemos suponer que $s_i \neq 0$ y este sistema finito es equivalente a :

$$\begin{array}{lcl}
x_1 + s_1 x_2 \geq s_1^2 & & x_1 \geq s_1^2 - s_1 x_2 \\
x_1 + s_2 x_2 \geq s_2^2 & \Leftrightarrow & x_1 \geq s_2^2 - s_2 x_2 \\
\text{.....} & & \text{.....} \\
x_1 + s_k x_2 \geq s_k^2 & & x_1 \geq s_k^2 - s_k x_2
\end{array}$$

Tenemos que $x_1 = \max\{s_j(s_j-2)\} < 0$, pero $x_2 = 2$ es una solución del subsistema con $x_1 < 0$. Por tanto (P_{12}) no satisface la cualificación de restricciones de Farkas-Minkowski.

Se puede comprobar que $F = \{x \in \mathbb{R}^2: x_1 \geq 0, x_1 + x_2 \geq 1\}$. Los puntos de F de la forma $x(\alpha) = \begin{bmatrix} 0 \\ \alpha \end{bmatrix}$ con $\alpha > 1$, no son regulares Lagrangianos, ya que $B(x)^* = \mathbb{R}^2$ y $T(F, x) = \{y \in \mathbb{R}^2: y_1 \geq 0\}$. Sin embargo, los puntos $x(\beta) = \begin{pmatrix} \beta \\ 1-\beta \end{pmatrix}$ con $\beta > 0$ si lo son, puesto que $z(\beta) = s(1-s)(s+\beta)$ y entonces $z(\beta) = 0$ si y sólo si, $s=0$ ó $s=1$. Luego $A(x(\beta)) = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\} \forall \beta > 0$ y $B(x)^* = T(F, x) = \{y \in \mathbb{R}^2: y_1 + y_2 \geq 0\}$. Entonces todos los puntos $x(\beta)$, $\beta > 0$ satisfacen la regularidad lagrangiana y todos son óptimos. En particular, para $\tilde{x} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ se tiene que $c = 1a(1)$, luego \tilde{x} es solución óptima.

Conclusiones.

Para poder garantizar que cada solución óptima del problema (PSIL) tiene asociado un punto de Kuhn-Tucker es necesario comprobar alguna cualificación de restricciones. Una posibilidad es probar si se satisface alguna de las cualificaciones globales, lo que podría resultar bastante complicado en algunos casos. La otra opción es comprobar una cualificación de restricciones local sobre la solución inicial posible.

Después de lo que acabamos de demostrar, si se tiene un punto posible con un número finito de restricciones activas, resolviendo un sencillo problema lineal sabremos si se satisface M-F. En caso afirmativo, podemos asegurar que la cualificación de Slater también se satisface y que cada óptimo tendrá asociado un punto de Kuhn-Tucker. Si el punto posible tiene un número infinito de restricciones activas probaremos M-F o buscaremos otra solución posible sobre la que sea más sencillo verificar si se cumple esa condición. Puesto que, siendo una cualificación de restricciones local, su equivalencia con Slater nos demuestra que juega el papel de una cualificación global.

Por otra parte, se sabe que la condición de optimalidad de Kuhn-Tucker es siempre necesaria para los problemas finitos con restricciones lineales, debido a que

siempre se cumple la condición de Abadie. Este no es el caso para el problema de PSIL, aunque debido a la estructura y propiedades de la cualificación M-F parece que esta se satisfará para una amplia clase de problemas y para ellos la condición de optimalidad establecida en los Teoremas 2.3 y 2.4 será necesaria y suficiente.

CAPITULO III.

UN ALGORITMO DE PURIFICACION

CAPITULO III. Un algoritmo de purificación.

Introducción.

En este capítulo describimos una forma de acceder desde una solución posible que no es un punto extremo del problema continuo de PSIL (P), a una solución básica cuyo valor de la función objetivo es al menos tan bueno como el de la solución inicial. Estos procedimientos son conocidos como algoritmos de purificación desde los trabajos desarrollados por Charnes y Kortanek (63) y Charnes, Kortanek y Raike (65) en PL finita. Asimismo, Kortanek y Strojwas (84) presentan un algoritmo de purificación para un programa lineal de maximización sobre un cuerpo ordenado arbitrario, que es aplicable al problema dual de PSIL y que permite detectar la no acotación del problema. Sin embargo, para el problema primal no había aparecido ningún algoritmo de purificación hasta hace poco tiempo. Este hecho puede justificarse porque hasta los trabajos de Nash (85) no había ninguna caracterización operativa para los puntos extremos del problema primal.

Lewis (85) ha desarrollado el esquema de un algoritmo de purificación para problemas acotados en programación lineal general. Sin embargo, es bien sabido que la existencia de direcciones de recesión en el conjunto factible no implica que el valor de la función objetivo sea no acotado. En la Sección 3.2 hemos desarrollado un procedimiento que puede utilizarse para cualquier problema semi-infinito, y hemos demostrado que tras aplicar este algoritmo ó bien encontramos un punto extremo, mejor o tan bueno como el punto de partida, ó detectamos la no acotación del problema.

En programación semi-infinita se han establecido algunas condiciones de recesión generales (ver p.e. Borwein (81) y Karney (81, 83)) con el propósito de garantizar diversas propiedades del problema (P): acotación, aproximación discreta, existencia de solución óptima, etc. En la Sección 3.1 analizamos el significado geométrico de algunas de estas condiciones y sus relaciones.

En la Sección 3.3 se muestra cómo puede obtenerse un punto extremo mejor a partir de cualquier solución posible y demostramos que el algoritmo de purificación que presentamos converge a lo sumo en n etapas. Nuestro algoritmo de purificación no resuelve (P), pero proporciona un buen punto extremo a partir del cual las

iteraciones del Simplex pueden acceder a la solución óptima como se verá en el Capítulo IV. Además la operación de purificación no supone ninguna dificultad para mantener la factibilidad.

Hay que destacar que en cualquier procedimiento tipo Simplex desarrollado para resolver (P) es necesario disponer de un algoritmo de purificación, ya sea el descrito por Anderson y Lewis (89) ó el que presentamos en el Capítulo IV. Dado que, no sólo se usan para determinar la solución posible básica inicial sino también en cada etapa del Simplex, ya que los criterios de avance permiten conservar la factibilidad de la solución actual pero no su carácter básico. Además, el hecho de que no aumente el valor de la función objetivo puede acelerar la convergencia del método.

En la Sección 3.4, por último se dan varios ejemplos en los que se puede comprobar el funcionamiento del método, incluyendo algún comentario sobre los problemas de tipo computacional que puede presentar su implementación, debidos al grado de precisión con el que se necesita trabajar.

Sección 3.1 Condiciones de recesión.

Consideremos el problema:

$$(P) \text{ Min } c^T x \\ \text{s.a. } a^T(s)x \geq b(s) \quad s \in S$$

donde S es un subconjunto compacto de \mathbb{R} y las funciones a y b son continuas sobre S . Asumimos que para cada punto posible x la correspondiente función de holgura tiene un número finito de ceros y que existen las derivadas hasta el orden correspondiente de estos ceros. En particular, podemos asegurar que se cumple este requisito cuando las funciones a_i , $i=1, \dots, n$ y b son analíticas en S .

Recordemos que el problema dual de (P) es:

$$(D) \text{ Max } \sum \lambda(s)b(s) \\ \text{s.a. } \sum \lambda(s)a(s) = c \\ \lambda \in \mathbb{R}^{(T)}_+$$

Siendo $\mathbb{R}^{(T)}_+$ el cono de las sucesiones finitas positivas generalizadas incluido en $\mathbb{R}^{(T)}$, donde $\mathbb{R}^{(T)} = \{ \alpha: T \rightarrow \mathbb{R} \mid \alpha(t) = 0 \ \forall t \text{ excepto para un subconjunto finito de } t \}$

En primer lugar, introduciremos algunos resultados de Teoría de la recesión que son necesarios para el desarrollo de esta sección.

Sea B un subconjunto no vacío y convexo de \mathbb{R}^n . El cono de recesión de B es $O^+B = \{ y \in \mathbb{R}^n : x + \lambda y \in B \text{ para todo } \lambda \geq 0 \text{ y } x \in B \}$.

Al conjunto $(O^+B) \cap (-O^+B)$ se le llama espacio de linealidad de B y se denota por $L(B)$.

Para cualquier función convexa f , se puede demostrar que $O^+(\text{epi } f)$ es, de hecho, el epigrafo de una cierta función que recibe el nombre de función de recesión de f y que denotamos por fO^+ . Por definición $O^+(\text{epi } f) = \text{epi}(fO^+)$ y, por tanto, fO^+ es una función convexa positivamente homogénea. Al conjunto de todos los vectores y , tales que $fO^+(y) \leq 0$ se le llama cono de recesión de f , y se le denota $\text{rec } f$.

Denotaremos por $C(f)$ el espacio de constancia de f , donde $C(f) = \{ y \in \mathbb{R}^n :$

$(fO^+)(y) \leq 0$ y $(fO^+)(-y) \leq 0$ }. Las direcciones de los vectores que se encuentran en el espacio de constancia de la función f , son las direcciones en las que f es constante.

En particular, cuando $f(x) = c^T x$, $fO^+(x) = f(x)$ y los elementos de $C(f)$ son los vectores ortogonales a c .

Con el propósito de encontrar las condiciones más débiles para asegurar la convergencia del algoritmo de purificación que hemos diseñado y que se presentará en la siguiente sección, hemos estudiado diferentes hipótesis que se han utilizado en Programación Semi-Infinita para garantizar que el problema (P) es resoluble.

En Anderson y Lewis (89) se presenta un algoritmo de purificación para (P) que converge en un número finito de etapas cuando se satisface la siguiente condición:

Definición 3.1

(R1) " $A := \{ x \in \mathbb{R}^n : c^T x \leq 0, a(s)^T x \geq 0, s \in S \} = \{0_n\}$ ".

(R1) se cumple trivialmente si F es acotado. Analizaremos su significado geométrico usando resultados de la Teoría de la recesión y estudiaremos su relación con otras condiciones similares que han sido establecidas en Programación semi-infinita para asegurar la existencia de solución óptima de (P).

Teorema 3.1

Supongamos que F es no vacío y que (P) satisface la propiedad (R1). Entonces (P) está acotado y la función objetivo alcanza su infimo sobre F . Recíprocamente, si (P) es resoluble entonces ó bien se da (R1) ó bien $c^T y = 0$ para todo vector no nulo $y \in A$.

Demostración.

El cono de recesión del conjunto posible es $O^+F = \{ x \in \mathbb{R}^n : a(s)^T x \geq 0, s \in S \}$. La función objetivo coincide con su función de recesión y por tanto $\text{rec } f = \{ x \in \mathbb{R}^n : c^T x \leq 0 \}$. Entonces $A = O^+F \cap \text{rec } f$. Se sabe que si el conjunto posible F y la función objetivo no tienen direcciones de recesión en común (Teorema 27.3 en Rockafellar (70)) entonces la función objetivo alcanza su mínimo sobre F .

Recíprocamente, si F es un conjunto acotado sabemos que $O^+F = \{0_n\}$ y (R1) se da trivialmente. En otro caso, sea x^* una solución óptima de (P) y sea $y \in A$, $y \neq 0_n$, entonces para todo $\lambda \geq 0$ se tiene que $x^* + \lambda y \in F$ y $c^T(x^* + \lambda y) = c^T x^* + \lambda c^T y \leq c^T x^*$,

por tanto $c^T y = 0$. ♦

Observese que los elementos no nulos de A serían aquellos vectores de \mathbb{R}^n que son direcciones de recesión de la función objetivo de (P) y, al mismo tiempo, direcciones de recesión de la región posible F .

Por otro lado, si tenemos una solución óptima para (P) y F es no acotado el conjunto de direcciones de A debe ser ortogonal a c .

El siguiente ejemplo pone de manifiesto el significado del recíproco.

Ejemplo 3.1

$$(P_1) \inf x_1 \\ \text{s.a } x_1 + s x_2 \geq 1, \quad s \in [0, 1]$$

Es fácil comprobar que $O^+F = \{x \in \mathbb{R}^2 : x_1 \geq 0, x_1 + x_2 \geq 0\}$, luego F es un conjunto no acotado. Nótese que los puntos $x_\lambda = (1, \lambda)^T$, para todo $\lambda \geq 0$, son soluciones óptimas de (P_1) .

Así pues $v(P_1) = 1$. Verificándose, por otra parte, que $A = \{\lambda(0, 1)^T, \lambda \geq 0\}$, y por tanto el conjunto A es ortogonal a c .

Para problemas semi-infinitos numerables consistentes, Karney (81) demuestra que la condición (R1) es suficiente para que su problema primal sea discretizable, i.e. pueda aproximarse mediante una sucesión de programas finitos $\{P_n\}$ con $v(P) = \lim v(P_n)$, lo que sabemos es equivalente a que no exista fallo de dualidad (Duffin y Karlowitz (65)). Asimismo, Goberna y López (87) han probado que un PSI con una cantidad no numerable de restricciones es resoluble y discretizable si se cumple (R1).

Karney (83) introduce otra condición de recesión para garantizar que no existe fallo de dualidad en el contexto de la programación semi-infinita convexa:

Definición 3.2

(R2) " $f \circ^+(y) > 0$ para todo $y \in O^+F \cap M^\perp$ distinto de cero "

donde M está definido de la siguiente forma : $M = L(F) \cap C(f)$; $C(f)$ es el espacio de constancia de la función objetivo f y F es la región posible del problema convexo.

Además, si el problema es consistente y se verifica (R2), existe una solución óptima para el problema primal. En nuestro caso tenemos que $M = \{x \in \mathbb{R}^n : c^T x = 0, a(s)^T x = 0, \text{ para todo } s \in S\}$.

Veamos la relación que existe entre las condiciones (R1) y (R2).

Lema 3.2

Si (P) es consistente, entonces (R1) implica (R2). Además, si $M = \{0_n\}$, ambas condiciones son equivalentes.

Demostración.

Sea $y \neq 0_n$, tal que $y \in O^+F \cap M^\perp$. Si $c^T y \leq 0$, entonces $y \in O^+F \cap \text{rec } f=A$, y esta contradicción implica que $c^T y > 0$.

Si $M = \{0_n\}$, entonces $M^\perp = \mathbb{R}^n$ y (R2) puede enunciarse " $y \in O^+F / \{0_n\}$ implica $c^T y > 0$ ", y se sigue que $A = \{0_n\}$. ♦

Nota. Si se reescribe (R1) como : " $f_{O^+}(y) > 0$ para todo $y \in O^+F$ no nulo ", el Lema 3.2 sigue siendo cierto para un programa semi-infinito convexo.

López (87) demuestra que la condición (R2) es equivalente a " c pertenece al interior relativo del cono convexo generado por el conjunto $\{a(s), s \in S\}$ ", lo que nos da la interpretación geométrica de (R2).

Veamos ahora un ejemplo en el que se pone de manifiesto que (R2) no implica (R1). Hemos considerado un problema para el que $M \neq \{0_n\}$, y que verifica la condición (R2), entonces dado que $M = \{x \in \mathbb{R}^n : c^T x = 0, a(s)^T x = 0, \text{ para todo } s \in S\} \subseteq A$, (R1) no puede darse.

Ejemplo 3.2.

Consideremos el problema siguiente:

$$(P_2) \quad \inf \{2x_1 + 2x_2 + x_3\}$$

$$\text{s.a. } x_1 + x_2 + sx_3 \geq 1 \quad s \in [0,1]$$

Puede comprobarse fácilmente que $L(F) = \{y \in \mathbb{R}^3 : y = (\alpha, -\alpha, 0), \alpha \in \mathbb{R}\}$ y $C(f) = \{y \in \mathbb{R}^3 : 2y_1 + 2y_2 + y_3 = 0\}$. Entonces $M = L(F)$.

Nótese que M es un subespacio de \mathbb{R}^3 , y que $\dim M = 1$, entonces $M^\perp = \{y \in \mathbb{R}^3 : y_1 - y_2 = 0\}$.

Por otra parte, se tiene que $O^+F \cap M^\perp = \{(\beta, \beta, \alpha) \in \mathbb{R}^3 : 2\beta + \alpha \geq 0, \beta \in [0, 1]\}$. Sea $y \in O^+F \cap M^\perp$, tal que $y \neq 0_n$. Tenemos que $\beta \geq 0$, $2\beta + \alpha \geq 0$ y que $c^T y = 4\beta + \alpha$. Hay dos posibilidades:

(i) Si $\beta > 0$, entonces $4\beta + \alpha = 2\beta + (2\beta + \alpha) > 0$.

(ii) Si $\beta = 0$, entonces $\alpha \geq 0$, para todo $\beta \in [0, 1]$, pero $\alpha \neq 0$, por tanto $\alpha > 0$ y $c^T y = \alpha > 0$.

En cualquier caso se tiene que $c^T y > 0$, para todo $y \in O^+F \cap M^\perp$ distinto de cero, es decir (R2) se verifica. Podemos, pues asegurar que $v(P) = v(D)$ y el problema (P) alcanza su mínimo sobre F .

Otra condición de recesión ha sido establecida por Borwein (81) para el problema de minimizar una función lineal sobre un conjunto convexo cerrado F de \mathbb{R}^n . Esta condición asegura la acotación del problema. Para enunciarla necesitamos introducir previamente una definición.

Definición 3.3

Sea B un conjunto convexo cerrado de \mathbb{R}^n y sea $y \in \mathbb{R}^n$, tal que $y^T x \geq 0$ para todo $x \in O^+B$. Diremos que y es positivo sobre O^+B si $\{x \in \mathbb{R}^n : y^T x = 0, x \in O^+B\} = L(B)$. El conjunto de todos estos elementos positivos se denota por $\text{pos } O^+B$.

Nótese que $\text{pos}(O^+B)$ es el conjunto integrado por los vectores del cono dual de O^+B , $(O^+B)^+$, que son "lo más positivo que es posible" sobre O^+B .

Además, dado un cono K convexo cerrado en \mathbb{R}^n , en general no es cierto que $\text{pos } K = K^+$. Este hecho se puede comprobar considerando $K = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 \geq 0, x_2 \geq 0\}$. Se tiene que $K^+ = \{(y_1, y_2, 0) \in \mathbb{R}^3 : y_1 \geq 0, y_2 \geq 0\}$, y que $K \cap (-K) = \{(0, 0, x_3) \in \mathbb{R}^3 : x_3 \in \mathbb{R}\}$. Se observa que el vector $(1, 0, 0) \in K^+$ pero no está en $\text{pos}(K)$.

En nuestro contexto, la condición de Borwein (81) puede enunciarse de la siguiente forma:

Definición 3.4

(R3) " $c \in \text{pos } O^+F$ ".

A continuación vamos a probar su equivalencia con la condición (R2).

Teorema 3.3

Si el problema (P) es consistente, entonces (R2) y (R3) son equivalentes.

Demostración.

Supongamos que se verifica (R2) entonces $c^T y \geq 0$ para todo $y \in O^+F$, puesto que en otro caso $v(P) = -\infty$, lo que contradiría la hipótesis inicial. Sea $y \in O^+F$, tal que $c^T y = 0$ y supongamos que existe $s' \in S$ con $a(s')^T y > 0$, es decir que $c \notin \text{pos } O^+F$. Por ser M un subespacio de \mathbb{R}^n , se tiene que $\mathbb{R}^n = M \oplus M^\perp$. Así que existirán $m_1 \in M$ y $m_2 \in M^\perp$ tales que $y = m_1 + m_2$. Como $a(s')^T m_1 = 0$, tenemos que $m_2 \neq 0$. A continuación vamos a comprobar que $m_2 \in O^+F$. Para todo $s \in S$, $a(s)^T y = a(s)^T m_2 \geq 0$, entonces m_2 es también una dirección de recesión del conjunto posible. Por tanto, por verificarse (R2), $c^T m_2 > 0$, pero $c^T y = c^T m_2 > 0$, lo que contradice la hipótesis inicial. Entonces, para todo $s \in S$, se tiene que $a(s)^T y = 0$, y (R3) se cumple.

Recíprocamente, supongamos que no se verifica (R2), entonces existirá un vector no nulo $y \in O^+F \cap M^\perp$ tal que $c^T y < 0$. Como $c \in \text{pos } O^+F$, $c^T y \geq 0$, así que $c^T y = 0$. Entonces $y \in -O^+F$, por consiguiente $y \in M$ y tendremos que $y \in M \cap M^\perp = \{0_n\}$, con lo que llegamos a una contradicción. ♦

Observese que ninguna de las condiciones (R_i), $i=1, 2, 3$, se verifica en el Ejemplo 3.1, puesto que $M = \{0_n\}$ y c está en la frontera del cono convexo generado por los vectores $\{a(s), s \in S\}$. Por otra parte, en el Ejemplo 3.2 se tiene que el rango de $\{a(s), s \in S\} = 2$, y que las condiciones (R2) y (R3) se dan, aunque no sucede lo mismo con (R1).

El siguiente resultado es una condición suficiente para la equivalencia de estas tres condiciones de recesión.

Lema 3.4

Si $\text{rg}\{a(s), s \in S\} = n$, entonces $M = \{0_n\}$.

Demostración.

Sea $\{a(s_i), s_i \in S, i = 1, \dots, n\}$ una subfamilia de vectores linealmente independientes. Sabemos que $d \in M$ si y sólo si $c^T d = 0$ y $a(s)^T d = 0$ para todo $s \in S$, así que $a(s_i)^T d = 0, i = 1, \dots, n$ y en consecuencia $d = 0_n$. ♦

En cualquier caso, no parece que vaya a resultar sencillo verificar el cumplimiento de las condiciones (Ri) $i=1, 2, 3$, sobre todo cuando aumenta la dimensión del problema, por lo que tendríamos que considerar la posibilidad de que el problema planteado no esté ni tan siquiera acotado. Para asegurar la convergencia del algoritmo de purificación que proponemos en la siguiente sección, impondremos la siguiente condición:

Definición 3.5

Cuando en un problema (P) se cumplan:

- (i) el conjunto de soluciones posibles no es vacío, y
- (ii) $\text{rg}\{a(s), s \in S\} = n$.

diremos que se satisface la condición (RA).

La hipótesis (i) se justifica por sí misma. Con respecto a (ii), diremos que es esencial para que existan soluciones básicas en el problema dual standard y que se exige como condición necesaria previa para la aplicación de los algoritmos tipo simplex para el problema semi-infinito dual (ver p.e. Glashoff-Gustafson (83) y Goberna y Jornet (88)).

Además, según hemos demostrado, la condición (RA) garantiza la equivalencia de las tres condiciones de recesión consideradas para el problema primal semi-infinito lineal. Si además se verifica alguna de las condiciones (Ri), $i=1, 2, 3$, tendremos que (P) es resoluble y $v(P)=v(D)$. En otro caso pueden darse diferentes situaciones.

Sabemos por la Teoría de la dualidad (ver p.e. Glashoff y Gustafson (83)) que puede presentarse cualquiera de las situaciones indicadas con SI en el diagrama. Vamos a comprobar mediante distintos ejemplos que la exigencia de que se cumpla

(RA) no predetermina de ninguna forma el comportamiento del par dual.

		<u>PRIMAL</u>	
		<u>Acotado</u>	<u>No acotado</u>
<u>DUAL</u>	<u>Inconsistente</u>	SI	SI
	<u>Acotado</u>	SI	NO

Nota. Hemos excluido el caso en que (P) es inconsistente porque estamos suponiendo que se cumple (RA).

Ejemplo 3.3

Consideremos el problema

$$(P_3) \quad \text{Inf} \quad x_1 + 2x_2$$

$$\text{s.a.} \quad x_1 + sx_2 \geq \exp(s), \quad s \in [0,1]$$

Se verifica (RA) porque $\text{rg}\{a(s), s \in [0,1]\} = 2$, y $x^0 = (e, 0)^T \in F$. Puede comprobarse fácilmente que $A = \{y \in \mathbb{R}^2 : y_1 + y_2 = 0, y_1 + 2y_2 \leq 0\}$, luego (R1) no se verifica. Si tomamos la dirección $d = (2, -3/2)^T \in A$, tenemos que $c^T d = -1 < 0$, con lo que $v(P_3) = -\infty$ y el problema dual standard es inconsistente.

Ejemplo 3.4 (en Glashoff y Gustafson (83))

$$(P_4) \quad \text{Inf} \quad x_1$$

$$\text{s.a.} \quad x_1 + sx_2 \geq \sqrt{s}, \quad s \in [0,1]$$

De nuevo se satisface (RA), aunque no (R1). En este caso, c pertenece a la frontera relativa del cono convexo generado por $\{a(s), s \in S\}$ y puede darse cualquier situación.

El problema dual tiene una solución única : $\lambda(s) = 0$ para $s \in (0,1]$ y $\lambda(0) = 1$, por tanto $v(D_4) = 0$.

Por otra parte, tenemos que $v(P_4) = 0$ puesto que, según se puede comprobar, para todo $\epsilon > 0$ el vector $x(\epsilon) = (\epsilon/2, 1/(2\epsilon))^T$ es una solución posible con $c^T x(\epsilon) < \epsilon$. Sin

embargo, no existe ninguna solución posible con $x_1=0$. Veámoslo, si existiera una solución posible de la forma $x=(0, x_2)^T$, debería cumplirse que $x_1 + sx_2 \geq \sqrt{s}$, $s \in [0,1]$ en particular para $s_n=1/n^2$, $n \geq 1$ tendríamos que $(1/n^2) x_2 \geq 1/n$, $n \geq 1$ y, por tanto, $x_2 \geq n$, $n \geq 1$, lo que es imposible. Así que el infimo de (P_4) no puede alcanzarse. Entonces tanto (P_4) como (D_4) están acotados y no hay fallo de dualidad, aunque no se alcanza $v(P_4)$.

Para este problema no puede cumplirse (R1) porque esto implicaría que (P_4) alcanza su mínimo en F y hemos visto que esto no sucede, de todas maneras puede comprobarse directamente que $A = \{y \in \mathbb{R}^2 : y_1 \leq 0, y_1 + sy_2 \leq 0\} = \{y \in \mathbb{R}^2 : y_1 = 0, y_2 \geq 0\} = \{\lambda \begin{bmatrix} 0 \\ 1 \end{bmatrix} : \lambda \geq 0\} \neq \{0_2\}$. La única dirección de recesión es $d = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ que es ortogonal a c .

En cuanto al problema considerado en el Ejemplo 3.1, satisface la condición (RA); (P_1) es un problema acotado y se alcanza el mínimo. El problema dual standard tiene una única solución $\lambda(s)=0$ para $s \in (0,1]$ y $\lambda(0)=1$, entonces $v(D_1)=1$ y no existe fallo de dualidad.

Por último, veamos un ejemplo en el que (P) es acotado, (D) es inconsistente y se satisface (RA).

Ejemplo 3.5

Consideremos el siguiente problema:

$$(P_5) \quad \text{Inf } x_1 \\ \text{s.a. } sx_1 + s^2x_2 \geq s^3, \quad s \in [0,1]$$

cuyo problema dual standard asociado es :

$$(D_5) \quad \text{Max } \sum \lambda(s) s^3 \\ \text{s.a. } \sum \lambda(s) \begin{bmatrix} s \\ s^2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ \lambda(s) \in \mathbb{R}^+(S)_+$$

Obviamente (D_5) es inconsistente, ya que el sistema:

$$\sum \lambda(s) s = 1$$

$$\sum \lambda(s) s^2 = 0$$

$$\lambda(s) \geq 0$$

no tiene solución. Pero en cambio (P_5) es acotado, $v(P_5)=0$ y se alcanza el óptimo, como vamos a probar a continuación:

$s x_1 + s^2 x_2 \geq s^3$, $s \in [0,1]$ implica que $x_1 + s x_2 \geq s^2$, $s \in (0,1]$. Tomando límites cuando s tiende a 0 por la derecha tendríamos que $x_1 \geq 0$, por tanto $v(P_5) \geq 0$.

Sea $x = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, se puede comprobar fácilmente que $x \in F$ y además su coste es 0, por tanto x es solución óptima para (P_5) y $v(P_5)=0$.

Este problema verifica (RA), basta considerar el conjunto de vectores

$\left\{ \begin{bmatrix} 0.5 \\ 0.25 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$ cuyo rango es 2. Desde luego no puede verificarse ninguna de las condiciones (Ri), porque $v(P) \neq v(D)$, pero también puede comprobarse directamente sin dificultad que $A \neq \{0_2\}$.

Sección 3.2 Algoritmo de purificación.

Con el propósito de escribir el problema (P) en la forma standard de la PL, y tal como se hacía en el Capítulo I, introducimos la función de holgura $z(s) = a(s)^T x - b(s)$, así que el problema (P) puede reescribirse de la siguiente forma:

$$\begin{aligned} \text{(P)} \quad & \text{Inf } c^T x \\ & \text{s.a. } a(s)^T x - z(s) = b(s) \quad s \in S \\ & \quad x \in \mathbb{R}^n, z(s) \geq 0 \quad s \in S \end{aligned}$$

Podemos tomar sin pérdida de generalidad $S=[0,1]$. Nótese que cuando las funciones a_i y b son analíticas sobre S , z es también una función analítica así que ó es idénticamente nula ó bien tiene un número finito de ceros en $[0,1]$.

Nuestro método sólo es aplicable a puntos posibles con un número finito de restricciones activas. Sin embargo, es fácil comprobar lo que sucede cuando la función de holgura $z(s)$ en un punto x es idénticamente nula, es decir si $b(s) = \sum_1^n x_i a_i(s)$ para todo $s \in S$. En este caso, tenemos el siguiente resultado:

Lema 3.5

Supongamos que $b(s) = \sum_1^n \alpha_i a_i(s)$ para algún $\alpha \in \mathbb{R}^n$. Entonces, el problema (P) es equivalente a un problema homogéneo de PSIL, que ó bien es no acotado ó tiene una solución óptima trivial .

Demostración.

Vamos a reescribir el problema (P) haciendo un cambio de variable. Por hipótesis, tenemos que $a(s)^T x - b(s) = \sum_1^n x_i a_i(s) - \sum_1^n \alpha_i a_i(s) = \sum_1^n (x_i - \alpha_i) a_i(s)$. Sea y el vector de \mathbb{R}^n cuya componente i -ésima y_i es $y_i = x_i - \alpha_i$, $i=1, \dots, n$, es decir $x = y + \alpha$. Entonces el problema (P) puede escribirse como :

$$\begin{aligned} \text{(P')} \quad & \text{Inf } \{c^T y + c^T \alpha\} \\ & \text{s.a. } a(s)^T y \geq 0 \quad s \in S \\ & \quad y \in \mathbb{R}^n \end{aligned}$$

Siendo $y = 0_n$ una solución posible para (P') . Hay dos posibilidades :

(a) Si existe $y \in \mathbb{R}^n$ tal que $c^T y < 0$, $a(s)^T y \geq 0$ para todo $s \in S$, entonces (P') es no acotado.

(b) Si para todo $y \in \mathbb{R}^n$ tal que $c^T y < 0$, existe algún $s \in S$ tal que $a(s)^T y < 0$, entonces 0_n es solución óptima de (P) . Por tanto, $x = \alpha$ es una solución óptima para (P) . ♦

Para un problema concreto de estas características, la cuestión que resta es comprobar si existe $y \in \mathbb{R}^n$ tal que $c^T y < 0$, $a(s)^T y \geq 0$ para todo $s \in S$.

Ejemplo 3.6

Consideremos el problema

$$(P_6) \quad \text{Inf } x_1 \\ \text{s.a. } sx_1 + s^2 x_2 \geq s - s^2, \quad s \in [0,1]$$

En este caso, $\alpha = (1, -1)^T$ y el problema queda:

$$(P_6') \quad \text{Inf } y_1 \\ \text{s.a. } sy_1 + s^2 y_2 \geq 0, \quad s \in [0,1]$$

siendo $y_1 = x_1 - 1$ e $y_2 = x_2 + 1$. No existe, pues, un vector $y \in \mathbb{R}^2$ tal que $c^T y < 0$, $a(s)^T y \geq 0$ para todo $s \in S$, es decir tal que $y_1 < 0$, y $sy_1 + s^2 y_2 \geq 0$ para todo $s \in [0,1]$. Por tanto, $x = \alpha$ es solución óptima para (P_6) , aunque no es única ya que todos los

vectores $x(\lambda) = \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \lambda \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ son soluciones óptimas para el problema.

Habría una única solución si, y solamente si, $\{y \in \mathbb{R}^n : c^T y \leq 0, a(s)^T y \geq 0, s \in S\} = \{0_n\}$, es decir si y sólo si (P') verifica (R1).

A continuación, vamos a describir nuestro algoritmo de purificación. Se trata de un procedimiento de mejora de soluciones posibles para el problema primal de PSIL, que termina cuando alcanza un punto extremo o detecta la no acotación de (P) . La notación que utilizamos fue ya introducida en la Sección 1.3.

Como ya hemos mencionado anteriormente, el algoritmo de purificación que

Anderson y Lewis (89) desarrollan, como especialización del esquema propuesto por Lewis (85) para problemas de PL general, exige el cumplimiento de la condición (R1). Esto implica, en particular, la acotación de (P). Dicho método pasa de una solución posible a otra con un coste menor o igual al de la anterior y la dirección de descenso que utiliza es el gradiente del coste proyectado sobre el subespacio V cambiada de signo.

Veamos primero el esquema general de Lewis para problemas acotados:

Partiendo de una solución posible x :

- (1) Si $B(x) \cap N(A) = \{0\}$, STOP (x es un punto extremo).
- (2) Tomar un vector no nulo $y \in B(x) \cap N(A)$ con $\langle c, y \rangle \leq 0$
- (3) Hacer $x^* := x + ky$ con k tan grande como sea posible, x^* es el nuevo punto posible.
- (4) Tomar $x := x^*$
- (5) Volver a (1)

Lewis (85) demuestra que $B(x^*) \cap N(A)$ está contenido estrictamente en $B(x) \cap N(A)$, entonces si el punto inicial x^0 es tal que $B(x^0) \cap N(A)$ es de dimensión finita m , tras m pasos, como mucho, el algoritmo encontrará un punto extremo con un valor de la función objetivo menor o igual que el de x^0 .

Vamos a presentar a continuación nuestro algoritmo de purificación basado también en esta caracterización de punto extremo, y después demostraremos que el algoritmo está bien definido bajo la hipótesis (RA), que es compatible con la no acotación de (P).

Inicialización. Partimos de (x^0, z) solución posible para (P). Determinar $\text{constr}(x^0)$, $d(i) \ i=1, \dots, m$, y V . Tomar $x := x^0$.

Paso 1. Si $V = \{0_n\}$, STOP, (x, z) es un punto extremo.

Paso 2. Sea $y^* \in V$, $y^* \neq 0$ una solución óptima para el problema lineal

$$\text{Inf } c^T y$$

$$\text{s.a } y \in V$$

$$-1 \leq y_i \leq 1 \quad i=1, \dots, n$$

Determinar $c^* = c^T y^*$ y $\beta_1 = \inf \{a(s)^T y^* : s \in [0, 1]\}$. Hacer $\beta = \beta_1$.

Paso 3. Si $c^* < 0$, ir al paso 6.

Paso 4. Si $\beta < 0$, ir al paso 7.

Paso 5. Hallar $\beta_2 = \min\{-a(s)^T y^* : s \in [0,1]\}$. Tomar $\beta = \beta_2$, $y^* = -y^*$ e ir al paso 7.

Paso 6. Si $\beta \geq 0$ imprimir "P no acotado". END.

Paso 7. Hallar $\lambda = \max\{g(s) : s \in [0,1]\}$ y $\sigma_1, \sigma_2, \dots, \sigma_p \in [0,1]$ tales que $\lambda = g(\sigma_k)$ $k=1, \dots, p$. Donde $g(s)$ se define como :

$$g(s) = \begin{cases} -a(s)^T y^* / z(s) & \text{si } s \in [0,1] / \text{constr}(x) \\ -a^{(d(i)+1)}(s_i)^T y^* / z^{(d(i)+1)}(s_i) & \text{si } s = s_i \in \text{constr}(x) \end{cases}$$

Hacer $x := x + (1/\lambda)y^*$, $\text{constr}(x) := \text{constr}(x) \cup \{\sigma_1, \dots, \sigma_p\}$,

$V := V \cap \{x \in \mathbb{R}^n : a^{(j)}(\sigma_k)^T x = 0 \quad 0 \leq j \leq d(k) \quad k=1, \dots, p\}$ e ir al paso 1.

Nótese que sólo se ejecuta el Paso 5 cuando $c^* = 0$ y $\beta_1 \geq 0$. En cuyo caso, haciendo $\beta = \beta_2$ tendríamos $\beta < 0$, como se demuestra en la siguiente proposición.

Lema 3.6

Supongamos que (P) satisface (RA) entonces según el esquema del algoritmo, si $c^* = 0$ se tiene que $\beta < 0$.

Demostración.

Supongamos que $c^* = 0$, entonces $c^T y \geq 0$, para todo $y \in V$. Por ser V un subespacio vectorial se verifica que $c^T y = 0$, para todo $y \in V$, en particular $c^T(-y^*) = 0$.

Sea $\beta_1 = \min\{a(s)^T y^* : s \in [0,1]\}$ y $\beta_2 = \min\{-a(s)^T y^* : s \in [0,1]\}$ entonces ó $\beta_1 < 0$ ó $\beta_2 < 0$. En primer lugar, nótese que, por construcción $\beta_1 + \beta_2 \leq 0$. Supongamos que $\beta_1 \geq 0$, entonces :

(i) Si $\beta_1 > 0$, necesariamente $\beta_2 < 0$.

(ii) Si $\beta_1 = 0$, se tendrá que $\beta_2 \leq 0$. Comprobaremos, por reducción al absurdo, que $\beta_2 < 0$. Supongamos que $\beta_2 = 0$. Como $\beta_1 = \beta_2 = 0$, serán ciertas las desigualdades siguientes: $a(s)^T y^* \geq 0$, $s \in [0,1]$ y $-a(s)^T y^* \geq 0$, $s \in [0,1]$. Así que $a(s)^T y^* = 0$, $s \in [0,1]$, lo que contradice la hipótesis (RA) puesto que $y^* \neq 0_n$. ♦

Lema 3.7

En el Paso 7 del algoritmo se tiene que $\lambda > 0$.

Demostración.

Si $\lambda \leq 0$, en particular tendríamos que $(-a(s)^T y^* / z(s)) \leq 0$ para todo $s \in [0,1]/\text{constr}(x)$. Entonces $a(s)^T y^* \geq 0$, para todo $s \in [0,1]/\text{constr}(x)$ y dado que $a(s_i)^T y^* = 0$, para $s_i \in \text{constr}(x)$, tendríamos como consecuencia que $\beta \geq 0$. Pero del Lema 3.6 se sigue que antes de ir al Paso 7 se tiene que $\beta < 0$. Por tanto $1/\lambda$ y x están bien definidos. ♦

En la página siguiente mostramos el esquema del algoritmo de purificación.

Sea $(x; z)$ una solución posible para (P) .
Hallar: $\text{constr}(x)$, $d(i)$ y V .

$V = \{0\}$

SI

NO

STOP
 $(x; z)$ es punto extremo.

Min $c^T y$
s.a. $y \in V$
 $-1 \leq y_k \leq 1 \quad k=1, 2, \dots, n$

Sea y^* la solución óptima, $c^* = c^T y^*$
y $\beta_1 = \min\{a^T(s)y^* : s \in S\}$.
Tomar $\beta = \beta_1$

$\beta \geq 0$

SI

$c^* < 0$

SI

STOP
 (P) no está acotado

NO

NO

Hallar $\beta_2 = \min\{-a^T(s)y^* : s \in S\}$
Tomar $\beta = \beta_1$ e $y^* = -y^*$

Hallar $\lambda = \max\{g(s) : s \in S\}$ y $\sigma_1, \sigma_2, \dots, \sigma_p$ tales que $\lambda = g(s_k) \quad k=1, 2, \dots, p$

$$g(s) = \begin{cases} -a(s)^T y^* / z(s) & s \in S - \text{constr}(x) \\ -a^{d(i)+1}(s_i)^T y^* / z^{d(i)+1}(s_i) & s = s_i \in \text{constr}(x) \end{cases}$$

Hacer $x := x + (1/\lambda) y^*$, $\text{constr}(x) := \text{constr}(x) \cup \{\sigma_1, \sigma_2, \dots, \sigma_p\}$ y
 $V := V \cap \{x \in \mathbb{R}^n : a^{(i)}(\sigma_k)^T x = 0, 0 \leq j \leq d(k), k=1, \dots, p\}$.

Sección 3.3. Análisis de la convergencia.

El algoritmo descrito en la sección anterior tiene claramente dos fases. En primer lugar, se determina una dirección de descenso óptima mediante la resolución de un programa lineal. Después nos movemos a lo largo de esta dirección mientras permanezcamos dentro de la región posible. Es decir, el algoritmo sigue el esquema típico de un método de direcciones posibles (ver por ejemplo Zoutendijk (76)). Además, tal y como determinamos la longitud de salto, reducimos la dimensión del siguiente subespacio, debido a que se añaden nuevas restricciones independientes, lo que nos permite asegurar que obtenemos un punto extremo tras un número finito de iteraciones.

Según hemos demostrado, al resolver el programa lineal en el Paso 2 del algoritmo, ó bien determinamos una dirección posible que es además dirección de recesión del conjunto de soluciones posibles (esto sucede cuando $c^* < 0$ y $\beta \geq 0$) ó bien obtenemos una dirección que nos conduce a un nuevo punto posible para (P), cuyo valor en la función objetivo es menor o igual que en el punto posible predecesor.

En primer lugar, veremos que los puntos obtenidos mediante nuestro algoritmo son puntos posibles, con la propiedad de que su función de holgura mantiene todos los ceros de la función de holgura previa. Además en el Paso 7 ó se obtiene un nuevo cero ó se aumenta la multiplicidad de uno antiguo.

Lema 3.8

Sea $x^1 = x + (1/\lambda)y^*$ y $z_1(s)$ su correspondiente función de holgura entonces x^1 es una solución posible, es decir, $z_1(s) \geq 0$, para $s \in [0,1]$. Además, $z_1^{(j)}(s_j) = 0$, $j=0,1,\dots,d(i)$, $i=1,\dots,m$.

Demostración.

Sabemos que $z_1(s) = z(s) + (1/\lambda)a(s)^T y^*$ con $\lambda > 0$. Para comprobar que $z_1(s) \geq 0$ para todo $s \in [0,1]$ consideramos dos casos:

- (i) Si $s \in \text{constr}(x)$, por definición de y^* , tenemos que $z_1(s) = 0$.
- (ii) Si $s \notin \text{constr}(x)$, entonces $z_1(s) \geq 0$ si y sólo si $\lambda z(s) \geq -a(s)^T y^*$, pero tal y como se determina λ tenemos que $\lambda \geq -a(s)^T y^* / z(s)$.

Para terminar la demostración, notese que para cada $s_j \in \text{constr}(x)$ se tiene que $z^{(k)}(s_j) = 0$, $k=0,1,\dots,d(i)$, por ser $y^* \in V$. ♦

Nota. El valor de λ se determina de manera que la longitud de salto en la dirección y^* sea máxima sin abandonar F . Veámoslo por reducción al absurdo, supongamos que existiera $\gamma > 1/\lambda$ tal que $x_\gamma \in F$, como $\lambda > 1/\gamma$, por definición de λ sabemos que debería existir $\sigma \in S$ tal que $\lambda > g(\sigma) > 1/\gamma > 0$. Por tanto, $\lambda > -a(\sigma)^T y^* / z(\sigma) > 1/\gamma$, entonces $\gamma > -z(\sigma) / a(\sigma)^T y^*$ y $-a(\sigma)^T y^* \gamma < -z(\sigma)$, dado que $z_\gamma(s) = z(s) + \gamma a(s)^T y^*$ se tendría $z_\gamma(\sigma) = z(\sigma) + \gamma a(\sigma)^T y^* < z(\sigma) - z(\sigma) = 0$, lo que contradice el supuesto inicial.

Lema 3.9

Supongamos que $\lambda = g(\sigma)$ para un cierto $\sigma \in S$.

- (i) Si $\sigma \notin \text{constr}(x)$, entonces $z_1(\sigma) = 0$, es decir σ es un nuevo cero.
- (ii) Si $\sigma = s_j$ para algún $j \in \{1, \dots, m\}$, entonces $z_1(s_j) = 0$ y $d^1(j) > d(j)$, donde $d^1(j)$ denota la multiplicidad de s_j como cero de $z_1(s)$.

Demostración.

Por definición de $g(s)$ tenemos que:

- (i) $\lambda = -a(\sigma)^T y^* / z(\sigma)$, entonces $z_1(\sigma) = z(\sigma) + (1/\lambda) a(\sigma)^T y^* = 0$.
- (ii) $\lambda = -a^{(d(j)+1)}(s_j)^T y^* / z^{(d(j)+1)}(s_j)$. Por el Lema 3.8 sabemos que $z_1^{(k)}(s_j) = 0$, $0 \leq k \leq d(j)$, así que $d^1(j) \geq d(j)$. Tomando derivadas a ambos lados de la igualdad $z_1(s) = z(s) + (1/\lambda) a(s)^T y^*$, tenemos que: $z_1^{(d(j)+1)}(s) = z^{(d(j)+1)}(s) + (1/\lambda) a^{(d(j)+1)}(s)^T y^*$, por tanto $z^{(d(j)+1)}(s_j) = 0$ y $d^1(j) > d(j)$. ♦

En el teorema que sigue a continuación demostramos la convergencia finita del método, y posteriormente damos un ejemplo en el que se detecta la no acotación de (P).

Teorema 3.10

Supongamos que (P) satisface (RA). Entonces, el algoritmo de purificación converge a un punto extremo, cuyo valor de la función objetivo es al menos tan bueno como el del punto inicial, en n etapas a lo sumo, o termina al detectar la no acotación de (P).

Demostración.

El algoritmo tiene dos criterios de parada

- (1) Si el algoritmo para con el mensaje "P no acotado", como $\beta \geq 0$ tendremos que $a(s)^T y^* \geq 0$, para $s \in [0, 1]$. Por tanto $y^* \in O^+F$ y la función objetivo disminuye

estrictamente a lo largo de la dirección y^* ya que $c^* < 0$ (notar que si $c^* = 0$ del Paso 5 se iría al Paso 7).

(2) Si no sucede esto, es decir si no se ejecuta el Paso 6, veamos que entonces el algoritmo termina en "algún" Paso 1. En este caso, en cada iteración, tanto si $c^* < 0$ como si $c^* = 0$ (por el esquema del algoritmo y el Lema 3.6) tenemos que $\beta < 0$, así que en el Paso 7 encontramos un nuevo punto posible $x^1 = x + (1/\lambda)y^*$ (ver Lema 3.8).

Sea V^1 el subespacio asociado a x^1 , como $y^* \in V/V^1$, el nuevo subespacio V^1 está contenido estrictamente en V . Como cada V es un subespacio de \mathbb{R}^n , en n iteraciones como máximo tendremos que $V = \{0_n\}$. ♦

Ejemplo 3.7

De nuevo consideremos el problema del Ejemplo 3.3:

$$(P_3) \quad \begin{array}{ll} \text{Inf} & x_1 + 2x_2 \\ \text{s.a.} & x_1 + sx_2 \geq \exp(s) \quad s \in [0,1] \end{array}$$

Recordemos que (P) satisface (RA). Tomamos $x = (e, 0)^T$ como punto de partida inicial, $z(s) = e - e^s : s \in [0,1]$ y $\text{constr}(x) = \{1\}$ con $d(1) = 0$. Como $V = \{x \in \mathbb{R}^2 : x_1 + x_2 = 0\} \neq \{0_2\}$, x no es un punto extremo.

Paso 2 : $y^* = (1, -1)^T$ es una solución óptima para el PL siguiente

$$\begin{array}{ll} \text{inf} & y_1 + 2y_2 \\ \text{s.a.} & y_1 + y_2 = 0 \\ & -1 \leq y_i \leq 1 \quad i=1,2 \end{array}$$

así que $c^* = -1 < 0$ y vamos al Paso 6. Como $\beta = \min\{1-s : s \in [0,1]\} = 0$, STOP: " P no acotado ".

Sección 3.4 Resultados numéricos.

Para implementar el algoritmo necesitamos utilizar algunas subrutinas matemáticas. En primer lugar un procedimiento que resuelva el problema de programación lineal generado en el Paso 2 y que nos proporcione una solución suficientemente precisa. Además necesitamos calcular máximos y mínimos locales de algunas funciones sobre un intervalo cerrado. Estas funciones pueden no ser continuas en todos los puntos. De hecho, $g(s)$ tiene un número finito de discontinuidades evitables, aunque $z(s)$ es continua en S .

Para resolver los problemas de programación lineal hemos utilizado el paquete XMP de Marsten (81).

Nótese que y^* debe calcularse con precisión porque en otro caso podríamos tener problemas en el cálculo de λ , y los ceros de la variable de holgura podrían cambiar de una iteración a la siguiente, lo que nos obligaría a remodelar V .

Todos los problemas que hemos considerado en esta Sección son de aproximación L_1 , que es, como ya se ha comentado, una de las aplicaciones clásicas de nuestro problema y sobre la que insistiremos algo más en el próximo capítulo.

Ejemplo 3.8

$$(P) \quad \text{Inf } \sum_1^n (1/i) x_i$$

$$\text{s.a. } \sum_1^n s^{i-1} x_i \geq 1/(1+s^2) \quad s \in [0, 1].$$

Este problema ha sido estudiado para diversos valores de n y los resultados se muestran a continuación.

$n=5$:

Iteración	x	constr(x), $d(i)$	y^*
1	(1,0,0,0,0)	{0}, $d(0)=1$	(0,0,-1,-1,-1)
2	(1,0,-1/6,-1/6,-1/6)	{0,1}, $d(0)=1, d(1)=0$	(0,0,-1,0,1)
3	(1,0,-2/3,-1/6,1/3)	{0,1}, $d(0)=1, d(1)=1$	(0,0,-1/2,1,1/2)
4	(1,0,-1,1/2,0)	{0,1}, $d(0)=2, d(1)=1$	****

$$\theta^* = -0.383333$$

$n=6$:

Iteración	x	constr(x), d(i)	y*
1	(1,0,-1,1/2,0,0)	{0,1}, d(0)=2, d(1)=1	(0,0,0,-1/2,1,-1/2)
2	(1,0,-1,1/4,1/2,-1/4)	{0,1}, d(0)=2, d(1)=2	****

$\theta^* = -0.466667$

$n=7$:

Iteración	x	constr(x), d(i)	y*
1	(1,0,-1,1/4,1/2,-1/4,0)	{0,1}, d(0)=2, d(1)=2	(0,0,0,-1/3,1,-1,1/3)
2	(1,0,-1,0,5/4,-1,1/4)	{0,1}, d(0)=3, d(1)=3	****

$\theta^* = -0.573809$

$n=9$:

Iteración	x	constr(x), d(i)	y*
1	(1,0,-1,0,5/4,-1,1/4,0,0)	{0,1}, d(0)=3, d(1)=3	(0,0,0,0,-1/6,2/3,-1,2/3,-1/6)
2	(1,0,-1,0,19/16,-3/4,-1/8,1/4,-1/16)	{0,1}, d(0)=3, d(1)=4	*****

$\theta^* = -0.737698$

Nota. Los casos $n=5, 6, 7$ y 9 han sido resueltos secuencialmente, utilizando cada vez la solución anterior como punto de partida inicial x para el siguiente problema.

En este ejemplo el algoritmo funciona según nuestras expectativas. Nótese que, en esta serie de problemas no hemos obtenido ningún cero de la variable de holgura dentro del intervalo $(0,1)$. Sólomente hemos obtenido un cero nuevo una vez (para $n=5$), en los demás casos se incrementaba la multiplicidad de un cero antiguo. Por supuesto, se mantenían cada vez los ceros antiguos.

El valor de la función objetivo disminuye estrictamente en cada iteración y hemos calculado el valor de θ^* para comprobar si en algún caso habíamos llegado a la solución óptima, pero no ha sido así.

En los ejemplos que siguen, aunque el algoritmo funciona bien, previene de las dificultades que pueden aparecer al aplicar el método a ciertos problemas.

En lo que sigue x^0 denotará el punto de partida inicial y x^* el punto extremo que se alcanza cuando termina la ejecución del algoritmo. C^* denota el conjunto de



restricciones activas en x^* y f^* el valor de la función objetivo en x^* .

Ejemplo 3.9 (en Anderson y Lewis (89))

$$(P_9) \quad \text{Inf} \sum_1^7 (1/i) x_i$$

$$\text{s.a.} \quad \sum_1^7 s^{i-1} x_i \geq -\sum_0^4 s^{2i} \quad s \in [0, 1].$$

Hemos aplicado nuestro método a este problema partiendo de la misma solución posible inicial que Anderson y Lewis (89). Tras seis pasos de su algoritmo, se alcanza un punto extremo tal que, aunque no indican exactamente la composición de $\text{constr}(x^*)$, hay un cero en el interior de grado 1 y los demás elementos del constr son 0 y 1 con $d(0)=2$ y $d(1)=1$. El valor de la función objetivo en la solución básica que encuentran es -1.785. En nuestro caso, también tras seis pasos, los resultados son los siguientes:

$$x^0=(10, 0, 0, 0, 0, 0, 0); \quad x^*=(-1, 0, 2.082883, -8.715998, 3.650722, 6.515089, -7.532669); \quad C^*=\{0, 1, 0.967\},$$

con $d(0)=1$, $d(1)=2$, $d(0.967)=1$ y $f^*=-1.744802$.

Nota. Dado que la función de holgura $z(s)$ asociada a un punto posible x es siempre mayor o igual que cero, determinamos el conjunto $\text{constr}(x)$ mediante el cálculo de los mínimos locales de esta. Cuando se tiene un mínimo local s^* en el interior de S puede suceder que sea un mínimo local débil, i.e. $z(s)=0$ en un entorno N de s^* , de tal forma que, en la siguiente iteración, el mínimo local de la nueva función de holgura $z_1(s)$ se alcanza en otro punto $s^1 \in N$, $s^1 \neq s^*$, aunque $z_1(s^*)=0$. Dependiendo del valor numérico de $z_1^{(1)}(s^*)$, s^* puede mantenerse, para evitar tener que remodelar V , o no.

En la práctica, en la quinta iteración del Ejemplo 3.9, se alcanza un mínimo local de $z(s)$ en 0.964, aunque $z(0.967)=11 \cdot 10^{-6}$. Como $z^{(1)}(0.967)=11 \cdot 10^{-5}$, hemos optado por conservar el antiguo cero. Sin embargo, en el Ejemplo 3.10, hemos remodelado V en cada iteración, lo que nos proporciona, como ventaja adicional, una mayor precisión en la definición de $g(s)$.

Ejemplo 3.10 (en Roleff (79))

$$(P) \quad \text{Inf } \sum_1^n (1/i) x_i$$

$$\text{s.a. } \sum_1^n s^{i-1} x_i \geq \tan(s) \quad s \in [0, 1].$$

Para este problema test de Roleff (79) conocemos los valores óptimos de cada uno de los problemas (pueden encontrarse en Coope y Watson (85)), y por tanto, podemos compararlos con el valor de la función objetivo f^* en los puntos extremos que hemos obtenido.

$n=3$

Iteración	x	constr(x), d(i)	y*
1	(tan(1),0,0)	{1}, d(1)=0	(-1,0,1)
2	(.210445,0,1.346963)	{.464,1}, d(.464)=1, d(1)=0	****

$\theta^* = -0.044966$

$n=6$

Iteración	x	constr(x), d(i)	y*
1	(.210445,0,1.346963,0,0,0)	{.464,1}, d(.464)=1 d(1)=0	(-.259,1,-.458,-1,-.193,1)
2	(.0604,.5799,1.0291,-.5799,-.1121,.5799)	{.472,.677,1} d(.472)=1=d(.677), d(1)=0	(-.024,.197,.633,1,.774,.235)
3	(0,1.0749,-.5641,1.9369,-2.0609,1.1706)	{0,.472,.677,1} d(.472)=1=d(.677), d(1)=0=d(0)	****

$\theta^* = -0.008744$

$n=8$

Iteración	x	constr(x), d(i)	y*
1	(0,1.0751,-.5655,1.9404,-2.0644,1.1719,0,0)	{0,.473,.677,1} d(.473)=1=d(.677) d(1)=0=d(0)	(0,-.052,.383,-1,.95,.261,-1,.459)
2	(0,1.0018,-.0313,.5462,-.7401,1.5355,-1.3942,.6395)	{0,.21,.471,.68,1} d(.471)=1=d(.68)=d(.21) d(1)=0=d(0)	*****

$\theta^* = -0.005768$

De nuevo los casos $n=3$, 6 y 8 han sido resueltos secuencialmente mediante el algoritmo de purificación. Aunque podemos ver que estos puntos extremos no son óptimos, queremos destacar la proximidad de los valores f^* y $v(P)$ especialmente en el caso $n=8$. En cualquier caso, no hay que olvidar que el propósito de un algoritmo de purificación, no es encontrar el óptimo de un problema sino proporcionar buenos puntos extremos.

<u>n</u>	<u>f^*</u>	<u>$v(P)$</u>	
3	.659433	.649042	f^* = valor de la función objetivo
6	.616560	.616085	en el punto extremo alcanzado
8	.615686	.615653	$v(P)$ = valor óptimo

CAPITULO IV.

REGLAS DE DESCENSO PARA EL PROBLEMA

DE PROGRAMACION SEMI-INFINTA LINEAL

CAPITULO IV. Reglas de descenso para el problema de PSIL.

Introducción.

Existen numerosos métodos que resuelven los problemas de programación semi-infinita, puede encontrarse una recopilación de algunos de ellos en Hettich (79), Glashoff y Gustafson (83) y otros. Para la programación semi-infinita lineal se habían planteado algoritmos para el problema dual, pero en los proceedings editados por Anderson y Philpott (85) aparecen los primeros trabajos que posibilitan desarrollar un simplex para el problema primal. Anderson y Nash (87) son los primeros que se plantean y desarrollan un algoritmo que trabaje directamente con los puntos extremos de la región factible del problema primal de PSIL y es, en este sentido, en el que puede considerarse como un método tipo Simplex.

El subconjunto de S en el que las restricciones son activas es crucial para los procedimientos primales, también lo es disponer de un algoritmo de purificación que permita pasar de una solución posible a un punto extremo. En el esquema de Anderson y Lewis (89), y para el caso no degenerado, la etapa de descenso toma una de las dos formas siguientes: ó se elimina un punto activo ó se perturba un punto activo hacia la derecha o hacia la izquierda, según convenga. En el caso degenerado se trabaja de otra forma, justificando este cambio de esquema. Ahora bien, en algunos casos, las reglas de descenso propuestas se alejan bastante del espíritu de la regla de pivote del método Simplex.

El significado geométrico de la regla de avance del Simplex es muy conocido: pasamos de un punto extremo a otro adyacente, a través de una dirección extrema, moviéndonos lo máximo posible en esa dirección. El nuevo punto extremo tiene un valor de la función objetivo menor o igual que el anterior. El método descubre también si el conjunto de soluciones posibles es vacío y si el problema no está acotado.

Nuestro objetivo en este capítulo es generalizar la regla de descenso del método Simplex de Programación Lineal para el caso de Programación Semi-infinita lineal. En el caso semi-infinito, según veremos a lo largo de este capítulo, no va a ser posible una extensión directa de este procedimiento, pero hemos intentado determinar unas reglas de avance que conserven, en la medida de lo posible, el estilo de la que se emplea en el método Simplex.

Como indicábamos antes distinguiremos dos casos: el de los puntos extremos no degenerados (Sección 4.1) y el de los puntos extremos degenerados (Sección 4.2). Para movernos a partir de un punto extremo no degenerado generamos una dirección

de manera que se mantengan activas todas las restricciones menos una que pasa a ser inactiva, y activamos alguna restricción que antes no lo era. Esta construcción no es posible para los puntos extremos degenerados, y ha sido necesario utilizar otra idea para obtener direcciones de descenso adecuadas. En ambos casos generamos direcciones factibles de descenso, si bien lo que no puede garantizarse es que se llegue a un punto extremo. Esta situación no supone ningún problema ya que disponemos de un eficiente algoritmo de purificación, desarrollado en el Capítulo III, que bajo condiciones poco restrictivas -condición (RA)- nos permite pasar de una solución posible a un punto extremo con un valor de la función objetivo menor o igual que el de partida y que puede detectar la no acotación del problema.

En la Sección 4.3 abordamos la cuestión de cómo encontrar una solución posible básica inicial, imprescindible para comenzar a aplicar el método. En realidad, es suficiente con disponer de alguna solución posible, pues siempre puede aplicarse el algoritmo de purificación para determinar un punto extremo mejor. Una situación de interés es la que se plantea cuando el problema artificial es no acotado. En este tipo de problemas se pone de manifiesto la importancia de disponer de un algoritmo de purificación aplicable tanto a problemas acotados como a problemas no acotados.

Sección 4.1. Reglas de descenso desde puntos extremos no degenerados.

Trabajamos con el problema continuo de PSIL :

$$\begin{aligned}
 \text{(PSI2) min } & c^T x \\
 \text{s.a. } & a^T(s)x - z(s) = b(s) \quad s \in S \\
 & x \in \mathbb{R}^n \quad z(s) \geq 0 \quad s \in S
 \end{aligned}$$

donde el conjunto de índices S es un intervalo cerrado de \mathbb{R} , y las componentes a_1, \dots, a_n de la función a y la función b son analíticas en S . Dado que la función de holgura z es analítica sobre S tendrá un número finito de ceros o será idénticamente nula.

Recordamos ahora aquella notación que es fundamental para el desarrollo del método. Para $(x^*; z^*)$ solución posible del (PSI2) tal que $\text{constr}(x^*) = \{s_1, s_2, \dots, s_m\}$ denotaremos por $d(s_i)$ o por d_i (indistintamente) el grado de s_i como cero de la función de holgura z^* .

Dado que una vez conocido el valor de $x^* \in \mathbb{R}^n$, automáticamente se tiene que su función de holgura asociada es $z^*(s) = a^T(s)x^* - b(s)$, por comodidad hablaremos también de x^* como solución posible del problema (PSI2). La matriz $n \times n$ \tilde{A} asociada a $(x^*; z^*)$ es la siguiente:

$$\tilde{A} = (a(s_1), \dots, a(s_m), a'(s_1), a''(s_1), \dots, a^{d(1)}(s_1), a'(s_2), a''(s_2), \dots, a^{d(2)}(s_2), \dots, a^{d(m)}(s_m)).$$

Siguiendo el enfoque del libro de Best y Ritter (85), denotamos la matriz $(\tilde{A}^{-1})^T = (g_1, g_2, \dots, g_m, b_1, b_2, \dots, b_{n-m})$.

Recordemos que un punto extremo no degenerado $(x; z)$ del (PSI2) es óptimo si,

y sólo si, $\lambda \geq 0_m$ y $\rho = 0_{n-m}$, siendo $\begin{bmatrix} \lambda \\ \rho \end{bmatrix} = \tilde{A}^{-1}c$. Supongamos que, tras aplicar la

caracterización anterior a una solución posible básica no degenerada $(x; z)$, detectamos su no optimalidad. Esta puede deberse a que ó bien existe algún $\lambda_k < 0$ ó bien existe algún $\rho_i \neq 0$. Vamos a comprobar que es posible llevar a cabo una extensión "directa" de la regla del Simplex de programación lineal cuando existe $\lambda_k < 0$. Sin embargo para el otro caso, $\rho_i \neq 0$, habrá que considerar una modificación de esta regla.

CASO 1. Sea $(x; z)$ una solución posible básica no degenerada tal que $\lambda_k < 0$ para algún $k \in \{1, \dots, m\}$. Vamos a generar una nueva solución posible $x^1 = x + \mu d$, siendo $d = g_k$ la k -ésima fila de \tilde{A}^{-1} .

Por construcción se tienen las igualdades siguientes:

- (i) $g_k^T a^{(h)}(s_j) = 0$ si $k \neq j$, para todo h , $0 \leq h \leq d(j)$
- (ii) $g_k^T a(s_k) = d^T a(s_k) = 1$
- (iii) $g_k^T a^{(h)}(s_k) = 0$, para todo h , $0 < h \leq d(k)$

Por conveniencia, y sin pérdida de generalidad, para cada $\mu \geq 0$ denotaremos $x(\mu) = x + \mu d$, y $z(s; \mu) = z(s) + \mu a(s)^T d$. Se tiene que $x(\mu) \in F$, si y sólo si, $z(s; \mu) \geq 0$, para todo $s \in S$. Podemos encontrar dos situaciones diferentes, siendo $d = g_k$:

- (1.a) Si $a(s)^T d \geq 0$ para todo $s \in S$, entonces d es una dirección de recesión a lo largo de la cual la función objetivo disminuye indefinidamente, es decir, el problema no está acotado.
- (1.b) Si, por el contrario, existe $s^0 \in S$, tal que $a(s^0)^T d < 0$, entonces $z(s; \mu) \geq 0$ para todo $s \in S$, si y sólo si, $\mu \leq \inf \{ -z(s)/(a(s)^T d) : a(s)^T d < 0, s \in S \}$. En el siguiente lema se demuestra que este ínfimo es positivo, y por tanto para un cierto $\mu > 0$, $x(\mu)$ será una nueva solución posible. Si este no es ya un punto extremo, podemos aplicar un algoritmo de purificación para obtenerlo, como se ha visto en el capítulo anterior.

Lema 4.1

Sea $\mu^* = \inf \{ -z(s)/(a(s)^T d) : a(s)^T d < 0, s \in S \}$, entonces $\mu^* > 0$.

Demostración.

Por reducción al absurdo. Supongamos que $\mu^* = 0$, entonces existiría $s^* \in S$ tal que $0 = \lim_{s \rightarrow s^*} -z(s)/(a(s)^T d)$, con $a(s^*)^T d \leq 0$. Veamos, en primer lugar, que $s^* \in \text{constr}(x)$.

Si $s^* \notin \text{constr}(x)$, entonces $z(s^*) > 0$ y por ser $z(s)$ continua existirían $\epsilon > 0$ y un intervalo cerrado de s^* , $U(s^*)$, tales que $z(s) > \epsilon$ para todo $s \in U(s^*)$. Consideramos las dos posibilidades para $a(s^*)^T d$:

- (a) Si $a(s^*)^T d = 0$ entonces $\lim_{s \rightarrow s^*} -z(s)/(a(s)^T d) = +\infty$
- (b) Si $a(s^*)^T d < 0$, existiría un entorno cerrado, $U'(s^*)$, tal que $a(s)^T d < 0$ para todo $s \in U'(s^*)$.

La función de s , $a(s)^T d$, es continua y, por tanto, acotada en $U'(s^*)$, luego existirá $M > 0$ tal que $|a(s)^T d| < M$. Así pues, tenemos que $z(s) > \epsilon$ y $-a(s)^T d < M$ para cada $s \in U(s^*) \cap U'(s^*)$, entonces $\lim_{s \rightarrow s^*} -z(s)/(a(s)^T d) \geq \epsilon/M > 0$, lo que contradice la hipótesis asumida.

En consecuencia, $s^* \in \text{constr}(x)$, es decir $s^* = s_j$ para algún índice j , $1 \leq j \leq m$, siendo $j \neq k$ puesto que $a(s_k)^T d = 1 > 0$.

Pero, por otra parte, se tiene que: $\lim_{s \rightarrow s_j} -z(s)/a(s)^T d = 0/0$, y aplicando la regla de L'Hopital, $\lim_{s \rightarrow s_j} -z(s)/a(s)^T d = \dots = -z^{d(i)+1}(s)/a^{d(i)+1}(s)^T d \neq 0$, llegamos a una contradicción y en consecuencia $\mu^* > 0$. ♦

Hemos obtenido un punto posible $x^1 = x + \mu^* d$ que mantiene todos los ceros excepto s_k . Ahora $\text{constr}(x^1) = \text{constr}(x) - \{s_k\} \cup J$, siendo $J = \{s \in S : \mu^* \text{ "se alcanza" en } s\}$. Escribimos "se alcanza", porque en rigor sólo se alcanzan los mínimos, pero en la práctica, cuando se aplica este procedimiento encontramos un valor concreto con una cierta precisión.

Nota. Anderson y Lewis (89) han utilizado también este procedimiento para moverse a partir de un punto extremo no degenerado no óptimo. Su demostración, basada en la aplicación de un lema de Hettich y Jongen (78), sólo validaba este criterio para puntos extremos cuyos índices activos están en el interior de S .

Veamos cómo funciona esta regla en el siguiente ejemplo, en el que se alcanza el óptimo en una etapa solamente.

Ejemplo 4.1

Consideremos el problema

$$(P_1) \text{ Min } x_1 + x_2 \\ \text{s.a. } sx_1 + s^3 x_2 \geq -1 \quad s \in [-3/2, 3/2]$$

Es fácil comprobar que $x^0 = \begin{pmatrix} 3/2 \\ -1/2 \end{pmatrix}$ es un punto posible con función de holgura $z^0(s) = 1 + 3/2s - 1/2s^3$, $\text{constr}(x^0) = \{-1\}$ y $d(-1) = 1$.

Se tiene que $\tilde{A} = \begin{pmatrix} -1 & 1 \\ -1 & 3 \end{pmatrix}$ y $\tilde{A}^{-1}c = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$, con $\lambda_1 < 0$. Generamos la dirección de descenso d , según la regla establecida en el CASO 1, y obtenemos: $d = (\tilde{A}^{-1})^T \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -3/2 \\ 1/2 \end{pmatrix}$. Calculamos la longitud de salto máxima y obtenemos $\mu^* = 2$ que se alcanza para $s = 1$. Entonces, $x^1 = x^0 + 2d = \begin{pmatrix} -3/2 \\ 1/2 \end{pmatrix}$ con $z^1(s) = 1 - 3/2s + 1/2s^3$, y $\text{constr}(x^1) = \{1\}$ con $d(1) = 1$. Además x^1 es óptima, pues $\tilde{A}^{-1}c = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

CASO 2. Sea $(x; z)$ una solución posible básica no degenerada no óptima con $\rho \neq 0_{n-m}$.

Denotaremos por $\rho_{k,r}$ la componente de $\begin{bmatrix} \lambda \\ \rho \end{bmatrix}$ asociada a la derivada r -ésima de $a(s)$ en s_k .

Vamos a comprobar que si $\rho_{k,r} \neq 0$ para un cierto k , $1 \leq k \leq m$, la regla clásica no puede extenderse. Supongamos que $\rho_{k,r} > 0$. Denotaremos por b^* la columna de $(\tilde{A}^{-1})^T$ que ocupa la misma posición que $a^{(r)}(s_k)$ en \tilde{A} . La dirección que obtendríamos aplicando la regla del CASO 1 sería b^* , pero esta es una dirección de ascenso y si tomamos $d = -b^*$, entonces puede ocurrir que $\mu^* = 0$. Como se comprueba en el ejemplo siguiente:

Ejemplo 4.2

Consideremos el problema

$$(P_2) \quad \text{Min } x_1 + 1/2x_2 \\ \text{s.a. } x_1 + sx_2 \geq 1 - s^2 \quad s \in [-1, 1]$$

Es inmediato que $x^0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ es una solución posible para (P_2) , con $z(s) = s^2$, $\text{constr}(x^0) = \{0\}$ y $d(0) = 1$. Se tiene que $\tilde{A} = I_2$ y $\tilde{A}^{-1}c = c$, y por tanto $\rho \neq 0$.

Si tomamos como dirección de mejora $d = -b^* = -(\tilde{A}^{-1})^T \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$, desde luego $c^T d = -1/2 < 0$. Pero al calcular μ^* , encontramos que $\mu^* = \inf\{s^2/s : -s < 0\} = 0$. No tendríamos, pues, una dirección factible.

Nota. En general, si para un elemento $s^* \in \text{constr}(x)$ se verifica que $a(s^*)^T d = 0$ y $a'(s^*)^T d < 0$, por ser la función $a(\cdot)^T d$ decreciente en s^* , se tendrá que :

$$\mu^* \leq \lim_{s \rightarrow s^*} -z(s)/a(s)^T d = 0/0 = \lim_{s \rightarrow s^*} -z'(s)/a'(s)^T d = 0/a'(s^*)^T d.$$

Anderson y Lewis (89) se enfrentan a este problema con un enfoque completamente diferente al que se propone en este capítulo. Primero parametrizan las soluciones posibles en un entorno de x^* como: $x(t) = A(t)^{-1}b(t)$, siendo $t = (t_1, t_2, \dots, t_m)$ un vector con tantas componentes como restricciones activas tiene x^* , $A(t)$ la matriz cuyas filas son

$$a(t_1), \dots, a(t_m), a'(t_1), a''(t_1), \dots, a^{d(1)}(t_1), a'(t_2), a''(t_2), \dots, a^{d(2)}(t_2), \dots, a^{d(m)}(t_m),$$

$$\text{y } b(t) = (b(t_1), \dots, b(t_m), b'(t_1), b''(t_1), \dots, b^{d(1)}(t_1), b'(t_2), b''(t_2), \dots, b^{d(2)}(t_2), \dots, b^{d(m)}(t_m))^T.$$

Si para un cierto k , $\rho_{k,1} \neq 0$, toman $h=e_k$ y para $\tau=(s_1, s_2, \dots, s_m)$ llevan a cabo una búsqueda lineal restringida hasta encontrar α^* que minimice $c^T x(\tau+\alpha h)$, sujeto a que $x(\tau+\alpha h)$ sea posible. Cuando $\rho_{k,j} \neq 0$, con $j>1$, dan otra clase de reglas.

Para las condiciones del CASO 2, nosotras proponemos un esquema para asegurar el movimiento, desde un punto extremo no degenerado, que está más cercano al espíritu del método Simplex.

CASO 2A. Sea $s_k \in \text{constr}(x)$ con $d(s_k) \geq 1$, tal que $\lambda_k \geq 0$ y $\rho_{k,1} > 0$. Resolviendo las siguientes ecuaciones, se genera una dirección factible d :

- (i) $d^T a^{(h)}(s) = 0$ para todo $s \in \text{constr}(x) - \{s_k\}$ $0 \leq h \leq d(s)$.
- (ii) $d^T a(s_k) = 1$.
- (iii) $d^T a'(s_k) = -1$.
- (iv) $d^T a^{(h)}(s_k) = 0$ para todo h tal que $1 < h \leq d(s_k)$.

Matricialmente, d es la dirección que satisface $d^T \tilde{A} = (0, \dots, 1, \dots, -1, \dots, 0)$ y, por tanto, $d^T = (0, \dots, 1, \dots, -1, \dots, 0) \tilde{A}^{-1}$. Entonces, se tiene que $c^T d = \lambda_k - \rho_{k,1}$ con $\lambda_k \geq 0$ y $\rho_{k,1} > 0$. Si $\lambda_k - \rho_{k,1} > 0$, necesariamente $\lambda_k > 0$ y podemos modificar ligeramente la dirección d , tomando como nueva dirección $d : d^T = (0, \dots, \rho_{k,1}/2\lambda_k, \dots, -1, \dots, 0) \tilde{A}^{-1}$. Así, pues, haremos $d^T = (0, \dots, \alpha, \dots, -1, \dots, 0) \tilde{A}^{-1}$, donde $\alpha = \rho_{k,1}/2\lambda_k$ si $\lambda_k - \rho_{k,1} \geq 0$, y $\alpha = 1$ si $\lambda_k - \rho_{k,1} < 0$.

Análogamente a la demostración desarrollada en el Lema 4.1 se comprueba que $\mu^* > 0$ y que $x^1 = x + \mu^* d$ es una solución posible. Más adelante probaremos que además es una solución mejor, pues el valor de la función objetivo decrece estrictamente a lo largo de d .

La regla de descenso que acabamos de introducir tiene un significado geométrico evidente. En el caso $\rho_{k,1} > 0$ hemos visto que el punto x^1 , al que vamos a acceder mediante esta regla de pivote, mantiene activas todas las restricciones excepto la correspondiente a s_k , cumpliéndose que $z_1(s_k) > 0$, $z_1'(s_k) < 0$, es decir en s_k la función de holgura z_1 es decreciente, por tanto el nuevo cero estará a la derecha de s_k . En la Sección 4.A del Apéndice se presentan algunas gráficas que permiten apreciar este hecho con detalle.

CASO 2B. Análogamente puede razonarse que cuando $\rho_{k,1} < 0$ el nuevo cero estará a la izquierda de s_k , y tendremos que escoger d de modo que se cumplan las siguientes condiciones:

- (i) $d^T a^{(h)}(s) = 0$ para todo $s \in \text{constr}(x) - \{s_k\}$ $0 \leq h \leq d(s)$.
- (ii) $d^T a(s_k) = \alpha$.
- (iii) $d^T a'(s_k) = 1$.
- (iv) $d^T a^{(h)}(s_k) = 0$ para todo h tal que $1 < h \leq d(s_k)$.

Matricialmente $d^T \tilde{A} = (0, \dots, \alpha, \dots, 1, \dots, 0)$ donde $\alpha = -\rho_{k,1}/2\lambda_k$ si $\lambda_k + \rho_{k,1} \geq 0$, y $\alpha = 1$ si $\lambda_k + \rho_{k,1} < 0$. Con lo que $c^T d = \alpha \lambda_k + \rho_{k,1}$.

Veamos algunos ejemplos en los que aplicamos estas reglas:

Ejemplo 4.3 (En Anderson-Nash (87))

- (P₃) Min $x_1 + x_2$
s.a $(1-s)x_1 + sx_2 \geq s(1-s)$ $s \in [0,1]$.

Si se parte de $x^0 = \begin{pmatrix} 1/9 \\ 4/9 \end{pmatrix}$ se tiene que $z^0(s) = s^2 - 2/3s + 1/9$ y $\text{constr}(x^0) = \{1/3\}$. La función de holgura del punto extremo no degenerado de partida tiene un único cero que se encuentra en el interior del conjunto S y, por tanto, tendrá grado impar.

Tenemos que $\tilde{A} = \begin{bmatrix} 2/3 & -1 \\ 1/3 & 1 \end{bmatrix}$ y $(\tilde{A})^{-1}c = \begin{pmatrix} 2 \\ 1/3 \end{pmatrix}$, luego x^0 no es una solución óptima pues $\rho \neq 0$. Además, notemos que $\lambda - \rho > 0$. Siguiendo el esquema del CASO 2A, tomamos la dirección $d = ((\tilde{A})^{-1})^T \begin{pmatrix} \rho/2\lambda \\ -1 \end{pmatrix} = \begin{pmatrix} 5/12 \\ -7/12 \end{pmatrix}$, calculamos $\mu^* = \inf \{ -[s^2 - 2/3s + 1/9] / (5/12 - s) : s \in (5/12, 1] \} = 1/3$, que se alcanza para $s=0.5$ y obtenemos el punto $x^1 = \begin{pmatrix} 1/4 \\ 1/4 \end{pmatrix}$ con $z^1(s) = s^2 - s + 1/4$ y $\text{constr}(x^1) = 0.5$.

La nueva matriz $\tilde{A} = \begin{bmatrix} 1/2 & -1 \\ 1/2 & 1 \end{bmatrix}$ y, por tanto, $(\tilde{A})^{-1}c = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$. Así pues, tenemos que x^1 es la solución óptima.

Ejemplo 4.4

$$(P_4) \text{ Min } x_1 + 1/2 x_2 \\ \text{s.a } x_1 + s x_2 \geq \sqrt{s} \quad s \in [0, 1].$$

Sea $x^0 = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}$ un punto extremo no degenerado con $z^0(s) = 1/2 + 1/2s - \sqrt{s}$ y $\text{constr}(x^0) = \{1\}$, $d(1) = 1$. Dado que, $(\tilde{A})^{-1}c = \begin{pmatrix} 1 \\ -1/2 \end{pmatrix}$, x^0 no es una solución óptima, pues $\rho < 0$. Aplicando el CASO 2B, como $\lambda + \rho > 0$, tomamos $d = ((\tilde{A})^{-1})^T \begin{pmatrix} \rho/2\lambda \\ 1 \end{pmatrix} = \begin{pmatrix} -3/4 \\ 1 \end{pmatrix}$. Calculamos $\mu^* = \inf \{ 1/2 + 1/2s - \sqrt{s} / (3/4 - s) : s \in (3/4, 1] \} = 1/6$, que se alcanza para $s = 0.562$, obteniéndose el punto posible $x^1 = \begin{pmatrix} 3/8 \\ 2/3 \end{pmatrix}$ con $z^1(s) = 3/8 + 2/3s - \sqrt{s}$ y $\text{constr}(x^1) = 0.562$.

Puede comprobarse fácilmente que x^1 no es una solución óptima. Hemos continuado aplicando las reglas de descenso y en tres iteraciones más hemos llegado al óptimo que es $x^* = \begin{pmatrix} 0.354070 \\ 0.706074 \end{pmatrix}$ con $v(P_4) = 0.707107$ y $\text{constr}(x^*) = \{0.5\}$.

Nota. Aunque $b(s) = \sqrt{s}$ no es analítica en todo S , pues $b'(s)$ no existe en $s=0$, como este no aparece como índice activo en ninguna iteración, no hay problema para aplicar la regla de pivote.

Ejemplo 4.5

$$(P_5) \text{ Min } -x_1 - x_2 \\ \text{s.a } s x_1 + s^3 x_2 \geq -1 \quad s \in [-1.5, 1.5].$$

Partiendo del punto $x^0 = \begin{pmatrix} -1 \\ 1/2 \end{pmatrix}$ con $z^0(s) = 1 - s + 1/2s^3$, cuyo $\text{constr}(x^0) = \emptyset$, y tras aplicar el algoritmo de purificación se alcanza el punto extremo no degenerado $x^1 = \begin{pmatrix} 1 \\ -4/27 \end{pmatrix}$. Calculamos $(\tilde{A})^{-1}c = \begin{pmatrix} 38.8125 \\ 12.65625 \end{pmatrix}$ y se comprueba que x^1 no es una solución óptima, pues $\rho > 0$. Como $\lambda - \rho > 0$, tomamos $d = ((\tilde{A})^{-1})^T \begin{pmatrix} \rho/2\lambda \\ -1 \end{pmatrix} = \begin{pmatrix} 15.352581 \\ -9.024457 \end{pmatrix}$. Hallamos $\mu^* = 0.019137$ que se alcanza para $s = -1.159$, obteniéndose el nuevo punto posible $x^2 = \begin{pmatrix} 1.293802 \\ -0.320849 \end{pmatrix}$ cuyo $\text{constr}(x^2) = \{-1.159\}$ y que es una solución básica no degenerada.

Aplicando de nuevo el test de optimalidad vemos que x^2 tampoco es una solución óptima, pero utilizando las reglas de descenso recién introducidas hemos llegado al óptimo en tres iteraciones más así $x^5 = \begin{pmatrix} 1.500302 \\ -0.500302 \end{pmatrix}$ cuyo $\text{constr}(x^5) = \{-1\}$, con $\theta^* = 0$ y $v(P_5) = -1$.

Si en un ejemplo concreto nos encontramos con una solución posible básica no óptima en la que $\lambda \geq 0_m$ y $\rho_{k,r} = 0$ $k=1, \dots, m$, pero para algún k , $\rho_{k,r} \neq 0$ con $r > 1$, necesitaremos definir una nueva regla de pivote. Nótese que el hecho de que s_k sea un cero de z situado en el interior o en la frontera de F no influye para nada en la dirección generada. Haremos una extensión directa del CASO 2, en la que de nuevo habrá que tener en cuenta el signo de $\rho_{k,r}$.

CASO 3. Si existe si existe $\rho_{k,r} \neq 0$ con $r > 1$, tomar d verificando:

- (i) $d^T a^{(h)}(s_j) = 0$ si $j \neq k$ para todo h $0 \leq h \leq d(j)$.
- (ii) $d^T a(s_k) = \alpha > 0$, siendo $\alpha = 1$ si $\lambda_k - |\rho_{k,r}| < 0$, y en otro caso $\alpha = \frac{|\rho_{k,r}|}{2\lambda_k}$.
- (iii) $d^T a^{(h)}(s_k) = 0$ para todo h tal que $0 < h \leq d(k)$ y $h \neq r$.
- (iv) $d^T a^{(r)}(s_k) = \beta$, donde $\beta = 1$ si $\rho_{k,r} < 0$ y $\beta = -1$ si $\rho_{k,r} > 0$.

Nota. La intuición geométrica de las reglas de pivote cuando $r=2$ también es sencilla. Supongamos, por ejemplo, que $\rho_{k,2} > 0$ entonces según (iv): $\beta < 0$. Así pues, en la nueva función de holgura s_k será un máximo local, ya que $z_1'(s_k) = 0$ y $z_1''(s_k) > 0$, aunque no sería un cero, pues la función de holgura ya no se anula en ese punto, i.e. $z_1(s_k) > 0$.

En las gráficas que se presentan en el Apéndice se aprecia el efecto de esta transformación de la función de holgura.

Resumiendo, tenemos las siguientes reglas para conseguir una dirección factible que veremos que es de descenso, por construcción. Dado $x \in F$ cuyo $\text{constr}(x) = \{s_1, s_2, \dots, s_m\}$, hallamos $(\lambda_1, \lambda_2, \dots, \lambda_m, \rho_{1,1}, \dots, \rho_{1,d(1)}, \dots, \rho_{m,1}, \dots, \rho_{m,d(m)})^T = (\tilde{A})^{-1} c$.

1) Si existe al menos un índice k tal que $\lambda_k < 0$, elegir d según CASO 1. Si hay varios valores $\lambda < 0$ elegiremos el menor de ellos para pivotar.

2) Si por el contrario $\lambda_k \geq 0$ para todo k , pasamos a considerar los valores de los ρ_k . Hay dos posibilidades para el esquema general, una es comenzar la exploración de

$\rho_{k,r}$ por los valores de r más altos, y la otra es comenzar por los más bajos, aplicando las reglas del CASO 3 y CASO 2, respectivamente.

Teorema 4.2

Sea $(x; z)$ una solución posible básica no degenerada no óptima de (PSI2). Sea d obtenida según CASO 1, 2 ó 3, entonces d es una dirección factible de mejora.

Demostración.

Dado que por construcción d es una dirección factible, se trata de comprobar que $x^1 = x + \mu^* d$ es un punto posible con un menor valor de la función objetivo. Veámoslo:

CASO 1: $c^T x^1 = c^T x + \mu^* c^T d = c^T x + \mu^* c^T (\tilde{A})^{-1} e_k = c^T x + \mu^* \lambda_k < c^T x$.

CASO 2A: $c^T d = (0, \dots, \alpha, \dots, -1, \dots, 0)$ $\tilde{A}^{-1} c = \alpha \lambda_k - \rho_{k,1}$, entonces $c^T d = \lambda_k - \rho_{k,1}$, si $\lambda_k - \rho_{k,1} < 0$ y $c^T d = (-1/2)\rho_{k,1}$, si $\lambda_k - \rho_{k,1} \geq 0$, siendo $\rho_{k,1} > 0$, luego d es una dirección de descenso.

CASO 2B: Análogamente al caso 2A se tiene que $c^T d = \alpha \lambda_k + \rho_{k,1} < 0$.

CASO 3: Se comprueba que $c^T d = \alpha \lambda_k + \beta \rho_{k,r} = \alpha \lambda_k - |\rho_{k,r}| < 0$. ♦

Hemos aplicado el algoritmo de purificación del Capítulo III y las anteriores reglas de avance para resolver con detalle el problema de aproximación L_1 siguiente:

Ejemplo 4.6

$$(P_6) \text{ Min } \sum_{i=1}^5 1/i x_i$$

$$\text{s.a. } \sum_{i=1}^5 x_i s_i^{-1} \geq 1/(1+s^2) \quad s \in [0,1]$$

Dado que una de las funciones $a_i(s)$ es una constante, concretamente $a_1(s)=1$, no tenemos ninguna dificultad para encontrar una solución posible inicial. Hallamos $\max\{b(s) : s \in S\}=1$, que se alcanza para $s=0$, y disponemos de una solución posible $x^0 = (1, 0, 0, 0, 0)^T$ cuyo $\text{constr}(x^0)=\{0\}$, $d(0)=0$ y $c^T x^0=1$.

Como x^0 no es un punto extremo, aplicamos el algoritmo de purificación y obtenemos el siguiente punto extremo $x^1=(1, 0, -1, 1/2, 0)^T$ cuyo $\text{constr}(x^1)=\{0,1\}$, $d(0)=2$, $d(1)=1$ y $c^T x^1=0.791667$. Calculamos $\tilde{A}^{-1}c=(3/5, 2/5, 3/20, -1/20, 1/60)^T=(\lambda_1, \lambda_2, \rho_{1,1}, \rho_{2,1}, \rho_{1,2})^T$, donde $s_1=0$, $s_2=1$, luego x^1 no es solución óptima, y tenemos ahora varias posibilidades para obtener una dirección de avance.

Hemos optado por trabajar con $\rho_{1,2} \neq 0$, entonces la dirección que generamos es (CASO 3): $d^T = (\alpha, 0, 0, 0, -1)$ \tilde{A}^{-1} siendo $\alpha = \rho_{1,2} / 2\lambda_1 = 1/72$. Determinamos la

longitud de salto máxima $\mu^* = 0.569445$ que se alcanza para $s = 0.425$ y obtenemos una

nueva solución posible $x^2 = \begin{pmatrix} 1.007909 \\ 0 \\ -1.284722 \\ 1.037809 \\ -0.260995 \end{pmatrix}$ con $\text{constr}(x^2) = \{0.425, 1\}$, $d_1 = 1$, $d_2 = 1$ y

$c^T x^2 = 0.786921$.

El punto x^2 no es punto extremo, por tanto hemos de aplicar el algoritmo de

purificación, y conseguimos la solución posible básica $x^3 = \begin{pmatrix} 1 \\ 0.053036 \\ -1.410856 \\ 1.162604 \\ -0.304783 \end{pmatrix}$ cuyo

$\text{constr}(x^3) = \{0, 0.425, 1\}$, $d_1 = 0$, $d_2 = 1$, $d_3 = 1$ y $c^T x^3 = 0.785927$. Calculamos ahora

$c^T A^{-1} = (0.153751, 0.576397, 0.297862, -0.014793, -0.028033)^T = (\lambda_1, \lambda_2, \lambda_3, \rho_{2,1},$

$\rho_{3,1})^T$ donde $s_1 = 0$, $s_2 = 0.425$ y $s_3 = 1$. Puesto que en valor absoluto $\rho_{3,1}$ es mayor que

$\rho_{2,1}$, decidimos "mover" s_3 , determinando para ello (CASO 2B): $d = (\tilde{A}^{-1})^T \begin{pmatrix} 0 \\ 0 \\ \alpha \\ 0 \\ 1 \end{pmatrix}$ donde

$\alpha = -\rho_{3,1} / 2\lambda_3 = 0.047057$. Se tiene que $\mu^* = 0.001047$ que se alcanza para $s = 0.917$.

Obtenemos $x^4 = \begin{pmatrix} 1 \\ 0.052611 \\ -1.408407 \\ 1.158129 \\ -0.302283 \end{pmatrix}$ con $\text{constr}(x^4) = \{0, 0.425, 0.917\}$, $d_1 = 0$, $d_2 = 1$,

$d_3 = 1$ y $c^T x^4 = 0.785911$. Tampoco es una solución óptima, aunque x^4 sí es un punto

extremo no degenerado tal que $\tilde{A}^{-1}c = (0.118602, 0.51923, 0.36217, -0.032335, -$

$0.020457)^T = (\lambda_1, \lambda_2, \lambda_3, \rho_{2,1}, \rho_{3,1})^T$. En este caso optamos por trasladar hacia la

izquierda a s_2 (CASO 2B), la nueva dirección de descenso es $d = (\tilde{A}^{-1})^T \begin{pmatrix} 0 \\ \alpha \\ 0 \\ 1 \\ 0 \end{pmatrix}$ donde $\alpha = -$

$\rho_{2,1} / 2\lambda_2 = 0.031137$. Ahora $\mu^* = 0.002350$ se alcanza para $s = 0.369$, luego $x^5 =$

$\begin{pmatrix} 1 \\ 0.04461 \\ -1.370725 \\ 1.104486 \\ -0.278222 \end{pmatrix}$ con $\text{constr}(x^5) = \{0, 0.369, 0.917\}$, $d_1 = 0$, $d_2 = 1$, $d_3 = 1$ y

$c^T x^5 = 0.785874$. Esta nueva solución es también un punto extremo.

A partir de ahora, sólo explicitaremos las soluciones que se van obteniendo, su conjunto de índices activos y el valor de la función objetivo en cada punto.

$$x^6 = \begin{pmatrix} 1 \\ 0.043345 \\ -1.362442 \\ 1.087461 \\ -0.267739 \end{pmatrix} \text{ constr}(x^6) = \{0, 0.369, 0.86\}, d_1=0, d_2=1, d_3=1 \text{ y } c^T x^6 = 0.785843$$

$$x^7 = \begin{pmatrix} 1 \\ 0.041123 \\ -1.351141 \\ 1.070193 \\ -0.259446 \end{pmatrix} \text{ constr}(x^7) = \{0, 0.356, 0.86\}, d_1=0, d_2=1, d_3=1 \text{ y } c^T x^7 = 0.785840$$

$$x^8 = \begin{pmatrix} 1 \\ 0.040658 \\ -1.347983 \\ 1.06346 \\ -0.255143 \end{pmatrix} \text{ constr}(x^8) = \{0, 0.356, 0.844\}, d_1=0, d_2=1, d_3=1 \text{ y}$$

$$c^T x^8 = 0.785838$$

Tenemos que x^8 es óptima, puesto que $\tilde{A}^{-1} c = (0.111107, 0.51249, 0.376403, -4.89 \cdot 10^{-4}, 3.58 \cdot 10^{-4})^T$ y consideramos que $\rho_{ij} = 0$ cuando en valor absoluto es menor que 10^{-3} . Aplicando el test de optimalidad que desarrollamos en el Capítulo III obtenemos $\theta^* = -4.07 \cdot 10^{-4}$.

Todas las soluciones que se han ido obteniendo a partir de x^3 y durante la aplicación del algoritmo son puntos extremos no degenerados, por lo que no ha sido necesaria ninguna iteración más del algoritmo de purificación. Además observando el conjunto de restricciones activas de cada uno de los puntos notamos que el algoritmo ha ido "ajustando" alternativamente cada uno de los ceros hasta llegar al óptimo.

Nota. El problema dual standard de (P_6) es el siguiente:

$$(D_6) \text{ Max } \sum \lambda(s) \frac{1}{1+s^2}$$

$$\text{s.a. } \sum \lambda(s) \begin{pmatrix} 1 \\ s \\ s^2 \\ s^3 \\ s^4 \end{pmatrix} = \begin{pmatrix} 1 \\ 1/2 \\ 1/3 \\ 1/4 \\ 1/5 \end{pmatrix}$$

$$\lambda(s) \in \mathbb{R}^+(s)$$

A la vista de la solución del problema (P_6) , deducimos que la solución óptima de (D_6) es $\hat{\lambda}^*(s)$, dada por $\hat{\lambda}^*(0) = 0.111107$, $\hat{\lambda}^*(0.356) = 0.51249$, $\hat{\lambda}^*(0.844) = 0.376403$ y $\hat{\lambda}^*(s) = 0$ para todos los demás valores de $s \in [0, 1]$.

Sobre este problema hemos utilizado el Simplex de Goberna y Jorner (88) para el dual con el fin de encontrar una cota inferior al valor óptimo de la solución del primal. Esta es, en principio, una buena estrategia porque nos permite parar cuando se alcanza la precisión deseada pero, de nuevo, ha aparecido el problema de los errores de cálculo y a partir de cierta iteración el valor de la supuesta cota excedía a $v(P_6)$.

Partiendo de la solución dual posible $T=\{0.1, 0.3, 0.4, 0.7, 1\}$ con $\lambda(0.1)=0.216029$, $\lambda(0.3)=0.178666$, $\lambda(0.4)=0.092505$, $\lambda(0.7)=0.416682$, $\lambda(1)=0.096118$ y $\lambda(s)=0$, para $s \in S/T$ que proporcionaba la primera de las cotas inferiores $l_1=0.785261$, y tras varias iteraciones del algoritmo hemos obtenido las cotas $l_2=0.785376$, $l_3=0.785694$, $l_4=0.785759$, $l_5=0.785803$ y $l_6=0.785849$. Como puede verse l_6 excede el valor de $v(P_6)=0.785838$. El valor de l_6 corresponde al de la función objetivo dual evaluada en $\lambda^6(s)$ dada por $\lambda^6(0)=0.109173$, $\lambda^6(0.3)=0.052301$, $\lambda^6(0.356)=0.453026$, $\lambda^6(0.807)=0.226286$, $\lambda^6(0.882)=0.159275$ y $\lambda^6(s)=0$ en otro caso. Por ser $\lambda^6(s)$ teóricamente una solución dual posible debería cumplirse que $\sum \lambda(s)a(s) = c$. Sin embargo, por los errores acumulados se tiene que $\sum \lambda(s)a(s) = c'$ con $\|c-c'\|=0.000136$. Trabajando con suficiente precisión quizá se podrían soslayar estas dificultades y, entonces, dispondríamos de un nuevo criterio de parada para nuestro algoritmo.

Sección 4.2 Reglas de descenso desde puntos extremos degenerados.

En Programación Lineal la forma de tratar los puntos extremos degenerados es llevar a cabo una serie de pivotes degenerados. Sin embargo, en Programación semi-infinita una generalización de este procedimiento puede no tener éxito porque los puntos extremos degenerados se agrupan en dominios de la frontera de la región posible (Nash (85), Jongen y Zwier (85)).

En términos generales podemos decir que la degeneración corresponde al caso en que hay "demasiadas restricciones activas", lo que además puede ser una dificultad bastante frecuente. Aunque hay algunas clases de problemas que no las padecen, por ejemplo los del tipo siguiente:

$$(P) \text{ Min } c^T x \\ \text{s.a. } \sum_{i=1}^n x_i s_i^{-1} \geq b(s) \quad s \in [0,1]$$

donde $b(\cdot)$ tiene la propiedad de que $b^{(n)}(\cdot)$ no tiene raíces en $[0,1]$. Pues, aplicando el Teorema de Rolle sucesivamente se comprueba que ninguna $z(\cdot)$ posible puede tener más de n raíces contadas con multiplicidad, así que cualquier punto extremo será no degenerado.

En la práctica, los puntos extremos degenerados nos plantean dos problemas fundamentalmente. El primero es que el coste reducido ya no nos proporciona un criterio de optimalidad sencillo, y el segundo es que ya no podemos mover independientemente los puntos activos como hacíamos en el caso no degenerado.

La primera cuestión ha quedado resuelta en el Capítulo II, pues cuando la solución posible $(x;z)$ verifica una cualificación de restricciones sabemos que la condición $c \in K\{a(s_i) : s_i \in \text{constr}(x)\}$ es necesaria y suficiente para la optimalidad. Por otra parte, si queremos mantener el esquema clásico de la programación lineal y mantener activas todas las restricciones excepto una, tendremos que introducir el criterio con el que escoger la restricción que va a dejar de ser activa. En el caso no degenerado este problema viene resuelto por el criterio de optimalidad elegido, como hemos visto. Pero la principal dificultad para los puntos extremos degenerados es que, incluso disponiendo de un criterio de selección de la restricción que pasará a ser inactiva, es posible que la dirección obtenida, manteniendo activas las demás, no sea de descenso. Es necesario, por tanto, cambiar de esquema y nos servirá de gran ayuda el test de optimalidad que se estableció en el Capítulo II. Vemos estos inconvenientes sobre un ejemplo:

Ejemplo 4.7

Consideremos de nuevo el problema del Ejemplo 4.6:

$$(P_6) \text{ Min } \sum_{i=1}^5 1/i x_i$$
$$\text{s.a. } \sum_{i=1}^5 x_i s_i^{-1} \geq 1/(1+s^2) \quad s \in [0,1]$$

Tras una serie de pasos de descenso hemos llegado al punto $x^0 = \begin{pmatrix} 1 \\ 0.064215 \\ -1.457953 \\ 1.224599 \\ -0.330861 \end{pmatrix}$ cuyo

constr(x^0) = {0, 0.551, 0.8, 1}, $d_1=0=d_4$, $d_2=1=d_3$, que es un punto extremo degenerado. Aplicando el test de optimalidad obtenemos que $\theta^* = -0.05796$ y sabemos que no se trata de una solución óptima, pues se cumple trivialmente la cualificación de restricciones de Mangasarian-Fromovitz..

Para determinar qué restricción dejará de ser activa, hemos optado por resolver por el método de mínimos cuadrados el sistema lineal superdeterminado: $B\lambda = c$, donde $B = (a(0), a(0.551), a(0.8), a(1))$, $B \in M_{5 \times 4}$. Es decir, hemos encontrado $\lambda \in \mathbb{R}^4$ tal que se minimice $\|B\lambda - c\|^2$ y hemos obtenido $\lambda = \begin{pmatrix} 0.196502 \\ 0.706271 \\ -0.006567 \\ 0.162895 \end{pmatrix}$. Nótese que x^0 sería óptimo

si $\lambda \geq 0$ y $\|B\lambda - c\|^2 = 0$. Ninguna de estas dos cosas sucede, pero parecería que disponemos de un criterio para decidir qué restricción deja de ser activa: aquella que corresponde a la tercera componente de λ que ha resultado ser negativa, así pues haremos inactivo $s_3 = 0.8$.

Nos planteamos, por tanto, mantener activas las restricciones asociadas a $s_1 = 0$, $s_2 = 0.551$ y $s_4 = 1$, y entonces la dirección d que buscamos debería cumplir las ecuaciones:

- (i) $a(s_i)^T d = 0$, para $i \in \{1, 2, 4\}$.
- (ii) $a(s_3)^T d = \alpha > 0$.
- (iii) $a'(s_2)^T d = 0$.

Sea A la matriz cuya traspuesta es $A^T = (a(s_1), a(s_2), a(s_3), a(s_4), a'(s_2))$, que es una matriz invertible (en otros ejemplos, pueden aparecer otras situaciones, como no disponer de una matriz invertible).

Las anteriores ecuaciones pueden escribirse: $d=A^{-1}(0, 0, \alpha, 0, 0)^T$. Tendríamos

que $d = \begin{pmatrix} 0 \\ 30.6046 \\ -141.693 \\ 211.894 \\ -100.806 \end{pmatrix} \alpha$, con $\alpha > 0$ pero $c^T d = 0.8836\alpha > 0$, y por tanto d no es una

dirección de descenso.

El test de optimalidad aplicado en el punto x^0 nos señala como dirección de búsqueda $d^*=(0, -0.297097, 1, -1, 0.297097)$, pero al calcular la longitud de salto resulta que $\mu^*=0$ y no nos podemos mover en esa dirección pues abandonaríamos el conjunto posible.

Veamos, a continuación el esquema que proponen Anderson y Lewis (89) para el caso degenerado. En primer lugar, estos autores amplian el espacio en que trabajan, definiendo el conjunto F de R^{n+k} :

$F=\{(x; t) : a(t_i)^T x \geq b(t_i), a'(t_i)^T x = b'(t_i) \ i=1, 2, \dots, k\}$, donde los puntos activos vienen especificados como $(t_1, \dots, t_k) \in R^k$ y demuestran el siguiente resultado :

Teorema 4.3

Supongamos que (ξ, ζ) es posible para el (PSI2) con puntos activos $\{s_1, \dots, s_k\}$ en $(0, 1)$ y $\zeta^{(2)}(s_i) > 0$ para cada $i=1, \dots, k$. Definimos $\tau=(s_1, \dots, s_k) \in R^k$. Entonces, existe un $\delta > 0$ tal que para todo $(x, z) \in N_\delta(\xi, \zeta)$, $(x, a(\cdot)^T x - b(\cdot))$ es posible para el (PSI2) si y sólo si $(x; \tau) \in F$.

Tratando a $\xi \in R^n$ como un vector fijo y considerando el problema finito:

(RP) Min $\{ c^T x \ / (x, t) \in F \}$

se tiene que (ξ, τ) es un mínimo local para (RP) si y sólo si (ξ, ζ) es un mínimo local y, por tanto, óptimo para el (PSI2). El espacio tangente a F en (ξ, τ) que denotan por M , viene dado por $M=\{(x, t) : a(s_i)^T x = 0 \text{ y } a'(s_i)^T x + \xi^{(2)}(s_i) t_i = 0 \ i=1, 2, \dots, k\}$.

El algoritmo propuesto por Anderson y Lewis (89) hace un paso de descenso moviéndose un poco en la dirección $-P_M(c, 0_k)$, donde P_M representa la proyección ortogonal sobre M , seguido de un paso de reajuste que permite volver a la región factible F .

En el caso de que $P_M(c, 0_k) = 0$, entonces se tiene que existen $\mu, v \in R^k$ tales que $c = \sum_{i=1}^k \mu_i a(s_i) + \sum_{j=1}^k v_j a'(s_j)$ y $v_j \xi^{(2)}(s_j) = 0 \ j=1, \dots, k$. Como, por hipótesis $\zeta^{(2)}(s_i) > 0 \ i=1, \dots, k$, entonces $c = \sum_{i=1}^k \mu_i a(s_i)$. Se pueden distinguir dos casos:

(a) si $\mu_k \geq 0$ entonces (ξ, ζ) es óptima, de lo contrario

(b) el algoritmo standard del gradiente proyectado borraría la restricción correspondiente al μ_j más negativo y se movería en la dirección del vector de coste proyectado sobre el subespacio determinado por las restantes restricciones activas cambiado de signo. En este caso, eliminar la restricción correspondiente a μ_j significa aumentar la holgura en t_j , y se pueden simplificar los cálculos trabajando en el conjunto más pequeño

$$F' = \{(x, t) : a(t_i)^T x \geq b(t_i), a'(t_i)^T x = b'(t_i) \quad i=1, 2, \dots, k \quad i \neq j\}.$$

Desde un punto de vista operativo el esquema propuesto no parece que vaya a ser muy efectivo, además sólo está justificado cuando los índices activos están en el interior del intervalo.

Vamos a discutir un método para generar direcciones factibles de descenso que, si bien es aplicable a cualquier solución posible x con un número finito de restricciones activas, sólo será aplicado sobre puntos extremos degenerados debido a que se pierden todas las restricciones activas del punto inicial y a que las reglas de descenso establecidas en la sección anterior trabajan bien para los puntos extremos no degenerados.

Supongamos que un punto $x \in F$ con $\text{constr}(x) = \{s_1, s_2, \dots, s_m\}$ no es óptimo. Sabemos, por el Capítulo II, que el problema lineal :

$$\begin{aligned} \text{(P1)} \quad & \text{Min } c^T d \\ & \text{s.a. } a(s_i)^T d \geq 0 \quad i=1, \dots, m \\ & \quad -1 \leq d_j \leq 1 \quad j=1, \dots, n \end{aligned}$$

cumple que $v(\text{P1}) < 0$. Además, si llamamos d^* a la solución óptima de (P1), tenemos que $c^T d^* < 0$ y $a(s_i)^T d^* \geq 0 \quad i=1, \dots, m$. Aunque el problema que nos puede aparecer con d^* es que no podamos avanzar a lo largo de esa dirección porque la longitud de salto correspondiente sea cero.

Ahora bien, si el problema verifica una cualificación de restricciones, como la de Mangasarian- Fromovitz, podemos garantizar que existe un vector $u \in R^n$ de manera que $a(s_i)^T u > 0, \quad i=1, \dots, m$. Para determinar, en la práctica, una dirección con estas características resolvemos el problema lineal siguiente:

$$\begin{aligned}
(P2) \quad & \text{Max } \epsilon \\
& \text{s.a. } c^T d \leq 0 \\
& a(s_i)^T d \geq \epsilon \quad i=1, \dots, m \\
& -1 \leq d_j \leq 1 \quad j=1, \dots, n
\end{aligned}$$

Sea (ϵ^*, g) una solución óptima para (P2), donde ϵ^* es el valor de la función objetivo.

Se comprueba en el lema siguiente que ϵ^* es un número estrictamente positivo.

Lema 4.4

Sea $(x; z)$ una solución posible no óptima de (PSI2) cuyo $\text{constr}(x) = \{s_1, \dots, s_m\}$. Supongamos que $(x; z)$ satisface la cualificación de restricciones de M-F, entonces $v(P2) > 0$.

Demostración.

Nótese en primer lugar que (P2) puede escribirse:

$$\begin{aligned}
(P2) \quad & \text{Max } \epsilon \\
& \text{s.a. } Ad - \epsilon e \geq 0_m \\
& c^T d \leq 0 \\
& d_i \leq 1 \quad 1 \leq i \leq n \\
& d_i \geq -1 \quad 1 \leq i \leq n
\end{aligned}$$

siendo $A^T = (a(s_1), a(s_2), \dots, a(s_m))$ y $e^T = (1, 1, \dots, 1)$. El dual de (P2) es el siguiente problema lineal (D2):

$$\begin{aligned}
(D2) \quad & \text{Min } \sum_{i=1}^n (\omega^2)_i - \sum_{i=1}^n (\omega^3)_i \\
& \text{s.a. } \alpha c + \omega^1 A + \omega^2 + \omega^3 = 0_n \\
& \sum_{i=1}^m (\omega^1)_i = -1 \\
& \alpha \geq 0, \omega^1 \leq 0_m, \omega^2 \geq 0_n, \omega^3 \leq 0_n
\end{aligned}$$

Ambos problemas tienen soluciones posibles, el vector nulo 0_{n+1} para (P2), y la que construiremos a continuación para (D2).

Tomar $\alpha = 0$, $(\omega^1)_1 = -1$ y $(\omega^1)_i = 0$, $i=2, \dots, m$, entonces llamando $\omega^1 A = (\omega^1)_1 a(s_1) = (p_1, \dots, p_m)$, si tomamos $\omega^2 = p^+$, $\omega^3 = p^-$ i.e. $(\omega^2)_i = \max\{p_i, 0\}$, $(\omega^3)_i = \min\{p_i, 0\}$, se satisfacen las condiciones y $(\alpha, \omega^1, \omega^2, \omega^3)$ es solución posible

de (D2). Por lo tanto, ambos tienen solución óptima y los valores de la función objetivo en el óptimo coinciden.

Veamos ahora que, para cualquier solución posible de (D2), se cumple que $\sum_{i=1}^m (\omega^2)_i - \sum_{i=1}^m (\omega^3)_i > 0$. Por reducción al absurdo, si existiera una solución tal que $\sum_{i=1}^m (\omega^2)_i - \sum_{i=1}^m (\omega^3)_i \leq 0$, necesariamente $\omega^2 = \omega^3 = 0_n$ y $\alpha c + \omega^1 A = 0_n$, $\sum_{i=1}^m (\omega^1)_i = -1$, $\omega^1 \leq 0_m$, $\alpha \geq 0$.

(i) Si $\alpha = 0$, tendríamos que $\omega^1 A = 0_n$, i.e. $0_n \in \text{co}\{a(s_1), \dots, a(s_m)\}$, lo que es imposible por verificarse la cualificación de restricciones de M-F, luego

(ii) $\alpha > 0$, entonces $(-1/\alpha) \omega^1 A = c$, y llamando $\lambda = (-1/\alpha) \omega^1 \geq 0_m$ se tendría que $c = \lambda A$ con $\lambda \geq 0_m$, pero esto no puede suceder porque $(x; z)$ no es una solución óptima del PSIL.

Entonces, en particular, $v(D2) > 0$ y por consiguiente $v(P2) > 0$. ♦

Definimos un nuevo problema lineal, basado en la solución del problema (P2), haciendo $\eta = 3/4 \epsilon^* > 0$:

$$(P3) \quad \begin{aligned} \text{Min} \quad & c^T g \\ \text{s.a.} \quad & a(s_i)^T g \geq \eta \quad i=1, \dots, m \\ & -1 \leq g_j \leq 1 \quad j=1, \dots, n \end{aligned}$$

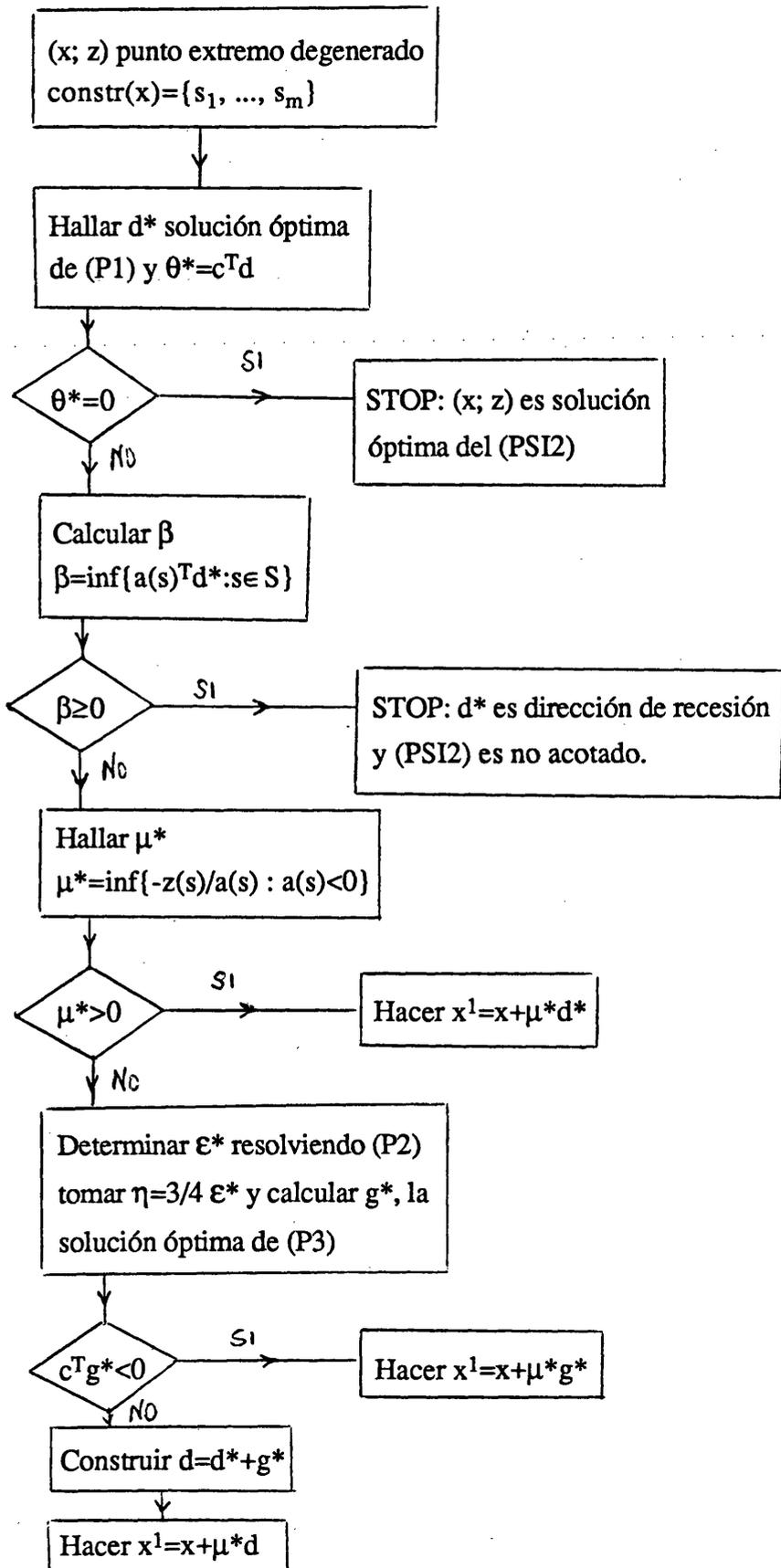
Teorema 4.5

Sea $(x; z)$ una solución posible no óptima de (PSI2) cuyo $\text{constr}(x) = \{a(s_1), \dots, a(s_m)\}$, y que satisface la cualificación de M-F. Sean d^* y g^* soluciones óptimas de los problemas (P1) y (P3), respectivamente. Entonces $d = d^* + g^*$ es una dirección factible de descenso en x .

Demostración.

Por construcción $v(P3) \leq 0$, pues el vector g en $(\epsilon^*; g)$, solución óptima de (P2) es una solución posible de (P3). Sea g^* una solución óptima para (P3), entonces g^* satisface $c^T g^* \leq 0$ y además $a(s_i)^T g^* > 0 \quad i=1, \dots, m$. Sea $d = d^* + g^*$, la solución óptima de (P1) satisface $c^T d^* < 0$ y $a(s_i)^T d^* \geq 0 \quad i=1, \dots, m$, luego el vector d verifica simultáneamente las dos condiciones que buscábamos: $c^T d < 0$ y $a(s_i)^T d > 0 \quad i=1, \dots, m$. ♦

De todo lo anterior y la discusión de la Sección 2.2, respecto a la solución óptima del test de optimalidad d^* , tenemos que la operación de pivote desde un punto extremo degenerado puede llevarse a cabo de acuerdo con el esquema siguiente:



Ejemplo 4.8

Consideremos el siguiente problema:

$$(P_8) \text{ Min } c_1x_1 + c_2x_2 \\ \text{s.a. } x_1 + sx_2 \geq -4s^4 + 12s^3 - 13s^2 + 6s \quad s \in [0,1]$$

El punto $x^0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ es una solución posible básica degenerada con $\text{constr}(x^0) = \{1/2, 1\}$, $d(1/2) = 1 = d(1)$ y $z^0(s) = 1 + 4s^4 - 12s^3 + 13s^2 - 6s \geq 0$ para $s \in [0,1]$.

En primer lugar nótese que, desde luego, el criterio de optimalidad de Anderson y Lewis (89) no puede utilizarse pues la matriz \tilde{A} asociada a x^0 no es invertible. Resolvemos el sistema:

$$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 1/2 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ \lambda_1, \lambda_2 \geq 0$$

y se tiene que x^0 es óptima para $c_1 \geq c_2 \geq c_1/2$, con $c_1 > 0$.

Vamos a considerar diferentes valores para los coeficientes c_i $i=1, 2$, lo que nos permitirá estudiar situaciones distintas:

4.8.a) Supongamos $c_1=1, c_2=2$. Si aplicamos el test de optimalidad en el punto x^0 tenemos que resolver:

$$(P1) \text{ Min } d_1 + 2d_2 \\ \text{s.a. } d_1 + 1/2d_2 \geq 0 \\ d_1 + d_2 \geq 0 \\ -1 \leq d_i \leq 1 \quad i=1, 2$$

Se obtiene $\theta^* = -1$, que nos confirma lo que ya sabíamos, que la solución x^0 no es óptima.

Veamos que ocurre con $d^* = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, aunque esta dirección puede no ser factible este no es el caso. Es fácil comprobar que $\beta \geq 0$ y, por lo tanto, d^* es una dirección de recesión y el problema (P₈) es no acotado.

En realidad, esto mismo sucede para cualquier par (c_1, c_2) con $0 \leq c_1 < c_2$.

4.8.b) Supongamos ahora, $c_1=1, c_2=1/4$. Para $x^0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $c^T x^0 = 1$.

(1) Resolviendo el problema lineal (P1), del correspondiente test de optimalidad, obtenemos $\theta^* = -0.25$ con $d^* = \begin{pmatrix} -1/2 \\ 1 \end{pmatrix}$. Además, d^* no es una dirección factible, pues $\mu^* = 0$. Hallamos las soluciones óptimas para los problemas (P2) y (P3) y se obtienen, respectivamente $\varepsilon^* = 0.25$ y $g^* = \begin{pmatrix} -0.3125 \\ 1 \end{pmatrix}$. Puesto que $c^T g^* < 0$, tenemos que g^* es dirección de descenso, pivotando llegamos a $x^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 2.046585 \begin{pmatrix} -0.3125 \\ 1 \end{pmatrix} = \begin{pmatrix} 0.360442 \\ 2.046585 \end{pmatrix}$

cuyo $\text{constr}(x^1) = \{0.205\}$ y $c^T x^1 = 0.872088$. Esta solución no es óptima, pero después de tres iteraciones más del método, en las que se aplican las reglas para puntos extremos no degenerados, lo alcanzamos.

(2) Aunque según el esquema algorítmico propuesto no sería necesario hacerlo, hemos construido la dirección $d = g^* + d^*$ para comprobar que también es de descenso, sabemos por el Teorema 4.5 que se trata de una dirección factible y pasamos a una nueva solución con menor valor de la función objetivo:

$$x^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 0.336434 \begin{pmatrix} -0.8125 \\ 2 \end{pmatrix} = \begin{pmatrix} 0.726647 \\ 0.672868 \end{pmatrix} \text{ y } c^T x^1 = 0.894864.$$

Dado que x^1 es un punto extremo no degenerado con $\text{constr}(x^1) = \{0.343\}$, podemos, pues, llevar a cabo un paso de pivote no degenerado, de acuerdo con las reglas dadas en la sección anterior puesto que no se trata de un óptimo y en dos iteraciones se consigue resolver el problema:

$$x^2 = \begin{pmatrix} 0.726647 \\ 0.672868 \end{pmatrix} + 0.75858 \begin{pmatrix} -0.2965 \\ 1 \end{pmatrix} = \begin{pmatrix} 0.501728 \\ 1.431448 \end{pmatrix}, c^T x^2 = 0.85959$$

$$x^3 = \begin{pmatrix} 0.501728 \\ 1.431448 \end{pmatrix} + 0.071667 \begin{pmatrix} -0.253 \\ 1 \end{pmatrix} = \begin{pmatrix} 0.483596 \\ 1.503115 \end{pmatrix}, \text{ cuyo } \text{constr}(x^3) = \{0.25\} \text{ y } c^T x^3 = 0.859375, \text{ que es solución óptima pues } \theta^* = 0.$$

Ejemplo 4.9

Consideremos de nuevo el problema del Ejemplo 3.8:

$$(P) \text{ Min } \sum_{i=1}^7 1/i x_i \\ \text{s.a. } \sum_{i=1}^7 x_i s^{i-1} \geq 1/(1+s^2) \quad s \in [0,1]$$

Tras aplicar el algoritmo de purificación del Capítulo III habíamos llegado al punto extremo degenerado $x^0 = (1, 0, -1, 0, 5/4, -1, 1/4)^T$ con $\text{constr}(x^0) = \{0, 1\}$, $d(0) = d(1) = 3$.

Resolvemos el problema (P1) y obtenemos $d^* = (0, -1, -1, -1, 1, 1, 1)$ y comprobamos que no podemos descender a lo largo de ella. El problema (P2) nos da como resultado $\varepsilon^* = 0.573809$, hacemos $\eta = 3/4\varepsilon^*$ resolvemos (P3) y encontramos $g^* = (0.430357, -1, -1, -1, 1, 1, 1)$, se cumple que $c^T g^* = -0.143452 < 0$, así pues g^* es dirección factible de mejora, determinamos la longitud de salto y esta resulta ser $\mu^* = 0.000029$.

Sección 4.3. La solución posible básica inicial.

Los dos métodos clásicos de obtención de soluciones posibles básicas, en Programación Lineal finita, son el método de las penalizaciones y el de las dos fases. El primero de ellos exige la no negatividad de la variable artificial, y el simplex que hemos diseñado no contempla esta situación. Por tanto, utilizaremos el método de las dos fases semi-infinito como se sugería también en Anderson y Lewis (89).

Dado un problema continuo de PSIL:

$$(P) \text{ Min } c^T x \\ \text{s.a. } a(s)^T x \geq b(s) \quad s \in S$$

A partir de (P) se construye el problema artificial

$$(PA) \text{ Min } y_0 \\ \text{s.a. } y_0 + a(s)^T x \geq b(s) \quad s \in S$$

Nótese que (PA) es un problema de programación semi-infinita sobre \mathbb{R}^{n+1} . Aplicando al problema (PA) el método simplex semi-infinito que hemos desarrollado trataremos de encontrar una solución posible $X = \begin{pmatrix} y_0 \\ x \end{pmatrix}$, no necesariamente óptima, con $y_0 \leq 0$.

Lema 4.6

El problema (P) es inconsistente si, y solo si, para toda solución posible de (PA) se tiene que $y_0 > 0$.

Demostración.

Notese que (PA) es consistente, pues disponemos de una solución posible que es $X^0 = \begin{pmatrix} K \\ 0_n \end{pmatrix}$, siendo $K > 0$ una cota superior de $b(s)$ en S . En particular, por ser $b(\cdot)$ continua en un compacto, para esta solución el conjunto de restricciones activas es no vacío.

Si para toda solución posible $X = \begin{pmatrix} y_0 \\ x \end{pmatrix}$ tenemos que $y_0 > 0$, es inmediato que $F = \emptyset$. El recíproco es también evidente. ♦

Nota. Podemos suponer $K > 0$, puesto que si $b(s) \leq 0$ para todo $s \in S$ tendríamos inmediatamente una solución posible para el problema original (P) que es $x^0 = 0_n$. Como sucede en el siguiente problema:

Ejemplo 4.10

(P) Min x_1+x_2

$$\text{s.a. } sx_1 + s^3x_2 \geq -1 \quad s \in [-3/2, 3/2]$$

Como $b(s) < 0$, para todo $s \in S$, hay una solución posible obvia para el problema original $x^0 = 0_2$, pues $z(s) = -b(s) > 0$ y $\text{constr}(x^0) = \emptyset$. Aplicamos el algoritmo de purificación del Capítulo III y, tras dos iteraciones, obtenemos $x^2 = \begin{pmatrix} -0.871795 \\ 0.091168 \end{pmatrix}$ cuyo $\text{constr}(x^2) = \{3/2\}$, $d(3/2) = 1$ y $c^T x^2 = -0.780627$, que es un punto extremo no degenerado que nos permitiría comenzar a aplicar las reglas de descenso de la Sección 4.1.

En los dos ejemplos siguientes, vemos cómo el método de las dos fases permite concluir que un problema (P) es inconsistente y que disponemos de una solución posible para (P), respectivamente.

Ejemplo 4.11

(P) Min $x_1 + 1/2x_2$

$$\text{s.a. } sx_1 - s^3x_2 \geq 1 + s^2 \quad s \in [-1, 1]$$

En este ejemplo $b(s) = 1 + s^2$ alcanza su máximo en $[-1, 1]$ para $s = \{-1, 1\}$, entonces $K=2$. Planteamos el problema artificial de la Fase I:

(PA) Min y_0

$$\text{s.a. } y_0 + sx_1 - s^3x_2 \geq 1 + s^2 \quad s \in [-1, 1]$$

La solución posible inicial que tenemos para (PA) es $X^0 = \begin{pmatrix} 2 \\ 0_2 \end{pmatrix}$, cuyo $\text{constr}(x^0) = \{-1, 1\}$, con $d(1) = d(-1) = 0$. Dado que no se trata de un punto extremo aplicamos el algoritmo de purificación, y tras una iteración obtenemos el punto extremo $X^1 = \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix}$ con $\text{constr}(X^1) = \{-1, 1\}$, $d(1) = 1$, $d(-1) = 0$. Además, si aplicamos el test de optimalidad, vemos que es una solución óptima para (PA), por tanto $v(\text{PA}) = 2 > 0$ y concluimos que el problema (P) es inconsistente.

Ejemplo 4.12

(P) Min $c_1x_1+c_2x_2$

$$\text{s.a. } \sin((\pi/2)s)x_1 + \cos((\pi/2)s)x_2 \geq 1-s \quad s \in [0,1]$$

Planteamos el problema artificial

(PA) Min y_0

$$\text{s.a. } y_0 + \sin((\pi/2)s)x_1 + \cos((\pi/2)s)x_2 \geq 1-s \quad s \in [0,1]$$

que tiene como solución posible $X^0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ cuyo $\text{constr}(X^0) = \{0\}$, $d(0) = 0$. No tenemos una solución posible básica, por lo que aplicaremos el algoritmo de purificación. En la primera iteración tenemos que el método ha encontrado una solución posible $X^1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, cuya primera componente es nula, así pues no es necesario continuar considerando el problema artificial.

Tenemos una solución posible para el problema original $x^0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, cuyo $\text{constr}(x^0) = \{0, 1\}$, con $d(0) = 0 = d(1)$ y que es un punto extremo no degenerado. Hallamos $\tilde{A}^{-1}c = \begin{pmatrix} c_2 \\ c_1 \end{pmatrix}$, entonces:

4.12 a) si $c_1, c_2 \geq 0$, x^0 es solución óptima,

4.12.b) si algún coeficiente es negativo, al aplicar la regla del CASO 1 para puntos extremos no degenerados encontramos al pivotar una dirección de recesión:

(i) si $c_2 < 0$, se tiene que $\lambda_1 < 0$ y $d = (0, 1)^T$, $a(s)^T d = \cos((\pi/2)s) \geq 0$, $s \in S$,

(ii) si $c_1 < 0$, se tiene que $\lambda_2 < 0$ y $d = (1, 0)^T$, $a(s)^T d = \sin((\pi/2)s) \geq 0$, $s \in S$.

Por último, es necesario considerar la posibilidad de que el problema artificial sea no acotado, como sucede en los problemas siguientes. En ambos casos se ha detectado la no acotación de (PA) a través del algoritmo de purificación, que se detiene al encontrar una dirección de recesión. No obstante, ello no es un obstáculo para encontrar la solución posible que buscábamos para (P).

Los algoritmos de purificación que siguen el esquema de Lewis (85) no pueden ser utilizados sobre esta clase de problemas, pues (PA) no satisface la condición (R1).

Ejemplo 4.13 (en Anderson y Nash (87))

$$(P) \text{ Min } x_1 + x_2 \\ \text{s.a. } (1-s)x_1 + sx_2 \geq s(1-s) \quad s \in [0,1]$$

El problema artificial de la Fase I es:

$$(PA) \text{ Min } y_0 \\ \text{s.a. } y_0 + (1-s)x_1 + sx_2 \geq s(1-s) \quad s \in [0,1]$$

$K=0.25$ es una cota superior para $b(s)$ en $[0, 1]$, por tanto $X^0 = \begin{pmatrix} 0.25 \\ 0 \\ 0 \end{pmatrix}$ cuyo $\text{constr}(X^0) = \{0.5\}$, $d(0.5) = 1$, es solución posible aunque no básica del problema (PA). Aplicando el algoritmo de purificación encontramos la dirección $d = (-1, 1, 1)^T$ que es una dirección de recesión, ya que $a(s)^T d = 0$ para todo s . Así pues, (PA) es no acotado.

Además, sabemos que:

$$X(\alpha) = \begin{pmatrix} 0.25 \\ 0 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0.25 - \alpha \\ \alpha \\ \alpha \end{pmatrix} \text{ es solución posible de (PA), siendo } \alpha \geq 0. \text{ Por}$$

consiguiente, $x(\alpha) = \begin{pmatrix} \alpha \\ \alpha \end{pmatrix}$ es solución posible para (P) cuando $\alpha \geq 0.25$. En particular, para $\alpha = 0.25$ se alcanza el óptimo de (P).

Como criterio general para este tipo de situaciones, construiremos la solución posible inicial para (P) a partir de aquella solución del problema artificial que tiene por primera componente 0.

Ejemplo 4.14

$$(P) \text{ Min } c_1 x_1 + c_2 x_2 \\ \text{s.a. } e^s x_1 + s x_2 \geq 1 - s^2 \quad s \in [0,1]$$

El problema artificial de la Fase I es:

$$(PA) \text{ Min } y_0 \\ \text{s.a. } y_0 + e^s x_1 + s x_2 \geq 1 - s^2 \quad s \in [0,1]$$

El punto $X^0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ cuyo $\text{constr}(X^0) = \{0\}$, $d(0) = 1$, es solución posible no básica y tras una iteración del algoritmo de purificación encontramos la dirección de recesión $d = (-1, 1, -1)^T$ del problema (PA).

Dado que $\{X(\alpha) = \begin{pmatrix} 1-\alpha \\ \alpha \\ -\alpha \end{pmatrix}, \alpha \geq 0\}$ es una familia de soluciones posibles para

(PA), tomaremos $\alpha = 1$ para construir una solución posible para (P). Así pues, $x^0 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ es una solución posible de (P), que será óptima o no según los valores c_i .

En general, el procedimiento que utilizaremos para determinar una solución posible inicial será el método de las dos fases, pues complementado con el algoritmo de purificación permite resolver con eficacia todo tipo de situaciones. Aunque en algunos problemas no será necesario usarlo debido a que la particular estructura del problema permite conocer inmediatamente una solución inicial para ellos. Consideremos el caso de los problemas de aproximación L_1 cuya forma es:

$$\begin{aligned} \text{Min } & \sum_{i=1}^n 1/i x_i \\ \text{s.a. } & \sum_{i=1}^n x_i s_i^{-1} \geq b(s) \quad s \in S \end{aligned}$$

Puesto que $a_1(s) = \text{cte}$, basta considerar $x^0 = \begin{pmatrix} K \\ 0_{n-1} \end{pmatrix}$, siendo K una cota superior de $b(s)$ en S . Esta idea es la que hemos venido utilizando para encontrar la solución posible inicial en la mayoría de los ejemplos de este tipo que hemos resuelto en la memoria.

Con respecto al método de las penalizaciones puede verse que los resultados teóricos del caso finito se trasladan fácilmente al semi-infinito. Aunque ya se ha dicho que este método no puede utilizarse en la práctica debido a como están diseñados el método Simplex de Anderson y Lewis (89) y el nuestro.

A partir de (P) el método de las penalizaciones construye el problema :

$$\begin{aligned} (P') \quad \text{Min } & c^T x + M x_0 \\ \text{s.a. } & x_0 + a(s)^T x \geq b(s) \quad s \in S \\ & x_0 \geq 0 \end{aligned}$$

Se trataría de encontrar, aplicando el simplex semi-infinito, una solución posible para el (P') con $x_0 = 0$, pero el problema es que debe mantenerse no negativa la variable x_0 .

Conclusiones

El método que hemos descrito para resolver el problema contínuo de PSIL, es un método de tipo primal y, como tal tiene la ventaja de que cada punto generado en la búsqueda es una solución posible. Así, si el proceso termina antes de alcanzar la solución óptima, el punto final al menos es una solución posible. Pero, claro está, presenta las desventajas también propias de esta clase de procedimientos: necesita un método de 2 fases para encontrar una solución posible inicial y, sobre todo, presenta dificultades computacionales derivadas de la necesidad de permanecer en la región factible, conforme avanza el método.

Se trata, asimismo, de un método de direcciones factibles. Uno de los problemas de estos métodos es que, para un problema general, no tenemos asegurado que tenga que existir una dirección posible. En nuestro caso esto no sucede, pues dadas las condiciones con que trabajamos las direcciones están claramente definidas. Un segundo inconveniente es que, si no se plantean adecuadamente, estos procedimientos pueden estar sujetos a zig-zag y, por tanto, no hay convergencia global. Para evitarlo suelen generarse las direcciones de manera que verifiquen alguna propiedad que permita asegurar la convergencia del método (ver p. e. Luenberger (73)).

Para nuestro procedimiento no hemos probado la convergencia global. En caso de que, tras un determinado número de iteraciones, la disminución en el valor de la función objetivo sea despreciable podríamos utilizar el punto final obtenido como punto de partida para un método con convergencia local (p.e. métodos Lagrangianos en Coope y Watson (85)).

La extensión del método Simplex que presentamos creemos que presenta no sólo un interés de tipo teórico, sino también práctico, pues puede ser fácilmente implementado, debido a la sencillez del esquema desarrollado para generar las direcciones de descenso. Además, el algoritmo de purificación del Capítulo III ha resultado ser un método potente y eficaz para encontrar puntos extremos.

Los problemas que puede plantear este método son los de tipo numérico -de precisión- que plantean, en general, los métodos de direcciones factibles, para los cuales no está probada la convergencia. Si bien al aplicarlo nos ha permitido llegar al

óptimo, cuando este existía, en todos los problemas que nos hemos planteado a lo largo del desarrollo de esta memoria. Mención aparte merece el Ejemplo 3.8, para $n=7$, en el que ligeras variaciones en una solución posible provocan cambios sustanciales en las iteraciones posteriores. En el Apéndice pueden encontrarse las soluciones de todos ellos.

APENDICE

APENDICE.

En este apéndice presentamos un resumen conciso de las características, propiedades y soluciones óptimas de los ejemplos que han ido apareciendo a lo largo del desarrollo de la memoria. Se ha hecho énfasis en la resolución numérica de los mismos, cuando ello ha sido posible, utilizando los resultados introducidos en los capítulos anteriores.

Cuando no se podía añadir más información sobre algún problema, respecto de lo que ya se había dicho en su momento, hemos optado por no citar el ejemplo correspondiente.

A.1 Ejemplos del Capítulo II.

Ejemplo 2.1

$$(P_1) \text{ Min } \{x_1+x_2\}$$
$$\text{s.a. } sx_1+s^2x_2 \geq 0 \quad s \in [0,1]$$

No se satisface la condición de Slater. El conjunto de soluciones óptimas es $\{x(\alpha) = \begin{pmatrix} \alpha \\ -\alpha \end{pmatrix} : \alpha \geq 0\}$ con $v(P_1)=0$, y todos estos puntos son regulares lagrangianos a excepción de $x(0)$.

Ejemplo 2.2 y 2.5 (en Glashoff y Gustafson (83))

$$(P_2) \text{ Min } x_1$$
$$\text{s.a. } sx_1+ s^2x_2 \geq s^2 \quad s \in [0,1]$$

No se satisface ninguna cualificación de restricciones. De hecho, el problema dual es inconsistente por lo que no puede haber puntos de Kuhn-Tucker. El conjunto de soluciones óptimas es $\{x(\alpha) = \begin{pmatrix} 0 \\ \alpha \end{pmatrix} : \alpha \geq 1\}$ con $v(P_2)=0$.

Ejemplo 2.4

$$(P_4) \text{ Min } \{x_1 + 1/2x_2\}$$

$$\text{s.a. } x_1 + sx_2 \geq 1/(1+s^2) \quad s \in [0,1]$$

Puesto que $b(s) > 0$ para todo $s \in S$, se satisface la cualificación de Mangasarian-Fromovitz. De la interpretación geométrica de este problema se obtiene $x = \begin{pmatrix} 1.103553 \\ -0.603553 \end{pmatrix}$ cuyo $\text{constr}(x) = \{0.41421, 1\}$, $d(0.41421) = 1$, $d(1) = 0$, según el Teorema 2.3 sabemos que es un punto extremo degenerado óptimo y $v(P_4) = 0.801776$.

También se puede llegar a esta solución aplicando las reglas de descenso del Capítulo IV y el algoritmo de purificación del Capítulo III:

Iterac.	x	constr(x), d(i)	dirección	μ^*
1	(1, 0)	{0}, d(0)=1	(0.25, -1) ^T	2/3
2	(5/3, -2/3)	{1}, d(1)=0	(-1, 1) ^T	0.063114
3	(1.103553, -0.603553)	{.414, 1}, d(0.414)=1, d(1)=0		

$$\theta^* = 0 \text{ y } v(P_4) = 0.801776.$$

Ejemplo 2.7

$$(P_7) \text{ Min } \{x_1 + (1/2)x_2\}$$

$$\text{s.a. } sx_1 + s^2x_2 \geq s^3 - 1 \quad s \in [-1, 1]$$

Resumimos en una tabla las soluciones que se van encontrando hasta llegar a la óptima. Señalamos con una D los puntos extremos degenerados.

Iterac.	x	constr(x), d(i)	dirección	μ^*
1	(1, -1) D	{1, -1}, d(1)=0, d(-1)=1	(-1, 1) ^T	4
2	(-3, 3) D	{1}, d(1)=2	(-0.625, 1) ^T	2
3	(-4.25, 5)	{0.5}, d(0.5)=1	****	

$$\theta^* = 0 \text{ y } v(P_7) = -1.75.$$

Ejemplo 2.10

$$(P_{10}) \text{ Min } \{x_1 + 1/2 x_2\}$$
$$\text{s.a. } x_1 + s^2 x_2 \geq 0 \quad s \in [-1, 1]$$

Se satisface la condición de Slater. El conjunto de soluciones óptimas es $\{x(\alpha) = \begin{pmatrix} \alpha \\ -\alpha \end{pmatrix} :$

$\alpha \geq 0\}$ con $v(P_1) = 0$.

A.2 Ejemplos del Capítulo III.

Ejemplo 3.1

$$(P_1) \text{ Min } x_1 \\ \text{s.a. } x_1 + sx_2 \geq 1, \quad s \in [0,1]$$

Se verifica la condición de Slater y el conjunto $\{x(\lambda) = (1, \lambda)^T, \text{ para } \lambda \geq 0\}$ es el de las soluciones óptimas de (P_1) . Así pues $v(P_1) = 1$.

Ejemplo 3.2

$$(P_2) \text{ Min } \{2x_1 + 2x_2 + x_3\} \\ \text{s.a. } x_1 + x_2 + sx_3 \geq 1 \quad s \in [0,1]$$

Los puntos posibles $x = (1, 0, 0)^T$ y $x' = (0, 1, 0)^T$ son óptimos, así como cualquier combinación convexa de ellos. Además son puntos regulares lagrangianos porque, para todos ellos, el cono de los momentos es $M_n = \{x = (1, 1, s)^T: s \in [0, 1]\}$ y este conjunto es cerrado.

Ejemplo 3.3 y 3.7

$$(P_3) \text{ Min } x_1 + 2x_2 \\ \text{s.a. } x_1 + sx_2 \geq \exp(s), \quad s \in [0,1]$$

Se verifica (RA). La dirección $d = (2, -3/2)^T$ es de recesión, por lo que el problema no está acotado. También se ha llegado a esta conclusión aplicando el algoritmo de purificación. Partiendo de la solución posible inicial $x = (e, 0)^T$ cuyo $\text{constr}(x) = \{1\}$, $d(1) = 0$, en la primera iteración obtenemos $d = (1, -1)^T$ con $\beta = 0$.

Ejemplo 3.5

$$(P_5) \text{ Min } x_1 \\ \text{s.a. } sx_1 + s^2x_2 \geq s^3, \quad s \in [0,1]$$

No se verifica la condición de Slater. Todos los elementos del conjunto $\{x(\alpha) = \begin{pmatrix} 0 \\ \alpha \end{pmatrix}; \alpha \geq 1\}$ son soluciones óptimas y $v(P_5) = 0$.

Ejemplo 3.8 Para $n=5$, resuelto en **Ejemplo 4.6**
 Para $n=7$ vease **Ejemplo 4.9**

Ejemplo 3.10 (en Roleff (79))

$$(P_6) \quad \text{Inf} \quad \sum_{i=1, \dots, 8} (1/i) x_i$$

$$\text{s.a} \quad \sum_{i=1, \dots, 8} s^{i-1} x_i \geq \tan(s) \quad s \in [0, 1].$$

Partiendo del punto extremo no degenerado que encontramos mediante el algoritmo de purificación llegamos a la solución óptima tras dos iteraciones más. Este punto es $x^0 = (0, 1.001832, -.031293, .546183, -.740157, 1.535467, -1.394181, .639555)^T$, cuyo $\text{constr}(x^0) = \{0, .21, .471, .68, 1\}$.

Tomamos como dirección de descenso de acuerdo con la regla del CASO 1 ($\lambda_{0,4} < 0$) $d_1 = (0, 42.9161, -688.644, 4317.37, -13453.8, 22026.6, -18016.1, 5771.72)^T$, y encontramos

$x^1 = (0, 1.003806, -.062971, .744782, -1.359032, 2.548691, -2.222922, .905054)^T$, cuyo $\text{constr}(x^1) = \{0, .21, .471, .829, 1\}$.

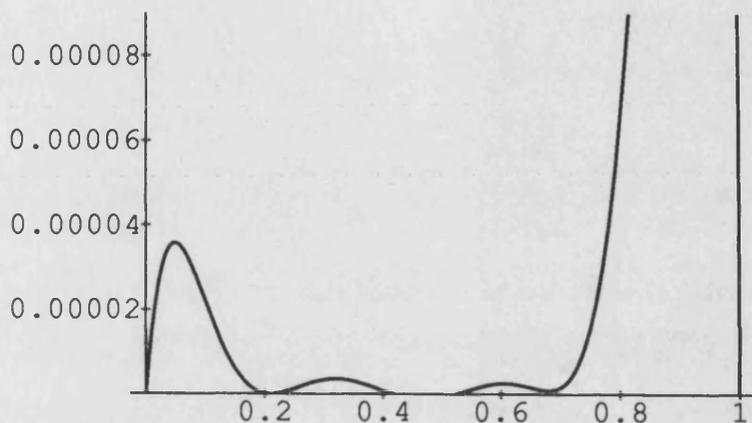
Finalmente, según la regla del CASO 2B ($\rho_{2,1} < 0$), generamos la dirección $d_2 = (0, -7.87317, 101.929, -498.842, 1213.68, -1570.39, 1034.52, -273.023)^T$, que nos conduce a la solución óptima

$x^2 = (0, 1.002137, -.041362, .639027, -1.101732, 2.215768, -2.003604, .847173)^T$ cuyo $\text{constr}(x^2) = \{0, .162, .471, .829, 1\}$ con $c^T x^2 = 0.615653$ ($\theta^* = -0.00053$).

En las gráficas siguientes puede verse como cambian los ceros de la función de holgura cuando se aplican las diferentes reglas de descenso.

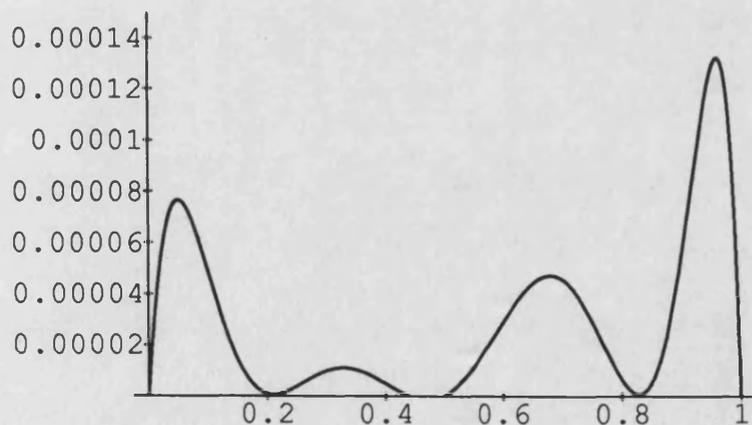
$$z_0[s] := 1.001832*s - .031293*s^2 + .546183*s^3 - .740157*s^4 + 1.535467*s^5 - 1.394181*s^6 + .639555*s^7 - \text{Tan}[s]$$

constr(x0)={0, .21, .471, .68, 1}, $\lambda_{4,0} < 0$: CASO 1



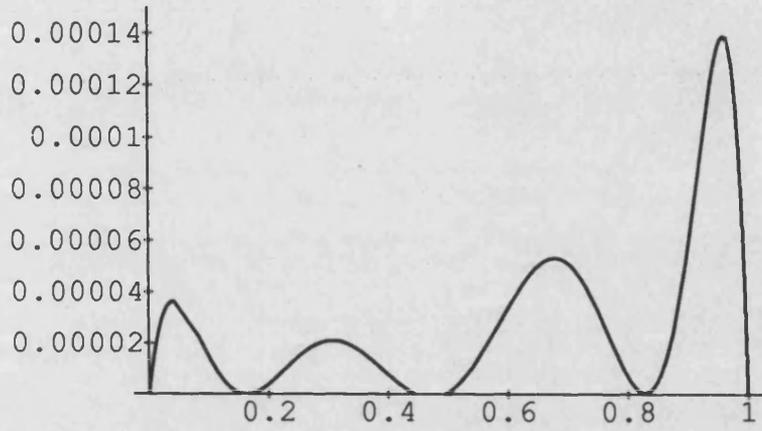
$$z_1[s] := 1.003806*s - .062971*s^2 + .744782*s^3 - 1.359032*s^4 + 2.548691*s^5 - 2.222922*s^6 + .905054*s^7 - \text{Tan}[s]$$

constr(x1)={0, .21, .471, .829, 1}, $\rho_{2,1} < 0$: CASO 2B



$$z_2[s] := 1.002137*s - .041362*s^2 + .639027*s^3 - 1.101732*s^4 + 2.215768*s^5 - 2.003604*s^6 + .847173*s^7 - \text{Tan}[s]$$

constr(x2)={0, .162, .471, .829, 1}, x2 es solución óptima.



A.3 Ejemplos del Capítulo IV.

Ejemplo 4.2

$$(P_2) \text{ Min } x_1 + 1/2x_2$$

$$\text{s.a. } x_1 + sx_2 \geq 1 - s^2 \quad s \in [-1, 1]$$

$x^0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ es una solución posible para (P_2) , con $z(s) = s^2$, $\text{constr}(x^0) = \{0\}$ y $d(0) = 1$. Se tiene que $\tilde{A} = I_2$ y $\tilde{A}^{-1}c = c$, y por tanto $\rho \neq 0$. Generamos $d = \begin{pmatrix} 1/4 \\ -1 \end{pmatrix}$, según la regla del CASO 2A, que nos conduce directamente al óptimo $x^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 1/4 \\ -1 \end{pmatrix} = \begin{pmatrix} 5/4 \\ -1 \end{pmatrix}$ cuyo $\text{constr}(x^0) = \{1/2\}$ y $d(1/2) = 1$.

Ejemplo 4.5

$$(P_5) \text{ Min } -x_1 - x_2$$

$$\text{s.a. } sx_1 + s^3 x_2 \geq -1 \quad s \in [-1.5, 1.5].$$

Partiendo del punto interior $x^0 = \begin{pmatrix} -1 \\ 1/2 \end{pmatrix}$ tras aplicar el algoritmo de purificación se alcanza el punto extremo extremo no degenerado $x^1 = \begin{pmatrix} 1 \\ -4/27 \end{pmatrix}$. A continuación resumimos las demás iteraciones.

Iterac.	x	constr(x), d(i)	dirección	μ^*
1	$(1, -.148148)^T$	$\{-3/2\}, d=1$	$(15.352581, -9.024457)^T$	0.019137
2	$(1.293802, -.320849)^T$	$\{-1.159\}, d=1$	$(4.0237904, -3.40437)^T$	0.042274
3	$(1.463904, -.464765)^T$	$\{-1.025\}, d=1$	$(2.237539, -2.181657)^T$	0.015374
4	$(1.498303, -.498305)^T$	$\{-1.001\}, d=1$	$(2.009021, -2.007011)^T$	0.000995
5	$(1.500302, -.500302)^T$	$\{-1\}, d=1$		

$\theta^* = 0$ y $v(P_5) = -1$.

Ejemplo 4.8

(P₈) Min $x_1 + 1/4 x_2$

$$\text{s.a. } x_1 + sx_2 \geq -4s^4 + 12s^3 - 13s^2 + 6s \quad s \in [0, 1]$$

El punto $x^0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ es una solución posible básica degenerada con $\text{constr}(x^0) = \{1/2, 1\}$, $d(1/2) = 1 = d(1)$ y $z^0(s) = 1 + 4s^4 - 12s^3 + 13s^2 - 6s \geq 0$ para $s \in [0, 1]$. Para x^0 , la dirección que se obtiene resolviendo el problema lineal del test de optimalidad, no es factible, pues $\mu^* = 0$. Hallamos $g^* = \begin{pmatrix} -0.3125 \\ 1 \end{pmatrix}$ que es dirección de descenso y pivotando llegamos a $x^1 = \begin{pmatrix} 0.360442 \\ 2.046585 \end{pmatrix}$ cuyo $\text{constr}(x^1) = \{0.205\}$. En resumen, se tiene:

Iterac.	x	constr(x), d(i)	dirección	μ^*
1	$(1, 0)^T$ D	$\{1/2, 1\}$ $d(1/2) = 1 = d(1)$	$(-.3125, 1)^T$	2.046585
2	$(0.360442, 2.046585)^T$	$\{.205\}$, $d=1$	$(.2275, -1)^T$	0.564423
3	$(0.488848, 1.482162)^T$	$\{.252\}$, $d=1$	$(-.253, 1)^T$	0.000577
4	$(0.488702, 1.482739)^T$	$\{.251\}$, $d=1$	$(-.2515, 1)^T$	0.007833
5	$(0.486732, 1.490572)^T$	$\{.25\}$, $d=1$		

$$\theta^* = 0 \text{ y } v(P_5) = 0.859375.$$

Ejemplo 4.9

(P) Min $\sum_{i=1, \dots, 7} (1/i) x_i$

$$\text{s.a. } \sum_{i=1, \dots, 7} s^{i-1} x_i \geq 1/(1+s^2) \quad s \in [0, 1].$$

El punto $x^1 = (1, 0, -1, 0, 1.25, -1, 0.25)^T$ cuyo $\text{constr}(x^1) = \{0, 1\}$, $d(0) = 3 = d(1)$, es un punto extremo degenerado con $c^T x^1 = 0.785714$. Aplicando la correspondiente regla de descenso para esta clase de puntos obtenemos la dirección $g^* = (.430357, -1, -1, -1, 1, 1, 1)^T$ y una longitud de salto μ^* . Según cual sea la precisión utilizada para determinar esta última magnitud, encontramos diferentes conjuntos de restricciones activas que conducen a direcciones de búsqueda diferentes. Veámoslo:

Iterac.	x	constr(x), d(i)	dirección	μ^*
1	$(1,0,-1,0,1.25,-1,0.25)^T$	D {0,1}, d=3	$(.430357,-1,-1,-1,1,1,1)^T$	0.00002941
2	$(1.000013,-.000029,-1.000029,-.000029,1.250029,-.999971,.250029)^T$	{.925}, d=1	$(-1,-.137371,1,1,1,-.923978-1)^T$ (p)	$c^T x^2 = .785711$ 0.000012
3	$(1.000001,-.000031,-1.000017,-.000017,1.250041,-.999982,.250017)^T$	{.034,.925} d(.034)=1=d(.925)	$(-.001103,.065714,-1,.613276,1,.302528,-1)^T$ (p)	$c^T x^3 = .785704$ 0.000613
4	$(1,.000009,-1.00063,.000359,1.250654,-.999796,.249404)^T$	{0,.034,.924,1} d(.034)=1=d(.924) d(1)=0=d(0)	$(0,-.000328,.020816,-.356506,.993022,-1,.342997)^T$ (p)	$c^T x^4 = .785679$ 0.004608
5	$(1,.000007,-1.000534,.001284,1.25523,-1.004404,.250984)^T$	{0,.034,.924,1} d(.034)=1=d(.924) d(1)=0=d(0)	(**)	$c^T x^5 = .785672$

$$\theta^* = -.161062$$

Escribimos (p) cuando se trata de una dirección generada por el algoritmo de purificación.

(**) En este punto que no es una solución óptima, no se ha encontrado un nuevo mínimo local para la función de holgura, por problemas de precisión, es por ello que no se puede continuar.

Ejemplo A.4

(P) Min $2x_1 + 2/3 x_3$

s.a. $x_1 + sx_2 + s^2x_3 \geq e^s \quad s \in [0,1]$

Para conseguir una solución posible inicial consideramos el polinomio de Taylor de grado 2 en $a=1$, $P_2(s) = e/2 + e/2 s^2$. Así pues, nuestro punto de partida será $x^0 = (e/2, 0, e/2)$ cuyo $\text{constr}(x) = \{1\}$, $d(1)=2$ y pivotaremos según la regla del CASO 3:

Iterac.	x	constr(x), d(i)	dirección	μ^*
1	$(1.359141, 0, 1.359141)^T$	$\{1\}, d=2$	$(-.166667, 1, -.5)^T$	1.372659
2	$(1.130364, 1.372659, .672811)^T$	$\{-.621\}, d=1$	$(-.3105, -1, -.805153)^T$ (p)	0.564423
3	$(1.063212, 1.156389, .498681)^T$	$\{-.621, 1\}$ $d(-.621)=1, d(1)=0$	$(-.278869, -.291544, .570414)^T$	0.138605
4	$(1.024559, 1.115980, .577743)^T$	$\{-.365, 1\}$ $d(-.365)=1, d(1)=0$	$(-.253032, -.471264, .724296)^T$	0.01474
5	$(1.020829, 1.109034, .588419)^T$	$\{-.334, 1\}$ $d(-.334)=1, d(1)=0$	$(-.250062, -.499375, .749438)^T$	0.000333
6	$(1.020786, 1.108949, .588547)^T$	$\{-.333, 1\}, d(-.333)=1, d(1)=0$		$\theta^*=0$

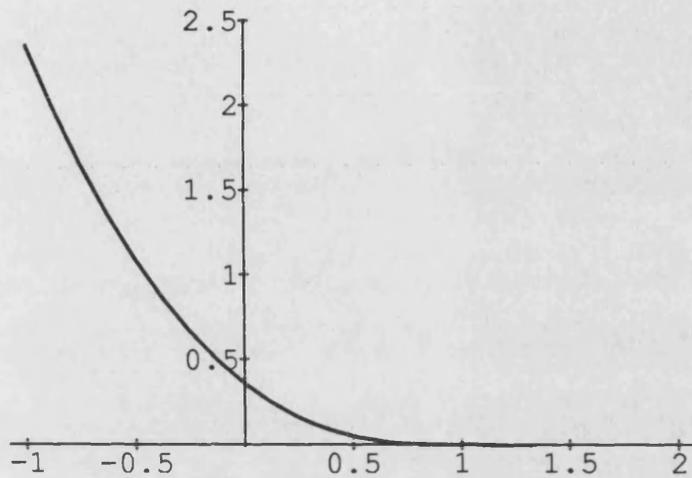
siendo $v(P) = 2.433937$.

Escribimos (p) cuando se trata de una dirección generada por el algoritmo de purificación.

En las gráficas siguientes puede verse como cambian los ceros de la función de holgura cuando se aplican las diferentes reglas de descenso.

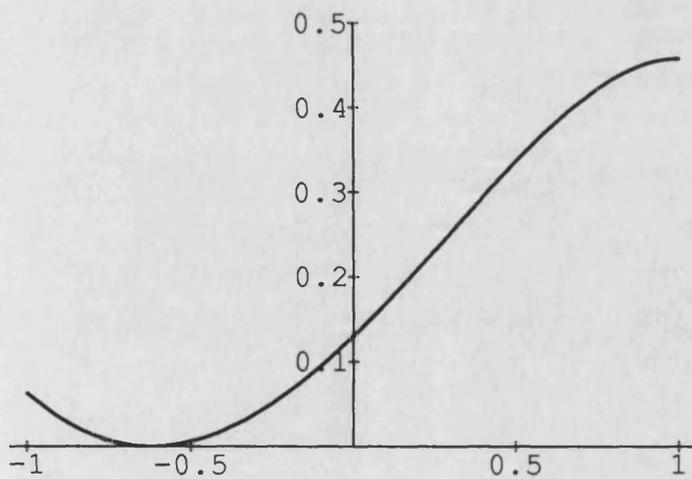
$$z_1(s) = 1.359141 + 1.359141 \cdot s^2 - \exp(s), \text{ constr}(x_1) = \{1\}, d(1) = 2$$

$\rho_2 \neq 0$: CASO 3



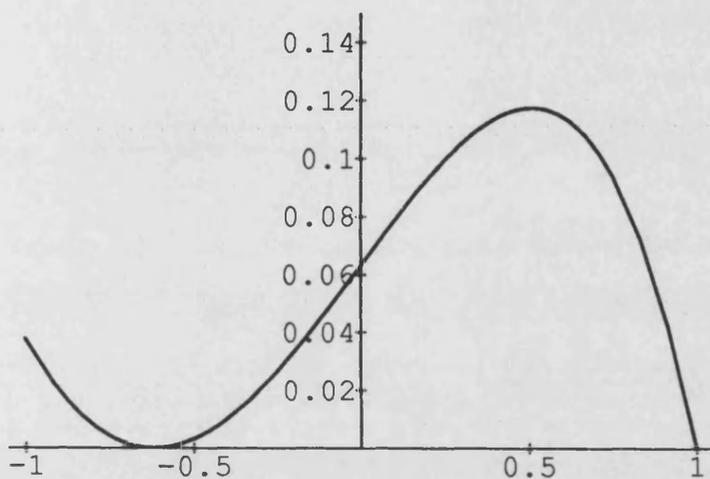
$$z_2(s) = 1.130364 + 1.372659 \cdot s + 0.672811 \cdot s^2 - \exp(s), \text{ constr}(x_2) = \{-.621\}, d(-.621) = 1$$

x_2 no es punto extremo : purificamos



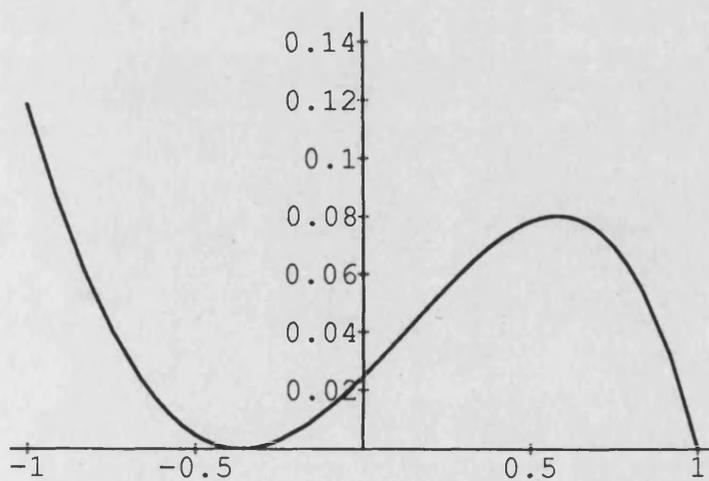
$$z_3(s) = 1.063212 + 1.156389*s + 0.498681*s^2 - \exp(s), \text{ constr}(x_3) = \{-.621, 1\}, d(-.621) = 1$$

$d(1) = 0, \rho_{1,1} > 0$: CASO 2A



$$z_4(s) = 1.024559 + 1.115980*s + 0.5777431*s^2 - \exp(s), \text{ constr}(x_4) = \{-.365, 1\}, d(-.365) = 1$$

$d(1) = 0, \rho_{1,1} > 0$: CASO 2A



BIBLIOGRAFIA

BIBLIOGRAFIA.

Anderson E.J. (85) "A new primal algorithm for semi-infinite linear programming", en *Infinite Programming*, E.J. Anderson y A. B. Philpott (eds), Springer Verlag, Berlin, 108-122.

Anderson E.J., Lewis A.S. (89) "An extension of the simplex algorithm for Semi-Infinite Linear Programming", *Mathematical Programming* 44, 247-269.

Anderson E.J., Lewis A.S., Wu S.Y. (89) "The capacity problem", *Optimization* 20, 725-742.

Anderson E.J., Nash P. (87) *Linear Programming in infinite-dimensional spaces*, John Wiley and Sons, Chichester.

Anderson E.J., Philpott A.B. (eds) (85), *Infinite Programming*, Springer Verlag, Berlin.

Bazaraa M.S., Shetty C.M. (76), *Foundations of Optimization*. Lecture Notes in Economics and Mathematical Systems 122, Springer Verlag, Berlin.

Ben-Tal A., Kerzner L., Slobec S. (80) "Optimality conditions for convex semi-infinite programming problems". *Naval Research Logistics Quarterly* 27, 412-435.

Best M.J., Ritter K. (85), *Linear Programming: Active Set Analysis and Computer Programs*, Prentice Hall, Englewood.

Borwein J.M. (81) "The Limiting Lagrangian as a Consequence of Helly's Theorem", *Journal of Optimization Theory and Applications* 33, 497-513.

Borwein J.M., Kortanek K.O. (83) "Fenchel duality and separably infinite programs". *Math. Op. Statist. Ser. Opt.* 14, 37-48.

Charnes A., Cooper W.W., Kortanek K.O. (62) "Duality Haar programs and finite sequence spaces". *Proc. Nat. Acad. Sci. US.* 48, 783-786.

Charnes A., Cooper W.W., Kortanek K.O. (63) "Duality in semi- infinite programs and some works of Haar and Caratheodory", *Management Science* 9, 209-228.

Charnes A., Cooper W.W., Kortanek K.O. (71), "Semi-infinite programming, differentiability and geometric programming, Part I: With examples and applications in Economics and Management Science", *J. Math. Sci. India* 6, 19-40.

Charnes A., Gribik P. R., Kortanek K.O. (80) "Separably-Infinite Programs", *Zeitschrift fur Operations Research* 24, 33-45.

Charnes A., Kortanek K.O. (63) "An opposite sign algorithm for purification to an extreme point solution", *ONR Research Memo* 84, Northwestern Univ., Evanston.

Charnes A., Kortanek K.O., Raiké W. (65) "Extreme point solutions in mathematical programming : An opposite sign algorithm", *Systems Research Memorandum* no.129, Northwestern University, Evanston.

Conn A.R., Gould N. (87) "An exact penalty function for semi-infinite programming", *Mathematical Programming* 37, 19-40.

Coope I.D., Watson G.A. (85) "A projected Lagrangian algorithm for semi-infinite programming", *Mathematical Programming* 32, 337-356.

Duffin R.J. (56) "Infinite programs", en *Linear Inequalities and Related systems*, Kuhn H.W., Tucker A. W. (eds), Princeton Univ. Press, Princeton N.J., 157-170.

Duffin R.J. (67) "An orthogonality theorem of Dines related to moment problems and linear programming", *J. Combin. Th.*, 2, 1-26.

Duffin, R., Karlovitz L.A. (65) "An infinite linear program with a duality gap", *Management Science* 12, 122-134.

Duffin, R., Karlovitz L.A. (68) "Formulation of linear programs in analysis", *SIAM J. Appl. Math.*, 16, 662-675.

Elzinga J., Moore T.G. (75) "A central cutting plane algorithm for the convex programming problem", *Mathematical Programming* 8, 134-145.

Ferris M.C., Philpott A.B. (89) "An interior point algorithm for Semi-infinite linear programming", *Mathematical Programming* 43, 257-276.

Fiacco A.V., Kortanek K.O. (83) (eds), *Semi-infinite Programming and Applications*, Springer Verlag, Berlin.

Gehner K.R. (74) "Necessary and sufficient optimality conditions for the Fritz-John problem with linear equality constraints", *SIAM J. Control* 12, 140-149.

Glashoff K. (79) "Duality theory of semi-infinite programming", en *Infinite Programming*, Hettich R. (ed), Springer Verlag, Berlin, 1-13.

Glashoff K., Gustafson S.A. (83) *Linear optimization and approximation*, Applied Mathematical Sciences vol 45, Springer Verlag, New York.

Goberna M.A., Jornet V. (88) "Geometric fundamentals of the simplex method in Semi-Infinite Programming", *O.R. Spektrum* 10, 145-152.

Goberna M.A., López M.A. (87) "Reduction and Discretization in Linear Semi-Infinite Programming", *Optimization* 18, 643-658.

Goberna M.A., López M.A., Pastor J. (81) "Farkas-Minkowski systems in semi-infinite programming". *Applied Mathematics and Optimization*, 7, 295-308.

Goberna M.A., López M.A., Pastor J., Vercher E. (84) "Alternative theorems for infinite systems with applications to semi-infinite games". *Nieuw Archief voor Wiskunde* (4) Vol 2, 218-234.

Gorr W., Gustafson S.A., Kortanek K.O. (72) "Optimal Control Strategies for Air Quality Standards and Regulatory Policies", *Environment and Planning* 4, 183-192.

Gribik P.R. (79) "A central-cutting plane algorithm for semi-infinite programming problems", en *Infinite Programming*, Hettich R. (ed), Springer Verlag, Berlin, 66-82.

Gribik P.R., Lee D.N. (78) "A comparison of central cutting-plane algorithms for Prototype Geometric Programming Problems". Presentado en III Symposium uber Operations Research, Mannheim.

Gustafson S.A. (70) "On the computational solution of a class of generalized moment problems", *SIAM J. on Numer. Anal.* 7, 343-357.

Gustafson S.A., Kortanek K.O. (73) "Numerical solution of a class of SIP problems", *Naval Research Logistics Quarterly* 20, 477-504.

Gustafson S.A., Kortanek K.O., Rom W. (70) "Non-Chebyshevian Moment Problems", *SIAM J. Numer. Anal* 7, 335-342.

Haar A. (24) "Uber lineare Ungleichungen", *Acta Math (Szeged)* 2, 1-14.

Hestenes M.R. (75), *Optimization theory-The finite dimensional case*, Wiley, New York.

Hettich R.(79) "A comparison of some numerical methods for Semi-infinite Programming", en *Semi-infinite Programming*, Hettich R. (ed), Springer Verlag, Berlin, 112-126.

Hettich R. (79) (ed), *Semi-infinite Programming*, Springer Verlag, Berlin.

Hettich R. (83) "A review of numerical methods for Semi-infinite Optimization", en *Semi-infinite Programming and Applications*, Fiacco A.V., Kortanek K.O.(eds), Springer Verlag, Berlin, 97-111.

Hettich R. (86) "An implementation of a discretization method for semi-infinite programming", *Mathematical Programming* 34, 354-361.

Hettich R., Jongen H. Th. (78) "Semi-infinite programming: conditions of optimality and applications", en *Optimization Techniques, Part 2.*, Stoer J. (ed), Springer Verlag, Berlin, 1-11.

Hettich R., Zencke P. (82) "Numerische Methoden der Approximation und semi-infiniten optimierung", Teubner, Stuttgart.

Hoffman K.H., Klostermair A. (76) "A semi-infinite linear programming procedure and application to approximation problems in Optimal Control", en *Approximation Theory II*, Lorentz et al (eds), Academic Press, 379-389.

John F. (48) "Extremum problems with inequalities as subsidiary conditions" en *Studies and Essays, Courant Anniversary Volume*, Interscience, 187-204.

Jongen H. Th., Zwier G. (85) "On the local structure of the feasible set in semi-infinite optimization", en *Parametric Optimization and Approximations*, B. Brosowski y F. Deutsch (eds.), Birkhauser Verlag, Basel, 185-202.

Karney D.F. (81) "Duality gaps in Semi-infinite Linear Programming. An Approximation Problem", *Mathematical Programming* 20, 129-143.

Karney D.F. (83) "A Duality theorem for Semi-infinite Convex Programs and their finite subprograms", *Mathematical Programming* 27, 75-82.

Kemperman J.H.B. (83) "On the role of duality in the theory of moments", en *Semi-infinite Programming and Applications*, Fiacco A.V., Kortanek K.O. (eds), Springer Verlag.

Kortanek K.O. (91) "Semi-infinite programming duality for order restricted statistical inference models". Presentado en 14th International Symposium on Mathematical Programming, Amsterdam.

Kortanek K.O., Strojwas H.M. (84) "An application of the Charnes-Kortanek-Raike purification algorithm to extreme points and extreme directions", Department of Mathematics, Carnegie Mellon Univ.

Kortanek K.O., Strojwas H.M. (85) "On constraint sets of infinite linear programs over ordered fields", *Mathematical Programming* 33, 146-161.

Kortanek K.O., Yamasaki M. (82) "Semi-infinite Transportation Problems", *J. Math. Anal. Appl.* 88 , 555-565.

Krabs W. (79), *Optimization and approximation* ,Wiley, Chichester.

Krein M.G. (51) "The ideas of P.L. Chebyshev and A.A. Markov in the theory of limiting values of integrals and their further developments", *Am. Math. Soc. Transl.*, ser 2,12, 11-22.

Kurcyusz S. (76)" On the existence and nonexistence of Lagrange multipliers in Banach spaces". *Journal of Optimization Theory and Applications* 20, 81-110.

León T., Vercher E. (91) " A purification algorithm for semi-infinite programming", *European Journal of Operational Research*, en prensa.

Lewis A.S. (85) "Extreme points and purification algorithms in general linear programming", en *Infinite Programming*, E.J. Anderson y A.B. Philpott (eds), Springer Verlag, Berlin, 123-135.

López M.A., Vercher E. (83) "Optimality conditions for nondifferentiable convex semi-infinite programming", *Mathematical Programming* 27, 307-319.

López M.A. (87) "A unified approach to the duality theory in linear and convex semi-infinite programming", Dpt. of Math. and Statistics, University of Alicante, Spain.

Lowe T.J., Hurter A.P. (76) "The generalized market area problem", *Management Science* 22, 1105-1115.

Luenberger D.G. (73), *Introduction to linear and nonlinear programming*, Addison-Wesley, Reading.

Mangasarian O.L. (69), *Nonlinear programming*, McGraw-Hill, New York.

Marsten R. (81) "The design of the XMP linear programming library", *ACM Transactions on Mathematical Software* 7, 481-497.

Nash P. (85) "Algebraic fundamentals of linear programming". en *Infinite Programming*, Anderson E.J., Philpott A.B. (eds.), Springer-Verlag, Berlin, 37-52.

Perold A.F.(81) "Extreme points and basic feasible solutions in continuous time linear programming", *SIAM J. Control and Optimization* 19, 52-63.

Polak E. (87) "On the mathematical foundations of nondifferentiable optimization in Engineering design", *SIAM Review* 1, 21-89.

Polak E., He L. (91), "Unified steerable Phase I-Phase II method of feasible directions for semi-infinite optimization", *Journal of Optimization Theory and Applications* 69, 83-107.

Rajgopal J., Bricker D.L.(90) "Posynomial Geometric Programming as a special case of Semi-infinite Programming", *Journal of Optimization Theory and Applications* 66, 455-475.

Rockafellar R.T. (70), *Convex Analysis*, Princeton University Press, Princeton, New Jersey.



Rogosinski W.E. (58) "Moments of non-negative mass", *Proc. Roy. Soc. (London)*. A, 245 (1958) 1-27.

Rogosinski W.E. (62) "Non-negative linear functionals, moment problems and extremum problems in polynomial spaces", en *Studies in Mathematical Analysis and related topics*, Szego G. (ed), Stanford University Press.

Roleff R. (79) "A stable multiple exchange algorithm for linear SIP", en *Semi-infinite Programming*, R. Hettich (ed), Springer Verlag, Berlin, 83-96.

Rudolph H. (91) "Semi-infinite programming in locally convex spaces: I The simplex algorithm", Institut fur Mathematik, Leipzig.

Tanaka Y., Fukushima M., Ibaraki T. (88) "A comparative study of several semi-infinite nonlinear programming algorithms", *European Journal of Operational Research* 36, 92-100.

Tichatschke R., Hettich R., Still G. (89) "Connections Between Generalized, Inexact and Semi-Infinite Linear Programming", *ZOR-Methods and Models of Operations Research* 33, 367-382.

Todd M.J. (78) "Solving the generalized market area problem", *Management Science* 24, 1549-1554.

Watson G.A. (75) "A multiple exchange algorithm for multivariate Chebyshev approximation", *SIAM J. Num. Anal.* 12, 46-52.

Zoutendijk G. (76), *Mathematical Programming Methods*, North Holland, Amsterdam.

Zowe J. (78) "A remark on a regularity assumption for the Mathematical Programming problem in Banach spaces", *Journal of Optimization Theory and Applications* 25, 375-381.



Reunido el Tribunal que suscribe, en el día de la fecha, acordó otorgar, por unanimidad, a esta tesis doctoral de D. M^{te} Teresa León Mendoza

la calificación de Apdo Cum Laude

Valencia, a 18 de Diciembre de 1991

El Secretario,



Presidente

[Handwritten signature]

[Handwritten signature]

