

Investigación Operativa

Álvaro García Tenorio ¹

17 de julio de 2017

¹alvgar14@ucm.es

Índice general

Prefacio	VI
I Programación lineal	1
1. Fundamentos teóricos	3
1.1. Planteamiento del problema	3
1.1.1. Forma estándar y formas canónicas	4
1.1.2. Terminología y casuística	7
1.2. Programación lineal y conjuntos convexos	7
1.2.1. Conjuntos convexos. Definición y propiedades	7
1.2.2. Puntos extremos	10
1.2.3. Direcciones extremas	13
1.3. Teoremas fundamentales	14
2. Algoritmo del Símplex	15
2.1. Implementación del algoritmo del símplex	15
2.2. Método de las dos fases	15
2.3. Método de las penalizaciones	15
2.4. Regla lexicográfica	15
2.5. Regla de Bland	15
3. Dualidad	17
3.1. Planteamiento del problema dual	17
3.2. Relaciones de dualidad	17
II Programación entera	19
4. Modelización	21
4.1. Modelización con variables binarias	21
4.2. Restricciones disyuntivas	21
4.3. Problemas con costes fijos	21
5. Ramificación y acotación	23
III Programación no lineal	25
IV Anexos	27
Índice de términos	29

Prefacio

Estas notas son una transcripción (libremente adaptada) de las clases de la asignatura “*Investigación Operativa*”, impartidas por María Inés Sobrón Fernández en el curso 2016–2017 a los cursos de cuarto y tercero de los dobles grados de Matemáticas – Física e Ingeniería Informática – Matemáticas (respectivamente) en la facultad de Ciencias Matemáticas de la Universidad Complutense de Madrid (UCM).

Cualquier aportación o sugerencia de mejora es siempre bienvenida.

Requisitos previos

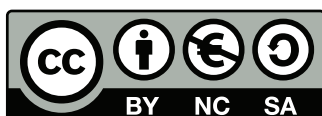
Para comprender estas notas en su totalidad es necesario tener soltura a la hora de trabajar con bases de espacios vectoriales de dimensión finita y comprender bien las aplicaciones lineales. También es bastante recomendable recordar algunos aspectos del cálculo diferencial en varias variables, no obstante, el texto es bastante autocontenido en ese aspecto.

Agradecimientos

La existencia de estas notas es debida a la amabilidad de Clara Rodríguez Núñez, quien me cedió sus apuntes tomados durante el curso, en los cuales se basa el núcleo de este texto.

Licencia

Esta obra está sujeta a la licencia Reconocimiento-NoComercial-CompartirIgual 4.0 Internacional de Creative Commons. Para ver una copia de esta licencia, visite <http://creativecommons.org/licenses/by-nc-sa/4.0/>.



Parte I

Programación lineal

Capítulo 1

Fundamentos teóricos

La disciplina de la programación lineal cogió fuerza en la segunda guerra mundial, pues era necesario optimizar los escasos recursos de los que se disponían para realizar operaciones militares. En un principio fue impulsada por el ejército británico, con un grupo de investigación dirigido por **Blackett**, quien más adelante ganaría el premio Nobel de física.

Cabe destacar la gran importancia que tuvo la programación lineal en países como Albania, debido, entre otras cosas, a su utilidad a la hora de optimizar los procesos de producción de carbón.

1.1. Planteamiento del problema

De ahora en adelante consideraremos fijadas las bases canónicas tanto de \mathbb{R}^n como de \mathbb{R} , a no ser que se especifique lo contrario.

Antes de plantear formalmente el problema de la programación lineal, introduzcamos unas definiciones.

Definición 1.1.1 (Vector no negativo). Un vector $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ se dice **no negativo** si todas sus componentes son, como su propio nombre indica, no negativas. Es decir, $x_i \geq 0$ para todo $i \in \{1, \dots, n\}$. Si un vector x es no negativo escribiremos $x \geq 0$.

Al hilo de esta definición, vamos a generalizar ese concepto que lleva presente en nuestras vidas desde que aquel docente de primaria nos dibujó dos rectas perpendiculares en la pizarra, los cuadrantes.

Definición 1.1.2 (2^n -ante positivo). Definimos el 2^n -**ante positivo** de \mathbb{R}^n como el conjunto de los vectores no negativos de \mathbb{R}^n .

Se deja como ejercicio al lector justificar el nombre de 2^n -ante.

Dicho esto, ya podemos formular el problema de la programación lineal.

Problema 1.1 (Formulación general). La disciplina de la **programación lineal** estudia procedimientos (implementables en ordenador) para calcular los extremos absolutos de una aplicación lineal $f : \mathbb{R}^n \rightarrow \mathbb{R}$ cuando restringimos su dominio a una variedad afín de \mathbb{R}^n cortada con el 2^n -ante positivo.



Figura 1.1: Ilustración del problema de programación lineal.

En el caso de la ilustración 1.1 suponemos que f es una aplicación lineal definida sobre todo el plano a la cual restringimos el dominio a la variedad afín \mathcal{L} cortada con el cuadrante positivo (línea azul). Nuestro deber es encontrar los puntos (si los hay) sobre la línea azul en los cuales f alcanza un máximo o un mínimo absoluto.

En esta sección nos dedicaremos simplemente a formular de diversas maneras el problema de la programación lineal, observando que todas son equivalentes entre sí. Comencemos pues, sin más dilación nuestro viaje por lo desconocido, un viaje del que probablemente no regresaremos.

A continuación vemos una sencilla caracterización de las aplicaciones lineales $f : \mathbb{R}^n \rightarrow \mathbb{R}$ a partir de su expresión analítica.

Observación 1.1.1 (Expresión analítica). Consideremos una aplicación lineal $f : \mathbb{R}^n \rightarrow \mathbb{R}$.

Por ser f una aplicación lineal, tendrá una única matriz asociada C . Como salta a la vista, C será una matriz fila con n columnas.

Entonces, dado un vector $x \in \mathbb{R}^n$ se tiene que

$$f(x) = f(x_1, \dots, x_n) = CX = (c_1 \quad \cdots \quad c_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = c_1x_1 + \cdots + c_nx_n \quad (1.1)$$

De esta forma se concluye que toda aplicación lineal $f : \mathbb{R}^n \rightarrow \mathbb{R}$ tiene una expresión analítica de la forma de la ecuación (1.1). Recíprocamente, toda función $g : \mathbb{R}^n \rightarrow \mathbb{R}$ con una expresión analítica como la de (1.1) es una ecuación lineal, para demostrarlo, basta con leer como los árabes la ecuación (1.1). \diamond

Antes de continuar es importante hacer un inciso acerca de la notación.

Observación 1.1.2 (Notación matricial). Siempre que escribamos matrices fila o columna lo haremos con letras minúsculas. Asimismo, siempre que estemos escribiendo una matriz fila añadiremos el símbolo de trasposición, para dejarlo claro implícitamente. De esta manera, a las funciones lineales las denotaremos por

$$f(x_1, \dots, x_n) = c^t x$$

donde c^t es la matriz asociada a f y x es el vector columna con las coordenadas de x respecto de la base canónica (nótese el pequeño abuso de notación). \diamond

1.1.1. Forma estándar y formas canónicas

Sea $f(x_1, \dots, x_n) = c_1x_1 + \cdots + c_nx_n = c^t x$ una función lineal a la que a partir de ahora llamaremos **función objetivo**. A la matriz c se la denomina **vector de coeficientes de la función objetivo**, mientras que x recibe el nombre de **vector de variables de decisión**.

Consideremos asimismo la variedad afín de \mathbb{R}^n dada por el conjunto de soluciones al sistema no necesariamente homogéneo $Ax = b$. Escrito de forma desarrollada

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases} \quad \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

A la matriz $A \in \mathfrak{M}_{m \times n}(\mathbb{R})$ usualmente la llamaremos **matriz de coeficientes de las restricciones**, además, a cada una de las ecuaciones del sistema las llamaremos **restricciones**. Asimismo el vector b será conocido como **vector de términos independientes**.

Expongamos ahora la llamada “*formulación estándar*” del problema de la programación lineal.

Problema 1.2 (Forma estándar). El problema de programación lineal en **forma estándar** consiste en hallar los vectores $x \geq 0$ tales que sean solución de $Ax = b$ y además sean mínimos absolutos de la función f . Usualmente se escribe de forma sintética (aunque no lo usaremos mucho)

$$\begin{array}{ll} \text{mín } c^t x \\ \text{Sujeto a:} & Ax = b, \quad x \geq 0 \end{array}$$

Nótese que la formulación del problema en forma estándar es exactamente la misma que la formulación general 1.1 con la única diferencia de que ya no buscamos extremos absolutos en general, sino únicamente mínimos.

La formulación estándar del problema de programación lineal puede parecer muy inflexible, en el sentido de que, a simple vista, no parece ser demasiado útil para modelizar problemas reales. Es esta aparente rigidez la que lleva a plantearse otras formulaciones más laxas del problema de programación lineal. Las llamadas “*formulaciones canónicas*” que exponemos a continuación.

Problema 1.3 (Primera forma canónica). Esta nueva formulación a la que llamamos **primera forma canónica** plantea el problema de hallar los vectores $x \geq 0$ tales que satisfagan la ecuación $Ax \leq b$ y que además sean máximos absolutos de la función objetivo f . Escrito de forma compacta

$$\begin{array}{ll} \text{máx } c^t x \\ \text{Sujeto a:} & Ax \leq b, \quad x \geq 0 \end{array}$$

Problema 1.4 (Segunda forma canónica). El problema de programación lineal en **segunda forma canónica** consiste en encontrar los vectores $x \geq 0$ tales que satisfagan la ecuación $Ax \geq b$ y que además sean mínimos absolutos de la función objetivo f . En resumen

$$\begin{array}{ll} \text{mín } c^t x \\ \text{Sujeto a:} & Ax \geq b, \quad x \geq 0 \end{array}$$

Una vez vistas todas las formulaciones hagamos una pequeña reflexión. Sería la apoteosis del tedio tener que demostrar resultados teóricos para cada una de las formulaciones. No obstante, no hay que dejar nunca de confiar en los dioses olímpicos, pues a veces son benévolo con nosotros. Esta benevolencia consiste en que todas las formulaciones son equivalentes en el sentido que veremos a continuación.

Maximización y minimización

Si se nos plantea el problema de hallar el máximo absoluto de una función objetivo f podemos transformar este problema en uno equivalente de forma que el objetivo de el nuevo problema sea hallar el mínimo de otra función objetivo g .

La utilidad de este hecho es que podemos transformar un problema de programación lineal en primera forma canónica (que surgen de forma natural a la hora de modelizar problemas reales) en un problema más parecido a la formulación estándar.

Lema 1.1.1 (Maximización–minimización). *Sea A un conjunto arbitrario y una función $f : A \rightarrow \mathbb{R}$ que alcanza el máximo y el mínimo. Entonces se cumple*

$$\text{máx } f = -\text{mín}(-f)$$

Demostración. Sea ξ el máximo de f , es decir, el único número de $f(A)$ que verifica que $f(x) \leq \xi$ para todo $x \in A$. Consideremos ahora la función $g \equiv -f$, que, por las propiedades de los números reales alcanza el mínimo, al que llamaremos η .

Evidentemente η verifica que $g(x) \geq \eta$ para todo $x \in A$, o equivalentemente $-g(x) \leq -\eta$ para todo $x \in A$. Pero $-g(x) = f(x)$, luego $f(x) \leq -\eta$ para todo $x \in A$. Por unicidad del máximo $-\eta = \xi$, como queríamos demostrar. ■

Observación 1.1.3 (Aplicación). Si tenemos un problema de programación lineal formulado en términos de maximización bastará con cambiar la función objetivo f por la función objetivo $-f$ para obtener un problema equivalente en términos de minimización cuyas soluciones deben ser cambiadas de signo. ◇

Igualdades y desigualdades

Una vez visto el apartado anterior, vamos a aprender a transformar restricciones de tipo desigualdad a restricciones de tipo igualdad, con lo cual ya podremos transformar cualquier problema planteado en alguna forma canónica a un problema equivalente planteado en forma estándar.

Lema 1.1.2 (Variables de holgura). *Sea una restricción de la forma*

$$a_{i1}x_1 + \cdots + a_{in}x_n \stackrel{(\geq)}{\leq} b_i \tag{1.2}$$

entonces la restricción

$$a_{i1}x_1 + \cdots + a_{in}x_n + x_i^h = b_i \quad (1.3)$$

es equivalente, en el sentido de que hay una biyección “natural” entre los conjuntos de vectores **no negativos** que verifican cada una de las restricciones.

A la variable x_i^h añadida en la restricción (1.3) se la denomina **variable de holgura**.

Demostración. Sean los conjuntos \mathcal{S} y \mathcal{S}' de vectores que verifican las restricciones (1.2) y (1.3) respectivamente. Consideremos la aplicación

$$\varphi : \begin{array}{c} \mathcal{S} \rightarrow \mathcal{S}' \\ (x_1, \dots, x_n) \mapsto (x_1, \dots, x_n, z) \end{array}$$

donde $z = b_i - a_{i1}x_1 - \cdots - a_{in}x_n$. Veamos que es una biyección.

La inyectividad es clara por definición, si dos vectores compartieran imagen todas sus componentes coinciden.

En cuanto a la sobreyectividad, dado un vector $(x_1, \dots, x_n, x_{n+1})$ que verifica la restricción (1.3), es claro que el vector resultante de eliminar la última componente, (x_1, \dots, x_n) , cumple la restricción (1.2).

La demostración para el caso \geq es totalmente simétrica. ■

Veamos que utilidad tiene lo que acabamos de demostrar.

Observación 1.1.4 (Aplicación). Si a la hora de modelizar un problema nos aparecen restricciones de tipo \geq o \leq basta con sustituirlas por restricciones equivalentes añadiendo variables de holgura.

Además, debemos cambiar la función objetivo f , extendiéndola de la siguiente manera

$$\hat{f} : \begin{array}{c} \mathbb{R}^{n+m} \rightarrow \mathbb{R} \\ (x_1, \dots, x_n, x_{n+1}, \dots, x_{n+m}) \mapsto f(x_1, \dots, x_n) \end{array}$$

donde n es el número total de restricciones originales y m el número de restricciones de tipo \geq o \leq . De esta forma obtenemos un problema equivalente cuya solución será el vector solución del nuevo problema quitando las componentes correspondientes a variables de holgura (¡compruébeselo!). ◇

Con lo que sabemos hasta el momento podemos transformar cualquier problema escrito en cualquiera de las formas canónicas en un problema en forma estándar.

Positividad y negatividad

Otro caso que se presenta de manera usual es que al modelizar un problema concreto sea necesario permitir que algunas componentes de los vectores candidatos a ser solución sean negativas. Esta es una situación que no contempla la formulación estándar, sin embargo, con un simple truco podemos meter esta clase de problemas dentro de nuestro marco de actuación.

Observación 1.1.5 (Truco). Si tenemos un problema en el que la componente i -ésima de los vectores puede ser negativa tomamos todas las restricciones

$$a_{j1}x_1 + \cdots + a_{ji}x_i + \cdots + a_{jn}x_n = b_j$$

que involucren a la componente x_i y las sustituimos por las restricciones

$$a_{j1}x_1 + \cdots + a_{ji}(x_i^+ - x_i^-) + \cdots + a_{jn}x_n = b_j$$

con x_i^+ y x_i^- no negativos.

Es claro que si un vector $(x_1, \dots, x_i, \dots, x_n)$ (donde x_i es negativo) verifica las restricciones originales, entonces el vector $(x_1, \dots, 0, -x_i, \dots, x_n)$ (por ejemplo) verifica las restricciones nuevas. Análogamente si x_i es positivo. Asimismo, si el vector $(x_1, \dots, x_i^+, x_i^-, \dots, x_n)$ cumple las restricciones nuevas, el vector $(x_1, \dots, x_i^+ - x_i^-, \dots, x_n)$ cumplirá las viejas, pudiendo concluir así que las restricciones viejas y nuevas son equivalentes en un sentido un poco más laxo que en el apartado anterior, ya que únicamente se tiene la sobreyectividad (pero esto es irrelevante).

De esta forma, si cambiamos la función objetivo f por la función

$$\bar{f}: \mathbb{R}^{n+1} \rightarrow \mathbb{R} \\ (x_1, \dots, x_i^+, x_i^-, \dots, x_n) \mapsto f(x_1, \dots, x_i^+ - x_i^-, \dots, x_n)$$

obtenemos un problema equivalente cuya solución será el vector $(x_1, \dots, x_i^+ - x_i^-, \dots, x_n)$ donde $(x_1, \dots, x_i^+, x_i^-, \dots, x_n)$ es el vector solución del problema con las restricciones nuevas. \diamond

Hechas todas estas disquisiciones iniciales cabe mencionar que a la hora de modelizar problemas, siempre que tanto la función objetivo como las restricciones sean lineales, podremos realizar transformaciones (las vistas en esta sección) para que el problema quede planteado en forma estándar, pudiendo ser así resuelto mediante el uso de toda la artillería teórica que veremos a continuación.

1.1.2. Terminología y casuística

Dado un problema de programación lineal en forma estándar que tiene a A como matriz de coeficientes de las restricciones y a b como vector de términos independientes. Exponemos las siguientes definiciones.

Definición 1.1.3 (Soluciones). Llamamos **solución** del problema P a cualquier vector $x \in \mathbb{R}^n$ (no necesariamente no negativo) que verifique $Ax = b$.

Sacando un poco de punta al asunto, diremos que si $x \in \mathbb{R}^n$ es solución del problema y además $x \geq 0$ entonces x es una **solución factible**.

Llamaremos **región factible** al conjunto de todas las soluciones factibles del problema. Normalmente la denotaremos con la letra \mathcal{S} .

Una vez resuelto el problema se pueden dar las siguientes situaciones

- La región factible posee al menos un punto en el cual la función objetivo alcanza el mínimo. En esta situación diremos que el problema posee **solución óptima**.
- La región factible del problema es el conjunto vacío. En este caso se dirá que el problema es **infactible**.
- La función objetivo alcanza valores tan bajos como queramos en la región factible. En este caso se dice que el problema posee una solución **no acotada**.

Teniendo en cuenta estas pequeñas observaciones podemos empezar a estudiar cómo son las regiones factibles, lo cual nos será muy útil para resolver el problema de programación lineal.

1.2. Programación lineal y conjuntos convexos

En esta sección demostraremos que las regiones factibles de un problema de programación lineal en forma estándar son siempre conjuntos convexos.

Asimismo, estudiaremos en profundidad las propiedades de estos conjuntos, explotándolas para obtener un procedimiento robusto con el cual resolver estos problemas.

1.2.1. Conjuntos convexos. Definición y propiedades

Comenzamos definiendo el concepto de conjunto convexo.

Definición 1.2.1 (Convexos). Sea \mathcal{S} un subconjunto de \mathbb{R}^n , se dice **convexo** si para cada par de puntos $x, y \in \mathcal{S}$ el segmento que tiene por extremos a x e y está contenido en \mathcal{S} . Es decir

$$\lambda x + (1 - \lambda)y \in \mathcal{S}$$

para todo $\lambda \in [0, 1]$. Nótese que el conjunto vacío se considera convexo.

La región factible de un problema de programación lineal en forma estándar es un conjunto convexo.

Lema 1.2.1 (Región convexa). *El conjunto de los vectores no negativos $x \in \mathbb{R}^n$ que verifican la condición $Ax = b$ para cierta matriz A y cierto vector b es un conjunto convexo.*

Demostración. Si la región factible \mathcal{S} es el conjunto vacío ya hemos terminado. En otro caso, basta considerar dos vectores x e y de la región factible y ver que $\alpha x + (1 - \alpha)y \in \mathcal{S}$ para todo $\alpha \in [0, 1]$, pero esto es inmediato ya que

$$A(\alpha x + (1 - \alpha)y) = \alpha Ax + (1 - \alpha)Ay = \alpha b + (1 - \alpha)b = b \quad \blacksquare$$

Una cuestión importante a destacar es que la intersección arbitraria de conjuntos convexos es convexo, tal y como muestra el siguiente lema.

Lema 1.2.2 (Intersección). *Sea \mathcal{F} una familia de conjuntos convexos. La intersección de la familia es un conjunto convexo.*

Demostración. Sean x e y dos puntos de la intersección de la familia, luego x e y están en todos los conjuntos de la familia. Como estos conjuntos son convexos, el segmento que une x e y estará contenido en todos los conjuntos de la familia simultáneamente, luego también en la intersección de la familia. \blacksquare

Un concepto por el que merece la pena pasar es el de “envoltura convexa” de un conjunto \mathcal{S} de \mathbb{R}^n , que viene a ser el menor conjunto convexo que contiene a \mathcal{S} .

Esta idea es especialmente recurrente en las matemáticas, por ejemplo, en álgebra lineal teníamos el subespacio generado por un subconjunto, en teoría de grupos el subgrupo generado por un subconjunto, en topología tenemos la adherencia de un conjunto, ... La lista es interminable.

Observación 1.2.1 (Justificación). Es fácil darse cuenta de que, como la intersección arbitraria de convexos es convexa, el menor conjunto convexo que contiene a uno dado (al que llamaremos \mathcal{S}) puede ser construido de la siguiente manera

$$\text{Conv}(\mathcal{S}) := \bigcap_{\mathcal{F} \supset \mathcal{S}} \mathcal{F}$$

En efecto, es un conjunto que contiene a \mathcal{S} , ya que todos los conjuntos de la familia a intersecar contienen a \mathcal{S} , además, es el menor de ellos, ya que, de haber uno más pequeño, pertenecería a la familia que se está intersecando, lo cual es absurdo (¡compruébese!).

De esta forma queda justificada la existencia del menor conjunto convexo que contiene a \mathcal{S} , no obstante, la construcción que hemos realizado nos da escasa información acerca de los elementos de la envoltura convexa. \diamond

Observación 1.2.2 (Curiosidad). A modo de curiosidad comentamos que el problema de calcular la envoltura convexa de un conjunto finito de n puntos en el plano o el espacio es uno de los problemas centrales de la **geometría computacional** por sus múltiples aplicaciones. \diamond

Veamos a continuación una caracterización útil de la envoltura convexa de un conjunto, para la cual necesitamos introducir una definición.

Definición 1.2.2 (Combinaciones convexas). Sea $\{z_1, \dots, z_r\}$ un conjunto finito de puntos. Diremos que z es **combinación lineal convexa** de dichos puntos si

$$z = \lambda_1 z_1 + \dots + \lambda_r z_r$$

donde $\lambda_i > 0$ para $i \in \{1, \dots, r\}$ y se cumple que $\sum_{i=1}^r \lambda_i = 1$.

Que no asuste la longitud de la demostración de la proposición pues realmente es muy sencilla.

Proposición 1.2.3 (Caracterización). *Dado un conjunto \mathcal{S} , su envoltura convexa es el conjunto de puntos que son una combinación lineal convexa de puntos de \mathcal{S} .*

Demostración. Llamemos $\mathcal{H}(\mathcal{S})$ al conjunto de puntos que son combinación lineal convexa de puntos de \mathcal{S} . Veamos en primer lugar que $\mathcal{H}(\mathcal{S})$ contiene a \mathcal{S} . Esto es claro, ya que todo punto x de \mathcal{S} es combinación lineal convexa de puntos de \mathcal{S} , en efecto, $x = 1 \cdot x$.

Veamos ahora que $\mathcal{H}(\mathcal{S})$ es convexo. Esto lo haremos simplemente usando la definición. Sean x e y dos puntos de $\mathcal{H}(\mathcal{S})$, veamos que $\alpha x + (1 - \alpha)y \in \mathcal{H}(\mathcal{S})$ para todo $\alpha \in [0, 1]$.

Por definición de $\mathcal{H}(\mathcal{S})$

$$x = \sum_{i=1}^r \lambda_i x_i \quad y = \sum_{j=1}^s \mu_j y_j$$

por tanto, si sustituimos esto en $\alpha x + (1 - \alpha)y$ obtenemos

$$\alpha x + (1 - \alpha)y = \sum_{i=1}^r (\alpha \lambda_i) x_i + \sum_{j=1}^s [(1 - \alpha) \mu_j] y_j$$

Haciendo los siguientes cambios de nombre

$$\xi_k := \begin{cases} \alpha \lambda_k & \text{si } k \in \{1, \dots, r\} \\ (1 - \alpha) \mu_{k-r} & \text{si } k \in \{r+1, \dots, r+s\} \end{cases} \quad z_k := \begin{cases} x_k & \text{si } k \in \{1, \dots, r\} \\ y_{k-r} & \text{si } k \in \{r+1, \dots, r+s\} \end{cases}$$

nos queda que

$$\alpha x + (1 - \alpha)y = \sum_{k=1}^{r+s} \xi_k z_k$$

luego por definición de $\mathcal{H}(\mathcal{S})$ solo queda comprobar que $\sum_{k=1}^{r+s} \xi_k = 1$, pero esto es evidente ya que

$$\sum_{k=1}^{r+s} \xi_k = \alpha \sum_{i=1}^r \lambda_i + (1 - \alpha) \sum_{j=1}^s \mu_j = \alpha + (1 - \alpha) = 1$$

Queda por ver pues que $\mathcal{H}(\mathcal{S})$ es el menor conjunto convexo que contiene a \mathcal{S} . Para ello, consideramos \mathcal{C} un conjunto convexo arbitrario que contiene a \mathcal{S} . Demostremos que $\mathcal{H}(\mathcal{S}) \subset \mathcal{C}$ por inducción sobre la “longitud” de la combinación lineal convexa de sus elementos.

El caso base es evidente, ya que los elementos de $\mathcal{H}(\mathcal{S})$ que son combinaciones lineales convexas de longitud unitaria son también elementos de \mathcal{S} y por tanto de \mathcal{C} . Supongamos cierto que los elementos de $\mathcal{H}(\mathcal{S})$ que son combinaciones lineales convexas de longitud n viven también en \mathcal{C} , y demostremos que lo mismo sucederá para los elementos que sean combinaciones de longitud $n+1$.

Consideremos $z \in \mathcal{H}(\mathcal{S})$ un elemento que es combinación de longitud $n+1$, es decir

$$z = \sum_{i=1}^{n+1} \lambda_i z_i \quad \text{con} \quad \sum_{i=1}^{n+1} \lambda_i = 1$$

Vamos a apañar la expresión de z para poder aplicar la hipótesis de inducción

$$z = \sum_{i=1}^n \lambda_i z_i + \lambda_{n+1} z_{n+1}$$

Es claro que $z_{n+1} \in \mathcal{C}$ por ser una combinación de longitud uno. Sin embargo, el primer sumando no es una combinación lineal convexa ya que $\sum_{i=1}^n \lambda_i = 1 - \lambda_{n+1} \neq 1$. Para arreglar esto vamos a multiplicar y a dividir el primer sumando por $1 - \lambda_{n+1}$

$$z = (1 - \lambda_{n+1}) \sum_{i=1}^n \frac{\lambda_i}{1 - \lambda_{n+1}} z_i + \lambda_{n+1} z_{n+1}$$

De esta forma el primer sumando (excluyendo el factor que lo multiplica) es una combinación lineal convexa de longitud n , ya que $\frac{1}{1 - \lambda_{n+1}} \sum_{i=1}^n \lambda_i = \frac{1 - \lambda_{n+1}}{1 - \lambda_{n+1}} = 1$.

Por hipótesis de inducción, tanto el primer sumando como el segundo (excluyendo los factores que los multiplican) son elementos de \mathcal{C} , y como \mathcal{C} es convexo $z \in \mathcal{C}$, como queríamos. ■

1.2.2. Puntos extremos

De ahora en adelante será de utilidad tener en mente un polígono regular cuando se nos hable de conjuntos convexos, pues nos dará una idea bastante intuitiva.

Para estrenar esta nueva buena costumbre se presenta la siguiente definición.

Definición 1.2.3 (Punto extremo). Sea \mathcal{C} un conjunto convexo. Diremos que $\bar{x} \in \mathcal{C}$ es un **punto extremo** de \mathcal{C} si se verifica que para cualquier par de puntos $x_1, x_2 \in \mathcal{C}$, \bar{x} **no** está en el segmento generado por dichos puntos (excluyendo los extremos). Es decir

$$\bar{x} = \lambda x_1 + (1 - \lambda)x_2 \iff \begin{cases} \lambda \in \{0, 1\} \\ x_1 = x_2 \end{cases}$$

La ilustración 1.2.2 marga en negro los puntos extremos de un polígono regular, para coger una idea intuitiva. A lo largo de este capítulo veremos la importancia que tienen estos puntos en la resolución del problema de programación lineal. El teorema 1.2.4 nos da una caracterización de los

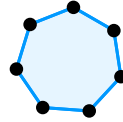


Figura 1.2: Ilustración de los puntos extremos de un heptágono regular.

puntos extremos de una región factible que, aunque muy potente, resulta poco intuitiva. Antes de enunciarlo recordemos un viejo resultado de álgebra lineal.

Observación 1.2.3 (Rango y submatrices). Es conocido desde tiempos inmemoriales que $A \in \mathfrak{M}_{m \times n}(\mathbb{K})$ con $m \leq n$ y $\text{rg}(A) = m$ entonces habrá al menos una submatriz cuadrada $B \in \mathfrak{M}_m(\mathbb{K})$ de A tal que $\text{rg}(B) = m$. De esta forma, reordenando las columnas de A si fuera necesario, podemos hacer la siguiente división por bloques de A .

$$A = (B|N)$$

Nótese que la igualdad se da solo cuando ya hemos reordenado A . ◇

A partir de ahora trabajaremos con matrices en las condiciones de la observación 1.2.3 a no ser que se especifique lo contrario. Nótese que esto no supone ninguna pérdida de generalidad en lo que se refiere al estudio de los problemas de programación lineal ya que si se nos presentara una matriz con más filas que columnas habrá restricciones redundantes.

Definición 1.2.4 (Bases). Dado un problema en forma estándar que tiene a A por matriz de coeficientes de las restricciones, llamaremos **base** de A a toda submatriz suya cuadrada y de rango máximo.

Teorema 1.2.4 (Caracterización). Dada una región factible \mathcal{S} , \bar{x} es un punto extremo de \mathcal{S} si y solo si hay alguna reordenación de las columnas de A de forma que

$$A = (B|N) \quad \bar{x} = \begin{pmatrix} \bar{x}_B \\ \bar{x}_N \end{pmatrix} \text{ con } \bar{x}_N = 0$$

donde B es una base de A y \bar{x}_B y \bar{x}_N son los “trozos” de \bar{x} acordes con la descomposición en bloques de A , es decir

$$A\bar{x} = (B|N) \begin{pmatrix} \bar{x}_B \\ \bar{x}_N \end{pmatrix} = B\bar{x}_B + N\bar{x}_N$$

Demostración. Veamos las dos implicaciones

⇐ Supongamos que \bar{x} no es punto extremo, entonces habrá dos puntos distintos $x_1, x_2 \in \mathcal{S}$ tales que $\bar{x} = \lambda x_1 + (1 - \lambda)x_2$ con $\lambda \in (0, 1)$. Desplegando la descomposición por bloques tendríamos que

$$(\bar{x}_B | \bar{x}_N)^t = \lambda(x_B^1 | x_N^1)^t + (1 - \lambda)(x_B^2 | x_N^2)^t$$

Como por hipótesis $\bar{x}_N = 0$ y tanto λ como $1 - \lambda$ son estrictamente positivos se tiene que $x_N^1 = x_N^2 = 0$, ya que x_1 y x_2 deben ser no negativos.

Como $x_1, x_2 \in \mathcal{S}$, entonces $Ax_i = b$ para $i \in \{1, 2\}$, por tanto

$$(B|N)(x_B^i | x_N^i)^t = (B|N)(x_B^i | 0)^t = Bx_B^i = b \implies x_B^i = B^{-1}b$$

Con lo que se concluye que $x_1 = x_2$, lo cual es absurdo.

⇒ Sea \bar{x} un punto extremo. Reordenemos el vector de forma que todas sus componentes nulas queden al final, reordenando solidariamente las columnas de la matriz A .

Podemos suponer pues que el vector \bar{x} tiene $p \neq 0$ componentes no nulas (si $p = 0$ el resultado es evidente). Vamos a demostrar que las p primeras columnas de A (ya reordenada) son linealmente independientes.

A partir de ahora denotaremos a la matriz A por sus columnas, es decir $A = (a_1 \cdots a_n)$ donde a_i representa la i -ésima columna de A .

Procedamos por reducción al absurdo, es decir, supongamos que hay ciertos números reales no todos nulos λ_i tales que $\sum_{i=1}^p \lambda_i a_i = 0$.

Definimos el vector $\lambda := (\lambda_1, \dots, \lambda_p, 0, \dots, 0) \in \mathbb{R}^n$. A continuación consideramos los vectores **no negativos** $x_1 := \bar{x} + \alpha \lambda$ y $x_2 := \bar{x} - \alpha \lambda$ tomando un α lo suficientemente pequeño para que ambos vectores sean precisamente no negativos, lo cual siempre puede hacerse al ser \bar{x} no negativo.

Los vectores x_1 y x_2 tienen la particularidad de que con ellos podemos obtener \bar{x} como combinación lineal convexa. En efecto $\bar{x} = \frac{1}{2}x_1 + \frac{1}{2}x_2$. Si resultara que x_1 y x_2 son puntos de la región factible estaríamos entrando en contradicción con la “extremalidad” de \bar{x} , y eso es exactamente lo que vamos a hacer. En efecto, con $i \in \{1, 2\}$ tenemos

$$Ax_i = A\bar{x} \pm \alpha A\lambda \stackrel{!}{=} A\bar{x} = b \implies x_i \in \mathcal{S}$$

donde la igualdad de la exclamación se debe a la composición de λ y la dependencia lineal de las p primeras columnas de A (por hipótesis). Con esto llegamos a un absurdo y concluimos que las p primeras columnas de A son linealmente independientes.

Además, de paso podemos concluir que $p \leq m$, ya que m es el rango máximo de A . Si además $p = m$ ya tenemos una base $B = (a_1 \cdots a_p)$ que cumple lo que queremos (¡compruébes!). En el caso de que $p < m$ la observación 1.2.3 nos dice que siempre podemos escoger columnas adicionales de A para completar una base y, tras una nueva reordenación de A si fuera necesario tendríamos lo que necesitamos. ■

Corolario 1.2.5 (Parte básica). *Dado un punto extremo $\bar{x} = (\bar{x}_B | 0)^t$ de una región factible \mathcal{S} se cumple que $\bar{x}_B = B^{-1}b$.*

Demostración. Basta con echar las cuentas. Como $\bar{x} \in \mathcal{S}$ es claro que $A\bar{x} = b$, luego

$$A\bar{x} = (B|N)(\bar{x}_B | 0)^t = B\bar{x}_B = b$$

y despejando se tiene que $\bar{x}_B = B^{-1}b$. ■

Al hilo de este teorema surge la siguiente definición.

Definición 1.2.5 (Soluciones básicas). A los puntos extremos de una región factible se les suele llamar **soluciones básicas factibles**. Esto es debido a que, por el teorema 1.2.4, están “asociados” a una base de la matriz asociada a la región factible del problema.

Observación 1.2.4 (Procedimiento constructivo). Nótese que si nos dan un punto extremo de una región factible y nos piden hallar la base asociada a dicho punto extremo no tenemos más que seguir la demostración de la implicación a la derecha del teorema 1.2.4 para encontrarla, es decir, tomar las columnas de A asociadas a las componentes no nulas del punto extremo y, en su caso, completar la base.

Esto nos lleva a deducir de forma evidente que cada punto extremo puede estar asociado a más de una base (salvo reordenaciones). \diamond

Una de las razones por la cual es estudio de los puntos extremos es importante es porque toda región factible de un problema de programación lineal en forma estándar tiene al menos un punto extremo.

Observación 1.2.5 (Finitud). Una región factible tiene un número finito de puntos extremos, esto es debido a que A tiene más bases que puntos extremos, ya que cada base B está asociada a un único candidato a punto extremo, el punto que tiene a $B^{-1}b$ como parte básica. Además, cada matriz tiene a lo sumo $\binom{n}{m}$ bases. \diamond

Teorema 1.2.6 (Existencia). *Toda región factible no vacía posee al menos un punto extremo.*

Demostración. Sea \bar{x} un punto arbitrario de la región factible. Reordenemos sus componentes de forma que las componentes nulas queden al final, asimismo reordenemos A de forma coherente con la reordenación de \bar{x} . Supongamos que \bar{x} posee p componentes no nulas (las p primeras tras la reordenación).

Ahora consideramos las columnas de A correspondientes a las p componentes no nulas de \bar{x} . Si estas columnas son linealmente independientes, por el teorema 1.2.4 habríamos terminado.

Si, por el contrario, las columnas (que serán las p primeras tras la reordenación) son linealmente dependientes, entonces existirán ciertos números reales λ_i no todos nulos tales que $\sum_{i=1}^p \lambda_i a_i = 0$, donde a_i representa a la i -ésima columna de A .

Lo que vamos a hacer es cambiar ligeramente el punto \bar{x} de forma que siga siendo de la región factible y conserve un subconjunto estrictamente menor de las componentes no nulas de \bar{x} .

En primer lugar cabe mencionar que siempre podemos suponer que existe al menos un $r \in \{1, \dots, p\}$ de forma que $\lambda_r > 0$, ya que en caso de que todos los números de la combinación lineal fueran negativos podemos considerar la combinación lineal alternativa $\sum_{i=1}^p -\lambda_i a_i = 0$, en la cual todos los escalares son positivos.

Definimos ahora el vector $\lambda := (\lambda_1, \dots, \lambda_p, 0, \dots, 0) \in \mathbb{R}^n$ y consideramos el punto $\bar{x}' := \bar{x} - \alpha \lambda$ para cierto $\alpha > 0$ que definiremos más adelante.

El punto \bar{x}' verifica la condición $Ax = b$, en efecto

$$A\bar{x}' = A\bar{x} - \alpha A\lambda = A\bar{x} = b$$

Luego únicamente queda elegir un α adecuado para que $\bar{x}' \geq 0$. Para ello elegiremos

$$\alpha := \min \left\{ \frac{\bar{x}_j}{\lambda_j} : \lambda_j > 0 \right\} = \frac{\bar{x}_l}{\lambda_l}$$

Veamos que \bar{x}' es no negativo. Para ello distinguiremos casos componente a componente.

- Si $\lambda_j > 0$ tenemos que $\bar{x}'_j = \bar{x}_j - \frac{\bar{x}_l}{\lambda_l} \lambda_j = \frac{\bar{x}_j}{\lambda_j} \lambda_j - \frac{\bar{x}_l}{\lambda_l} \lambda_j = \lambda_j \left(\frac{\bar{x}_j}{\lambda_j} - \frac{\bar{x}_l}{\lambda_l} \right)$ y este número es no negativo por definición de α .
- Si $\lambda_j \leq 0$ evidentemente $\bar{x}'_j = \bar{x}_j - \alpha \lambda_j$ es no negativo por ser la suma de números no negativos.

Cabe destacar que si $\bar{x}_j = 0$ entonces $\bar{x}'_j = 0$, además $\bar{x}_l > 0$ mientras que $\bar{x}'_l = 0$, luego \bar{x}' tiene a lo sumo $p-1$ componentes no nulas, con lo que, al volver al principio de la demostración el conjunto de columnas que queremos que sean linealmente independientes se ha visto reducido en una columna como poco. Por tanto, tras un número finito de pasos encontraremos un punto extremo. Esto proporciona un algoritmo para encontrar un punto extremo de una región factible. ■

Culminamos este teorema con una pequeña observación.

Observación 1.2.6 (Extremalidad del 0). Si una región factible contiene al vector 0, este es siempre un punto extremo, basta con aplicar el teorema 1.2.6 para comprobarlo (el conjunto vacío es linealmente independiente). \diamond

El siguiente teorema termina de mostrar la brutal importancia de los puntos extremos de las regiones factibles en los problemas de programación lineal, y es que, si un problema tiene solución óptima, esta se alcanza en un punto extremo o solución básica factible.

Teorema 1.2.7 (Optimalidad). *Si un problema de programación lineal en forma estándar tiene solución óptima, esta se alcanza en un punto extremo.*

Demostración. Procedemos de análogamente a como hicimos en la demostración del teorema 1.2.6.

Consideremos \bar{x} un punto donde se alcanza la solución óptima al problema en cuestión. Reordenemos las coordenadas de \bar{x} dejando las componentes nulas al final, reordenando también A solidariamente.

Suponiendo que \bar{x} tiene p componentes no nulas y que las columnas de A correspondientes a dichas columnas son linealmente dependientes (en caso contrario hemos terminado), entonces habrá ciertos números reales no todos nulos (de hecho podemos suponer que alguno será positivo) tales que $\sum_{i=1}^n \lambda_i a_i = 0$. Definimos el vector $\lambda \in \mathbb{R}^n$ de la manera usual y consideramos los puntos

$$x_1 = \bar{x} + \alpha_1 \lambda \quad x_2 = \bar{x} - \alpha_2 \lambda$$

siendo α_1 y α_2 coeficientes positivos lo suficientemente pequeños para que x_1 y x_2 sean no negativos.

Es evidente que tanto x_1 como x_2 cumplen la condición $Ax = b$, esto se debe a la dependencia lineal de las p primeras columnas y a la definición de λ , como ya hemos hecho en otras ocasiones.

Evaluemos los puntos x_1 y x_2 por la función objetivo

$$\begin{aligned} c^t x_1 &= c^t \bar{x} + \alpha_1 c^t \lambda \\ c^t x_2 &= c^t \bar{x} - \alpha_2 c^t \lambda \end{aligned}$$

Por la optimalidad de \bar{x} se tiene que

$$c^t \bar{x} \leq c^t \bar{x} + \alpha_1 c^t \lambda \quad (1.4)$$

$$c^t \bar{x} \leq c^t \bar{x} - \alpha_2 c^t \lambda \quad (1.5)$$

De la ecuación (1.4) se deduce que $c^t \lambda \geq 0$ y de la ecuación (1.5) se desprende que $c^t \lambda \leq 0$, luego $c^t \lambda = 0$, por tanto tanto x_1 como x_2 son puntos en los que también se alcanza la solución óptima. Centrémonos en x_2 . Si tomamos α_2 como

$$\alpha_2 := \min \left\{ \frac{\bar{x}_j}{\lambda_j} : \lambda_j > 0 \right\}$$

sabemos que x_2 sigue perteneciendo a la región factible (como demostramos en el teorema 1.2.6) luego es una elección válida. Además, por las mismas razones que en el teorema 1.2.6, el vector x_2 tiene a lo sumo $p - 1$ componentes positivas, luego, tras un número finito de pasos encontraremos un punto extremo en el cual se alcance la solución óptima. \blacksquare

1.2.3. Direcciones extremas

El siguiente es el último concepto relativo a convexidad que veremos, se trata del concepto de “dirección”. En palabras llanas podríamos decir que una “dirección” de un conjunto convexo es una flecha tal que si partimos de cualquier punto del conjunto y la seguimos nunca saldremos del conjunto. la siguiente definición formaliza este concepto.

Definición 1.2.6 (Dirección). Dado un conjunto convexo \mathcal{C} y un vector $d \in \mathbb{R}^n$ diremos que d es una **dirección** de \mathcal{C} si para todo punto $x \in \mathcal{C}$ se verifica que

$$x + \lambda d \in \mathcal{C}$$

para todo λ no negativo. Asimismo, dadas dos direcciones d_1 y d_2 diremos que son **equivalentes** si son vectores proporcionales.

Observación 1.2.7 (No acotación). Nótese que si un conjunto convexo es acotado no puede tener direcciones (compruébese). Luego cuando hablemos de direcciones será recomendable tener en mente conjuntos como el que representa la ilustración 1.2.3 (que continúa infinitamente hacia arriba y hacia la derecha). \diamond

Veamos a continuación un concepto análogo al de los puntos extremos pero en el ámbito de las direcciones.

Definición 1.2.7 (Dirección extrema). Sea d una dirección. Diremos que d es una **dirección extrema** si no puede escribirse como combinación lineal positiva de dos direcciones no equivalentes. De forma sintética

$$d = \lambda d_1 + \mu d_2 \text{ con } \lambda, \mu > 0 \iff d_1 \sim d_2$$

donde \sim denota la equivalencia de direcciones.



Figura 1.3: Ilustración de una dirección y una dirección extrema

Veamos una caracterización muy sencilla de las direcciones extremas para regiones factibles.

Lema 1.2.8 (Caracterización débil). *Dada una región factible \mathcal{S} , d es una dirección de \mathcal{S} si y solo si d es un vector no negativo y no nulo que verifica que $Ad = 0$.*

Demostración. Veamos ambas implicaciones.

\implies Sea d una dirección, por tanto, dado $x \in \mathcal{S}$ y $\lambda \geq 0$ se cumple que $x + \lambda d \in \mathcal{S}$, luego $A(x + \lambda d) = b$, por tanto

$$A(x + \lambda d) = Ax + \lambda Ad = b + \lambda Ad = b \iff Ad = 0$$

Queda comprobar que d es un vector no negativo, para ello supongamos que tiene alguna componente negativa, digamos la i -ésima. Entonces por hipótesis $x_j + \lambda d_j \geq 0$ para todo $\lambda \geq 0$, sin embargo, tomando un λ lo suficientemente grande llegamos a un absurdo.

\impliedby Es claro que $x + \lambda d$ es un vector no negativo para todo $x \in \mathcal{S}$ (por hipótesis). Comprobemos que está en la región factible. En efecto

$$A(x + \lambda d) = Ax + \lambda Ad = Ax = b \quad \blacksquare$$

El lema anterior nos abre las puertas a una caracterización de las direcciones extremas mucho más potente y muy parecida al teorema 1.2.4, es decir, muy poco intuitiva.

Teorema 1.2.9 (Caracterización fuerte). *Dada una región factible \mathcal{S} , d es una dirección extrema de \mathcal{S} si y solo si, tras una conveniente reordenación, hay una base B de A tal que*

$$A = (B|N) \quad d = \alpha \begin{pmatrix} -B^{-1}a_k \\ e_t \end{pmatrix}$$

con $k = m + t$, $\alpha > 0$, $B^{-1}a_k \leq 0$ y siendo e_t el t -ésimo vector de la base canónica.

Demostración. Veamos ambas implicaciones.

\impliedby

\implies \blacksquare

1.3. Teoremas fundamentales

Capítulo 2

Algoritmo del Símplex

- 2.1. Implementación del algoritmo del símplex
- 2.2. Método de las dos fases
- 2.3. Método de las penalizaciones
- 2.4. Regla lexicográfica
- 2.5. Regla de Bland

Capítulo 3

Dualidad

3.1. Planteamiento del problema dual

3.2. Relaciones de dualidad

Parte II

Programación entera

Capítulo 4

Modelización

- 4.1. Modelización con variables binarias
- 4.2. Restricciones disyuntivas
- 4.3. Problemas con costes fijos

Capítulo 5

Ramificación y acotación

Parte III

Programación no lineal

Parte IV

Anexos

Índice de términos

- 2^n -ante positivo, 3
- base, 10
- combinación
 - lineal convexa, 8
- conjunto
 - convexo, 7
- dirección, 13
 - equivalente, 13
 - extrema, 14
- forma
 - canónica primera, 5
 - canónica segunda, 5
 - estándar, 4
- función
 - objetivo, 4
- geometría
 - computacional, 8
- matriz
 - de coeficientes de las restricciones, 4
- problema
 - infactible, 7
- programación
 - lineal, 3
- punto
 - extremo, 10
- región
 - factible, 7
- restricciones, 4
- solución, 7
 - óptima, 7
 - básica factible, 11
 - factible, 7
 - no acotada, 7
- variable
 - de holgura, 6
- vector
 - de coeficientes de la función objetivo, 4
 - de términos independientes, 4
 - de variables de decisión, 4
 - no negativo, 3