1A
2ej

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

ESCUELA NACIONAL DE ESTUDIOS PROFESIONALES

ACATLAN

UN METODO PARA LA OPTIMIZACION DE SISTEMAS LINEALES CON ALMACENAMIENTO

T E S I S

QUE PARA OBTENER EL TITULO DE

A C T U A R I O

P R E S E N T A

LUIS ANGEL BARRIENTOS HUESCA

ACATLAN, EDO. DE MEX. AGOSTO DE 1994

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

ESCUELA NACIONAL DE ESTUDIOS PROFESIONALES "ACATLAN"

DIVISION DE MATEMATICAS E INGENIERIA
PROGRAMA DE ACTUARIA Y M.A.C.

SR. LUIS ANGEL BARRIENTOS HUESCA
Alumno de la carrera de Actuaría
P r e s e n t e .

De acuerdo a su solicitud presentada con fecha 16 de noviembre de 1993, me complace notificarle que esta Jefatura tuvo a bien asignarle el siguiente tema de tesis: "UN METODO PARA LA OPTIMIZACION DE SISTEMAS LINEALES CON ALMACENAMIENTO", el cual se desarrollará como sigue:

INTRODUCCION
CAP. I. El problema de la Programación Matemática
CAP. 2. Condiciones de Optimalidad
CAP. 3. Programación Lineal
CAP. 4. Programación Dinámica
CAP. 5. Optimización de Sistemas Lineales con almacenamiento
CONCLUSIONES Y RECOMENDACIONES
BIBLIOGRAFIA

Asimismo fué designado como Asesor de Tesis el M. en C. Lucio Pérez Rodríguez, Profesor de esta Escuela.

Ruego a usted tomar nota que en cumplimiento de lo especificado en la Ley de Profesiones, deberá presentar servicio social durante un tiempo mínimo de seis meses como requisito básico para sustentar examen profesional así como de la disposición de la Coordinación de la Administración Escolar en el sentido de que se imprima en lugar visible de los ejemplares de la tesis el título del trabajo realizado. Esta comunicación deberá imprimirse en el interior de la tesis.

E.N.E.P. ACATLAN

A T E N T [illegible] E
"POR MI R[illegible]A EL ESPIRITU"
Acatlán, [illegible] julio 11 de 1994.

ACT. LAURA [illegible] RIVERA BECERRA
Jefe de[illegible] Actuaría
y M.A.C.

LMRB.cg'

"Si tú no haces lo mejor con lo que llegues a tener, nunca harás lo mejor que tú podrías haber hecho con lo que tú deberías haber tenido"

R. Aris (1964)

Agradecimientos

A mis padres, Agustín y Leonor, por enseñarme a ser hombre, por esa singular manera de dar cariño y su confianza ciega en mi persona.

A mi tía Susy, mi segunda madre, por su inagotable paciencia, su cariño y cuidados que han hecho feérica mi estancia en México.

A mis hermanos, Agustín Pépe y Arturo, por el cariño, entusiasmo y muy significativo apoyo que siempre me han ofrecido.

A Vero por todo el amor, paciencia, apoyo y comprensión que han favorecido a mi existencia y por ser mi fácula.

A mi Abue por regalarme esta familia.

A mis tías Ludy, Margarita y Lety por su cariño incomparable.

A Isabel (Chelita) por esta bonita amistad y por ser un especial soporte en mi estancia universitaria.

A Juan por su amistad, apoyo y el mundo de cosas que nos une.

A Ramón, Sandy, Marisol, Aristóteles y Paco por su sincero afecto y su entusiasta convicción para hacer de esta etapa profesional una época inolvidable.

Índice

Glosario

Conjunto.- Colección de objetos llamados puntos o elementos.

Conjunto abierto.- Un subconjunto S de R^n se dice conjunto abierto si para cualquier vector x que pertenece a S se puede encontrar un número positivo e, tal que:

$\{ z : | x - z | < e \}$ esta incluido en S.

Conjunto acotado.- Sea x que pertenece a S, un subconjunto de R^n. Entonces S está acotado si para algún $M > 0$, setiene que $| x | \leq M$ para toda x que pertenece a S.

Conjunto cerrado.- Un conjunto sedice cerrado sí y sólo sí su complemento en R^n es un conjunto abierto .

Conjunto compacto.- Se dice que un subconjunto de R^n es compacto sí y sólo sí el conjunto es cerrado y acotado (véase *Conjunto cerrado y Conjunto acotado*).

Conjunto compatible.- Es aquel conjunto no vacio formado por la intersección de conjuntos.

Conjunto convexo poliédrico.- Se define como la intersección de un número finito de semiespacios cerrados (véase *Semiespacio cerrado*).

Dimensión de un espacio vectorial.- Se define como el número máximo de vectores linealmente independientes que existen en un espaciovectorial (véase *Espacio vectorial y Linealmente independientes*).

Ecuación de continuidad.- Se dice así a la relación que guarda el nivel de almacenamiento al principio del k-ésimo intervalo y la tasa de flujo durante este periodo con el nivel de almacenamiento al final del intervalo.

Espacio vectorial.- Se define como el conjunto de puntos, llamados vectores, sobre el campo R junto con las operaciones: adición vectorial y multiplicación escalar.

Estado.- Es un resumen de la información de los procesos a priori que están suficientemente detallados para facilitar la evaluación de las alternativas actuales.

Función continua.- Se dice que f es una función continua en x que pertenece a X si $f(x_k)$ tiende a f(x) para cualquier x_k que tiende a x.

Función continuamente diferenciable.- Sea f una función valuada en los reales, f: X a R (léase como la función f que mapea al conjunto X en los números reales) donde X que pertenece a R^n es un conjunto abierto, se dice que f es continuamente diferenciable si las derivadas parciales $\delta f(x) / \delta x_i$ para i = 1,..., n existen para cada x que pertenece a X y son funciones continuas de x sobre X (véase *Función Continua*).

Intervalo.- Un subconjunto S de R se llama intervalo si $ax + (1-a)y$ está en S para toda x, y en S y toda $0 < a < 1$.

Linealmente dependientes.- Sean $x_1, \ldots, x_n$ vectores de un espacio vectorial. Entonces, se dice que $x_1, \ldots, x_n$ son linealmente dependientes si existe un x_j para algún j (= 1, ..., n) y $a_j \neq 0$ tal que:

$$x_j = (a_1 x_1 + \ldots + a_{j-1} x_{j-1} + a_{j+1} x_{j+1} + \ldots + a_n x_n) / a_j$$

Linealmente independientes.- Sean $x_1, \ldots, x_n$ vectores de un espacio vectorial. Entonces, $x_1, \ldots, x_n$ son linealmente independientes si y sólo si:

$$a_1 x_1 + \ldots + a_n x_n = 0 \quad \text{entonces } a_1 = \ldots = a_n = 0$$

(véase *Espacio vectorial*).

Matriz hesssiana (o de Hesse).- Sea f una función 2 veces diferenciable y continua en X (subconjunto de los números reales). La matriz hessiana de f en x se define como la matriz simétrica de nxn, teniendo en la ij-ésima posición al elemento $\delta^2 f(x) / \delta x_i \delta x_j$ para i,j = 1,..., n.

Semiespacio cerrado.- Se define así al conjunto en el espacio euclidiano que cumple con:

$$\{ x \in E^n : a_1 x_1 + \ldots + a_n x_n \leq b \}$$

Sistemas de baja memoria.- Es aquel sistema en donde la producción de comodidad en un intervalo de tiempo no se ve afectada por la producción en otro intervalo de tiempo.

Subespacio vectorial o simplemente subespacio.- Se define como el subconjunto S de un espacio vectorial tal que la adición vectorial en S y la multiplicación escalar caen en el conjunto S (véase *Espacio vectorial*).

Tasa de flujo.- Se dice así a la cantidad de comodidad que se desplaza entre el sistema de producción y el almacén.

Tasa de flujo promedio neto.- Se dice así a la cantidad promedio de comodidad que se desplaza entre el sistema de producción y el almacén en un intervalo de tiempo.

Unidad marginal.- Es el incremento adicional en la función que se obtiene por cada unidad incrementada.

Vector gradiente.- Sea f una función con primeras derivadas parciales en X. Entonces el vector gradiente de f en x que pertenece a X se define como:

$$f^{(1)}(x) = (\delta f(x) / \delta x_1, \ldots, \delta f(x) / \delta x_n)$$

Lista de símbolos

x_i	léase como "variable equis i"
$\mathbf{x}$	vector columna $\mathbf{x}$
$\mathbf{x}'$	vector transpuesto $\mathbf{x}$
E^n	espacio euclidiano de n dimensiones
R^n	espacio euclidiano de n dimensiones de los números reales
$f(\mathbf{x})$	función "f" con respecto al vector $\mathbf{x}$
$f(x_1, x_2, \ldots, x_n)$	función "f" con respecto a las variables $x_1, x_2, \ldots, x_n$
$\max_{\mathbf{x}} f(\mathbf{x})$	maximizar la función "f" con respecto al vector $\mathbf{x}$
$\mathbf{X}$	conjunto solución de $\mathbf{x}$
ε	pertenece a
$bf(\mathbf{x})$	producto del escalar b por la función "f" con respecto al vector $\mathbf{x}$
$=$	igual a
$\neq$	diferente a
$>$	mayor que
$\geq$	mayor o igual que
$<$	menor que
$\leq$	menor o igual que
$/$	divide
Σ	suma
$!$	factorial
$g_i(\mathbf{x})$	función "g" subíndice i con respecto al vector $\mathbf{x}$
b_i	escalar b subíndice i
$\mathbf{g}(\mathbf{x})$	vector columna de funciones "g" subíndice i con respecto al vector $\mathbf{x}$
$\mathbf{b}$	vector columna de los escalares b subíndice i

$g_i((x_1, x_2, \ldots, x_n)$	función "g" subíndice i con respecto a las variables $x_1, x_2, \ldots, x_n$
$\mathbf{0}$	vector columna cero
$\mathbf{c}$	vector fila c
c_i	escalar c subíndice i
$\mathbf{Ax}$	producto matricial de la matriz **A** y el vector **x**
$\mathbf{h}$	vector fila h
h_i	escalar h subíndice i
$\mid h_i \mid$	valor absoluto del escalar h subíndice i
$N_e(\mathbf{x}_o)$	entorno de $\mathbf{x_o}$ para cierto número positivo e
$\mathbf{x}_i$	vector **x** subíndice i
$\mathbf{H}$	matriz hessiana
$\mathbf{x}^*$	vector óptimo **x**
x_i^*	variable óptima x subíndice i
i.e.	es decir
h^2	escalar h elevado al cuadrado
$f^{(1)}(*)$	.vector gradiente de la función f
$f^{(2)}(*)$	matriz hessiana de la función f
$f_i^{(1)}(*)$	primera derivada parcial de la función f con respecto a la i-ésima variable
$f_{ij}^{(2)}(*)$	segunda derivada parcial de la función f con respecto a la i, j-ésima variable
$\min\limits_{\mathbf{x}} f(\mathbf{x})$	minimizar la función "f" con respecto al vector **x**
$\delta f(\mathbf{x})$	diferencial de la función f
δx	diferencial de x
$f_{\mathbf{x}}^{(1)}$	vector gradiente de la función f con respecto a **x**
$\mathbf{J}_{mxm}$	matriz J de m filas y m columnas
$\mathbf{C}_{mx(n-m)}$	matriz C de m filas y (n-m) columnas
$\mathbf{J}^{-1}$	inversa de la matriz J
$f_c^{(1)}$	vector gradiente restringido de f con respecto a z
$\delta g_i(\mathbf{x}^*)/\delta x_j$	primera derivada parcial de la función g subíndice i con respecto a la variable x subíndice j
$\phi(\mathbf{v})$	función del vector v
$L(\mathbf{x},\mathbf{y})$	función lagrangiana con respecto a **x** y **y**
$d\mathbf{g}/d\mathbf{x}$	derivada total del vector g con respecto al vector x

Γ^{i}	función Γ superíndice i
I	matriz identidad
L'	función lagrangiana prima
A^{ij}	i, j-ésima partición de la matriz A
x^{i}	i-ésima partición del vector x
y^{i}	i-ésima partición del vector y
b^{i}	i-ésima partición del vector b
c^{i}	i-ésima partición del vector c
X^{k}_{min}	límite inferior de la variable x en el periodo k
X^{k}_{max}	límite superior de la variable x en el periodo k
U^{k}_{min}	límite inferior de la variable u en el periodo k
U^{k}_{max}	límite superior de la variable u en el periodo k
t^{k}	k-ésimo intervalo de tiempo
x^{k}	nivel de almacenamiento en el periodo t^{k}
u^{k}	tasa promedio del flujo de comodidad en el periodo t^{k}
$f^{k}(u^{k})$	costo neto por unidad de producción de comodidad en el periodo t^{k}
U_{j}^{k}	j-ésima pendiente de $f^{k}(*)$ en el periodo t^{k}
Δu_{j}^{k}	j-ésima función de la tasa de flujo de u^{k}
g_{j}^{k}	diferencia de U_{j-1}^{k} y U_{j}^{k}
c_{j}^{k}	pendiente negativa de la función costo de producción
π^{k}	carga total en la producción de comodidad
$v^{k}(x^{k})$	función costo de terminación
Δx_{i}^{k}	i-ésima función de la tasa de flujo de x^{k}
h_{i}^{k}	diferencia de X_{j-1}^{k} y X_{j}^{k}
d_{i}^{k}	pendiente negativa de la función costo de terminación
ρ^{k}	k-ésimo multiplicador de continuidad de almacenamiento
$\delta v^{k-1}(x^{k-1})/ \delta x^{k-1}$	derivada de la función v^{k-1} con respecto a x^{k-1}
$\Theta(\rho^{k})$	conjunto de x^{k-1} que cumplen con la proposición 2

Lista de figuras y gráficas

Resumen

Este documento expone un método que resuelve el problema de optimización de sistemas lineales con almacenamiento. Este procedimiento lleva a una soluciónexacta ya que se basa en las condiciones de Kuhn-Tucker.

Para el desarrollo del método fue necesario comprender algunas bases de la programación lineal y dinámica, esto es, porque el método exponela acción combinada de las programaciones antes mencionadas.

Este sistema debe cumplir con algunos supuestos para poder utilizar este procedimiento, tales condiciones son: que el sistema consista de un almacén y una producción de comodidad, el cual se describe por medio de una función de costo de producción lineal y convexa. La función debe cumplir con ser determinística, es decir, la función es constante para cada intervalo del periodo de producción bajo estudio.

Esto se presenta en la sección 1 del capítulo 5, después en las secciones 5.2 y 5.3 se plantea el sistema con la programación lineal y la programación dinámica y se expone la teoría que sustenta al método, con lo cual se puede decir que se logra la solución del problema en forma exacta y con un número de cálculos reducido. Por último se da el algoritmo del método que resume los pasos a seguir paraoptimizar el sistema.

Abstract

This document shows a method, that resolves the problem of optimization of lineal systems with storage. This process has an exact solutionbased upon the conditions of Kuhn-Tucker.

To develop the method, it was necessary to understand some basic concepts of lineal and dynamic programation because the method shows the combined action of the programation mentioned above.

This system must comply with some suppositions in order to be able to use this procedure, such as: that the system must has a storage unit and acomodity production, that is described by a function of the linear and convex production cost. The function must be deterministic, this means, that the function is constant for each interval of production period under study.

This is shown in section 1 chapter 5. later in sections 5.2 and 5.3 the system states the linear and dynamic programation and shows the theory that supports the method, and because of this you can get an accurate solution with a minimum of calculation. To complete the process you use the algorithmof method that summarizes the steps to follow in order to optimize the system.

Introducción

A través de la historia, el hombre ha empleado rituales elaborados para tomar decisiones. Él ha preparado pócimas, sacrificado animales, interpretado las estrellas y observado el vuelo de los pájaros. Se ha basado en proverbios, refranes y dichos para hacer algunas conjeturas de la vida. Después, con el tiempo se desarrolló la Investigación de Operaciones, en la II Guerra Mundial, ésta se constituyó por un grupo de científicos encargados de investigar las operaciones militares. Se abordaron problemas relacionados directamente con la milicia, como el de minimizar el riesgo de fallarle al objetivo, o afines, como los de abastecimiento del ejército, transporte de tropa, arsenal y el de la dieta óptima para el soldado. Al término de la guerra, este cuerpo de científicos se vieron desempleados, más muy pronto encontraron similitudes entre los problemas analizados en la guerra y los que se dan en una empresa. La Investigación de Operaciones tiene cada vez mayor importancia en la administración y selección de estrategias o políticas de decisión. Por mucho tiempo (posterior a la II Guerra Mundial) y en muchos casos se ha empleado la programación lineal, ya que, acerca el problema a resolver a un programa matemático y tiene la posibilidad de encontrar la solución, en general, eficientemente.

Sin embargo, hay situaciones en las que es necesario aplicar métodos que sean más representativos de estos problemas, como por ejemplo con: programación entera, programación dinámica, programación no lineal, programación convexa, etc. Pero, dada la gama de problemas a optimizar, es necesario desarrollar nuevos mecanismos que permitan optimizar los diferentes tipos de problemas. De aquí, el compromiso del *ACTUARIO*, aplicado en la Investigación de Operaciones, ante la sociedad para ayudar a resolver estos problemas.

La diversidad de situaciones en las que el investigador de operaciones se emplea es cada vez más amplia, en donde cada problema a optimizar presenta diferentes condiciones, características, objetivos, etc. y de igual forma el investigador de operaciones debe estar preparado para utilizar los conocimientos oportunos para emplear la técnica adecuada para resolver dicho problema. Luego, la necesidad del investigador de operaciones de conocer, comprender y desarrollar nuevos métodos de optimización y con esto, emplear de mejor manera los métodos más representativos del problema.

Un problema típico de empresas fabriles es el concerniente a la producción en diferentes periodos de tiempo y almacenamiento del producto en estos periodos con el fin de optimizar sus recursos. Para este tipo de problemas, la soluciónse complica si se trata de resolver por medio de la programación lineal, pues, el número de restricciones será proporcional al número de periodos; de aquí que la solución del problema con programación lineal resulte limitativo.

Si se pretendiera resolver el mismo problema haciendo uso de la programación dinámica es posible que se topara con el "problema de la dimensionalidad"[1] que dificultaría los cálculos. Luego, la necesidad de buscar otras alternativas que den solución a este problema en particular; este trabajo intenta desarrollar un método que optimice el problema arriba mencionado utilizando en forma combinada la programación lineal y la programación dinámica.

Para desarrollar este método, se delimita el problema que se pretende resolver. El sistema lineal con almacenamiento debe estar formado por dos componentes: un almacén y un sistema de producción de comodidad que tenga la característica de que la producción de comodidad en un periodo, no se vea afectada por la producción en algún otro periodo. Además, se necesita del conocimiento en cada periodo de: límites físicos del almacén, límites en el flujo de comodidad del almacén y costos de producción; es decir, se pretende resolver problemas donde es válida la aproximación determinística para los distintos intervalos de tiempo.

1 R. Bellman (1957)

Por tanto, de lo anterior se puede decir que, para ofrecer otras alternativas de solución de programas matemáticos, se obliga al investigador de operaciones a conocer la teoría de la programación matemática, esto es, el investigador de operaciones debe comprender la teoría de la programación matemática para utilizarla de la manera más conveniente, de tal forma que desarrolle nuevos métodos que sean una representación más fiel del problema real y de igual forma lo resuelva.

En los capítulos 1, 2, 3 y 4 se dan las bases necesarias para desarrollar el método; esta parte del trabajo se recomienda para el lector que tenga conocimientos en: álgebra lineal, cálculo diferencial de una y más variables y, nociones generales de investigación de operaciones. El capítulo 5 es el desarrollo del método, en las secciones 1, 2 y 3 se da la descripción del sistema y el planteamiento con programación lineal y programación dinámica, después, tomando en cuenta algunas proposiciones se concluye el método; en la sección 4 se da el algoritmo, ésta sección se recomienda para el lector con nociones de matemáticas y modelación de problemas de investigación de operaciones.

Este documento se inicia en su capítulo 1 con el planteamiento general del problema de la programación matemática, con esto se intenta que el lector distinga las similitudes y diferencias existentes entre las formulaciones de las programaciones: clásica, lineal y no lineal. Enseguida, en un segundo apartado se dan algunos conceptos básicos, que son necesarios para obtener las condiciones de optimalidad.

El capítulo 2 se divide en dos secciones en donde se estudian por separado las condiciones de optimalidad para la programación clásica y la no lineal. La importancia de este capítulo es el desarrollo que se sigue para obtener las condiciones Kuhn-Tucker y entender su significado y relevancia.

El capítulo 3 hace referencia a la programación lineal, éste es un caso particular de la programación no lineal. Se da un desarrollo teórico de las condiciones de optimalidad (Kuhn-Tucker) de dicha programación, además de, una visión general del algoritmo simplex y se expone el planteamiento de un problema de producción y almacenamiento por medio de la programación lineal.

Se intenta en el capítulo 4 meter al lector en la teoría básica de la programación dinámica, para esto, se desarrolla la teoría para la resolución de problemas de etapas múltiples por medio de la programación dinámica y las condiciones Kuhn-Tucker para un programa dinámico con ciertas características. También se reconocen algunas deficiencias de esta programación, y se expone el planteamiento de un problema de producción y almacenamiento por medio de la programación dinámica.

En el último capítulo, el 5, se expone en un principio la descripción del sistema y sus características especiales, luego, se da la formulación del sistema utilizando programación lineal y dinámica, aquí se exponen los argumentos que sustentan al método, para finalizar, se da el algoritmo del método.

En la sección que se presenta al final, conclusiones y recomendaciones, se exponen las ventajas que tiene el método y un breve resumen de la estructura del algoritmo.

Finalmente, quisiera agradecer el apoyo que tuve de la E.N.E.P. Acatlán y los profesores: Luis Recoder, Mario Arriaga, Enrique Peña, Manuel Valadez, Ricardo Aparicio y Blanca Elena del Pozo. También de aquellos profesores que apartándose del cumplimiento de su deber dentro de las aulas de estudio, me despertaron la inquietud en la Investigación de Operaciones y orientaron para llevar acabo este documento. De manera especial quiero agradecer al profesor Lucio, por su apoyo y notas prestadas para realizar este trabajo.

Luis Angel Barrientos Huesca

Verano 1994

Capítulo 1. El problema de la Programación Matemática

La Programación Matemática es un término adoptado por Robert Dorfman allá por 1950, y se aplica en muchas áreas, además, que promete tener un uso más amplio en el futuro. En la actualidad este término es muy general, puesto que, abarca a: la Programación Clásica, Programación No Lineal, Programación Lineal, Programación Entera, Programación Convexa, la Teoría deFlujos en Redes, Programación Dinámica y Programación Bajo Incertidumbre.

El problema de la Programación Matemática trata de la "distribución o asignación de recursos escasos entre fines competitivos en un momento determinado" (Intrilligator M. D., 1972). La interpretación matemática del problema deprogramación matemática es la de encontrar los valores de ciertas variables sujetas a un conjunto de restricciones determinado en cuanto a sus valores posibles, de modo que se optimice una función dada.

1.1 Planteamiento del problema

Formalmente un planteamiento del problema de programación matemática se compone de variables de decisión, la función objetivo y un conjunto de restricciones. La problemática se encuentra en la elección de valores para *n* variables, $x_1, x_2, \ldots, x_n$, llamadas variables de decisión. Las variables se agruparán en el vector:

$$\mathbf{x}' = (x_1, x_2, \ldots, x_n) \qquad (1.1.1)$$

llamado vector decisional (o vector de decisión), este vector pertenece al espacio euclidiano de dimension n, E^n.

El vector decisional x se dirá factible o admisible si satisface todas las restricciones del problema y el conjunto de todos los vectores factibles (admisibles) se llamará el conjunto factible o conjunto solución de x, que es un subespacio de E^n. Una vez establecido el problema, se trata de elegir un vector decisional del conjunto factible dex. El espacio factible de cualquier problema cuyas restricciones sean compatibles o consistentes forma un espacio no vacío y contiene al menos dos puntos distintos.

La finalidad del problema descrito en forma matemática se define mediante la función objetivo:

$$f(x) = f(x_1, x_2, \ldots, x_n) \tag{1.1.2}$$

la que se supone dada y continuamente diferenciable.

Luego, el problema general de la programación matemáticase puede resumir en la elección de un vector decisional del conjunto factible, de manera que la función objetivo sea maximizada (en general, optimizada), esto es:

$$\begin{array}{ll} & \max_{x} f(x) \\ \text{sujeta a} & x \varepsilon X \end{array} \tag{1.1.3}$$

Algunas anotaciones de interés pueden ser:

Dado max f(x); los escalares aditivos y factores escalares positivos que intervienen en la función objetivo no afectan al problema, i.e., si se búsca max f(x), al sumarle cualquier escalar y/o multiplicar la función por alguna constante positiva el problema conserva su sentido (maximizar), pero, si la función objetivo se multiplica por algún escalar negativo el problema de maximización se convierte en un problema de minimización. Esta última propiedad se puede utilizar para pasar del problema de maximización a uno de minimización y viceversa.

En principio, los casos especiales que se distinguen por su importancia en la programación matemática[1] son: la Programación Clásica, la Programación no Lineal y la Programación Lineal.

- La Programación Clásica consiste de restricciones del tipode igualdad y son las *m* igualdades:

$$
\begin{aligned}
g_1(x) &= b_1 \\
g_2(x) &= b_2 \\
\vdots \quad & \quad \vdots \\
g_m(x) &= b_m
\end{aligned}
\qquad (1.1.4)
$$

en donde las funciones $g_1(x)$, $g_2(x)$,...,$g_m(x)$ son *m* funciones continuamente diferenciables y definidas por las variables de decisión, llamadas funciones de restricción y; las constantes b_1, b_2, ..., b_m son *m* números reales conocidos, llamados constantes de restricción. En forma vectorial las restricciones se denotan por:

$$g(x) = b \qquad (1.1.5)$$

en donde g(x) y b son vectores columna de dimensión *m*. O bien, en forma desarrollada:

$$g(x) = \begin{pmatrix} g_1(x_1, x_2, \ldots, x_n) \\ g_2(x_1, x_2, \ldots, x_n) \\ \vdots \\ g_m(x_1, x_2, \ldots, x_n) \end{pmatrix} \qquad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \qquad (1.1.6)$$

De lo arriba expuesto, se puede decir que una vez planteado el problema de Programación Clásica se busca maximizar la función objetivo sujeta a un conjunto de restricciones de igualdad conocidas.

$$
\begin{aligned}
&\max_{x} f(x) \\
&\text{sujeta a} \quad g(x) = b
\end{aligned}
\qquad (1.1.7)
$$

1 Intrilligator, M.D., (1972)

- La Programación no Lineal se distingue por dos *tipos* de restricción, que son: restricciones de no negatividad para las variables de decisión:

$$\mathbf{x} \geq \mathbf{0} \qquad (1.1.8)$$

y restricciones de desigualdad para las funciones de restricción:

$$\mathbf{g}(\mathbf{x}) \leq \mathbf{b} \qquad (1.1.9)$$

(escritas en forma vectorial) donde, **0** es un vector columna de ceros, **g(x)** y **b** representados de igual forma que en (1.1.6). Las funciones de restricción, al igual que en el caso de la programación clásica, $g_1(x)$, $g_2(x)$,...,$g_m(x)$ se suponen continuamente diferenciables, y también, se supone que son números reales dados las constantes de restricción b_1, b_2, ..., b_m. A semejanza con el problema de programación clásica el problema de programación no lineal trata de maximizar la función objetivo, la cual esta sujeta a un conjunto de restricciones de: no negatividad para las variables de decisión y de desigualdad para las funciones de restricción, esto es:

$$\max_{\mathbf{x}} f(\mathbf{x})$$

$$\text{sujeta a} \quad \mathbf{g}(\mathbf{x}) \leq \mathbf{b}, \quad \mathbf{x} \geq \mathbf{0} \qquad (1.1.10)$$

- En la Programación Lineal la función objetivo se escribe como una combinación lineal de las variables de decisión:

$$f(\mathbf{x}) = c_1x_1 + c_2x_2 + \ldots + c_nx_n = \mathbf{cx} \qquad (1.1.11)$$

donde **c** es el vector fila de *n* parámetros dados:

$$\mathbf{c} = (c_1, c_2, \ldots, c_n) \qquad (1.1.12)$$

y las restricciones son de dos tipos: restricciones de desigualdad lineales, para las funciones de restricción:

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &\leq b_1 \\
a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &\leq b_2 \\
\vdots \qquad\qquad\qquad\qquad & \quad \vdots \\
a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &\leq b_m
\end{aligned}
\qquad (1.1.13)
$$

y restricciones de no negatividad para el vector de decisión:

$$x_1, x_2, \dots, x_n \geq 0 \qquad (1.1.14)$$

escribiendo en forma vectorial el conjunto de las funciones de restricción:

$$\mathbf{Ax} \leq \mathbf{b}, \quad \mathbf{x} \geq \mathbf{0} \qquad (1.1.15)$$

donde **A** es la matriz de *mxn* dada por:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \qquad (1.1.16)$$

De aquí se concluye que, el problema de Programación Lineal es el problema de maximizar una función lineal (la función objetivo) sujeta a un conjunto de restricciones de desigualdad de funciones lineales (las funciones restricción) y la no negatividad de las variables de decisión, o sea:

$$
\begin{aligned}
&\max_{\mathbf{x}} f(\mathbf{x}) = \mathbf{cx} \\
&\text{sujeta a} \quad \mathbf{Ax} \leq \mathbf{b}, \quad \mathbf{x} \geq \mathbf{0}
\end{aligned}
\qquad (1.1.17)
$$

Se puede notar que la Programación Lineal es un caso especial de la Programación no Lineal para el cual la función objetivo y las funciones de restricción son todas lineales.

1.2 Algunas definiciones y teoremas básicos

En esta sección se exponen algunos conceptos que se utilizarán en los siguientes capítulos. Primero se dan las definiciones de: puntos extremos, máximos y mínimos locales y globales, estrictos. Después, se expresan algunas condiciones para encontrar puntos extremos, en esta sección se exhibe el teorema fundamental de la Programación Matemática.

Definición.- Un punto extremo de una función $f(\mathbf{x})$ es un máximo de la función, esto es, un punto $\mathbf{x}_0 = (x_1, x_2, ..., x_n)$ es un máximo si:

$$f(\mathbf{x}_0+\mathbf{h}) \leq f(\mathbf{x}_0) \qquad (1.2.1)$$

para toda $\mathbf{h} = (h_1, h_2, ..., h_n)$ tal que $| h_j |$ es suficientemente pequeña para toda j, es decir, $\mathbf{x}_0$ es un máximo si el valor de f en cada punto del entorno no excede el valor de $f(\mathbf{x}_0)$.

Definición.- Un punto extremo de una función $f(\mathbf{x})$ es un mínimo de la función, esto es, un punto $\mathbf{x}_0 = (x_1, x_2, ..., x_n)$ es un mínimo si:

$$f(\mathbf{x}_0+\mathbf{h}) \geq f(\mathbf{x}_0) \qquad (1.2.2)$$

para toda $\mathbf{h} = (h_1, h_2, ..., h_n)$ tal que $| h_j |$ es suficientemente pequeña para toda j, es decir, $\mathbf{x}_0$ es un mínimo si el valor de f en cada punto del entorno excede el valor de $f(\mathbf{x}_0)$.

Definición.- Un máximo global $\mathbf{x}_0$ es aquel en donde la función toma un valor mayor o igual al obtenido por cualquier otro vector $\mathbf{x}_i$, i.e.:

$$\mathbf{x}_0 \in \mathbf{X} \quad \text{y} \quad f(\mathbf{x}_0) \geq f(\mathbf{x}_i) \quad \text{para todo} \quad \mathbf{x}_i \in \mathbf{X} \qquad (1.2.3)$$

Definición.- El máximo global $\mathbf{x}_0$ es un máximo global estricto si el valor de la función en $\mathbf{x}_0$ es estrictamente mayor que en cualquier otro punto:

$$\mathbf{x}_0 \in \mathbf{X} \quad \text{y} \quad f(\mathbf{x}_0) > f(\mathbf{x}_i) \quad \text{para todo} \quad \mathbf{x}_i \in \mathbf{X}, \quad \mathbf{x}_0 \neq \mathbf{x}_i \qquad (1.2.4)$$

De un razonamiento análogo se pueden expresar las definiciones de mínimo global y mínimo global estricto, esto es:

Definición.- Un mínimo global $\mathbf{x}_0$ es aquel en donde la función toma un valor menor o igual al obtenido por cualquier otro vector $\mathbf{x}_i$, i.e.:

$$\mathbf{x}_0 \in \mathbf{X} \quad \text{y} \quad f(\mathbf{x}_0) \leq f(\mathbf{x}_i) \quad \text{para todo} \quad \mathbf{x}_i \in \mathbf{X}, \qquad (1.2.5)$$

Definición.- El mínimo global $\mathbf{x}_0$ es un mínimo global estricto si el valor de la función en $\mathbf{x}_0$ es estrictamente mayor que en cualquier otro punto:

$$\mathbf{x}_0 \in \mathbf{X} \quad \text{y} \quad f(\mathbf{x}_0) < f(\mathbf{x}_i) \quad \text{para todo} \quad \mathbf{x}_i \in \mathbf{X}, \quad \mathbf{x}_0 \neq \mathbf{x}_i \qquad (1.2.6)$$

Es importante destacar que un máximo global estricto es único, pues si $\mathbf{x}_1$ y $\mathbf{x}_2$ fueran dos máximos globales estrictos distintos entonces de (1.2.4) sedebe cumplir que $f(\mathbf{x}_1) > f(\mathbf{x}_2)$ y también $f(\mathbf{x}_2) > f(\mathbf{x}_1)$, pero estasproposiciones no se pueden cumplir a la vez. De forma análoga, si existe un mínimo global estricto este es único.

Definición.- El vector $\mathbf{x}_0$ es un máximo local si la función toma un valor mayor o igual al obtenido con cualquier otro vector suficientemente próximo a él:

$$\mathbf{x}_0 \in \mathbf{X} \quad \text{y} \quad f(\mathbf{x}_0) \geq f(\mathbf{x}_i)$$
$$\text{para todo} \quad \mathbf{x}_i \in \mathbf{X} \quad \text{intersección con} \quad N_e(\mathbf{x}_i) \qquad (1.2.7)$$

donde $N_e(\mathbf{x}_0)$ es el entorno para algún número positivo e, por pequeño que sea, y se forma por el conjunto de todos los **x** tales que:

$$| \mathbf{x} - \mathbf{x}_0 | < e \qquad (1.2.8)$$

Definición.- Un máximo local x_0 es un máximo local estricto si el valor de la función para x_0 supera al que toma dicha función para cualquier otro vector suficientemente próximo a él:

$$x_0 \varepsilon X \quad y \quad f(x_0) > f(x_i)$$

$$\text{para todo} \quad x_i \varepsilon X \quad \text{intersección con} \quad N_e(x_0),\ x_0 \neq x_i \qquad (1.2.9)$$

De lo mencionado anteriormente se puede decir que un máximo global es también un máximo local, pero no lo contrario, esto es, pueden existir otros máximos locales que sustituidos en la función den un valor todavía más alto.

Definición.- El vector x_0 es un mínimo local si la función toma un valor menor o igual al obtenido con cualquier otro vector suficientemente próximo a él:

$$x_0 \varepsilon X \quad y \quad f(x_0) \leq f(x_i)$$

$$\text{para todo} \quad x_i \varepsilon X \quad \text{intersección con} \quad N_e(x_i) \qquad (1.2.10)$$

donde $N_e(x_0)$ es el entorno para cierto e positivo, por pequeño que sea, y se forma por el conjunto de todos los x tales que:

$$| x - x_0 | < e \qquad (1.2.11)$$

Definición.- Un mínimo local x_0 es un mínimo local estricto si el valor de la función para x_0 toma un valor menor al obtenido con cualquier otro vector suficientemente próximo a él:

$$x_0 \varepsilon X \quad y \quad f(x_0) > f(x_i)$$

$$\text{para todo} \quad x_i \varepsilon X \quad \text{intersección con} \quad N_e(x_0),\ x_0 \neq x_i \qquad (1.2.12)$$

Luego, de las definiciones de un mínimo global y local se tiene que un mínimo global es un mínimo local, pero lo contrario no se puede asegurar, i.e., pueden existir otros mínimos locales que valuados en la función f den un valor menor.

Ahora se dan algunos conceptos en los cuales se expresan las condiciones necesarias y suficientes para encontrar los puntos extremos. Se dan primero unas definiciones que se emplearán para enunciar los siguientes 3 teoremas. Los teoremas que siguen solamente se redactarán y se explicarán pero no se darán sus demostraciones. Las demostraciones se pueden encontrar en Intrilligator (1972).

Definición.- Sea g una función que mapea un intervalo S en los números reales. La función g se llama convexa en S si para puntos cualesquiera $\mathbf{x}$, $\mathbf{y} \in S$, y $0 \leq a \leq 1$ se tiene que:

$$g[a\,\mathbf{x} + (1-a)\,\mathbf{y}] \leq a g(\mathbf{x}) + (1-a) g(\mathbf{y}) \qquad (1.2.13)$$

Definición.- La función $f(\mathbf{x})$ es cóncava sí y sólo sí $-f(\mathbf{x})$ es convexa.

Definición.- Sea $\mathbf{x} \in S$, un elemento de un subconjunto propio S de E^n y $\mathbf{y} \geq \mathbf{0}$ para toda y. Defínase a $L(\mathbf{x},\mathbf{y})$ como la función lagrangiana, tal que sea convexa con respecto a $\mathbf{x}$ y cóncava con respecto a $\mathbf{y}$. Entonces $(\mathbf{x}^*,\mathbf{y}^*)$ es un punto de silla[2] si:

$$\textit{i.-}\quad L(\mathbf{x}^*,\mathbf{y}^*) \leq L(\mathbf{x},\mathbf{y}^*) \quad \text{para toda } \mathbf{x} \in S \qquad (1.2.14)$$
$$\textit{ii.-}\quad L(\mathbf{x}^*,\mathbf{y}^*) \geq L(\mathbf{x}^*,\mathbf{y}) \quad \text{para toda } \mathbf{y} \geq \mathbf{0} \qquad (1.2.15)$$

Teorema .- Una condición necesaria para que $\mathbf{x}_0$ sea un punto extremo de $f(\mathbf{x})$ es:

$$f^{(1)}(\mathbf{x}_0) = \mathbf{0} \qquad (1.2.16)$$

De aquí se puede concluir que, para cualquier punto extremo, la condición $f^{(1)}(\mathbf{x}_0) = \mathbf{0}$ se debe satisfacer (i.e., el gradiente en este vector debe ser nulo). Pero esta condición también se cumple para los puntos de silla por lo que esta condición es necesaria mas no suficiente para identificar los puntos extremos. Por lo tanto, sedefinen a los puntos que cumplan con esta condición como puntos estacionarios.

En el problema general de la programación matemática, el vector de decisión $\mathbf{x}_0$ es un máximo global (o solución) si esta dentro del conjunto solución y hace que valuado en la función objetivo cumpla con las condiciones definidas para un máximo global. Así que, para conseguir esta solución se necesita del teorema fundamental de la Programación Matemática, llamado el teorema de Weiestrass:

2 Del inglés "saddle point"

Teorema.- Si el conjunto de decisiones **X** es compacto (i.e., cerrado y acotado, dado que **X** es un subconjunto de un espacio euclidiano de *n* dimensiones) y no vacío, y la función objetivo $f(x_s)$ es continua en **X**, entonces $f(x_s)$ tiene un máximo global o bien en el interior o en el contorno de **X**.

La importancia del teorema de Weiestrass es que da condiciones suficientes para la existencia de un máximo global.

Otro teorema fundamental de la Programación Matemática, es el teorema local-global, que da las condiciones suficientes para que un máximo local sea un máximo global.

Teorema.- Si el conjunto de decisiones **X** es un conjunto compacto no vacío y $f(x)$ es una función continua cóncava respecto a **X**, entonces un máximo local es un máximo global, y el conjunto de puntos en los cuales se obtiene el máximo es convexo. Si además, se supone que $f(x)$ es estrictamente cóncava entonces la solución es única, es decir, existe un (único) máximo global estricto.

Teorema.- Una condición suficiente y necesaria para que un punto estacionario x_0 sea extremo es que la matriz hessiana **H** evaluada en x_0 sea:

a) definida (o semidefinida) negativa cuando x_0 es un punto máximo, y
b) definida (o semidefinida) positiva cuando x_0 es un punto mínimo.

Más adelante, en la sección 2.1, se dará una demostración de este teorema para el inciso (a) y de manera semejante se puede demostrar (b). Se han dado, con el teorema anterior condiciones suficientes para que puntos estacionarios se identifiquen como puntos extremos.

Se ha observado la importancia que tiene la solución de problemas de programación matemática para una distribución y asignación de recursos en forma óptima. Para esto, se han enunciado los planteamientos generales de tres grupos de problemas de Optimización Matemática, que resuelven un gran número de situaciones en donde es necesario optimizar los recursos.

Las condiciones de optimalidad que deben cumplir cada programación representan un campo importante de la Investigación de Operaciones, luego, la relevancia de distinguir las características especiales de cada programación, para resolver el problema de optimización.

Capítulo 2. Condiciones de Optimalidad

En este capítulo se hace un estudio de las condiciones de optimalidad para la programación clásica y la programación no lineal. En cada programación se observan los casos de la presencia o ausencia de restricciones en el problema, así como también algunos métodos para obtener dichas condiciones.

A. Programación Clásica

El problema de la programación clásica consiste en obtener los valores de ciertas variables que optimicen una función dada (la función objetivo) sujeta a un conjunto de restricciones de igualdad dado. El problema de hallar un máximo en la programación clásica se plantea, de igual manera, como se planteó en el capítulo anterior:

$$\max_{\mathbf{x}} f(\mathbf{x}) \qquad (2.0.1)$$

$$\text{sujeta a } \mathbf{g}(\mathbf{x}) = \mathbf{b}$$

donde:

$$\mathbf{x}' = (x_1, x_2, \ldots, x_n) \tag{2.0.2}$$

y

$$\mathbf{g}(\mathbf{x}) = \begin{pmatrix} g_1(x_1, x_2, \ldots, x_n) \\ g_2(x_1, x_2, \ldots, x_n) \\ \vdots \\ g_m(x_1, x_2, \ldots, x_n) \end{pmatrix} \qquad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \tag{2.0.3}$$

Las n variables $x_1, x_2, \ldots, x_n$ son las variables dedecisión y se agrupan en el vector **x**. La función objetivo se denota por $f(^*)$, y las m funciones $g_1(^*), g_2(^*), \ldots, g_m(^*)$ son las funciones de restricción, reunidas en el vector columna $\mathbf{g}(^*)$. Las constantes de restricción se describen por los números reales $b_1, b_2, \ldots, b_m$ representadas en el vector columna **b**.

En la programación clásica se supone que el número devariables de decisión, n, y el número de restricciones de igualdad, m, son números dados y finitos, además, se debe cumplir que el número de variables sea mayor que el número de restricciones ($n > m$). Como se asumió en el capítulo anterior lasfunciones $f(^*)$, $g_1(^*), g_2(^*), \ldots, g_m(^*)$ son continuamente diferenciables y no contienen elementos aleatorios; el vector **b** se forma de escalares reales conocidos y **x** es cualquier vector que consta de números reales que cumplen con las m restricciones del problema.

2.1 Caso sin restricciones

El caso sin restricciones es el caso en donde $m = 0$, con el vector de decisiones $\mathbf{x}$ que consta de $n \geq 1$ variables. El problema (2.0.1) consiste en elegir un vector $\mathbf{x}$ con elementos reales que maximice la función $f(\mathbf{x})$, entonces, si se supone que existe un máximo local en el punto $\mathbf{x}^*$, se tiene:

$$f(\mathbf{x}^*) \geq f(\mathbf{x}^*+h\Delta\mathbf{x}^*)\,, \qquad (2.1.1)$$

o, lo que es lo mismo

$$f(x_1^*, x_2^*, \ldots, x_n^*) \geq f(x_1^* + h\Delta x_1^*, x_2^* + h\Delta x_2^*, \ldots, x_n^* + h\Delta x_n^*) \qquad (2.1.2)$$

donde h es un número pequeño positivo y arbitrario; Δx_j es una variación arbitraria en x_j, $j = 1, 2, \ldots, n$; y $\Delta\mathbf{x} = (\Delta x_1, \Delta x_2, \ldots, \Delta x_n)$ es una dirección de movimiento en E^n.

Al asumir que $f(^*)$ es dos veces continuamente diferenciable con derivadas parciales continuas y finitas, la función del segundo miembro de (2.1.1) se puede tomar como una función de h que, al desarrollarla en serie de Taylor entorno a $h = 0$, se tiene:

$$f(\mathbf{x}^* + h\Delta\mathbf{x}) = f(\mathbf{x}^*) + hf^{(1)}(\mathbf{x}^*)\Delta\mathbf{x} + (1/2!)h^2(\Delta\mathbf{x})'\, f^{(2)}(\mathbf{x}^* + \theta h\Delta\mathbf{x})(\Delta\mathbf{x})$$

$$0 < \theta < 1 \qquad (2.1.3)$$

donde $f^{(1)}(^*)$ es el vector gradiente y $f^{(2)}(^*)$ la matriz hessiana (o de Hesse):

$$f^{(1)}(\mathbf{x}) = (f_1^{(1)}(\mathbf{x}), f_2^{(1)}(\mathbf{x}), \ldots, f_n^{(1)}(\mathbf{x})) \qquad (2.1.4)$$

$$f^{(2)}(\mathbf{x}) = \begin{pmatrix} f_{11}^{(2)}(\mathbf{x}) & f_{12}^{(2)}(\mathbf{x}) & \cdots & f_{1n}^{(2)}(\mathbf{x}) \\ f_{21}^{(2)}(\mathbf{x}) & f_{22}^{(2)}(\mathbf{x}) & \cdots & f_{2n}^{(2)}(\mathbf{x}) \\ \vdots & \vdots & & \vdots \\ f_{n1}^{(2)}(\mathbf{x}) & f_{n2}^{(2)}(\mathbf{x}) & \cdots & f_{nn}^{(2)}(\mathbf{x}) \end{pmatrix} \qquad (2.1.5)$$

Al desarrollar (2.1.3) se llega a:

$$f(x_1^* + h\Delta x_1^*, x_2^* + h\Delta x_2^*, \ldots, x_n^* + h\Delta x_n^*) = f(x_1^*, x_2^*, \ldots, x_n^*) + h\sum_{j=1}^{n} f_j^{(1)}(x_1^*, x_2^*, \ldots, x_n^*)\Delta x_j +$$

$$+ \tfrac{1}{2}h^2 \sum_{j=1}^{n} \sum_{k=1}^{n} f_{jk}^{(2)}(x_1^* + \theta h\Delta x_1^*, x_2^* + \theta h\Delta x_2^*, \ldots, x_n^* + \theta h\Delta x_n^*)\Delta x_j \Delta x_k$$

$$(2.1.6)$$

Ahora, al sustituir (2.1.3) en (2.1.13) y eliminar términos semejantes se obtiene la "desigualdad fundamental":

$$hf^{(1)}(\mathbf{x}^*)\Delta\mathbf{x} + (1/2!)h^2(\Delta\mathbf{x})'\, f^{(2)}(\mathbf{x}^* + \theta h\Delta\mathbf{x})(\Delta\mathbf{x}) \leq 0 \qquad (2.1.7)$$

que debe cumplirse para todas las direcciones $\Delta\mathbf{x}$ y para todo número positivo pequeño h. Ahora, si se dividen ambos miembros por h y se toma el límite cuando h tiende a cero, entonces, se tiene que la desigualdad fundamental requiere como condición necesaria de primer orden que el vector gradiente se anule en $\mathbf{x}^*$:

$$f^{(1)}(\mathbf{x}^*) = \mathbf{0} \qquad (2.1.8)$$

De este desarrollo se puede decir que un máximo local debe tener lugar en un punto estacionario en el cual se anulen todas las derivadas parciales de primer orden. Luego, como consecuencia, es condición necesaria de segundo orden que la matriz hessiana de la desigualdad fundamental sea definida negativa o semidefinida negativa en el punto máximo local:

$$(\Delta\mathbf{x})'\, f^{(2)}(\mathbf{x}^*)(\Delta\mathbf{x}) \leq 0 \ \text{ para toda } \Delta\mathbf{x} \qquad (2.1.9)$$

Esta manera de solucionar el problema de programación por medio del cálculo presenta grandes dificultades prácticas. El igualar el gradiente de la función objetivo a cero, que es la condición necesaria para que un punto sea un óptimo local, genera *n* ecuaciones que hay que resolver, esto en general es difícil, existen ocasiones en que $f^{(1)}(x) = 0$ no puede expresarse en forma explícita, sino más bien implícita. Además, si se puediera resolver el sistema de ecuaciones generado por $f^{(1)}(x) = 0$, no es posible garantizar que ese punto sea un óptimo local, pues puede darse el caso de un punto de silla. Por lo tanto, se requieren de otros métodos de solución.

Como no es el fin de este documento el profundizar en sistemas de programación clásica no sujetos a restricciones, solamente se mencionarán algunos métodos existentes para la solución de este tipo de problemas[1]:

a) Métodos de búsqueda para una sola variable:

1.- Fibonacci

2.- Sección de oro

3.- Interpolación cúbica

4.- Interpolación cuadrada

5.- Newton-Raphson

b) Métodos basados en el uso de derivadas para funciones de varias variables:

1.- Método de gradientes

2.- Método basado en segundas derivadas

3.- Método de direcciones conjugadas

4.- Método de variable métrica

c) Métodos de búsqueda para funciones de varias variables, no diferenciables.

1.- Método de Powell

1 Prawda, J., (1990)

2.2 Caso sujeto a restricciones

En esta sección se presentan dos métodos para optimizar funciones sujetas a restricciones de igualdad . Primero se presenta el método de Jacobi. Seguido después, del método de los multiplicadores de Lagrange, se podrá notar que este último está muy relacionado con el método jacobiano. En esta sección se dará además, una interpretación de los multiplicadores lagrangianos.

2.2.1 Método de Jacobi

El desarrollo de este método es parecido al caso no sujeto a restricciones, con la diferencia que el conjunto factible debe cumplir con las restricciones del problema.

Considere el problema

$$\min_{\mathbf{x}} f(\mathbf{x}) \tag{2.2.1.1}$$

$$\text{sujeta a } \mathbf{g}(\mathbf{x}) = \mathbf{b}$$

donde

$$\mathbf{x} = (x_1, x_2, \ldots, x_n)$$

y

$$\mathbf{g}' = (g_1, g_2, \ldots, g_m) \tag{2.2.1.2}$$

Como antes, las funciones $f(\mathbf{x})$ y $g_i(\mathbf{x})$, para $i = 1, 2, \ldots, m$, se suponen continuamente diferenciables.

Se utilizan derivadas restringidas para resolver (2.2.1.1) con el fin de encontrar una expresión para las primeras derivadas parciales de $f(x)$ en todos los puntos que satisfacen $g(x) = b$. Los puntos en donde $f^{(1)}(x)$ se anula son los puntos estacionarios. Las condiciones de suficiencia de puntos extremos dadas anteriormente se pueden utilizar para identificar a los puntos estacionarios[2].

De acuedo con la serie de Taylor para los puntos factibles $x+\Delta x$ en el entorno factible de x, se tiene:

$$f(x+h\Delta x) - f(x) = hf^{(1)}(x)\,\Delta x + (1/2!)h^2(\Delta x)'\,f^{(2)}(x^* + \theta\Delta x)h(\Delta x) \qquad (2.2.1.3)$$

y

$$g(x+h\Delta x) - g(x) = hg^{(1)}(x^*)\,\Delta x + (1/2!)h^2(\Delta x)'g^{(2)}(x^* + \theta h\Delta x)(\Delta x) \qquad (2.2.1.4)$$

$$0 < \theta < 1$$

Si se dividen ambos miembros por h y luego se hace tender el límite de h a cero en (2.2.1.3) y (2.2.1.4) estos se pueden reducir a :

$$\delta f(x) = f^{(1)}(x)\,\delta x \qquad (2.2.1.5)$$

$$\delta g(x) = g^{(1)}(x)\,\delta x \qquad (2.2.1.6)$$

De (2.2.1.1) se tiene que $g(x) = b$, por tanto $\delta g(x) = 0$, con esto y las dos ecuaciones anteriores se tiene que

$$\delta f(x) - f^{(1)}(x)\,\delta x = 0 \qquad (2.2.1.7)$$

$$g^{(1)}(x)\,\delta x = 0 \qquad (2.2.1.8)$$

Ahora bien, (2.2.1.7) y (2.2.1.8) forman un sistema de $(m+1)$ ecuaciones con $(n+1)$ incógnitas; las incógnitas están dadas por $\delta f(x)$ y δx. La incógnita $\delta f(x)$ se determina, en cuanto se conoce δx. Por tanto, se puede decir que existen m ecuaciones con n incógnitas.

2 Taha,H. A. (1989)

El supuesto de que existan más variables de decisión que funciones de restricción ($n > m$) se basa en el siguiente razonamiento: si $m > n$, al menos (m-n) ecuaciones son redundantes. Después de eliminar esta redundancia, el sistema se reduce a un conjunto de ecuaciones independientes, tal que $m \leq n$. Para el caso donde $m = n$, puesto que no hay entorno factible, el conjunto de soluciones consta de un sólo punto y la solución es $\delta\mathbf{x} = \mathbf{0}$; tal caso no es de interés para la programación matemática. El caso restante es cuando $m < n$, en donde el conjunto solución consta de al menos dos puntos. Tomando esto como base, se desarrolla el método de Jacobi, en el cual se tiene como primer paso el particionar al vector **x**.

$$\mathbf{x} = (\mathbf{u}, \mathbf{v}) \qquad (2.2.1.9)$$

donde

$$\mathbf{u} = (u_1, u_2, \ldots, u_m) \quad \text{y} \quad \mathbf{v} = (v_1, v_2, \ldots, v_{n-m}) \qquad (2.2.1.10)$$

y **u** es el vector de las variables dependientes y **v** el vector de variables independientes, que corresponden al vector **x**. Si se escriben los vectores gradiente de f y g sustituyendo a **x** por **u** y **v**, se tiene que:

$$f^{(1)}(\mathbf{u},\mathbf{v}) = (f_{\mathbf{u}}^{(1)}, f_{\mathbf{v}}^{(1)}) \qquad (2.2.1.11)$$

$$\mathbf{g}^{(1)}(\mathbf{u},\mathbf{v}) = (\mathbf{g}_{\mathbf{u}}^{(1)}, \mathbf{g}_{\mathbf{v}}^{(1)}) \qquad (2.2.1.12)$$

Si se define:

$$\mathbf{J} = \mathbf{g}_{\mathbf{u}}^{(1)} = \begin{pmatrix} g_{1u}^{(1)} \\ \vdots \\ g_{mu}^{(1)} \end{pmatrix} \qquad (2.2.1.13)$$

$$\mathbf{C} = \mathbf{g}_{\mathbf{v}}^{(1)} = \begin{pmatrix} g_{1v}^{(1)} \\ \vdots \\ g_{mv}^{(1)} \end{pmatrix} \qquad (2.2.1.14)$$

A J_{mxm} se le llama la matriz jacobiana y a $C_{mx(n-m)}$ se le conoce como la matriz de control, ya que contiene a las variables independientes. Se supone que J es no singular, dado que se tiene el supuesto de *m* ecuaciones independientes, entonces, las variables se pueden volver a numerar, si es necesario, de modo que las *m* primeras columnas de $g_u^{(1)}$ tengan un determinante distinto de cero; otra forma de ver la no singularidad de J es por el teorema de la función implícita (véase Bertsekas,D.P., (1982) pp. 11-12).

Si se utilizan las ecuaciones de (2.2.1.11) a (2.2.1.14), se puede escribir (2.2.1.7) y (2.2.1.8) como:

$$\delta f\,(u,v) = f_u^{(1)}\delta u + f_v^{(1)}\delta v \qquad (2.2.1.15)$$

y

$$J\,\delta u = -\,C\,\delta v \qquad (2.2.1.16)$$

Dado que J es no singular, existe su inversa J^{-1}, entonces:

$$\delta u = -\,J^{-1}C\,\delta v \qquad (2.2.1.17)$$

El conjunto de ecuaciones de (2.2.1.17) relaciona el efecto de variación en δv sobre δu . Si se sustituye por δu en la ecuación (2.2.1.15),se tiene a $\delta f\,(u,v)$ como una función de δv , i.e.:

$$\delta f\,(u,v) = (\, f_v^{(1)} - f_u^{(1)}\, J^{-1}C\,)\,\delta v \qquad (2.2.1.18)$$

De esta ecuación, se obtiene la derivada restringida con respecto a v, que es la derivada valuada en los puntos donde se cumplen las restricciones del problema:

$$f_C^{(1)}(u,v) = (\,\delta_C f\,(u,v))/\,\delta_C v = f_v^{(1)} - f_u^{(1)}\, J^{-1}C \qquad (2.2.1.19)$$

donde $f_c^{(1)}(\mathbf{u},\mathbf{v})$ denota al vector gradiente restringido de f con respecto a **v**. Luego, $f_c^{(1)}(\mathbf{u},\mathbf{v})$ debe ser nulo en los puntos estacionarios.

Las condiciones de suficiencia son semejantes a las expuestas anteriormente. Pero ahora la matriz hessiana corresponderá al vector independiente **v**. Mientras tanto los elementos de la matriz hessiana son las segundas derivadas restringidas, esto es, como en (2.2.1.15):

$$f^{(1)}(\mathbf{u},\mathbf{v}) = f_{\mathbf{v}}^{(1)} - \mathbf{Z}\mathbf{C} \tag{2.2.1.20}$$

donde:

$$\mathbf{Z} = f_{\mathbf{u}}^{(1)} \mathbf{J}^{-1} \tag{2.2.1.21}$$

Se tiene que el renglón i para $i = 1, ..., n\text{-}m$ de la matriz hessiana (restringida) es $\delta f_c^{(1)} / \delta v_i$. Nótese que **Z** está en función de **u** y **u** está en función de **v**, esto es:

$$\delta \mathbf{u} = -\mathbf{J}^{-1}\mathbf{C}\, \delta \mathbf{v} \tag{2.2.1.22}$$

Por tanto, al tomar la derivada parcial de $f_c^{(1)}$ con respecto a v_i, se debe aplicar a **Z** la regla de la cadena. Luego, se tiene que:

$$\delta u_j / \delta v_i = (\delta u_j / \delta u_i)(\delta u_i / \delta v_i) \tag{2.2.1.23}$$

Así que para obtener un punto extremo se debe cumplir que la matriz hessiana sea definida (o semidefinida) positiva para un mínimo y en caso de un máximo que sea definida (o semidefinida) negativa.

2.2.2 Método de los multiplicadores de Lagrange

2.2.2.1 Desarrollo del método de los multiplicadores de Lagrange

Enseguida, se presenta el método de los multiplicadores de Lagrange. Este método es, en muchos casos, recurrido por la programación clásica por su eficiencia. Se destaca la importancia de este método porque a la vez que se le utiliza como enfoque de base de casi todos los problemas de optimización, también, aporta información sobre el grado de sensibilidad del valor optimal de la función objetivo respecto a los cambios en las constantes de restricción[3]. El principio del desarrollo de este método se encuentra estrechamente relacionado con el método de Jacobi.

Sea el problema

$$\max_{\mathbf{x}} f(\mathbf{x}) \qquad (2.2.2.1.1)$$

$$\text{sujeta a } g(\mathbf{x}) = \mathbf{b}$$

Se inicia este método tomando los mismos supuestos del método de Jacobi. Entonces, de acuerdo con lo anterior, en la sección 2.2.1 se encontró una expresión (ecuación 2.2.1.17) de **u** en función de **v**, esto es:

$$\mathbf{u} = \phi(\mathbf{v}) \qquad (2.2.2.1.2)$$

donde ϕ es un vector columna de *m* funciones; ahora, el problema puede escribirse así:

$$\max_{\mathbf{v}} \Phi(\mathbf{v}) = f(\phi(\mathbf{v}), \mathbf{v}) \qquad (2.2.2.1.3)$$

3 Intrilligator, M.D. (1972)

es decir a un problema no sujeto a restricciones y dado (2.2.1.2) se tiene que una condición necesaria para un máximo local es:

$$(\delta\Phi / \delta v) = (\delta f / \delta v) + (\delta f / \delta u)(\delta\phi / \delta v) = 0 \qquad (2.2.2.1.4)$$

en donde $\delta\Phi / \delta v$ es un vector (1x(*n-m*)) y $\delta\phi / \delta \mathbf{v}$ es una matriz de (*m*x(*n-m*)). Ahora, al sustituir (2.2.2.1.2) en la matriz de funciones de restricción $\mathbf{g(x)}$:

$$\mathbf{g}(\phi(\mathbf{v}), \mathbf{v}) = \mathbf{b} \qquad (2.2.2.1.5)$$

cuya derivada con respecto a v es

$$(\delta\mathbf{g} / \delta v) + (\delta\mathbf{g} / \delta u)(\delta\phi / \delta v) = 0 \qquad (2.2.2.1.6)$$

donde $\delta\mathbf{g} / \delta u$ es la matriz jacobiana de (*mxm*). Por tanto, (2.2.2.1.6) se puede expresar también como:

$$(\delta\phi / \delta v) = -(\delta\mathbf{g} / \delta u)^{-1}(\delta\mathbf{g} / \delta\mathbf{v}) \qquad (2.2.2.1.7)$$

de aquí que las condiciones (2.2.2.1.4) se pueden expresar así:

$$(\delta f / \delta v) - (\delta f / \delta u)(\delta\mathbf{g} / \delta u)^{-1}(\delta\mathbf{g} / \delta v) = 0 \qquad (2.2.2.1.8)$$

Puesto que $\delta\mathbf{g} / \delta u$ es no singular, $\delta f / \delta u$ se puede escribir como:

$$(\delta f / \delta u) - (\delta f / \delta u)(\delta\mathbf{g} / \delta u)^{-1}(\delta\mathbf{g} / \delta u) = 0 \qquad (2.2.2.1.9)$$

Ahora si se denota a y como:

$$\mathbf{y} = (\delta f / \delta u)(\delta\mathbf{g} / \delta u)^{-1} = (y_1, y_2, \ldots, y_m) \qquad (2.2.2.1.10)$$

las condiciones necesarias para puntos extremos de f se pueden escribir:

$$(\delta f / \delta\mathbf{x}) - \mathbf{y}(\delta\mathbf{g} / \delta\mathbf{x}) = 0 \qquad (2.2.2.1.11)$$

Otra forma de encontrar las condiciones necesarias con las restricciones iniciales, es derivando esta función:

$$f(\mathbf{x}) - \mathbf{y}(\mathbf{b} - \mathbf{g}(\mathbf{x})) \qquad (2.2.2.1.12)$$

con respecto a **x** y a **y**. Esto es la base del método de los multiplicadores de Lagrange.

Por tanto, para aplicar el método de los multiplicadores de Lagrange al problema de la programación clásica:

$$\max_{\mathbf{x}} f(\mathbf{x}) \qquad (2.2.2.1.13)$$

$$\text{sujeta a } \mathbf{g}(\mathbf{x}) = \mathbf{b}$$

Se introduce primero, un vector fila de m variables:

$$\mathbf{y} = (y_1, y_2, \ldots, y_m) \qquad (2.2.2.1.14)$$

Estas variables se conocen como los multiplicadores de Lagrange. El siguiente paso es definir la función lagrangiana como la función objetivo más el producto interno del vector y de multiplicadores de Lagrange, y el vector columna, formado por las diferencias de las constantes de restricción y las funciones de restricción:

$$L(\mathbf{x},\mathbf{y}) = f(\mathbf{x}) - \mathbf{y}(\mathbf{b} - \mathbf{g}(\mathbf{x})) \qquad (2.2.2.1.15)$$

o en forma desarrollada:

$$L(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m) = f(x_1, x_2, \ldots, x_n) - \sum_{i=1}^{m} y_i (b_i - g_i(x_1, x_2, \ldots, x_n))$$

(2.2.2.1.16)

Por último se busca el punto (x^*, y^*) en donde se anulen todas las derivadas parciales de primer orden de la lagrangiana, esto es:

$$\delta L(\mathbf{x}^*, \mathbf{y}^*) / \delta \mathbf{x} = (\delta f (\mathbf{x}^*) / \delta \mathbf{x}) - \mathbf{y}^*(\delta \mathbf{g} (\mathbf{x}^*) / \delta \mathbf{x}) = 0$$

$$\delta L(\mathbf{x}^*, \mathbf{y}^*) / \delta \mathbf{y} = \mathbf{b} - \mathbf{g}(\mathbf{x}^*) = 0 \qquad (2.2.2.1.17)$$

Se puede observar que el primer conjunto de restricciones (2.2.2.1.17), es el mismo que (2.2.2.1.11), por lo que el vector gradiente de la función objetivo debe ser lo mismo que el vector de multiplicadores de Lagrange multiplicado por el gradiente de las funciones de restricciones:

$$(\delta f (\mathbf{x}^*) / \delta \mathbf{x}) = \mathbf{y}^*(\delta \mathbf{g} (\mathbf{x}^*) / \delta \mathbf{x}) \qquad (2.2.2.1.18)$$

o, en forma desarrollada:

$$\delta f(x_1^*, x_2^*, \ldots, x_n^*)/\delta x_j = \sum_{i=1}^{m} y_i(\delta g_i(x_1^*, x_2^*, \ldots, x_n^*)/\delta x_j)$$

$$j = 1, 2, \ldots, n \qquad (2.2.2.1.19)$$

El segundo sistema de ecuaciones de (2.2.2.1.17) son *m* condiciones que expresan las restricciones del problema:

$$\mathbf{g}(\mathbf{x}^*) = \mathbf{b} \qquad (2.2.2.1.20)$$

Al resolver las *n+m* ecuaciones de (2.2.2.1.17) se obtienen las soluciones para las *m+n* incógnitas: las variables de decisión $\mathbf{x}^* = (x_1^*, x_2^*, \ldots, x_n^*)'$ y los multiplicadores de Lagrange $\mathbf{y}^* = (y_1^*, y_2^*, \ldots, y_m^*)$. Si se toma al punto $\mathbf{x}^*$ como una solución local al problema de programación clásica se puede concluir que el máximo de la función lagrangiana en el punto $(\mathbf{x}^*,\mathbf{y}^*)$ y la función objetivo son iguales, ya que, al satisfacerse las restricciones en $\mathbf{x}^*$ estas se anulan quedando únicamente la función objetivo maximizada, esto es:

$$L(\mathbf{x}^*, \mathbf{y}^*) = f(\mathbf{x}^*) \qquad (2.2.2.1.21)$$

puesto que se satisfacen las restricciones (2.2.2.1.20).

Ahora bien, las condiciones necesarias de segundo orden son que la matriz hessiana de las derivadas parciales de segundo orden de la lagrangiana con respecto a $\mathbf{x}$:

$$\delta^2 L/\delta \mathbf{x}^2 = \begin{pmatrix} \delta^2 L/\delta x_1^2 & \delta^2 L/\delta x_1 \delta x_2 & \cdots & \delta^2 L/\delta x_1 \delta x_n \\ \delta^2 L/\delta x_2 \delta x_1 & \delta^2 L/\delta x_2^2 & \cdots & \delta^2 L/\delta x_2 \delta x_n \\ \vdots & \vdots & & \vdots \\ \delta^2 L/\delta x_n \delta x_1 & \delta^2 L/\delta x_n \delta x_2 & \cdots & \delta^2 L/\delta x_n^2 \end{pmatrix} \qquad (2.2.2.1.22)$$

sea definida negativa o semidefinida negativa para el punto máximo local $(\mathbf{x}^*, \mathbf{y}^*)$ sujeta a las m condiciones, i.e., que cumpla las restricciones del problema:

$$d\mathbf{g} = (\delta \mathbf{g}(\mathbf{x}^*) / \delta \mathbf{x})\, d\mathbf{x} = \mathbf{0} \qquad (2.2.2.1.23)$$

Si una matriz hessiana es definida negativa sujeta a las condiciones (2.2.2.1.23), entonces las condiciones de primer orden (2.2.2.1.18) y $\mathbf{g}(\mathbf{x}) = \mathbf{b}$ son suficientes para un máximo local.

2.2.2.2 Interpretación de los multiplicadores de Lagrange

La relevancia de las condiciones (2.2.2.1.18) no termina al encontrar un vector de decisiones localmente óptimo, $\mathbf{x}^*$, sino que también proporciona el vector de multiplicadores de Lagrange, $\mathbf{y}^*$, y puesto que $\mathbf{J}$ es no singular, $\mathbf{y}^*$ es única para el óptimo local $\mathbf{x}^*$. La importancia de los multiplicadores de Lagrange se basa en la información que aportan respecto al problema, ya que miden el grado de sensibilidad del valor optimal de la función objetivo a las variaciones en las constantes de restricción $\mathbf{b}$:

$$\mathbf{y}^* = \delta f(\mathbf{x}^*) / \delta \mathbf{b} \qquad (2.2.2.2.1)$$

o, lo que es lo mismo:

$$y_i^* = \delta f(\mathbf{x}^*) / \delta b_i, \quad i = 1, 2, \dots, m \qquad (2.2.2.2.2)$$

Es posible llegar a las igualdades anteriores si se toman las b_i como variables y si se consideran las ecuaciones (2.2.2.1.18) y (2.2.2.1.20), se llega a:

$$\Gamma^1(\mathbf{b}, \mathbf{y}, \mathbf{x}) = \mathbf{b} - \mathbf{g}(\mathbf{x}) = \mathbf{0} \qquad (2.2.2.2.3)$$

$$\Gamma^2(\mathbf{b}, \mathbf{y}, \mathbf{x}) = (\delta f(\mathbf{x}) / \delta \mathbf{x}) - \mathbf{y}(\delta \mathbf{g}(\mathbf{x}) / \delta \mathbf{x}) = \mathbf{0}$$

que es un **sistema** de $m+n$ ecuaciones con $2m+n$ variables ($\mathbf{b}, \mathbf{y}, \mathbf{x}$). El gradiente de este sistema de ecuaciones es:

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} & -(\delta \mathbf{g}/\delta \mathbf{x}) \\ \mathbf{0} & -(\delta \mathbf{g}/\delta \mathbf{x})' & \delta^2 L/\delta \mathbf{x}^2 \end{pmatrix} \qquad (2.2.2.2.4)$$

donde $\mathbf{I}$ es la matriz identidad $m\times m$. Esta matriz, como se definió antes, es de rango de fila completo.

Si se despejan las variables de decisión, **x**, y los multiplicadores de Lagrange, **y**, como funciones de las constantes de restricción, **b**, se puede resolver el sistema formado por las $m+n$ condiciones de primer orden:

$$\mathbf{y} = \mathbf{y}(\mathbf{b})$$

$$\mathbf{x} = \mathbf{x}(\mathbf{b}) \qquad (2.2.2.2.5)$$

Al sustituir (2.2.2.2.5) en la función lagrangiana, ésta queda en función de las constantes de restricción, i.e.:

$$L(\mathbf{b}) = f(\mathbf{x}(\mathbf{b})) + \mathbf{y}(\mathbf{b})[\mathbf{b} - \mathbf{g}(\mathbf{x}(\mathbf{b}))] \qquad (2.2.2.2.6)$$

Si se deriva con respecto a **b**:

$$(\delta L / \delta \mathbf{b}) = (\delta f / \delta \mathbf{x})(\delta \mathbf{x} / \delta \mathbf{b}) + \mathbf{y} - \mathbf{y}(\delta \mathbf{g} / \delta \mathbf{x})(\delta \mathbf{x} / \delta \mathbf{b}) + (\delta \mathbf{y} / \delta \mathbf{b})(\mathbf{b} - \mathbf{g}(\mathbf{x}))$$

$$= [(\delta f / \delta \mathbf{x}) - \mathbf{y}(\delta \mathbf{g} / \delta \mathbf{x})](\delta \mathbf{x} / \delta \mathbf{b}) + (\delta \mathbf{y} / \delta \mathbf{b})(\mathbf{b} - \mathbf{g}(\mathbf{x})) + \mathbf{y}$$

(2.2.2.2.7)

De las condiciones de primer orden (2.2.2.1.17), los primeros dos términos de (2.2.2.2.7) se anulan cuando se encuentra la solución $(\mathbf{x}^*, \mathbf{y}^*)$, de modo que el cambio en la lagrangiana es igual al vector de multiplicadores de Lagrange. Pero en $(\mathbf{x}^*, \mathbf{y}^*)$, el valor de la lagrangiana es el valor optimal de la función objetivo (2.2.2.1.21):

$$\delta L(\mathbf{x}^*, \mathbf{y}^*) / \delta \mathbf{b} = \delta f(\mathbf{x}^*) / \delta \mathbf{b} = \mathbf{y}^* \qquad (2.2.2.2.8)$$

El método de multiplicadores de Lagrange, por tanto, además de resolver el problema de la programación clásica, proporciona también un análisis de sensibilidad que permite ver en los valores de los multiplicadores de Lagrange, ede que manera se afecta el valor optimal de la función objetivo, a los cambios en las constantes de restricción (referencias que tratan el análisis de sensibilidad, por ejemplo: Taha (1989), Intrilligator (1972), Bertsekas (1982)).

B. Programación no Lineal

En Ingeniería, Economía, Ciencias Sociales, Administración y Ciencias Físicas, así como en otros campos, el analista se enfrenta con frecuencia al problema de optimizar complejos: arreglos de equipo, operaciones, circuitos, procesos, etc. Él desea maximizar o minimizar alguna función objetivo (un costo, un peso, un beneficio, etc.), por medio de la selección de ciertas variables de decisión (parámetros de diseño, ajustes del controlador, lecturas instrumentales, cantidades de recursos, etc.), sujeto a ciertas restricciones (requerimientos técnicos, condiciones de operación, capacidades de flujo, factores de seguridad, etc.) inherentes en el proceso. Cuando se formulan en forma matemática, una amplia clase de estos problemas de optimización se pueden agrupar en una categoría llamada "problema de Programación no Lineal". Los métodos de solución de tales problemas se llaman Programación no Lineal. Como se explicará más adelante estos problemas también incluyen a los de Programación Clásica.

Así que, el problema de la Programación no Lineal es el de obtener valores no negativos de ciertas variables llamadas variables de decisión, de tal manera que se optimice una función dada, llamada función objetivo, sujeta a un conjunto de restricciones de desigualdad ya establecido. Al emplear la notación del capítulo 1, el problema del máximo en Programación no Lineal se expresa en forma vectorial como:

$$\max_{\mathbf{x}} f(\mathbf{x})$$

$$\text{sujeta a} \quad \mathbf{g}(\mathbf{x}) \leq \mathbf{b}, \qquad \mathbf{x} \geq \mathbf{0} \tag{2.0.4}$$

o, en forma desarrollada:

$$\max_{x_1, x_2, \ldots, x_n} f(x_1, x_2, \ldots, x_n)$$

sujeta a

$$
\begin{aligned}
g_1(x_1, x_2, \ldots, x_n) &\leq b_1 \\
g_2(x_1, x_2, \ldots, x_n) &\leq b_2 \\
&\vdots \\
g_m(x_1, x_2, \ldots, x_n) &\leq b_m \\
x_1, x_2, \ldots, x_n &\geq 0
\end{aligned}
\tag{2.0.5}
$$

Al igual que como se definió en el capítulo precedente, las n variables $x_1, x_2, \ldots, x_n$, son las variables de decisión, agrupadas en el vector **x**. La función f (*) es la función objetivo, y las m funciones $g_1(*)$, $g_2(*)$, ...,$g_m(*)$ son las funciones de restricción, reunidas en el vector columna **g**(*). Los números reales dados b_1, b_2, ...,b_m, son las constantes de restricción, reunidas en el vector columna **b**. Se supone que n y m son finitas, también que están dadas las $m+1$ funciones f (*), $g_1(*)$, $g_2(*)$, ...,$g_m(*)$, y que son continuamente diferenciables y no contienen elementos aleatorios. Además, **x** puede ser cualquier vector real sujeto únicamente a las $m+n$ restricciones contenidas en (2.0.4).

Algunas cuestiones interesantes en relación con el problema de la Programación no Lineal son:

1. No existen restricciones en cuanto a los valores de m y n, a diferencia de lo que ocurre con el supuesto en el problema de la Programación Clásica.

2. El sentido de las desigualdades es sólo cuestión de convenio.

3. Una restricción de igualdad, puede reemplazarse por dos restricciones de desigualdad.

4. Las restricciones de no negatividad en cuanto a las variables decisionales no son limitativas.

De lo anterior, se puede decir que el problema de la programación clásica es un caso especial de la Programación no Lineal, en donde las restricciones de desigualdad pueden combinarse para formar restricciones de igualdad y no existe ninguna restricción de no negatividad.

Otras cuestiones interesantes son las hipótesis de convexidad, ya que, al recordar el teorema local-global, un máximo local de una función situado en (o sobre el contorno de) el conjunto factible, es un máximo global y el conjunto de puntos en donde tiene lugar un máximo global es convexo si se supone que las funciones de restricción son convexas y la función objetivo es cóncava. Además, si se asume que la función objetivo es estrictamente cóncava entonces la solución es única.

2.3 Caso no sujeto a restricciones de desigualdad

Cuando no existen restricciones de desigualdad, el problema (2.0.4) se convierte en maximizar una función eligiendo valores no negativos de las variables:

$$\max_{\mathbf{x}} f(\mathbf{x})$$
$$\text{sujeta a} \quad \mathbf{x} \geq \mathbf{0} \qquad (2.3.1)$$

El problema (2.1.1) se puede resolver de igual manera a la utilizada en el problema de programación clásica sin restricciones en la sección 2.1, esto es, ocupando la serie de Taylor como sigue: supóngase que existe un máximo local, $\mathbf{x}^*$, para (2.3.1), entonces se debe cumplir para todos los puntos vecinos, $\mathbf{x}^* + \Delta\mathbf{x}$:

$$f(\mathbf{x}^*) \geq f(\mathbf{x}^* + h\Delta\mathbf{x}) \qquad (2.3.2)$$

donde $\Delta\mathbf{x}$ una dirección de movimiento de E^n y h un número positivo arbitrariamente pequeño. Ahora, asúmase que $f(\mathbf{x})$ es dos veces continuamente diferenciable, luego, si la función del segundo miembro de (2.3.2) se desarrolla en serie de Taylor entorno a $\mathbf{x}^*$:

$$f(\mathbf{x}^* + h\Delta\mathbf{x}) = f(\mathbf{x}^*) + h\, f_x^{(1)}(\mathbf{x}^*)\, \Delta\mathbf{x} + (1/2)\, h^2\, \Delta\mathbf{x}'\, f_x^{(2)}(\mathbf{x}^* + \theta h\Delta\mathbf{x})\, \Delta\mathbf{x}$$

$$0 < \theta < 1 \qquad (2.3.3)$$

Al sustituir (2.3.3) en (2.3.2) y eliminar términos semejantes, se obtiene la desigualdad fundamental:

$$h\, f_x^{(1)}(\mathbf{x}^*)\, \Delta\mathbf{x} + (1/2)\, h^2\, \Delta\mathbf{x}'\, f_x^{(2)}(\mathbf{x}^* + \theta h\Delta\mathbf{x})\, \Delta\mathbf{x} \leq 0 \qquad (2.3.4)$$

como se mostró en la sección 2.1, la condición necesaria para un máximo local en $\mathbf{x}^*$.

Ahora bien, si $\mathbf{x}^*$ es una solución interior, i.e. $\mathbf{x}^* > \mathbf{0}$, entonces la desigualdad fundamental debe cumplirse para todas las direcciones $\Delta\mathbf{x}$, puede observarse que son las mismas condiciones de primer orden de la Programación Clásica, esto es, la anulación de todas las derivadas parciales de primer orden. Si se supone que una de las variables está situado en el contorno, $x_j^* = 0$, y que todos los otros incrementos son nulos, la desigualdad fundamental (2.3.4) implica que en $x_j^* = 0$ la única dirección permitida es aquella para la cual $\Delta x_j \geq 0$, esto es:

$$f_{x_j}^{(1)}(\mathbf{x}^*)\Delta x_j \leq 0 \qquad (2.3.5)$$

De aquí se desprende que la desigualdad fundamental requiere como condición necesaria de primer orden que:

$$f_{x_j}^{(1)}(\mathbf{x}^*) \leq 0 \qquad \text{si } x_j^* = 0 \qquad (2.3.6)$$

Así que, al tomar la primera derivada con respecto a x_j, ésta se anula si x_j es una solución interior, $x_j > 0$, y si x_j es una solución de contorno, $x_j = 0$, se tiene que la primera derivada con respecto a x_j es menor igual que cero, como se expuso en (2.3.6). Luego, o la variable correspondiente toma el valor de cero (solución de contorno) o la derivada toma el valor de cero (solución interior), se tiene que el producto de estas dos es siempre cero, simbólicamente:

$$f_{x_j}^{(1)}(\mathbf{x}^*)x_j^* = 0 \qquad j = 1, \ldots, n \tag{2.3.7}$$

Al sumar todos los productos (2.3.7) de las n variables se tiene:

$$f_{\mathbf{x}}^{(1)}(\mathbf{x}^*)\mathbf{x}^* = \sum_{j=1}^{n} f_{x_j}^{(1)}(\mathbf{x}^*)x_j^* = 0 \tag{2.3.8}$$

Dicho de otra forma, esta condición de la anulación de la suma de los productos implica, que cada uno de los sumandos se anule (i.e., implica que (2.3.7) se cumpla para toda j) a causa de las condiciones de que todas las variables son no negativas y todas las derivada parciales de primer orden no positivas.

Se puede determinar el máximo local, $\mathbf{x}^*$, utilizando las $2n + 1$ condiciones de primer orden:

$$\begin{aligned} &f_{\mathbf{x}}^{(1)}(\mathbf{x}^*) \leq \mathbf{0} \\ &f_{\mathbf{x}}^{(1)}(\mathbf{x}^*)\mathbf{x}^* = 0 \\ &\mathbf{x}^* \geq \mathbf{0} \end{aligned} \tag{2.3.9}$$

Estas condiciones implican que cada una de las derivadas parciales de primer orden se anula si la variable correspondiente es positivo, y es no positivo si la variable es cero:

$$\begin{aligned} &f_{x_j}^{(1)}(\mathbf{x}^*) = 0 \qquad \text{si } x_j^* > 0 \\ &f_{x_j}^{(1)}(\mathbf{x}^*) \leq 0 \qquad \text{si } x_j^* > 0 \end{aligned} \tag{2.3.10}$$

para $j = 1, 2, \ldots, n$

2.4 Caso sujeto a restricciones de desigualdad

Dado el problema de Programación no Lineal:

$$\max_{\mathbf{x}} f(\mathbf{x})$$
$$\text{sujeta a} \quad \mathbf{g}(\mathbf{x}) \leq \mathbf{b}, \qquad \mathbf{x} \geq \mathbf{0} \tag{2.4.1}$$

Para resolver este problema, primero se convierten las restricciones de desigualdad en restricciones de igualdad añadiéndole un vector no negativo **s** de *m* variables, llamadas variables de holgura, así:

$$\mathbf{s} = \mathbf{b} - \mathbf{g}(\mathbf{x}) = (s_1, s_2, \ldots, s_m)' \tag{2.4.2}$$

ahora, el problema puede formularse:

$$\max_{\mathbf{x}} f(\mathbf{x})$$
$$\text{sujeta a} \quad \mathbf{g}(\mathbf{x}) + \mathbf{s} = \mathbf{b}, \qquad \mathbf{x} \geq \mathbf{0}, \ \ \mathbf{s} \geq \mathbf{0} \tag{2.4.3}$$

El problema (2.4.3) un problema de Programación Clásica si no tuviera las $m + n$ restricciones de no negatividad de las variables, luego, se podría resolver dicho problema utilizando la función lagrangiana, como se hizo en la Programación Clásica de la siguiente forma:

$$L' = f(\mathbf{x}) + \mathbf{y}(\mathbf{b} - \mathbf{g}(\mathbf{x}) - \mathbf{s}) \tag{2.4.4}$$

donde $\mathbf{y} = (y_1, y_2, \ldots, y_m)$ es un vector de multiplicadores de Lagrange, como el utilizado en la Programación Clásica.

Entonces, de resultados anteriores (sección 2.2.2.1), las condiciones necesarias de primer orden se obtendrían como las condiciones de que todas las derivadas parciales de primer orden de L' con respecto a **x**, **y** y **s** se anulen. Puesto que se trata de un problema de Programación no Lineal existen vectores no negativos, **x** y **s**, en que las condiciones de las derivadas de primer orden con respecto **x** y **s**, se sustituyen por las condiciones (2.3.9), esto es, las condiciones necesarias de primer orden para un máximo local son:

$$
\begin{aligned}
&(L')_{x}^{(1)} = f_{x}^{(1)} - y\, g_{x}^{(1)} \leq 0 \\
&(L')_{x}^{(1)}(x) = (f_{x}^{(1)} - y\, g_{x}^{(1)})(x) = 0 \\
&x \geq 0 \\
&(L')_{y}^{(1)} = b - g(x) - s = 0 \\
&(L')_{s}^{(1)} = -y \leq 0 \\
&(L')_{s}^{(1)}(s) = -y\, s = 0 \\
&s \geq 0
\end{aligned}
\tag{2.4.5}
$$

en donde las funciones y derivadas se valuan en x^*, y^* y s^*. El siguiente resultado, llamado las condiciones Kuhn-Tucker, se deduce al sustituir el vector de variables de holgura, **s**, por $b - g(x)$, esto es:

$$
\begin{aligned}
&f_{x}^{(1)} - y\, g_{x}^{(1)} \leq 0 \\
&(f_{x}^{(1)} - y\, g_{x}^{(1)})(x) = 0 \\
&x \geq 0 \\
&b - g(x) \geq 0 \\
&y \geq 0 \\
&y\,(b - g(x)) = 0
\end{aligned}
\tag{2.4.6}
$$

Otra forma de obtener las condiciones Kuhn-Tucker es definir la función lagrangiana para el problema original (2.4.1), así:

$$L(x,y) = f(x) + y(b - g(x)) \qquad (2.4.7)$$

esta función lagrangiana, en palabras, significa la suma de la función objetivo más el producto de los multiplicadores de Lagrange por la diferencia entre las constantes de restricción y las funciones de restricción. Luego, se evalúa (2.4.6) en x^* y y^* óptimos:

$$(L')_x^{(1)}(x^*,y^*) = f_x^{(1)}(x^*) - y^* g_x^{(1)}(x^*) \leq 0 \qquad (2.4.8)$$

$$(L')_x^{(1)}(x^*,y^*)(x^*) = (f_x^{(1)}(x^*) - y^* g_x^{(1)}(x^*))(x^*) = 0 \qquad (2.4.9)$$

$$x^* \geq 0 \qquad (2.4.10)$$

$$(L')_y^{(1)}(x^*,y^*) = b - g(x^*) \geq 0 \qquad (2.4.11)$$

$$y^*(L')_y^{(1)}(x^*,y^*) = y^*(b - g(x^*)) = 0 \qquad (2.4.12)$$

$$y^* \geq 0 \qquad (2.4.13)$$

Obsérvese que las condiciones de no negatividad y las restricciones de desigualdad del problema original de la Programación no Lineal, están dadas en (2.4.10) y (2.4.11) respectivamente. También, obsérvese que una vez establecidas las restricciones de signo en (2.4.8) y (2.4.10) cada uno de los términos de la suma de (2.4.9) se anula, esto es:

$$f_{x_j}^{(1)} - y g_{x_j}^{(1)} = 0 \quad \text{o} \quad x_j = 0 \quad (\text{o ambas}) \quad j = 1, \ldots, n \qquad (2.4.14)$$

es decir, o bien se cumple la condición de la primera derivada (punto interno) como una igualdad, o bien la variable está en el contorno o ambas cosas a la vez, entonces:

$$f_{x_j}^{(1)} - \sum_{i=1}^{m} y_i (g_{x_j}^{(1)}) \leq 0, \qquad \text{pero} = 0 \quad \text{si} \;\; x_j^* > 0$$

$$x_j \geq 0, \qquad \text{pero} = 0 \qquad \text{si} \;\; f_{x_j}^{(1)} - \sum_{i=1}^{m} y_i (g_{x_j}^{(1)}) < 0 \tag{2.4.15}$$

$$j = 1, 2, \ldots, n$$

Igualmente debido a las restricciones de signo de (2.4.11) y (2.4.13), cada término de la suma en (2.4.12) se anula. O bien

$$y_i = 0 \qquad \text{ó} \qquad g_i(x^*) = b_i \qquad \text{(o ambas)}$$

$$i = 1, 2, \ldots, m \tag{2.4.16}$$

es decir, o bien se anula el multiplicador de Lagrange, o bien, se satisface la restricción de desigualdad como igualdad estricta, o ambas cosas a la vez:

$$g_i(x^*) \leq b_i \qquad \text{pero} = b_i \qquad \text{si} \quad y_i > 0$$

$$y_i^* \geq 0 \qquad \text{pero} = 0 \quad \text{si} \quad g_i(x^*) < b_i \tag{2.4.17}$$

$$i = 1, 2, \ldots, m$$

Las condiciones (2.4.15) y (2.4.17) son llamadas "las condiciones de holgura complementaria" y es otra forma de interpretar las condiciones Kuhn-Tucker.

De la condición (2.4.12), el valor de la lagrangiana para la solución, es el valor optimal de la función objetivo:

$$L\,(x^*, y^*) = f\,(x^*) + y^*(b - g(x^*)) = f\,(x^*) \tag{2.4.18}$$

2.5 El teorema de Kuhn-Tucker

Otra vez, el problema general de la Programación no Lineal :

$$\max_{x} f(x)$$
$$\text{sujeta a} \quad g(x) \le b, \quad x \ge 0 \qquad (2.5.1)$$

como se desarrolló en la sección anterior, el teorema de Kuhn- Tucker consiste en introducir un vector $y = (y_1, y_2, \ldots, y_m)$; donde y es el vector de multiplicadores de Lagrange, y definir la función lagrangiana como:

$$L(x,y) = f(x) + y(b - g(x)) \qquad (2.5.2)$$

Las condiciones Kuhn-Tucker como se describió de (2.4.8) a (2.4.13), son:

$$\begin{array}{ll} L_x^{(1)}(x^*,y^*) \le 0 & L_y^{(1)}(x^*,y^*) \ge 0 \\ L_x^{(1)}(x^*,y^*)x^* = 0 & y^* L_x^{(1)}(x^*,y^*) = 0 \\ x^* \ge 0 & y^* \ge 0 \end{array} \qquad (2.5.3)$$

Obsérvese el sentido de las desigualdades y de acuerdo con (2.3.9), se tiene que el punto (x^*,y^*) forma un punto de silla en la función lagrangiana, donde maximiza la función respecto a las variables del vector x y minimiza respecto a los multiplicadores de Lagrange, y; con x, y vectores no negativos. Ahora bien de lo dicho en los renglones anteriores se cumple:

$$L(x,y^*) \le L(x^*,y^*) \le L(x^*,y) \quad \text{para todo } x,y \ge 0 \qquad (2.5.4)$$

Al hecho de encontrar vectores no negativos (x^*,y^*) que satisfagan (2.5.4) se conoce como el problema de punto de silla.

Teorema de Kuhn-Tucker (condición de suficiencia de optimalidad).- Sea $f(x)$ cóncava, $g(x)$ convexa. Si (x^*,y^*) es solución al problema del punto de silla entonces x^* resuelve el problema de Programación no Lineal.

Es decir, si (x^*,y^*) es un punto de silla en la función lagrangiana entonces x^* resuelve el problema de la programación no lineal. Luego, si se supone que (x^*,y^*) es un punto de silla, como x^* maximiza la lagrangiana (respecto a todo $x \geq 0$):

$$f(x) + y^*(b - g(x)) \leq f(x^*) + y^*(b - g(x^*)) \qquad (2.5.5)$$

y puesto que y^* minimiza:

$$f(x^*) + y^*(b - g(x^*)) \leq f(x^*) + y(b - g(x^*)) \qquad (2.5.6)$$

(2.5.6) también puede escribirse:

$$(y - y^*)(b - g(x^*)) \geq 0, \qquad y \geq 0 \qquad (2.5.7)$$

ahora como los componentes de y pueden ser arbitrariamente grandes, se tiene que x^* satisface las restricciones de desigualdad:

$$g(x^*) \leq b \qquad (2.5.8)$$

Pero, si se toma a $y = 0$ en (2.5.7), se tiene que :

$$y^*(b - g(x^*)) = 0 \qquad (2.5.9)$$

dado que $y \geq 0$ y $b - g(x^*) \geq 0$.

Ahora, si se observa (2.5.5), y utilizando (2.5.9), se tiene que:

$$f(x^*) \geq f(x) + y^*(b - g(x)), \qquad x \geq 0 \qquad (2.5.10)$$

Dado que y^* es no negativa, si x es admisible se tendrá que:

$$f(x^*) \geq f(x) \qquad (2.5.11)$$

de modo que x^* maximiza $f(^*)$, en el conjunto de los vectores admisibles, por tanto el problema de la Programación no Lineal queda resuelto.

Teorema de Kuhn-Tucker (condiciones necesarias de optimalidad).- Sean $f(^*)$, función objetivo, una función cóncava y $g(^*)$, vector de restricciones de desigualdad, un vector de funciones convexas que cumple con la condición de calificación de restricción[3]. Entonces, x^* resuelve el problema de Programación no Lineal si y sólo si existe una y^* tal que (x^*,y^*) resuelve el problema del punto de silla.

La primera parte del problema del punto de silla consiste en maximizar $L(x^*,y^*)$ mediante la elección de variables no negativas de x. Entonces, de los obtenidos en (2.3.9) se pueden aplicar y se llega a las condiciones:

$$\begin{aligned} &L_x^{(1)}(x^*,y^*) \leq 0 \\ &L_x^{(1)}(x^*,y^*)\, x^* = 0 \qquad (2.5.12) \\ &x^* \geq 0 \end{aligned}$$

La segunda parte del problema del punto de silla, que consiste en minimizar $L(x^*,y^*)$ eligiendo multiplicadores de Lagrange, y, no negativos, da las condiciones:

$$\begin{aligned} &L_y^{(1)}(x^*,y^*) \geq 0 \\ &y^*\, L_y^{(1)}(x^*,y^*) = 0 \qquad (2.5.13) \\ &y^* \geq 0 \end{aligned}$$

Donde (2.5.12) y (2.5.13) son las condiciones de Kuhn-Tucker, expuestas en (2.5.3).

3 Calificación de restricción.- es la condición de que exista algún punto en el conjunto de restricciones de desigualdad como desigualdad estricta, esto es, existe un vector x tal que $x \geq 0$ y $g(x) < b$.

Para ejemplos acerca de la utilización de las condiciones Kuhn-Tuckerse pueden revisar: Luenberger (1989), Hadley (1964), Intrilligator (1972).

Se ha desarrollado la teoría de la Programación Clásica yProgramación no Lineal, en cuanto a sus condiciones de optimalidad, y se ha hecho hincapié en el papel que juega las condiciones de Kuhn-Tucker para saber si el valor obtenido cumple conser un óptimo. En las secciones anteriores a las condiciones Kuhn-Tucker se plantearon las bases para llegar a esta última.

Las condiciones Kuhn-Tucker se han desarrollado para trabajar con la visión general que guarda, y así, poder entrar a un caso particular de la Programación No Lineal, la Programación Lineal, en esta última es donde se interpretan las condiciones Kuhn-Tucker y su relevancia para obtener un valor optimal del problema a resolver.

El siguiente capítulo trata de la Programación Lineal y la importancia de las condiciones Kuhn-Tucker, además, se plantea un problema de esta programación.

Capítulo 3. Programación Lineal

Los modelos de asignación de recursos muchas veces se pueden formular razonablemente usando sólo relaciones lineales, por tanto este tipo de problemas se puede resolver utilizando Programación Lineal, otros campos en los que se ha aplicado con éxito esta programación es en los modelos de transporte y redes; programación de la producción, diseño estructural y análisis económico.

Esta programación tuvo un gran desarrollo de 1947 a 1967 en que muchos artículos se redactaron y se codificaron programas para computadoras para resolver problemas grandes. El método simplex, propuesto por G. B. Dantzig en 1947, es un método efectivo para resolver problemas de Programación Lineal.

Como se expresó en el capítulo 1, para resolver el problema de la Programación Lineal se deben elegir valores no negativos de las variables de decisión, de tal forma que maximice o minimice una función lineal dada sujeta a un conjunto de restricciones de desigualdad lineal dado.

Entonces, un problema de Programación Lineal se puede escribir en forma vectorial, así:

$$\max_{x} f(x) = cx$$

sujeta a

$$Ax \leq b, \quad x \geq 0 \tag{3.0.1}$$

o, en forma desarrollada:

$$\max_{x_1, x_2, \ldots, x_n} f(x_1, x_2, \ldots, x_n) = c_1x_1 + c_2x_2 + \ldots + c_n x_n$$

sujeta a

$$\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n \leq b_1 \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n \leq b_2 \\ \quad\vdots \qquad\qquad\qquad \vdots \\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n \leq b_m \end{array} \tag{3.0.2}$$

$$x_1, x_2, \ldots, x_n \geq 0$$

La notación de arriba indica que, el vector columna **x** tiene por elementos a las *n* variables de decisión, $x_1, x_2, \ldots, x_n$ del problema. El vector fila **c** esta constituido por las *n* constantes objetivo, $c_1, c_2, \ldots, c_n$; la matriz **A** de *mxn* tiene por elementos en sus entradas a las constantes del programa, llamadas constantes coeficientes $a_{11}, a_{12}, \ldots, a_{1n}; a_{21}, a_{22}, \ldots, a_{2n}; a_{m1}, a_{m2}, \ldots, a_{mn}$; por último, el vector columna **b** formado por las constantes de restricción $b_1, b_2, \ldots, b_n$. Para el estudio que se hará se supone que *n* y *m* son finitos, los elementos de **A**, **b**, **x** y **c** son números reales; además, los elementos de **x** están restringidos por las *m+n* restricciones de (3.0.1).

Se puede observar que este problema es un caso especial de la Programación no Lineal; en donde la función objetivo y las funciones de restricción son lineales.

Cada una de las *n* restricciones de no negatividad:

$$x_j \geq 0, \quad j = 1, 2, \ldots, n \tag{3.0.3}$$

forma un semiespacio cerrado y la intersección de todos los semiespacios definidos en (3.0.3) se encuentran en el ortante no negativo del espacio euclidiano de dimensión *n*, E^n. De igual forma, cada una de las *m* restricciones de desigualdad:

$$a_{j1}x_1+a_{j2}x_2+\ldots+a_{jn}x_n \leq b_j, \qquad j = 1, 2, \ldots, m \tag{3.0.4}$$

constituye un semiespacio cerrado de E^n que es el conjunto de puntos situados en el hiperplano definido por:

$$\{ x \in E^n / \; a_{j1}x_1+a_{j2}x_2+...+a_{jn}x_n \leq b_j, \; j = 1, 2, ..., m \} \qquad (3.0.5)$$

Se conoce como espacio o conjunto solución aquel conjunto formado por todos los vectores **x** que satisfacen las $m + n$ restricciones de desigualdad y de no negatividad de (3.0.1), esto es:

$$\{ \mathbf{x} \in E^n / \; \mathbf{A\,x} \leq \mathbf{b}, \; \mathbf{x} \geq \mathbf{0} \} \qquad (3.0.6)$$

(3.0.6) es un conjunto poliédrico convexo cerrado en el ortante no negativo en el espacio euclidiano de dimensión n. Dado que, el conjunto solución es convexo y la función objetivo es lineal, por el teorema local-global mencionado en el capitulo 1, la solución a parte de ser un máximo local también es máximo global. Así que, si un vértice en el conjunto solución produce ningún valor menor que cualesquiera otros vértices próximos, entonces el problema tiene solución.

Si existen más variables de decisión que restricciones, $n > m$, entonces existe un vértice del conjunto solución en donde $n - m$ ó más de las variables se anulan, es decir, debe haber por lo menos una solución al problema que consta de una solución al problema que consta de un vector **x** con a lo más m de sus n elementos distintos de cero.

De lo mencionado antes, esto es, la función objetivo es continua y el conjunto solución es cerrado y de acuerdo con el teorema de Weiestrass el problema de Programación Lineal tiene solución si el conjunto solución es no vacío y acotado. Luego, se pueden presentar dos casos en los cuales el problema no tenga solución:

1. Las restricciones son incompatibles de modo que el conjunto solución es vacío.

2. El conjunto solución no es acotado y la función objetivo puede crecer sin límite en este conjunto.

Ahora bien, si el conjunto solución es no vacío y acotado la solución debe ser una solución de contorno.

Por tanto, se puede decir que existen tres casos posibles en la solución del problema de la Programación Lineal:

i) Una solución única (en un vértice).
ii) Infinitas soluciones (entre dos vértices o más).
iii) Ninguna solución (si el conjunto solución es vacío o no acotado).

3.1 La relación del problema primal y su dual

El desarrollo de las condiciones de optimalidad en los problemas de Programación Lineal se debe a la relación que a cada problema de Programación Lineal corresponde un problema dual.

Para esto, suponga que el problema original, conocido como problema primal, es el problema del máximo de la Programación Lineal (3.0.1):

$$\max_{x} F = cx$$

sujeta a

$$Ax \leq b, \qquad x \geq 0 \tag{3.1.1}$$

entonces se define el problema dual como el problema del mínimo dela Programación Lineal:

$$\min_{y} G = yb$$

sujeta a

$$\mathbf{yA} \geq \mathbf{c}, \qquad \mathbf{y} \geq \mathbf{0} \qquad (3.1.2)$$

en donde **y** es el vector fila:

$$\mathbf{y} = (y_1, y_2, \ldots, y_m) \qquad (3.1.3)$$

O en forma desarrollada (3.1.2):

$$\min_{y_1, y_2, \ldots, y_m} G(y_1, y_2, \ldots, y_m) = b_1 y_1 + b_2 y_2 + \ldots + b_m y_m$$

sujeta a

$$\begin{aligned} a_{11}y_1 + a_{21}y_2 + \ldots + a_{m1}y_m &\geq c_1 \\ a_{12}y_1 + a_{22}y_2 + \ldots + a_{m2}y_m &\geq c_2 \\ \vdots \qquad\qquad & \quad \vdots \\ a_{1n}y_1 + a_{2n}y_2 + \ldots + a_{mn}y_m &\geq c_n \end{aligned} \qquad (3.1.4)$$

$$y_1, y_2, \ldots, y_m \geq 0$$

Al observar (3.1.1) y (3.1.2) se tiene que los dos problemas buscan una solución de una función lineal mediante la elección de variables no negativas y sujeta a restricciones de desigualdad lineal; el problema primal y su dual tienen cada uno $m + n$ restricciones de desigualdad; ambos utilizan los mismos parámetros dados por lamatriz **A**, el vector columna **b** y el vector fila **c**.

Además, mientras que el problema dual elige m variables contenidas en **y** el problema primal debe elegir n variables agrupadas en **x**; por otro lado si el problema original consiste en maximizar y utiliza desigualdades del tipo ≤, entonces el problema dual minimiza y utiliza desigualdades del tipo ≥; por último las constantes de restricción de cada problema se transforman en las constantes objetivo del otro.

Otra situación que se presenta, es que el dual del problema dual es el problema original, esto es, si se transformara una vez más el problema dual se llegaría nuevamente al problema primal. Por tanto, si se conoce cualquiera de los dos problemas (primal o dual) se puede llegar al otro (dual o primal);de aquí que ninguno de los dos problemas sea el problema fundamental.

3.2 Teoremas fundamentales de Programación Lineal

Los teoremas fundamentales de la Programación Lineal están basados en la función lagrangiana y las condiciones Kuhn-Tucker. De aquí, nace la importancia de los problemas duales, puesto que, las variables duales se utilizan como los multiplicadores lagrangianos del problema primal.

Suponga que el problema primal es el problema de máximo:

$$\max_{x} f(\mathbf{x}) = \mathbf{cx}$$

sujeta a

$$\mathbf{Ax} \leq \mathbf{b}, \qquad \mathbf{x} \geq \mathbf{0} \tag{3.2.1}$$

El vector columna $\mathbf{x}^*$ es una solución de (3.2.1), según el teorema de Kuhn-Tucker, si existe un vector fila $\mathbf{y}^*$ tal que la función de Lagrange se defina como:

$$L(\mathbf{x},\mathbf{y}) = \mathbf{cx} + \mathbf{y}(\mathbf{b} - \mathbf{Ax}) = \mathbf{cx} + \mathbf{yb} - \mathbf{yAx} \tag{3.2.2}$$

y las condiciones de Kuhn-Tucker se cumplen para los vectores $\mathbf{x}^*$, $\mathbf{y}^*$:

$$\begin{gathered}
\delta L / \delta \mathbf{x} = \mathbf{c} - \mathbf{yA} \leq \mathbf{0} \\
(\delta L / \delta \mathbf{x})\, \mathbf{x} = \mathbf{cx} - \mathbf{yAx} = \mathbf{0} \\
\mathbf{x} \geq \mathbf{0} \qquad\qquad (3.2.3) \\
\delta L / \delta \mathbf{y} = \mathbf{b} - \mathbf{Ax} \geq \mathbf{0} \\
\mathbf{y}(\delta L / \delta \mathbf{y}) = \mathbf{y}(\mathbf{b} - \mathbf{Ax}) = \mathbf{0} \\
\mathbf{y} \geq \mathbf{0}
\end{gathered}$$

Ahora, si el problema primal fuera el problema del mínimo:

$$\min_{y} G = yb$$

sujeta a

$$yA \geq c, \qquad y \geq 0 \qquad (3.2.4)$$

Según el teorema de Kuhn-Tucker, el vector fila y^* es una solución si existe un vector columna x^* tal que, la función lagrangiana se defina como:

$$L(x,y) = yb + (c - yA)x = yb + cx - yAx \qquad (3.2.5)$$

y las condiciones de Kuhn-Tucker se cumplen para los vectores x^*, y^* :

$$\delta L / \delta y = b - Ax \geq 0$$

$$y(\delta L / \delta y) = y(b - Ax) = 0$$

$$y \geq 0 \qquad (3.2.6)$$

$$\delta L / \delta x = c - yA \leq 0$$

$$(\delta L / \delta x)\, x = cx - yAx = 0$$

$$x \geq 0$$

Luego, no importa cual se tome como el problema primalo dual se tiene el mismo sistema de ecuaciones y el mismo número de variables, así que se tendrá el mismo resultado.

Enseguida, se dan tres teoremas fundamentales de la Programación Lineal, los cuales son cada uno necesarios y suficientes y están basados en el análisis de los multiplicadores de Lagrange y las condiciones Kuhn-Tucker:

1. Teorema de existencia.- Existe una solución en el problema de Programación Lineal sí y sólo sí los conjuntos solución tanto del problemaprimal como de su dual son no vacíos.

Para probar la primera parte de este teoremase debe demostrar que si existen vectores admisibles para los problemas primal y dual, entonces existe solución para ambos. Por tanto, si se toman las restricciones de desigualdad de losproblemas primal y dual:

$$\mathbf{Ax} \leq \mathbf{b} \tag{3.2.7}$$

$$\mathbf{yA} \geq \mathbf{c} \tag{3.2.8}$$

y se multiplica a la izquierda de (3.2.7) por el vector no negativo **y** , se llega a:

$$\mathbf{yAx} \leq \mathbf{yb} = G(\mathbf{y}) \tag{3.2.9}$$

y después se multiplica a la derecha de (3.2.8) por el vector no negativo **x**, se tiene:

$$F(\mathbf{x}) = \mathbf{cx} \leq \mathbf{yAx} \tag{3.2.10}$$

Luego, si **x** y **y** son vectores admisibles de los conjuntos solución de los respectivos problemas, se tiene:

$$F(\mathbf{x}) \leq G(\mathbf{y}) \tag{3.2.11}$$

o en palabras, el valor de la función objetivo en el problema de maximizar no puede superar el valor de la función objetivo en el problema dual de minimizar.

Así que, si se supone que $\mathbf{x}^o$ y $\mathbf{y}^o$ son vectores admisibles para ambos problemas, entonces, el conjunto solución para el problema primal es no vacío y la función objetivo es acotada, esto es:

$$F(\mathbf{x}) \leq G(\mathbf{y}^o) \text{ para todo vector admisible } \mathbf{x} \tag{3.2.12}$$

según el teorema de Weiestrass, existe una solución para el problemaprimal. De igual forma, para el problema dual, puesto que el conjunto solución contiene ay^o y la función objetivo es acotada:

$$F(x^o) \leq G(y) \text{ para todo vector admisible } y \qquad (3.2.13)$$

según el teorema de Weiestrass, el problema dual tiene una solución.

Ahora falta demostrar que la existencia de la solución a un problema de Programación Lineal implica que el conjunto solución tanto del problema mismo como de su dual son no vacíos, para esto, supóngase que x^* resuelve el problema del máximo; entonces, el problema del máximo tiene un vector admisible x^* y dado que la función objetivo es cóncava y las funciones de restricción convexas (por definición, las funciones lineales son a la vez cóncavas y convexas) y puesto que se cumple la condición de calificación de restricción, por las condiciones Kuhn-Tucker, si x^* resuelve el problema del máximo entonces existe un vector fila y^* que satisface las condiciones (3.2.3), i.e.:

$$y^*A \geq c, \qquad y^* \geq 0 \qquad (3.2.14)$$

de tal forma y^* es un vector admisible. Por tanto, existe una solución al problema de Programación Lineal sí, y sólo sí, los problemas primal y dual tienen vectores admisibles.

2. Teorema de dualidad.- Un vector admisible es una solución al problema de Programación Lineal sí y sólo sí existe un vector admisible para los problemas primaly dual para los cuales los valores de las funciones objetivo de ambos problemas son iguales.

Para demostrar este teorema, hay que probar que si el vector columnax^* es una solución para el problema del máximo, entonces existe un vector filay^* en el conjunto solución del problema dual tal que las funciones objetivo de ambos problemas valuados en x^* y y^*, respectivamente, tienen el mismo valor. Para esto, el vectory^* es admisible como se probo en el teorema anterior en (3.2.14), esto es:

$$y^*A \geq c, \qquad y^* \geq 0 \tag{3.2.15}$$

y de las condiciones Kuhn-Tucker en (3.2.3):

$$(c - y^*A)x^* = 0$$

$$y^*(b - Ax^*) = 0 \tag{3.2.16}$$

se llega a la siguiente igualdad por transitividad:

$$F(x^*) = cx^* = y^*Ax^* = y^*b = G(y^*) \tag{3.2.17}$$

de esta forma se puede probar la igualdad de los valores de la función objetivo. Análogamente, se puede probar este teorema para el caso en que el problema primal sea el problema del mínimo haciendo uso de (3.2.6).

Ahora, para demostrar que existen vectores admisibles para ambos problemas en que los valores de las funciones objetivo son iguales, entonces estos vectores resuelven los dos problemas, supóngase que x^* y y^* son vectores admisibles:

$$F(x^*) = cx^* = y^*b = G(y^*) \tag{3.2.18}$$

De (3.2.11), si x y y son admisibles:

$$F(x^*) \leq G(y^*) \tag{3.2.19}$$

dado que y^* es admisible, se tiene:

$$F(x) \leq G(y^*) \tag{3.2.20}$$

y por (3.2.18):

$$F(x) \leq F(x^*) \ \text{ para todo vector admisible } x \tag{3.2.21}$$

Por lo tanto, x^* es una solución del problema del máximo. De manera similar:

$$G(y^*) \leq G(y) \text{ para todo vector admisible } y \tag{3.2.22}$$

Con esto se prueba el teorema, i.e., si x^* es admisible, resuelve el problema del máximo sí, y sólo sí, existe un vector fila y^* admisible para el problema dual en que se cumpla (3.2.18). Ahora bien, al observar (3.2.18), (3.2.21) y (3.2.22) se tiene:

$$F(x) \leq F(x^*) = G(y^*) \leq G(y) \tag{3.2.23}$$

Esto es, mientras que F sea menor o igual que G, el maximizar F mediante la elección de x y el minimizar G mediante la elección de y hacen que el valor de F aumente y que el valor de G disminuya hasta que se igualan al momento que se llega a la solución del problema.

3. Teorema de holgura complementaria.- Los vectores admisibles x^*, y^* resuelven los problemas duales si satisfacen las condiciones de holgura complementaria:

$$(c - y^*A)x^* = 0$$

$$y^*(b - Ax^*) = 0 \tag{3.2.24}$$

Se puede explicar la necesidad de estas condiciones pues salen directamente de las condiciones Kuhn-Tucker. Con respecto a la suficiencia de este teorema, se puede verificar con el teorema de la dualidad, si se supone que x^*, y^* son admisibles, entonces de (3.2.24):

$$F(x^*) = cx^* = y^*Ax^* = y^*b = G(y^*) \tag{3.2.25}$$

puesto que los valores de las funciones objetivo son iguales, x^* y y^* son soluciones. Al desarrollar las condiciones de holgura complementaria, éstas requieren que:

$$(c_j - \sum_{i=1}^{m} a_{ij} y_i^*) x_j^* = 0 \qquad j = 1,\ 2,\ \ldots,\ n$$

$$y_i^* (b_i - \sum_{j=1}^{n} a_{ij} x_i^*) = 0 \qquad i = 1,\ 2,\ \ldots,\ m \qquad (3.2.26)$$

y si se desarrolla esto tomando en cuenta las restricciones de admisibilidad:

$$x_j^* \geq 0 \qquad \text{pero} = 0 \quad \text{si} \quad \sum_{i=1}^{m} a_{ij} y_i^* > c_j$$

$$\sum_{i=1}^{m} a_{ij} y_i^* \geq c_j \qquad \text{pero} = c_j \quad \text{si} \quad x_j^* > 0$$

$$\text{con} \quad j = 1, 2, \ldots, n \qquad (3.2.27)$$

$$y_i^* \geq 0 \qquad \text{pero} = 0 \quad \text{si} \quad \sum_{j=1}^{n} a_{ij} x_j^* < b_i$$

$$\sum_{j=1}^{n} a_{ij} x_j^* \leq b_i \qquad \text{pero} = b_i \quad \text{si} \quad y_i^* > 0$$

$$\text{con} \quad i = 1, 2, \ldots, m$$

Esto es, cualquier restricción que satisface la solución como desigualdad estricta, la variable dual correspondiente es cero en la solución y; si una variable es positiva en la solución entonces la restricción de desigualdad correspondiente en el problema dual se satisface como igualdad.

Note que es posible expresar las condiciones de admisibilidad y de holgura complementaria utilizando el teorema de holgura complementaria, donde losvectores $\mathbf{x}^*$ y $\mathbf{y}^*$ resuelven los problemas duales del máximo y del mínimo sí, y solamente sí:

$$\mathbf{Ax^*} \leq \mathbf{b}, \qquad \mathbf{x^*} \geq 0, \qquad \mathbf{y^*(b - Ax^*) = 0}$$

$$\mathbf{y^*A} \geq \mathbf{c}, \qquad \mathbf{y^*} \geq 0, \qquad \mathbf{(c - y^*A)x^* = 0} \qquad (3.2.28)$$

Si se suma el vector de variables de holgura $\mathbf{s} = (s_1, s_2, ..., s_m)'$ al problema del máximo y el vector fila de variables de holgura $\mathbf{r} = (r_1, r_2, ..., r_n)$ al problema del mínimo, de (3.2.28) se tiene que una solución al problema puede expresarse como:

$$\mathbf{A}\mathbf{x}^* + \mathbf{s}^* = \mathbf{b}, \qquad \mathbf{x}^* \geq 0, \qquad \mathbf{s}^* \geq 0, \qquad \mathbf{y}^*\mathbf{s}^* = 0$$

$$\mathbf{y}^*\mathbf{A} = \mathbf{c} + \mathbf{r}^* \qquad \mathbf{y}^* \geq 0, \qquad \mathbf{r}^* \geq 0, \qquad \mathbf{r}^*\mathbf{x}^* = 0 \tag{3.2.29}$$

en donde la no negatividad de cada una de las variables lleva a que se satisfagan las restricciones de desigualdad de ambos problemas y la anulación de los productos punto de las variables de holgura y de las variables duales determina que se satisfacen las condiciones de holgura complementaria.

3.3 Análisis de sensibilidad

Esta sección trata del análisis de sensibilidad y para su estudio se destaca la importancia de las variables duales, que son los multiplicadores de Lagrange del problema primal.

Luego, para este estudio supóngase que en la solución a los problemas duales (3.1.1) y (3.1.2) se satisfacen como igualdades m_1 de las m restricciones de desigualdad del problema del máximo mientras que $(m - m_1)$ se satisfacen como desigualdades estrictas, en tanto que n_1 de las n restricciones de desigualdad de los problemas del mínimo se satisfacen como igualdades y $(n - n_1)$ se satisfacen como desigualdades estrictas. Si es necesario, se renumeran otra vez las restricciones de tal forma quelas primeras m_1 restricciones del problema del máximo y las primeras n_1 del problema del mínimo sean las que se satisfagan como igualdades, sin olvidar que esta reordenación de las restricciones pide una permutación igual de las variables del problema dual. Por tanto, la matriz de coeficientes, los vectores fila y columna de los parámetros pueden particionarse de la siguiente manera:

$$\mathbf{c} = (\mathbf{c}^1 \quad \mathbf{c}^2)$$

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}^{11} & \mathbf{A}^{12} \\ \mathbf{A}^{21} & \mathbf{A}^{22} \end{pmatrix} \qquad \mathbf{b} = \begin{pmatrix} \mathbf{b}^1 \\ \mathbf{b}^2 \end{pmatrix} \qquad (3.3.1)$$

en donde $\mathbf{c}^1$ es el vector fila con n_1 entradas, $\mathbf{b}^1$ es un vector columna con m_1 elementos y $\mathbf{A}^{11}$ es la submatriz $m_1 x n_1$. Entonces los vectores fila y columna de las variables **x** y **y** se pueden particionar así:

$$\mathbf{x}' = (\mathbf{x}^1 \quad \mathbf{x}^2) \qquad \mathbf{y} = (\mathbf{y}^1 \quad \mathbf{y}^2) \qquad (3.3.2)$$

en donde $\mathbf{x}^1$ agrupa n_1 elementos e $\mathbf{y}^1$ contiene m_1 entradas. Conforme a las suposiciones iniciales de la sección, la solución de los problemas duales $\mathbf{x}^*$, $\mathbf{y}^*$ es:

$$\mathbf{A}^{11}\mathbf{x}^{1*} + \mathbf{A}^{12}\mathbf{x}^{2*} = \mathbf{b}^1$$

$$\mathbf{A}^{21}\mathbf{x}^{1*} + \mathbf{A}^{22}\mathbf{x}^{2*} < \mathbf{b}^2 \qquad (3.3.3)$$

$$\mathbf{y}^{1*}\mathbf{A}^{11} + \mathbf{y}^{2*}\mathbf{A}^{21} = \mathbf{c}^1$$

$$\mathbf{y}^{1*}\mathbf{A}^{12} + \mathbf{y}^{2*}\mathbf{A}^{22} > \mathbf{c}^2$$

Por el teorema de holgura complementaria se tiene:

$$\mathbf{x}^{1*} \geq 0, \qquad \mathbf{x}^{2*} = 0$$

$$\mathbf{y}^{1*} \geq 0, \qquad \mathbf{y}^{2*} = 0 \qquad (3.3.4)$$

de (3.3.4) se pueden expresar las igualdades de (3.3.3) como:

$$\mathbf{A}^{11}\mathbf{x}^{1*} = \mathbf{b}^1$$

$$\mathbf{y}^{1*}\mathbf{A}^{11} = \mathbf{c}^1 \qquad (3.3.5)$$

De (3.3.5) se puede concluir que ambos problemas tienen solución única si $\mathbf{A}^{11}$ es una matriz cuadrada y no singular, si así sucede:

$$(\mathbf{A}^{11})^{-1}\,\mathbf{b}^1 = \mathbf{x}^{1*}, \qquad \mathbf{x}^{2*} = 0$$

$$\mathbf{c}^1\,(\mathbf{A}^{11})^{-1} = \mathbf{y}^{1*}, \qquad \mathbf{y}^{2*} = 0 \qquad (3.3.6)$$

se resuelven los problemas primal y dual en términos de las $m_1n_1 + m_1 + n_1$ parámetros de $\mathbf{A}^{11}$, $\mathbf{b}^1$ y $\mathbf{c}^1$. Los valores óptimos correspondientes de las funciones objetivo son:

$$F\,(\mathbf{x}^*) = \mathbf{c}^1\mathbf{x}^{1*} = \mathbf{c}^1(\mathbf{A}^{11})^{-1}\,\mathbf{b}^1 = \mathbf{y}^{1*}\mathbf{b}^1 = G\,(\mathbf{y}^*) \qquad (3.3.7)$$

donde los valores óptimos de ambos problemas son iguales como se indica en el teorema de la dualidad y; $F\,(\mathbf{x}^*)$ y $G\,(\mathbf{y}^*)$ son funciones de los $m_1n_1 + m_1 + n_1$ parámetros del problema.

El análisis de sensibilidad que se hará se relaciona solamante con los efectos de cambiar estos parámetros tanto en las soluciones (3.3.6) como en los valores óptimos (3.3.7), para esto se separa en los siguientes casos:

1.- Considérese primero el efecto de cambiar $\mathbf{b}$ sobre el valor óptimo $F\,(\mathbf{x}^*)$:

$$\delta F\,(\mathbf{x}^*)\,/\,\delta\mathbf{b}^1 = \mathbf{c}^1\,(\mathbf{A}^{11})^{-1} = \mathbf{y}^{1*}, \quad \delta F\,(\mathbf{x}^*)\,/\,\delta\mathbf{b}^2 = 0 \qquad (3.3.8)$$

Así:

$$\delta F\,(\mathbf{x}^*)\,/\,\delta\mathbf{b} = \mathbf{y}^* \qquad (3.3.9)$$

De igual forma, considérese el efecto de cambiar **c** del problema dual:

$$\delta G\ (y^*)\ /\ \delta c^1 = (A^{11})^{-1}\ b^1 = x^{1*}, \quad \delta G\ (y^*)\ /\ \delta c^1 = 0 \qquad (3.3.10)$$

Así:

$$\delta G\ (y^*)\ /\ \delta c = x^* \qquad (3.3.11)$$

De manera que, la sensibilidad del valor óptimo de la función objetivo a los cambios en las constantes de restricción viene medida por el valor óptimo de lavariable dual correspondiente.

2.- Considérese el efecto de cambiar ligeramente el valor de la constante de restricción, tal que dual correspondiente es cero, i.e., si la restricción no es limitativa:

$$\delta x^{1*}\ /\ \delta b^2 = 0, \qquad \delta x^{2*}\ /\ \delta b^2 = 0$$

$$\delta y^{1*}\ /\ \delta c^2 = 0, \qquad \delta y^{2*}\ /\ \delta c^2 = 0 \qquad (3.3.12)$$

El valor óptimo de una función objetivo es independiente de una constante de restricción si la variable dual correspondiente es nula, o dicho de otra forma, alcambiar el valor de la constante de restricción no limitativa no afecta la solución del problema.

3.- Considérense las sensibilidades de las soluciones a los cambios en las constantes de restricción limitativas, esto se puede obtener derivando (3.3.6) así:

$$\delta x^{1*}\ /\ \delta b^1 = (A^{11})^{-1}$$

$$\delta y^{1*}\ /\ \delta c^1 = (A^{11})^{-1} \qquad (3.3.13)$$

se observa que los elementos correspondientes a estas matrices son iguales. Por tanto, las sensibilidades de los valores óptimos de la función objetivo a los cambios en las restricciones objetivo son:

$$\delta F(\mathbf{x}^*) / \delta \mathbf{c}^1 = (\mathbf{A}^{11})^{-1}\, \mathbf{b}^1 = \mathbf{x}^{1*}$$

$$\delta G(\mathbf{y}^*) / \delta \mathbf{b}^1 = \mathbf{c}^1 (\mathbf{A}^{11})^{-1} = \mathbf{y}^{1*} \qquad (3.3.14)$$

4.- Considérese ahora los efectos de los cambios en la matriz de coeficientes. Todos los elementos de **A** que no son los de la submatriz de $m_1 x n_1$ **A** no tienen efecto sobre los valores óptimos. En cuanto a la submatriz $\mathbf{A}^{11}$, si se deriva (3.3.7) se obtiene:

$$\delta F(\mathbf{x}^*) / \delta \mathbf{A}^{11} = -(\mathbf{A}^{11})^{-1}\, \mathbf{b}^1\, \mathbf{c}^1\, (\mathbf{A}^{11})^{-1} = \delta G(\mathbf{y}^*) / \delta \mathbf{A}^{11} \qquad (3.3.15)$$

si se utiliza (3.3.14) y se realiza el producto se tiene:

$$\delta F(\mathbf{x}^*) / \delta a_{ij} = -x_j^*\, y_i^* = \delta G(\mathbf{y}^*) / \delta a_{ij}$$

$$\text{para } i = 1, 2, \ldots, m_1; \quad j = 1, 2, \ldots, n_1 \qquad (3.3.17)$$

3.4 Descripción del algoritmo simplex

El algoritmo simplex es un método algebraico iterativo diseñado para resolver programas lineales. Se trata de un algoritmo inicialmente restringido, que parte de cualquier vértice del conjunto de oportunidades, que se supone no vacío. Se desplaza desde este vértice a un vértice vecino en una dirección a lo largo de la cual se incrementa la función objetivo, continua de esta forma de vértice en vértice hasta que se alcanza un vértice para el cual no exista incremento en el valor de la función objetivo cuando se desplaza a cualquier vértice vecino; este vértice es una solución global. Ahora bien, si un movimiento a cualquier vértice vecino hace decrecer el valor de la función objetivo entonces la solución es única; si un movimiento a algún vértice vecino no disminuye el valor de la función objetivo, entonces la solución no es única y todos estos vértices (y todos los puntos intermedios entre ellos) son soluciones. Dado que existe solamente un número finito de vértices del conjunto solución, el métodosimplex o bien halla la solución o bien indica que la función objetivo es ilimitada en un número finito de pasos.

El algoritmo simplex no sólo determina los valores óptimos de las variables, sino que también da información sobre el análisis de sensibilidad y permite una interpretación del problema.

Dado que este algoritmo es un proceso iterativo, generalmente se hace necesario de herramientas tales como las computadoras digitales para realizar los cálculos.

Ahora se expone un ejemplo de un sistema que involucra entre sus componentes a un sistema de producción de comodidad y un almacén.

Una compañía tiene como componentes en su sistema de producción, para cubrir su demanda: 2 unidades de producción y un depósito de almacenamiento del producto.

De acuerdo a las especificaciones de las unidades, en cuanto a su capacidad de producción y el costo de la materia prima para los próximos 5 meses, se estima que las capacidades y los costos de producción para este periodo son, las que se muestran en la tabla 1.

$t^k \backslash u_i$	u_1		u_2	
k	c_1^k	p_1^k	c_2^k	p_2^k
1	80	0.07	45	0.09
2	70	0.06	45	0.08
3	65	0.10	40	0.12
4	50	0.12	40	0.16
5	50	0.06	40	0.08

Tabla 1. Capacidades y costos de producción unitarios de cada unidad de producción para el periodo de estudio.

donde k es el número del mes bajo estudio, u_i la unidad de producción i, c_i^k y p_i^k son las capacidades y costos de producción en el periodo k, k =1, 2, 3, 4, 5, de la unidad de producción i, i = 1, 2.

Se cuenta con un deposito, donde se almacena el producto. La capacidad física de almacenamiento tiene como límites 110 y 0 unidades del producto en cualquier instante del tiempo. El flujo promedio del producto entre el depósito y las unidades de producción se calcula que tiene como límites físicos 65 unidades por mes.

El depósito se vacía cada 5 meses para su mantenimiento. Al principio del periodo de estudio se acababa de dar servicio al depósito.

Se cree que la demanda (s^k) del producto en los próximos 5 meses será:

k	s^k
1	90
2	50
3	80
4	100
5	95

Tabla 2. Demanda del producto en el periodo de estudio.

La política de la empresa tiene como principio que toda la demanda será satisfecha puntualmente. Si se desea optimizar el costo del sistema en este periodo ¿Cómo debe operar la compañía los próximos 5 meses?

Para resolver este problema, denótese por $v^k(*)$ la función de costo total a optimizar al tiempo t^k (la función objetivo) y; sea $v^K(x^K)$ la función de costo terminal, que describe las condiciones finales del sistema. El sistema se optimizará si se minimizan los costos del sistema en los periodos k+1, k+2, ..., K.

Una manera de escribir el programa, es mediante la parametrización de variables, para esto:

Sea π^k la tasa de producción promedio de la planta de producción al instante k (t^k), η^k la demanda que se satisface al tiempo k; entonces, $0 \leq \pi^k \leq c^k_1 + c^k_2$ y $0 \leq \eta^k \leq s^k$ y; $u^k = \eta^k - \pi^k$ la diferencia de, la demanda que se satisface y la tasa de producción de comodidad al tiempo k, para k = 0, 1, ..., 5;luego, $u^k \varepsilon [-(c^k_1 + c^k_2), s^k]$ y denótese por U^k_{max} y U^k_{min} los límites superior e inferior de u^k. A u^k se le llamará la tasa de flujo promedio neto al tiempo k. Nótese que la demanda puede ser mayor, menor o igual que el total de comodidad que produce la planta, la diferencia, $u^k = \eta^k - \pi^k$, se encuentra en la tasa de flujo de la conexión con el almacén.

Ahora bien, denótese por x^k el nivel de almacenamiento al tiempo k, entonces, para para k = 1, 2, ..., K, x^k queda determinado por:

$$x^k = x^{k-1} + u^k \Delta t^k \qquad (3.4.1)$$

donde Δt^k es la duración del periodo k, i.e., la diferencia en tiempo de k y k-1 ($\Delta t^k = t^k - t^{k-1}$, para k = 1, 2, ...,5 = K); los límites físicos del almacén restringen los valores de x^k, por tanto, esta variable deberá estar acotada por:

$$X^k_{min} \leq x^k \leq X^k_{max} \qquad (3.4.2)$$

donde X^k_{max} y X^k_{min} indican los límites físicos superior e inferior del almacén en el k-ésimo intervalo .

Para cumplir con la restricción de satisfacer la demanda, se tendrá que en elcaso de no satisfacer dicha demanda en el periodo correspondiente se incurrirá en un costo M; tal que M sea un número muy grande de tal formaque no exista posibilidad de tomar esta opción para optimizar el sistema; supóngase que M = 10.

De lo anterior, se puede identificar una función $f^k(u^k)$ que denota el costo de producción por unidad de tiempo en el k-ésimo intervalo de tiempo.

Por tanto, ahora el programa se puede escribir así:

$$v^{k-1}(x^{k-1}) = \min\left\{ \sum_{m=k}^{K} [f^m(u^m)\Delta t^m] + v^K(x^K) \right\}$$

sujeto a:

$$U^m_{min} \leq u^m \leq U^m_{max} \qquad (3.4.3)$$

$$x^m = x^{m-1} - u^m \Delta t^m$$

$$X^m_{min} \leq x^m \leq X^m_{max} \qquad \text{para } m = k, k+1, \ldots, K$$

Así que, se busca una sucesión $u^0, u^1, \ldots, u^K$ que optimice el sistema para el intervalo de tiempo en estudio.

Se ha establecido las condiciones generales para resolver el problema de Programación Lineal; así como también, la importancia de las condiciones Kuhn-Tucker que dan los supuestos suficientes y necesarios para caracterizar la solución del problema. Con estas condiciones se asegura que el valor óptimo de la función objetivo es exacto.

Al final del capítulo se mostró un problema típico de sistemas que involucran una planta de producción y un almacén. Este problema es posible plantearlo como uno de Programación Lineal, pero, este procedimiento se complica al ver involucrado un almacén, pues, enlaza las operaciones de cada periodo.

En el siguiente capítulo se da la teoría básica de la Programación Dinámicaque se necesitará conocer para desarrollar el método de optimización. En ese capítulo se establecen las condiciones necesarias y suficientes para caracterizar la solución del problema, mediante las propiedades del problema por etapas. Además, se interpretael ejemplo del sistema de producción con un almacén pero con Programación Dinámica.

Capítulo 4. Programación Dinámica

Ha transcurrido más de un cuarto de siglo desde que se publicó el libro deRichard Bellman (1957): Dynamic Programming. Con sulibro y otros artículos suyos se estimuló el dessarrollo de la Programación Dinámica y la importancia que tiene en la Investigación de Operaciones. Esta programación fue diseñada con la ideade facilitar los cálculos de algunos problemas de Programación Matemática. La idea, entonces, es descomponer el problema en problemas más pequeños de tal forma que sea más sencillo de resolver, esto es, el problema se resuelve particionándolo en etapas, en donde cada etapa tiene únicamente un vector (vector de control) a optimizar; después, se utiliza una función recurrente que enlaza los resultados de cada etapa yasí obtener un óptimo a todo el problema.

En sus principios la Programación Dinámica se inclinó a los problemas en losque la toma de decisiones tenía relación con el tiempo (por ejemplo: problemas de inventario), pero, con el desarrollo que ha tenido a través del tiempo laProgramación Dinámica ha empezado a emplearse en otras áreas en los quela toma de decisiones no está relacionado con el tiempo. Por esta razón, se ha propuesto un nombre alternativo, como lo comenta Hamdy A. Taha (1989), éste es Programación de Etapa Múltiple, puesto que el procedimiento se resuelve en etapas. Pero, como la teoría necesaria para desarrollar el método en este trabajo se basa enla toma de decisiones relacionada con el tiempo, únicamente se hará referencia a la variable tiempo para entender los conceptos de problemas de etapa múltiple.

La teoría de la Programación Dinámica se basa en el principio de optimalidad, que dice, según R. Bellman (1957)[1]:

1. Princple of Optimality: An optimal policy has the property that whatever the initial state and initial decision arethe remaining decisiona must constitute an optimal policy with regard to the state resulting from the first transition.

"una política óptima tiene la propiedad de que, cualesquiera que sean el estado y decisión iniciales, *las decisiones* restantes deben constituir una política óptima con respecto al estado resultante de *la* primera decisión".

4.1 Solución de Problemas de Optimización de Etapa Múltiple

Existen problemas, en los cuales el tiempo aparece como una variable discreta en lugar de continua y tales problemas pueden resolverse con Programación Dinámica.

En esos problemas de Optimización de Etapa Múltiple, en los que la variable tiempo toma los valores discretos:

$$t^0, t^1, t^2, \ldots, t^K \qquad (4.1.1)$$

El estado del sistema en el tiempo t^k, viene dado porel vector $\mathbf{x}^k$ y el control en el tiempo t^k por el vector $\mathbf{u}^k$. El estado en el tiempo t^{k+1} viene dado por la ecuación de continuidad:

$$\mathbf{x}^{k+1} = \mathbf{g}^k(\mathbf{x}^k, \mathbf{u}^k) \quad \text{para } k = 0, 1, 2, \ldots, K-1 \qquad (4.1.2)$$

donde $\mathbf{g}^k(*, *)$ es un vector de funciones continuamente diferenciables delestado actual y de las variables de control. El estado inicial es $\mathbf{x}^0$,que se supone conocido. La función objetivo es:

$$v = \min\left(\sum_{k=0}^{K-1} f^k(\mathbf{x}^k, \mathbf{u}^k) + F(\mathbf{x}^K, t^K)\right) \qquad (4.1.3)$$

donde $F(\mathbf{x}^K, t^K)$ se conoce como la condición de contorno (o función total terminal) y da las condiciones finales del sistema. Entonces, (4.1.3) se minimiza eligiendo una sucesión de vectores de control:

$$\{\mathbf{u}^0, \mathbf{u}^1, \ldots, \mathbf{u}^{K-1}\} \qquad (4.1.4)$$

sujeta a la restricción de que esta sucesión pertenezca a un conjunto de controles dado:

$$u^k \in U^k \quad \text{para} \quad k = 0, 1, 2, ..., K-1 \tag{4.1.5}$$

El enfoque por Programación Dinámica consiste en incorporar el problema a resolver en un conjunto de problemas caracterizado por ciertos parámetros y emplear, luego, el principio de optimalidad para obtener una relación de recurrencia. Si se toman como parámetros del problema de Optimización de Etapa Múltiple el estado $\mathbf{x}^k$ y el tiempo t^k, la función óptimal es:

$$v(\mathbf{x}^k,t^k)^* = v^k(\mathbf{x}^k)^* \tag{4.1.6}$$

que es el valor óptimo de la función objetivo para un problema que comienza en el estado $\mathbf{x}^k$ y en el tiempo t^k. Ahora bien, si se toman como parámetros el estado y el tiempo iniciales, la solución del problema es:

$$v^0(\mathbf{x}^0)^* \tag{4.1.7}$$

Del principio de optimalidad se tiene que:

$$v^k(\mathbf{x}^k)^* = \min_{u^k}[f^k(\mathbf{x}^k,\mathbf{u}^k) + v^{k+1}(\mathbf{x}^{k+1})^*] \tag{4.1.8}$$

esto es, el valor óptimo de la función objetivo, comenzandoen el estado $\mathbf{x}^k$ y tiempo t^k, consiste en la suma óptima de la cantidad añadida en el tiempo t^k, $f^k(\mathbf{x}^k, \mathbf{u}^k)$, y el valor óptimo restante, $v^{k+1}(\mathbf{x}^{k+1})^*$. De acuerdo a la ecuación (4.1.2), la relación de recurrencia es:

$$v^k(\mathbf{x}^k)^* = \min_{u^k}[f^k(\mathbf{x}^k,\mathbf{u}^k) + v^{k+1}(\mathbf{g}^k(\mathbf{x}^k,\mathbf{u}^k))^*] \tag{4.1.9}$$

Por lo tanto, se tiene de la condición de contorno que:

$$v^K(\mathbf{x}^K)^* = F(\mathbf{x}^K, t^K) \tag{4.1.10}$$

i.e., el valor óptimo de la función objetivo, a partir de $\mathbf{x}^K$ y t^K, es el valor de la función terminal en este estado y en este tiempo.

Otra forma de tratar el problema de Optimización de Etapa Múltiple consiste en caracterizar el problema no por el estado inicial y el tiempo inicial, sino por el estado inicial y por la cantidad de tiempo que resta por transcurrir en el problema. La función óptimal de este problema es:

$$v_k(x^{K-k})^* \tag{4.1.11}$$

Esto es, el valor óptimo de la función objetivo para un problema de longitud k, comenzando del estado x^{K-k}. La solución al problema para $k = K$ es $v_K(x^0)^*$. En este caso, la Programación Dinámica resuelve el problema retrocediendo desde el tiempo terminal t^K a través de una sucesión de soluciones. El primer término de esta sucesión es $v_0(x^K)^*$ que es el valor óptimo de la función objetivo para un problema de longitud cero partiendo y permaneciendo en x^K. Pero el valor óptimo para este problema es simplemente el valor de la función objetivo terminal.

$$v_0(x^K)^* = F(x^K, t^K) \tag{4.1.12}$$

Ahora, sea $v_1(x^{K-1})^*$ el valor óptimo de la función objetivo para el problema de longitud 1, iniciando en x^{K-1}, esto se conoce como la "primera etapa". Este problemade longitud 1, implica que la elección del vector de control u^{K-1}, se optimiza al maximizar la parte de la función objetivo relacionada con este tiempo $f^{K-1}(x^{K-1}, u^{K-1})$ más el valor óptimo del problema que parte en t^K:

$$v_1(x^{K-1})^* = \min_{u^{K-1}}[f^{K-1}(x^{K-1}, u^{K-1}) + v_0(x^K)^*] \tag{4.1.13}$$

o, con (4.1.2) se tiene:

$$v_1(x^{K-1})^* = \min_{u^{K-1}}[f^{K-1}(x^{K-1}, u^{K-1}) + v_0(g^{K-1}(x^{K-1}, u^{K-1}))^*] \tag{4.1.14}$$

Esta elección de control de la etapa uno, cumple conel principio de optimalidad, ya que el control u^{K-1} es óptimo con respecto al estado x^{K-1} resultante de las primeras K-1 elecciones del vector de control $u^0, u^1, \ldots, u^{K-2}$. De igual manera para la segunda etapa.

con dos unidades de tiempo por pasar:

$$v_2(x^{K-2})^* = \min_{u^{K-2}}[f^{K-2}(x^{K-2},u^{K-2}) + v_1(g^{K-2}(x^{K-2},u^{K-2}))^*] \qquad (4.1.15)$$

La relación de recurrencia para la etapa k es:

$$v_k(x^{K-k})^* = \min_{u^{K-k}}[f^{K-k}(x^{K-k},u^{K-k}) + v_{k-1}(g^{K-k}(x^{K-k},u^{K-k}))^*] \qquad (4.1.16)$$

El problema se resuelve con $v_K(x^0)^*$, último valor óptimo encontrado en la sucesión de problemas de optimización de una sola etapa, descritos por las ecuaciones (4.1.16) para k = 1, 2, ..., K, con la condición de contorno (4.1.12). El problema de Optimización de Etapa Múltiple, queda reducido, mediante este procedimieto matemático, a una sucesión de problemas de optimización de una sola etapa. Ejemplos de estos procedimientos se encuentran en: Bertsekas (1987), Denardo (1987), Hadley (1964), Intrilligator (1972), Taha (1989), Kaufmann (1967), Nemhauser (1966).

Enseguida se presenta un procedimiento alternativo para caracterizar la solución del problema de minimización con función objetivo seccionalmente lineal y convexa mediante las condiciones Kuhn-Tucker. Para esto, sea:

$$v^{k-1}(x^{k-1}) = \min\left\{\sum_{m=k}^{K}[f^m(u^m)\Delta t^m] + v^K(x^K)\right\}$$

sujeto a:

$$U^m_{min} \le u^m \le U^m_{max} \qquad (4.1.17)$$

$$x^m = x^{m-1} - u^m \Delta t^m$$

$$X^m_{min} \le x^m \le X^m_{max}$$

para m = k, k+1, ..., K

donde v^K (x^K) son las condiciones terminales del sistema, se supone seccionalmente lineal y convexa, U^m_{min}, U^m_{max}, X^m_{min}, X^m_{max}, para m= k, k+1, ..., K son las restricciones del sistema y $x^m = x^{m-1} - u^m \Delta t^m$ es la ecuación de continuidad del problema. Sea $f^m(*)$ una función seccionalmente lineal y convexa, de tal forma que se puede escribir como:

$$f^m(u^m) = f_0^m - \sum_{j=1}^{J^m} c_j^m \Delta u_j^m \qquad (4.1.18)$$

donde

$$u^m = U_0^m - \sum_{j=1}^{J^m} \Delta u_j^m \qquad (4.1.19)$$

y

$$0 \leq \Delta u_j^m \leq g_j^m \qquad \text{para } j = 1, 2, ..., J^m \qquad (4.1.20)$$

En el capítulo 5 se explicará al detalle esta formulación; por el momento suponga que lo anterior es posible representarlo así. Entonces (4.1.17) se puede escribir de lasiguiente forma:

$$v^{k-1}(x^{k-1}) = \min\left\{\sum_{m=k}^{K}[f_0^m - \sum_{j=1}^{J^m} c_j^m \Delta u_j^m]\Delta t^m + v^K(x^K)\right\}$$

sujeto a:

$$0 \leq \Delta u_j^m \leq g_j^m \qquad \text{para } j = 1, 2, ..., J^m \qquad (4.1.21)$$

$$x^m = x^{m-1} - \left[U_0^m + \sum_{j=1}^{J^m} \Delta u_j^m\right]\Delta t^m$$

$$X^m_{min} \leq x^m \leq X^m_{max}$$

para m = k, k+1, ..., K

Se puede observar que (4.1.21) es un programa lineal en que las condiciones Kuhn-Tucker caracterizan la solución, encontrando en forma simultánea todas las u^m y con un proceso de atraso se pueden encontrar las x^m. Ahora bien, suponga que $v^k(\cdot)$ es la función de costo total desde el tiempo K hasta k y que es seccionalmente linealy convexa (de modo que es posible denotarlo en forma semejante a $f^m(\cdot)$) yque (4.1.21) se quiere resolver con Programación Dinámica; por tanto, de las condiciones de optimalidad el problema se puede expresar como sigue:

$$v^{k-1}(x^{k-1}) = \min\left\{[f_0^m - \sum_{j=1}^{J^m} c_j^m \Delta u_j^m]\Delta t^m + [v_0^m - \sum_{i=1}^{I^m} d_i^m \Delta x_i^m]\right\}$$

sujeto a:

$$0 \le \Delta u_j^k \le g_j^k \qquad \text{para } j = 1, 2, \ldots, J^k \tag{4.1.22}$$

$$X_0^k + \sum_{i=1}^{I^k} \Delta x_i^k = x^{k-1} - \left[U_0^k + \sum_{j=1}^{J^k} \Delta u_j^k\right]\Delta t^k$$

$$0 \le \Delta x_i^k \le h_i^k \qquad \text{para } i = 1, 2, \ldots, I^k$$

donde

$$x^k = X_0^k + \sum_{i=1}^{I^k} \Delta x_i^k \tag{4.1.23}$$

y

$$v^k(x^k) = v_0^k - \sum_{i=1}^{I^k} d_i^k \Delta x_i^k \tag{4.1.24}$$

Esta notación se explica en el capítulo 5 sección 3 con más detenimiento.

Dado que se puede interpretar como un problema de Programación Lineal entonces la solución se puede caracterizar mediante las condiciones Kuhn-Tucker, para esto hágase lo siguiente: divida las restricciones de Δu_j^k y Δx_i^k, para $j = 1, 2, \ldots, J^k$, $i = 1, 2, \ldots, I^k$, respectivamente, como restricciones de no negatividad y acotadas superiormente, i.e.:

$$0 \leq \Delta u_j^k \qquad \Delta u_j^k \leq g_j^k \quad \text{para } j = 1, 2, \ldots, J^k$$
$$0 \leq \Delta x_i^k \qquad \Delta x_i^k \leq h_i^k \quad \text{para } i = 1, 2, \ldots, I^k \tag{4.1.25}$$

Luego, defina la función lagrangiana, así:

$L(\Delta u_1^k, \ldots, \Delta u_{J^k}^k, \Delta x_1^k, \ldots, \Delta x_{I^k}^k, \rho^k, \gamma_1^k, \ldots, \gamma_{J^k}^k, \theta_1^k, \ldots, \theta_{I^k}^k) =$

$$= [f_0^m - \sum_{j=1}^{J^m} c_j^m \Delta u_j^m] \Delta t^m + [v_0^m - \sum_{i=1}^{I^m} d_i^m \Delta x_i^m] + \rho^k \left[X_0^k + \sum_{i=1}^{I^k} \Delta x_i^k - x^{k-1} + \left[U_0^k + \sum_{j=1}^{J^k} \Delta u_j^k \right] \Delta t^k \right] -$$
$$- \sum_{j=1}^{J^k} \gamma_j^k \left(g_j^k - \Delta u_j^k \right) - \sum_{i=1}^{I^k} \theta_i^k \left(h_i^k - \Delta x_i^k \right)$$

sujeto a (4.1.26)

$$0 \leq \Delta u_j^k \quad \text{para } j = 1, 2, \ldots, J^k$$
$$0 \leq \Delta x_i^k \quad \text{para } i = 1, 2, \ldots, I^k$$

Derivando con respecto a las variables del lagrangiano:

$$\delta L / \delta \Delta x_i^k = -d_i^k + \rho^k + \theta_i^k \qquad \text{para } i = 1, 2, \ldots, I^k \tag{4.1.27}$$

$$\delta L / \delta \Delta u_j^k = -c_j^k \Delta t^k + \rho^k \Delta t^k + \gamma_j^k \qquad \text{para } j = 1, 2, \ldots, J^k \tag{4.1.28}$$

$$\delta L / \delta \rho^k = X_0^k + \sum_{i=1}^{I^k} \Delta x_i^k - x^{k-1} + \left[U_0^k + \sum_{j=1}^{J^k} \Delta u_j^k \right] \Delta t^k \tag{4.1.29}$$

$$\delta L / \delta \theta_i^k = h_i^k - \Delta x_i^k \qquad \text{para } i = 1, 2, \ldots, I^k \tag{4.1.30}$$

$$\delta L / \delta \gamma_j^k = g_j^k - \Delta u_j^k \qquad \text{para } j = 1, 2, \ldots, J^k \tag{4.1.31}$$

Entonces se deben cumplir las condiciones necesarias de primer orden en que las derivadas parciales (4.127)-(4.1.29) sean = 0 y que las de (4.1.30) y (4.1.31) sean ≤ 0. De acuerdo a las condiciones Kuhn-Tucker, las condiciones de holgura complementaria:

$$(\delta L / \delta \Delta x_i^k)\Delta x_i^k = 0, \qquad (\delta L / \delta \Delta u_j^k)\Delta u_j^k = 0$$

$$(\delta L / \delta \theta_i^k)\theta_i^k = 0, \qquad (\delta L / \delta \gamma_j^k)\gamma_j^k = 0 \qquad (\delta L / \delta \rho^k)\rho^k = 0 \tag{4.1.32}$$

Por tanto de estas anotaciones, para un óptimo se tiene que las condiciones de holgura complementaria son[2]:

para $j = 1, ..., J^k$:

$$\gamma_j^k \begin{cases} <= 0 & \Leftrightarrow \quad \Delta u_j^k = 0 \\ = 0 & \Leftrightarrow \quad 0 < \Delta u_j^k < g_j^k \\ >= 0 & \Leftrightarrow \quad \Delta u_j^k = g_j^k \end{cases} \tag{4.1.33}$$

y para $i = 1, ..., I^k$

$$\theta_i^k \begin{cases} <= 0 & \Leftrightarrow \quad \Delta x_i^k = 0 \\ = 0 & \Leftrightarrow \quad 0 < \Delta x_i^k < h_i^k \\ >= 0 & \Leftrightarrow \quad \Delta x_i^k = h_i^k \end{cases} \tag{4.1.34}$$

Para finalizar esta sección se expone el planteamiento de un sistema de producción con un almacén con Programación Dinámica. Para tal efecto se tomará el ejemplo mostrado en el capítulo anterior y la parametrización hecha, pero en este caso, se utiliza el principio de optimalidad para desarrollar una relación de recurrencia para $v^{k-1}(\cdot)$ una vez que se conoce $v^k(\cdot)$ para $k \varepsilon \{1, 2, ..., 5 = K\}$. Puesto que $v^K(x^K)$ es la función del costo total terminal (esta función se obtiene del contexto del problema), entonces, con la relación de recurrencia se puede caracterizar a la función $v^k(\cdot)$, para $k = K-1, ..., 1, 0$ y así, encontrar las tasas de flujo óptimo $u^0, u^1, ..., u^K$.

2. G. Hadley, (1964)

Entonces el programa se puede representar como:

$$v^{k-1}(x^{k-1}) = \min \{ f^k(u^k)\, \Delta t^k + v^k(x^k) \}$$

sujeto a:.

$$U^k_{min} \leq u^k \leq U^k_{max} \qquad (4.1.35)$$

$$x^k = x^{k-1} - u^k\, \Delta t^k$$

$$X^k_{min} \leq x^k \leq X^k_{max}$$

En donde se optimizan los costos de operación, de acuerdo a las restricciones físicas del sistema. Obsérvese como aquí, otra vez, aparece $x^k = x^{k-1} - u^k \Delta t^k$ ésta es la ecuación de continuidad del sistema.

4.2 La dimensionalidad de los estados en la Programación Dinámica

Supóngase que en el problema de Programación Dinámica, el vector de estados está conformado por n (>1) variables, en tal caso, se define el problema de Programación Dinámica con vector de estados multidimensional.

Ahora bien, un problema que se tiene en la Programación Dinámica es cuando se incrementa el número de variables de estados (o sea, la dimensión de **x**), pues, esto significa un mayor número de evaluaciones para las diferentes alternativas en cada etapa. Esto se puede observar, al hacer los cálculos de Programación Dinámica por computadora en que el aumento de variables de estados ocupa más memoria de la computadora y por tanto, aumenta el tiempo de cómputo; como ejemplo de este problema se tiene el estudio que hizo M. D. Intrilligator, sobre la memoria necesaria dela computadora:

"Las exigencias de almacenamiento temporal requiere Q^n localizaciones dememoria en la computadora, siendo Q el tamaño de la rejilla, o sea, el número de puntosdiscretos tomados por cada una de las variables de estado. Si, por ejemplo, cada variable de estado se divide en 100 puntos discretos y n = 4, entonces las necesidades de memeoria es de 100^4 de localizaciones."

A este problema se le conoce como el problema de la dimensionalidad o como lo llamó R. Bellman: "plaga de la dimensionalidad" y; como se puede observar es una desventaja que se tiene al resolver problemas de tamaño "mediano" o "grande" con Programación Dinámica.

A pesar de la basta teoría de la Programación Dinámica, en este capítulo sólo se hace mención a conceptos que se consideran importantes para comprender el mecanismo de resolver problemas de Investigación de Operaciones por medio de la Programación Dinámica. Se establece las condiciones Kuhn-Tuker como condiciones necesarias y suficientes para caracterizar la solución de un problema propuesto que más adelante será la base del método.

Además, se plantea un ejemplo con Programación Dinámica. Obsérvese que este ejemplo se va complicando en cuanto aumenta el número de estados en cada etapa, hasta llegar a presentar severas dificultades de resolución en tiempo y cálculo, esto es, existe la posibilidad de tener el problema de la dimensionalidad al buscar la solución del programa.

Nótese como el ejemplo planteado en el capítulo 3 con Programación Lineal y vuelto a plantear en este capítulo con Programación Dinámica presenta dificultadespara optimizar este tipo de problemas, pero es posible resolver este sistema si se toman en cuenta los beneficios de las dos programaciones y con esto eliminar, en lo posible, sus deficiencias.

También, en el capítulo se han expuesto algunos fundamentos de la Programación Dinámica que serán de ayuda para el desarrollo del método de optimización de sistemas lineales con almacenamiento. Es importante observar la trascendencia de las propiedades del problema por etapas.

Capítulo 5. Optimización de Sistemas Lineales con Almacenamiento

Hasta el momento se ha desarrollado la teoría de la Programación Matemática y la importancia de las condiciones Kuhn-Tucker en estos problemas; además, se han dado las bases de la Programación Lineal y Dinámica, también se ha dado un ejemplo de un problema de almacenamiento y producción. Con los fundamentos de los capítulos anteriores se desarrolla un método que permite optimizar sistemas lineales de baja memoria[1].

El manejo del almacenamiento de comodidades se hace de acuerdo a su costo y/o disponibilidad que cambian con el tiempo; así que, se tienen beneficios económicos al almacenar comodidades cuando:

- su costo es menor, o
- la demanda para ésta disminuye,

para uso en otras etapas cuando:

- el costo es mayor, o
- está en demanda.

Los sistemas considerados aquí consisten de un sólo almacén y un sistema de producción de comodidad. Luego, se desarrolla una técnica de optimización que combina conceptos de Programación Lineal y Dinámica, esta técnica optimiza la operación de almacenamiento resolviendo una sucesión de pequeños problemas de Programación Lineal y utilizando las características de problemas de una sola etapa. Este método resuelve problemas determinísticos con un sólo almacén.

1 Sistemas de baja memoria: el concepto se da más adelante.

5.1 Descripción del sistema

Los dos componentes de los cuales consiste el sistema a optimizar son los siguientes: un almacén y en general, el proceso de producción de comodidad de "baja memoria" (i.e., la producción de comodidad en un intervalo detiempo no se ve afectada por la producción en otro intervalo de tiempo), el cual se alimenta de material de dentro y fuera del almacén, como se muestra en la figura1. El periodo de tiempo bajo estudio, $t^0 \leq t \leq T$, se particiona en K intervalos, no necesariamente iguales la longitud de los intervalos. En cada intervalo se supone que la tasa de flujo de material, el estadode la producción de comodidad y las restricciones en los límites físicos del almacén son constantes.

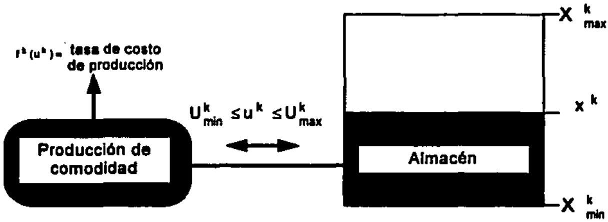


Figura 1. Modelo de producción de comodidad y almacenamiento

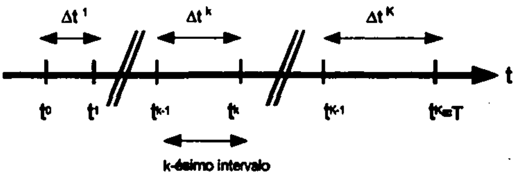
Así que, si la partición del periodo en estudio es:

$$t^0 < t^1 < ... < t^k < ... < t^K = T \qquad (5.1.1)$$

entonces, el k-ésimo intervalo de tiempo, $t^{k-1} \leq t < t^k$, para k = 1,..., K, tiene duración (longitud del intervalo) de:

$$\Delta t^k = t^k - t^{k-1} \tag{5.1.2}$$

como se muestra en la gráfica 1.



Gráfica 1. Partición del periodo T en K intervalos

El estado de almacén se denota por x^k, el nivel en el periodo k. Las restricciones en el almacén son:

$$X^k_{min} \leq x^k \leq X^k_{max} \tag{5.1.3}$$

donde los límites superior e inferior X^k_{max} y X^k_{min} pueden cambiar con el intervalo k.

La tasa de flujo promedio neta durante el k-ésimo intervalo es u^k. Los límites de la tasa de flujo promedio se expresan por:

$$U^k_{min} \leq u^k \leq U^k_{max} \tag{5.1.4}$$

Si el flujo neto durante el k-ésimo intervalo va del almacén al proceso de producción, entonces $u^k > 0$; si por otro lado, el proceso de producción suministra material al almacén, entonces $u^k < 0$.

Se puede establecer una "ecuación de continuidad" que relaciona el nivel de almacenamiento al principio del k-ésimo intervalo y la tasa durante este periodo con el nivel de almacenamiento al final del intervalo:

$$x^k = x^{k-1} - u^k \Delta t^k \qquad (5.1.5)$$

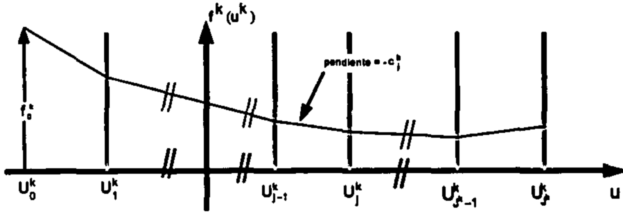
Habrá ocasiones en que valores de x^k serán no factibles, i.e., podrán existir algunas m > k tal que el nivel de almacenamiento para la sucesión de flujo permitido $u^{k+1},\ldots, u^m$ (i.e., $U^n_{min} \leq u^n \leq U^n_{max}$ para n = k+1,...,m), en el periodo t^m que se obtiene por la aplicación repetida de (5.1.5), x^m no obedecerá las restricciones de los niveles de almacenamiento (5.1.3), o sea:

$$x^m > X^m_{max} \quad ó \quad x^m < X^m_{min} \qquad (5.1.6)$$

Ahora bien, dado que la tasa de flujo neta de almacenamiento durante el k-ésimo intervalo es u^k, se denota por $f^k(u^k)$ el costo neto por unidad de tiempo incurrido en la producción de comodidad. El costo de producción total neto durante el k-ésimo intervalo es $f^k(u^k)\Delta t^k$.

El método para su desarrollo supone que para cada intervalo k = 1,..., K; la función de costo de producción es seccionalmente lineal y convexa con respecto a u^k, como se ilustra en la gráfica 2. Las líneas seccionadas en la pendiente de la función seetiquetan con U_j^k, j = 0, ..., J^k; donde:

$$U^k_{min} = U_0^k < U_1^k < \ldots < U_j^k < \ldots < U^k_{J^k} = U^k_{max} \qquad (5.1.7)$$



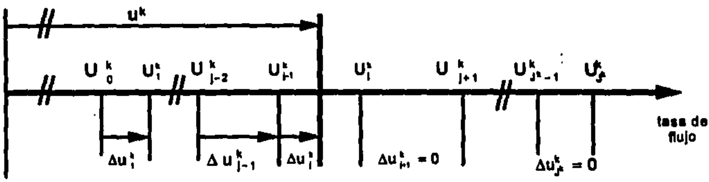
Gráfica 2. Función de costo de producción $f^k(^*)$

Se replantea ahora $f^k(u^k)$, esto se hace con la finalidad de ocupar en forma más conveniente las condiciones Kuhn-Tucker; para esto se definen las variables Δu^k_j, $j = 1,\ldots,J^k$ como las funciones de la tasa de flujo u^k:

$$\Delta u_j^k = \begin{cases} 0 & \text{si} \quad u^k \le U_{j-1}^k \\ u^k - U_{j-1}^k & \text{si} \quad U_{j-1}^k < u^k \le U_j^k \\ g_j^k & \text{si} \quad U_j^k < u^k \end{cases} \qquad (5.1.8)$$

(ver la gráfica 3) donde:

$$g_j^k = U_j^k - U_{j-1}^k \qquad (5.1.9)$$



Gráfica 3. Definición de Δu^k_j

La variable de control de almacenamiento u^k, ahora, se puede escribir como:

$$u^k = U_0^k + \sum_{j=1}^{J^k} \Delta u_j^k \qquad (5.1.10)$$

Las ecuaciones (5.1.8) y (5.1.10) muestran que las variables Δu^k_j, $j = 1,\ldots,J^k$ son representaciones equivalentes de la tasa de flujo u^k y; por tanto, se puede hacer un cambio de variables.

Como se muestra en la gráfica 2, se denota:

Como se muestra en la gráfica 2, se denota:

$$f_0^k = f^k(U^k_{min}) \tag{5.1.11}$$

Sea c_j^k la pendiente negativa de la función de costo de producción $f^k(\bullet)$ en el j-ésimo segmento ($U_{j-1}^k < u^k < U_j^k$), esto es:

$$c_j^k = -\left(f^k(U_j^k) - f^k(U_{j-1}^k)\right) / \left(U_j^k - U_{j-1}^k \right) \tag{5.1.12}$$

De lo anterior, la función costo de producción también se puede denotar como:

$$f^k(u^k) = f_0^k - \sum_{j=1}^{J^k} c_j^k \Delta u_j^k \tag{5.1.13}$$

Dado el supuesto de convexidad de la función de costo de producción, se tiene que c^k_j decrece en cuanto aumenta j, esto es:

$$c^k_j > c^k_{j+1} \qquad \text{para } j = 1,\ldots, J^k-1 \tag{5.1.14}$$

Para optimizar el sistema son necesarias las condiciones inicial y final, que en general se deben conocer. Por tanto, dadas las condiciones iniciales, es posible conocer el nivel de almacenamiento al principio del intervalo de estudio.

Ahora, se necesita una función de costo terminal denotada por $v^K(\bullet)$ para describir las condiciones finales con la restricción en los niveles de almacenamiento permitidos $X^K_{min} \leq x^K \leq X^K_{max}$. Al igual que $f^k(\bullet)$ se supondrá que la función de costo terminal es seccionalmente lineal y convexa, así que, se puede representar en forma similar a la función de costo de producción $f^k(\bullet)$ en (5.1.13).

Aún cuando, se pretende que el sistema sea lo más general posible, es necesario para la optimización del mismo suponer lo siguiente:

1.- El problema es determinístico en el k-ésimo intervalo de tiempo, esto es, todos los valores subsecuentes de la función de costo de producción $f^m(\bullet)$ y límites U^m_{min}, U^m_{max}, X^m_{min} y X^m_{max} son conocidos, para $m = k+1,\ldots, K$;

2.- El costo de producción no depende del nivel de almacenamiento;

3.- Las pérdidas del almacén no dependen del nivel de almacenamiento;

4.- Hay únicamente un recurso de almacenamiento.

En la formulación anterior, la operación de producción de comodidad se presenta por la función de costo de producción $f^k(*)$ y los límites en la tasa de flujo u^k. Con este modelo es posible capturar una amplia variedad de sistemas de producción de comodidad formada por conexiones de componentes con:

a. Límites en la tasa de flujo en la salida, entrada o en ambas partes;

b. Costos variables, los cuales pueden cambiar como funciones lineales o seccionalmente lineales en la salida, entrada o en ambas partes; y

c. Las características operacionales en cada intervalo de tiempo no se ven afectadas por operaciones en otros intervalos de tiempo.

Para una tasa de flujo dado desde el almacén, u^k, una solución de costo menor en la optimización del problema de almacenamiento requiere que los costos de operación en el proceso de producción también se minimicen. Para cada valor u^k en el conjunto factible, $f^k(u^k)$ es la suma de las variables de costo de cada componente deproducción, dado esto, la comodidad se maneja de tal manera que minimice la variable de costo total, en donde la tasa de flujo del almacén es u^k. Luego, cálculos de la función $f^k(*)$ requerirá la solución de una familia de problemas de optimización separados, parametrizada por una tasa de flujo u^k. Enseguida se presenta un ejemplo que muestra en forma general el sistema a optimizar.

1.- El proceso de producción consiste de la demanda de un producto en particular y N unidades de producción, cada una con diferentes capacidades de producción y tasas de costo. Se supone que la tasa de la variable de costo de cada unidad no varía con el tiempo de salida, i.e., la tasa de costo se conserva igual en esa partición del periodo. La demanda varía sobre el tiempo, pero de manera seccionalmente constante, esto es, la demanda no cambia en cada partición del intervalo en estudio. También se supone que la demanda debe ser satisfecha por completo en cada partición del intervalo. La demanda puede ser menor, igual o mayor que el total de unidades de producción de salida, la diferencia se encuentra en la tasa de flujo con el almacén. Este modelo puede representar un sistema de fuerza eléctrica con una sola bomba para almacenamiento.

Para un k dado y u^k, la carga total en las unidades de producción, π^k, se puede hallar como la diferencia entre la demanda y u^k. La suma de las variables de costo de las unidades de producción se minimiza, al utilizar las unidades por orden meritorio, esto es, se comienza con la unidad con la menor tasa en la variable de costo, las unidades se utilizan en orden creciente de la tasa de la variable de costo hasta que el total de salida es π^k. Si u^k se incrementa ligeramente, tal que π^k sea ligeramente reducida, $f^k(u^k)$ será reducida en la misma tasa de la variable de costo de la última unidad que se utilizó (la unidad marginal). De aquí que, la función de costo de producción es seccionalmente lineal.

La tasa máxima de flujo de salida del almacén, U^k_{max}, será determinado por el mínimo del proceso de producción y la capacidad de flujo de salida de la conexión entre el almacén y el proceso de producción.

La tasa mínima de flujo, U^k_{min}, se determina por la capacidad disponible (i.e., la capacidad total de las N unidades menos la demanda) o la capacidad de entrada de flujo de la conexión.

Cuando la función costo de producción es más compleja en cuanto a sus restricciones (conservando siempre la linealidad de éstas) este problema puede resolverse con la parametrización de las variables a optimizar, como lo comenta Taha p. 307 (1991).

5.2 Formulación del sistema con Programación Lineal

Suponga cualquier periodo t^{k-1} para algún $k \varepsilon \{1, 2, ..., K\}$, y que el nivel de almacenamiento es x^{k-1}. Lo que se pretende es encontrar una sucesión $u^k, u^{k+1}, ..., u^K$ de tasa de flujo la cual minimice los costos totales de operación hasta el tiempo t^K. La función de costo terminal $v^K(*)$ se supone dada. El problema se puede expresar como un programa matemático, así:

$$v^{k-1}(x^{k-1}) = \min\left\{\sum_{m=k}^{K}[f^m(u^m)\Delta t^m] + v^K(x^K)\right\}$$

sujeto a:

$$U^m_{min} \leq u^m \leq U^m_{max} \qquad (5.2.1)$$

$$x^m = x^{m-1} - u^m \Delta t^m$$

$$X^m_{min} \leq x^m \leq X^m_{max}$$

para $m = k, k+1, ..., K$

Usando (5.1.8), (5.1.10) y (5.1.13) para expresar u^m y $f^m(u^m)$ en términos de Δu_j^m para $j = 1, 2, ..., J^m$, (5.2.1) se puede sustituir por:

$$v^{k-1}(x^{k-1}) = \min\left\{\sum_{m=k}^{K}[f_0^m - \sum_{j=1}^{J^m} c_j^m \Delta u_j^m]\Delta t^m + v^K(x^K)\right\}$$

sujeto a:

$$0 \leq \Delta u_j^m \leq g_j^m \qquad \text{para } j = 1, 2, \ldots, J^m \qquad (5.2.2)$$

$$x^m = x^{m-1} - \left(U_0^m + \sum_{j=1}^{J^m} \Delta u_j^m \right) \Delta t^m$$

$$X^m_{min} \leq x^m \leq X^m_{max}$$

$$\text{para } m = k, k+1, \ldots, K$$

La sucesión en la tasa de flujo óptima u^m para $m = k, k+1, \ldots, K$, se da como:

$$u^m = U_0^m + \sum_{j=1}^{J^m} \Delta u_j^m \qquad (5.2.3)$$

donde Δu_j^m resuelve a (5.2.2).

La ecuación (5.2.2) es un programa lineal tal que las condiciones Kuhn-Tucker caracterizan completamente su solución. El número de restricciones en (5.2.2) es proporcional al número de periodos K - k. Dado que el tiempo para solucionar programas lineales son aproximadamente proporcionales al cubo del número de restricciones[2], hay severas restricciones en el número de periodos de los que puede componerse si este programa lineal se resuelve directamente.

En la sección siguiente, el programa lineal (5.2.2) se reformula como un programa dinámico de una sola etapa para manejar la optimización del costo de la producción como un subproblema.

2. G.L. Nemhauser (1966)

5.3 Formulación del sistema con Programación Dinámica

Ahora, se plantea el sistema como uno de Programación Dinámica, el principio de optimalidad se usa para desarrollar una relación para la función de costo total $v^{k-1}(*)$ para $k = 1, 2, ..., K$ arbitraria, definida en (5.2.2) en términos de la función de costo total $v^k(*)$ para el siguiente intervalo. Se inicia con la función de costo total terminal $v^K(*)$, ya que se supone dada, y al utilizar recurrentemente la relación antes mencionada, se encuentra $v^k(*)$ para $k = K-1, ... 1, 0$. La sucesión óptima de las tasasde flujo $u^0, u^1, ..., u^K$ se resuelve utilizando la fase de avance.

La relación recurrente entre $v^{k-1}(x^{k-1})$ y $v^k(x^k)$ es:

$$v^{k-1}(x^{k-1}) = \min\left\{[f_0^k - \sum_{j=1}^{J^k} c_j^k \Delta u_j^k]\Delta t^k + v^k(x^k)\right\}$$

sujeto a:

$$0 \le \Delta u_j^k \le g_j^k \qquad \text{para } j = 1, 2, ..., J^k \qquad (5.3.1)$$

$$x^k = x^{k-1} - \left(U_0^k + \sum_{j=1}^{J^k} \Delta u_j^k\right)\Delta t^k$$

$$X^k_{min} \le x^k \le X^k_{max}$$

Luego, la fase de avance consiste en calcular la u^k óptima una vez que se conoce x^{k-1}. Esto se puede hacer así:

$$u^k = U_0^k + \sum_{j=1}^{J^k} \Delta \underline{u}_j^k \qquad (5.3.2)$$

donde $\Delta \underline{u}_j^k$ resuelve a (5.3.1).

La solución del programa dinámico se puede resolver como un programa matemático, si se conoce $v^k(\cdot)$. Pero si se resuelveel problema tomando en cuenta las siguientes proposiciones, es posible también encontrar el óptimo del problema.

Proposición 1. Si las funciones costo de producción $f^k(\cdot)$ para k = 1, 2, ..., K y la función de costo total terminal $v^K(\cdot)$ son seccionalmente lineales y convexas, entonces para cada k = K-1, ..., 1, 0 la función de costo total de la Programación Dinámica $v^k(\cdot)$ es también seccionalmente lineal y convexa.

Para probar esto, sean $0 \leq \alpha, \beta \leq 1$ tal que:

$$v^{k-1}(\cdot) = f^k[\alpha u + (1 - \alpha)w] + v^k[\beta x + (1 - \beta)y] \tag{5.3.3}$$

Dado que $f^k(\cdot)$ y $v^k(\cdot)$ son convexas y van de los números reales a los reales, se tiene por (1.2.13):

$$f^k[\alpha u + (1 - \alpha)w] \leq \alpha f^k(u) + (1 - \alpha) f^k(w) \tag{5.3.4}$$

$$v^k[\beta x + (1 - \beta)y] \leq \beta v^k(x) + (1 - \beta) v^k(y) \tag{5.3.5}$$

luego, al sumar (5.3.4) y (5.3.5) se llega a:

$$v^{k-1}(\cdot) = f^k[\alpha u + (1 - \alpha)w] + v^k[\beta x + (1 - \beta)y] \leq$$

$$\leq \alpha f^k(u) + (1 - \alpha) f^k(w) + \beta v^k(x) + (1 - \beta) v^k(y) \tag{5.3.6}$$

Dado que $f^k(\cdot)$ y $v^k(\cdot)$ son seccionalmente lineales su suma por segmentos dara otra función lineal, por tanto $v^{k-1}(\cdot)$ es seccionalmente lineal y convexa.

De la proposición 1, se tiene que para cada k = 0, 1, ..., K las funciones $v^k(\cdot)$se pueden expresar en forma similar a la utilizada para $f^k(\cdot)$ en (5.1.7) - (5.1.1 4).

Como se muestra en la gráfica 4, las líneas seccionadas se etiquetan con X_i^k,

$i = 0, 1, ..., I^k$, donde:

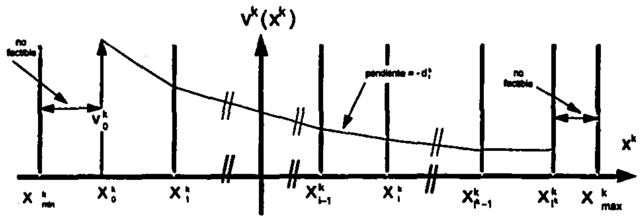
$$X_{min}^k \le X_0^k < X_1^k < ... < X_i^k < ... < X_{I^k}^k \le X_{max}^k \qquad (5.3.7)$$

Las regiones $X_{min}^k \le x^k \le X_0^k$ y $X_{I^k}^k \le x^k \le X_{max}^k$ son regiones no factibles. Las variables Δx_i^k, $i = 1, 2, ..., I^k$ se definen como funciones de almacenamiento de x^k por:

$$\Delta x_i^k = \begin{cases} 0 & \text{si} \quad x^k \le X_{i-1}^k \\ x^k - X_{i-1}^k & \text{si} \quad X_{i-1}^k < x^k \le X_i^k \\ h_i^k & \text{si} \quad X_i^k < x^k \end{cases} \qquad (5.3.8)$$

donde:

$$h_i^k = X_i^k - X_{i-1}^k \qquad (5.3.9)$$



Gráfica 4. Función de costo total $v^k(*)$

Luego, x^k se puede expresar así:

$$x^k = X_0^k + \sum_{i=1}^{I^k} \Delta x_i^k \qquad (5.3.10)$$

Las ecuaciones (5.3.8) y (5.3.10) establecen que la variable Δx_i^k, $i = 1, 2, ..., I^k$ son una representación equivalente del nivel de almacenamiento x^k.

Denótese, como se muestra en la gráfica 4:

$$v_0^k = v^k(X_0^k) \tag{5.3.11}$$

Sea d_i^k la pendiente negativa de la función de costo total $v^k(^*)$ en el i-ésimo segmento ($X_{i-1}^k < x^k < X_i^k$), o sea:

$$d_i^k = -(v^k(X_i^k) - v^k(X_{i-1}^k)) / (X_i^k - X_{i-1}^k) \tag{5.3.12}$$

De aquí que la función de costo total se puede escribir como:

$$v^k(x^k) = v_0^k - \sum_{i=1}^{I^k} d_i^k \Delta x_i^k \tag{5.3.13}$$

Como se estableció antes, dado el supuesto sobre la convexidad de esta función implica que d_i^k decrece cuando aumenta i. Por lo tanto:

$$d_i^k > d_{i+1}^k \quad \text{para } i = 1, 2, \ldots, I^k-1 \tag{5.3.14}$$

La siguiente proposición establece las condiciones necesarias y suficientes de Kuhn-Tucker para la solución del problema de Programación Dinámica de una sola etapa.

Proposición 2. Para cada $k \in \{1, \ldots, K\}$, son condiciones necesarias y suficientes para la solución del problema de una sola etapa de Programación Dinámica en (5.3.1) que existan: ρ^k, Δu_j^k para $j = 1, \ldots, J^k$ y Δx_i^k para $i = 1, \ldots, I^k$ tal que estas tres condiciones se cumplan:

I. Para $j \in \{1, \ldots, J^k\}$

$$\begin{array}{lll} \text{si } c_j^k < \rho^k & \text{entonces} & \Delta u_j^k = 0 \\ \text{si } c_j^k = \rho^k & \text{entonces} & 0 < \Delta u_j^k < g_j^k \\ \text{si } c_j^k > \rho^k & \text{entonces} & \Delta u_j^k = g_j^k \end{array} \tag{5.3.15}$$

II. Para $i \varepsilon \{1, ..., I^k\}$

$$\begin{array}{llll} \text{si} & d_i^k < \rho^k & \text{entonces} & \Delta x_i^k = 0 \\ \text{si} & d_i^k = \rho^k & \text{entonces} & 0 < \Delta x_i^k < h_i^k \\ \text{si} & d_i^k > \rho^k & \text{entonces} & \Delta u_j^k = h_i^k \end{array} \qquad (5.3.16)$$

III. La relación de continuidad para el intervalo k, esto es:

$$X_0^k + \sum_{i=1}^{I^k} \Delta x_i^k + \left(U_0^k + \sum_{j=1}^{J^k} \Delta u_j^k \right) \Delta t^k = x^{k-1} \qquad (5.3.17)$$

Es posible probar esto, si se utiliza las ecuaciones (5.3.10) y (5.3.13) para sustituir a x^k y $v^k(x^k)$ respectivamente, entonces el programa (5.3.1)se puede escribir como sigue:

$$v^{k-1}(x^{k-1}) = \min\left\{ [f_0^k + \sum_{j=1}^{J^k} c_j^k \Delta u_j^k]\Delta t^k + [v_0^k - \sum_{i=1}^{I^k} d_i^k \Delta x_i^k] \right\}$$

sujeto a:

$$0 \leq \Delta u_j^k \leq g_j^k \qquad \text{para } j = 1, 2, ..., J^k \qquad (5.3.18)$$

$$X_0^k + \sum_{i=1}^{I^k} \Delta x_i^k = x^{k-1} - \left(U_0^k + \sum_{j=1}^{J^k} \Delta u_j^k \right) \Delta t^k$$

$$0 \leq \Delta x_i^k \leq h_i^k \qquad \text{para } i = 1, ..., I^k$$

Ahora bien, las condiciones Kuhn-Tucker son que existen Δu_j^k, γ_j^k para $j = 1, ..., J^k$, Δx_i^k, θ_i^k para $i = 1, ..., I^k$ y ρ^k tal que:

$$\{-c_j^k \Delta t^k + \rho^k \Delta t^k + \gamma_j^k = 0 \quad \text{para} \quad j = 1, \ldots, J^k$$

$$-d_i^k + \rho^k + \theta_i^k = 0 \qquad \text{para} \quad i = 1, \ldots, I^k \tag{5.3.19}$$

$$X_0^k + \sum_{i=1}^{I^k} \Delta x_i^k - x^{k-1} + \left(U_0^k + \sum_{j=1}^{J^k} \Delta u_j^k \right) \Delta t^k = 0\}$$

con las condiciones de holgura complementarias para $j = 1, \ldots, J^k$:

$$\gamma_j^k \begin{cases} \leq 0 & \Leftrightarrow \quad \Delta u_j^k = 0 \\ = 0 & \Leftrightarrow \quad 0 < \Delta u_j^k < g_j^k \\ \geq 0 & \Leftrightarrow \quad \Delta u_j^k = g_j^k \end{cases} \tag{5.3.20}$$

y para $i = 1, \ldots, I^k$

$$\theta_i^k \begin{cases} \leq 0 & \Leftrightarrow \quad \Delta x_i^k = 0 \\ = 0 & \Leftrightarrow \quad 0 < \Delta x_i^k < h_i^k \\ \geq 0 & \Leftrightarrow \quad \Delta x_i^k = h_i^k \end{cases} \tag{5.3.21}$$

Dado que $\Delta t^k > 0$, hay que demostrar que (5.3.15)-(5.3.17) son equivalentes a (5.3.19)-(5.3.21). Para esto, se analiza cada caso:

Para I se tiene:

Si $c_j^k < \rho^k$ de (5.3.19) entonces $\gamma_j^k < 0$; ahora bien de (5.3.20) se tiene que $\gamma_j^k \leq 0$ sí y sólo sí $\Delta u_j^k = 0$. Por lo tanto:

si $c_j^k < \rho^k$ entonces $\Delta u_j^k = 0$

Si $c_j^k = \rho^k$ de (5.3.19) entonces $\gamma_j^k = 0$; ahora bien de (5.3.20) se tiene que $\gamma_j^k = 0$ sí y sólo sí $0 < \Delta u_j^k < g_j^k$. Por lo tanto:

si $c_j^k = \rho^k$ entonces $0 < \Delta u_j^k < g_j^k$

Si $c_j^k > \rho^k$ de (5.3.19) entonces $\gamma_j^k > 0$; ahora bien de (5.3.20) se tiene que $\gamma_j^k \geq 0$ sí y sólo sí $\Delta u_j^k = g_j^k$. Por lo tanto:

si $c_j^k > \rho^k$ entonces $\Delta u_j^k = g_j^k$

Para II se tiene:

Si $d_i^k < \rho^k$ de (5.3.19) entonces $\theta_i^k < 0$; ahora bien de (5.3.21) se tiene que $\theta_i^k \leq 0$ sí y sólo sí $\Delta x_i^k = 0$. Por lo tanto:

si $d_j^k < \rho^k$ entonces $\Delta x_i^k = 0$

Si $d_i^k = \rho^k$ de (5.3.19) entonces $\theta_i^k = 0$; ahora bien de (5.3.21) se tiene que $\theta_i^k = 0$ sí y sólo sí $0 < \Delta x_i^k < h_i^k$. Por lo tanto:

si $d_i^k = \rho^k$ entonces $0 < \Delta x_i^k < h_i^k$

Si $d_i^k > \rho^k$ de (5.3.19) entonces $\theta_i^k > 0$; ahora bien de (5.3.21) se tiene que $\theta_i^k \geq 0$ sí y sólo sí $\Delta x_i^k = h_i^k$. Por lo tanto:

si $d_i^k > \rho^k$ entonces $\Delta x_i^k = h_i^k$

Para III se puede observar que al sumar $-x^{k-1}$ a ambos miembrosde la ecuación (5.3.17) se llega a la ecuación de continuidad en (5.3.19). Con esto queda demostrada la proposición.

Proposición 3. Para $k = 1, ..., K$ $\delta v^{k-1}(x^{k-1}) / \delta x^{k-1} = - \rho^k$ casi en cualquier lugar.

Para probar esto, tómese en cuenta (5.3.18) ysustituya a $v^{k-1}(x^{k-1})$ y x^{k-1} por sus respectivas igualdades, semejantes a las de (5.3.10) y (5.3.13); entonces, se puede derivar parcialmente con respecto a cada una de las variables que constituye a x^{k-1}, o sea:

$$\delta\, v^{k-1}(x^{k-1}) \,/\, \delta\Delta x_i^k = -\, d_i^k \qquad \text{para } i = 1, \ldots, I^k$$

$$\delta\, v^{k-1}(x^{k-1}) \,/\, \delta\Delta u_j^k \Delta t^k = -\, c_j^k \qquad \text{para } j = 1, \ldots, J^k \qquad (5.3.22)$$

Ahora bien, de las restricciones del problema ($0 \leq \Delta x_i^k \leq h_i^k$ y $0 \leq \Delta u_j^k \leq g_j^k$), la igualdad de la proposición se cumplirá sólo cuando las condiciones de Δx_i^k y Δu_j^k sean desigualdades estrictas en (5.3.20) y (5.3.21); esto es, las condiciones de holgura complementaria cumplan que $\gamma_j^k = \theta_i^k = 0$ si $0 < \Delta x_i^k < h_i^k$ y $0 < \Delta u_j^k < g_j^k$; luego, tomando en cuenta (5.3.19), se sustituyen el valor de d_i^k para $i \in \{1, \ldots, I^k\}$ ó c_j^k para $j \in \{1, \ldots, J^k\}$ por ρ^k y se obtiene la igualdad propuesta.

Luego, el valor marginal de un incremento adicional del contenido de almacenamiento en el tiempo t^{k-1} es dado por el multiplicador de continuidad de almacenamiento ρ^k.

Se tiene de lo anterior una interpretación de los multiplicadores de continuidad de almacenamiento como sigue: de las condiciones Kuhn-Tucker para (5.2.1) ρ^k sólo cambia cuando la trayectoria de almacenamiento choca con un límite, esto es:

a. si $x^k = X_{min}^k$ entonces $\rho^{k+1} \leq \rho^k$

b. si $X_{min}^k < x^k < X_{max}^k$ entonces $\rho^{k+1} = \rho^k$ (5.3.23)

c. si $x^k = X_{max}^k$ entonces $\rho^{k+1} => \rho^k$

En palabras, se puede notar que la siguiente etapa para:

a. si el nivel del almacén es X_{min}^k en el tiempo t^k implica que el valor marginal de un incremento adicional del contenido del almacén para el tiempo t^k es mayor o igual que al tiempo t^{k-1};

b. si el nivel del almacén se mantiene dentro de los límites, los costos marginales por incremento del nivel del almacenamiento sonlos mismos en los periodos k y k-1;

c. si el nivel del almacén topa con el máximo nivel permisible en t^k, entonces, indica que la producción en ese periodo es más económica que en el periodo anterior.

5.4 El algoritmo solución

Para llegar a la parte central de este trabajo ha sido necesario entender y conocer la teoría fundamental de la Programación Matemática, puesto que, aún cuando el algoritmo se basa prácticamente en las proposiciones anteriores, para utilizarlas fue necesario comprender la teoría que las sustentan.

El algoritmo genera la función de costo total $v^{k-1}(\bullet)$ dado una función de costo total $v^k(\bullet)$ y una función costo de producción $f^k(\bullet)$, ambas seccionalmente lineales y convexas. Este procedimiento se puede aplicar recurrentementepara k = K, K-1, ...,1 tal que cada función $v^k(\bullet)$ se caracterice. Las tasas de flujo y niveles de almacenamiento óptimos se pueden encontrar utilizando una fase de avance, de forma similar al usado en uno de Programación Dinámica.

Entonces, para cada valor factible x^{k-1} existe un ρ^k que satisface las condiciones de la proposición 2. Para cada valor posible de ρ^k se define:

$\Theta(\rho^k) = \{ x^{k-1}$: ρ^k satisface las condiciones de la proposición 2 para $x^{k-1} \}$ (5.4.1)

El conjunto $\Theta(\rho^k)$ consistirá de un punto o un intervalo. El algoritmo procedepor identificación de este conjunto para cada valor posible de ρ^k.

Hay cuatro casos posibles:

Caso 1

$$\rho^k \varepsilon \{ c_j^k, d_i^k : j = 1, ..., J^k ; i = 1, ..., I^k \}.$$

Caso 2

$\rho^k = c_{j^*}^k$ para algún $j^* = 1, ..., J^k$, y

$\rho^k \varepsilon \{ d_i^k : i = 1, ..., I^k \}$.

Caso 3

$\rho^k = d_{i^*}^k$ para algún $i^* = 1, ..., I^k$, y

$\rho^k \varepsilon \{ c_j^k : j = 1, ..., J^k \}$.

Caso 4

$\rho^k = c_{j^*}^k$ para algún $j^* = 1, ..., J^k$, y

$\rho^k = d_{i^*}^k$ para algún $i^* = 1, ..., I^k$.

Ahora, cada caso se examina.

Caso 1

$$\rho^k \varepsilon \{ c_j^k, d_i^k : j = 1, ..., J^k; i = 1, ..., I^k \}.$$

De (5.1.14), existe $j^* \varepsilon \{1, ..., J^k\}$ tal que $c_{j^*}^k > \rho^k > c_{j^*+1}^k$, así que:

$$\begin{aligned} c_j^k > \rho^k &\quad \text{para} \quad j = 1, ..., j^*, \ \text{y} \\ \rho^k > c_j^k &\quad \text{para} \quad j = j^*+1, ..., J^k \end{aligned} \qquad (5.4.2)$$

Se sigue de (5.3.15) que:

$$\Delta u_j^k = g_j^k \quad \text{para} \quad j = 1, \ldots, j^*, \text{ y} \quad .$$
$$\Delta u_j^k = 0 \quad \text{para} \quad j = j^*+1, \ldots, J^k \qquad (5.4.3)$$

Así que:

$$u^k = U_{j^*}^k \qquad (5.4.4)$$

De manera similar, de (5.3.14) existen $i^* \varepsilon \{1, \ldots, I^k\}$ tal que $d_{i^*}^k > \rho^k > d_{i^*+1}^k$, así que:

$$x^k = X_{i^*}^k \qquad (5.4.5)$$

Luego, por (5.3.17):

$$x^{k-1} = X_{i^*}^k + U_{j^*}^k \Delta t^k \qquad (5.4.6)$$

Hay por tanto solamente un valor de x^{k-1} correspondiente a valores de ρ^k en el intervalo:

$$\max \{ c_{j^*+1}^k, d_{i^*+1}^k \} < \rho^k < \min \{ c_{j^*}^k, d_{i^*}^k \} \qquad (5.4.7)$$

Se puede encontrar trazando una línea recta en el plano *t-x* desde el punto $(t^k, X_{i^*}^k)$ a una pendiente de $- U_{j^*}^k$.

Caso 2

$$\rho^k = c_{j^*}^k \quad \rho^k \varepsilon \{ d_i^k : i = 1, \ldots, I^k \}.$$

Si se usa (5.3.15) y (5.1.10), se sigue que:

$$\Delta u_j^k = g_j^k \quad \text{para} \quad j = 1, \ldots, j^*-1$$

$$0 \leq \Delta u_{j^*}^k \leq g_{j^*}^k \qquad (5.4.8)$$

$$\Delta u_j^k = 0 \quad \text{para} \quad j = j^*+1, \ldots, J^k$$

Así que:

$$U_{j^*-1}^k \leq u^k \leq U_{j^*}^k \qquad (5.4.9)$$

De (5.3.14) existe $i^* \varepsilon \{1, \ldots, I^k\}$ tal que:

$$d_{i^*}^k > \rho^k > d_{i^*+1}^k \qquad (5.4.10)$$

Así que de (5.3.16) y (5.3.13), se tiene:

$$x^k = X_{i^*}^k \qquad (5.4.11)$$

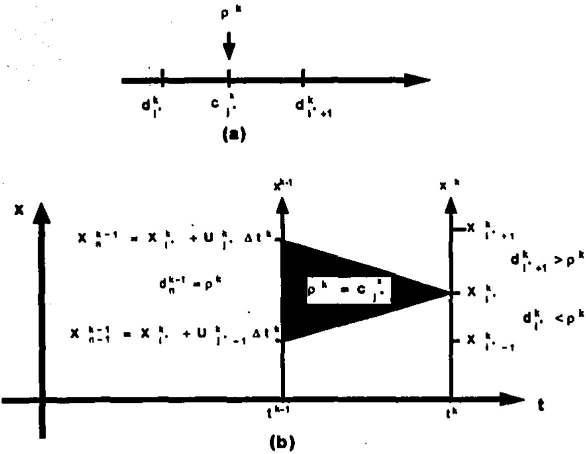
Se sigue de (5.3.17) que tales valores de ρ^k corresponden a valores de x^{k-1} en el intervalo:

$$X_{i^*}^k + U_{j^*-1}^k \Delta t^k \leq x^{k-1} \leq X_{i^*}^k + U_{j^*}^k \Delta t^k \qquad (5.4.12)$$

Este intervalo se puede construir en un plano *t-x* como la región delimitada por las dos pendientes $- U_{j^*-1}^k$ y $- U_{j^*}^k$, ambas parten del punto (t^k, $X_{i^*}^k$), como se muestra en la gráfica 5. Con esta región, por el supuesto del caso 2 ($\rho^k = c_{j^*}^k$), se tiene:

$$\delta v^{k-1}(x^{k-1}) \,/\, \delta x^{k-1} = -\rho^k = -c_{j^*}^k \qquad (5.4.13)$$

Luego, el intervalo definido en (5.4.12) corresponde a un segmento de la función $v^{k-1}(^*)$ con una pendiente $-c_{j^*}^k$.



Gráfica 5. (a) Definición del caso 2; (b) Construcción de la region $\Theta(\rho^k)$

Caso 3

$$\rho^k = d_{j^*}^k \quad \rho^k \varepsilon \{ c_j^k : j = 1, \ldots, J^k \}.$$

Si se usa (5.3.16) y (5.3.10), se sigue que:

$$X_{j^*-1}^k \leq x^k \leq X_{j^*}^k \qquad (5.4.14)$$

De (5.1.14) existe $j^* \varepsilon \{1, \ldots, J^k\}$ tal que:

$$c_{j^*}^k > \rho^k > c_{j^*+1}^k \qquad (5.4.15)$$

Así que de (5.3.15) y (5.1.10):

$$u^k = U_{j^*}^k \tag{5.4.16}$$

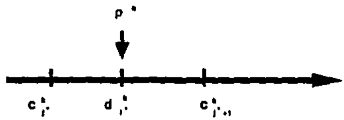
Se sigue de (5.3.17) que tales valores de ρ^k corresponden a valores de x^{k-1} en el intervalo:

$$X_{j^*-1}^k + U_{j^*}^k \Delta t^k \leq x^{k-1} \leq X_{j^*}^k + U_{j^*}^k \Delta t^k \tag{5.4.17}$$

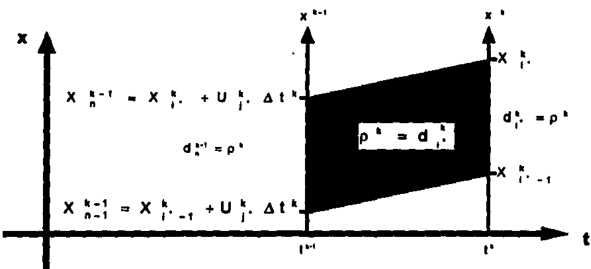
Este intervalo puede ser construido en un plano *t-x* como la región delimitada por las dos líneas paralelas de pendientes $U_{j^*}^k$, una parte del punto (t^k, $X_{j^*}^k$), la otra de (t^k, $X_{j^*-1}^k$), como se muestra en la gráfica 6. Con esta región y de $\rho^k = d_{j^*}^k$, se sigue:

$$\delta\, v^{k-1}(x^{k-1}) \,/\, \delta\, x^{k-1} = -\, \rho^k = -\, d_{j^*}^k \tag{5.4.18}$$

Luego, el intervalo definido en (5.4.17) corresponde a un segmento de la función $v^{k-1}(^*)$ con una pendiente $-\, d_{j^*}^k$.



(a)



(b)

Gráfica 6. (a) Definición del caso 3; (b) Construcción de la region $\Theta(\rho^k)$

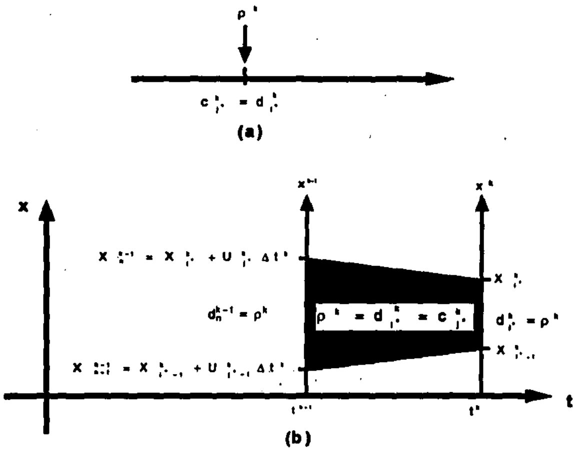
Caso 4

$$\rho^k = c_{j^*}^k = d_{i^*}^k$$

Al continuar con la misma lógica como para los casos 2 y 3, esto corresponde a valores de x^{k-1} en el intervalo:

$$X_{i^*-1}^k + U_{j^*-1}^k \Delta t^k \le x^{k-1} \le X_{i^*}^k + U_{j^*}^k \Delta t^k \qquad (5.4.19)$$

La región en el plano *t-x* se ilustra en la gráfica 7 y con ella, la pendiente de la función $v^{k-1}(\cdot)$ será $-c_{j^*}^k = -d_{i^*}^k$.



Gráfica 7. (a) Definición del caso 4; (b) Construcción de la region $\Theta(\rho^k)$

Luego, las líneas en el plano *t-x* que corresponde al caso 1 dividen el plano en regiones que corresponden a los casos 2, 3 y 4. La intersección de estas regiones y la línea $t = t^{k-1}$ formará los segmentos $X_{n-1}^{k-1} < x^{k-1} < X_n^{k-1}$, en la cual la pendiente de la función $v^{k-1}(\cdot)$ es constante, $-\rho^k = -d_n^{k-1}$. Basado en esta observación, los valores X_{n-1}^{k-1}, $n = 0, 1, 2, \ldots, I^{k-1}$, y d_n^{k-1}, $n = 1, 2, \ldots, I^{k-1}$, se pueden encontrar con el siguiente procedimiento.

Algoritmo: Procedimiento de etiquetación para $v^{k-1}(\cdot)$

Paso 0 Inicie $j^* := 0$ y $i^* := 0$;

Inicie $n^* := 0$;

Inicie $X_0^{k-1} := X_0^k + U_{min}^k \Delta t^k$;

Defina $c_{J^k+1}^k$ y $d_{I^k+1}^k$ como cualquier número más pequeño que min $(c_{J^k}^k, d_{I^k}^k)$.

Paso 1 Si $j^* = J^k$ e $i^* = I^k$ entonces: ir al Paso 6;

sino: ir al Paso 2.

Paso 2 Sea $n^* := n^* + 1$;

Si $c_{j^*+1}^k > d_{i^*+1}^k$ entonces: ir al Paso 3;

Si $c_{j^*+1}^k < d_{i^*+1}^k$ entonces: ir al Paso 4;

Si $c_{j^*+1}^k = d_{i^*+1}^k$ entonces: ir al Paso 5.

Paso 3 Sea $d_{n^*}^{k-1} = c_{j^*}^k$; (Caso 2)

Sea $j^* := j^* + 1$;

Sea $X_{n^*}^{k-1} := X_{i^*}^k + U_{j^*}^k \Delta t^k$;

Ir al Paso 1.

Paso 4 Sea $d_{n^*}^{k-1} = d_{i^*}^k$; (Caso 3)

Sea $i^* := i^* + 1$;

Sea $X_{n^*}^{k-1} := X_{i^*}^k + U_{j^*}^k \Delta t^k$;

Ir al Paso 1.

Paso 5 Sea $d_{n^*}^{k-1} = c_{j^*}^k$; (Caso 4)

Sea $j^* := j^* + 1$; $i^* = i^*+1$;

Sea $X_{n^*}^{k-1} := X_{i^*}^k + U_{j^*}^k \Delta t^k$;

Ir al Paso 1.

Paso 6 Sea $l^{k-1} := n^*$;

Pare.

Si $f^m(^*)$, U_{max}^m, U_{min}^m, X_{max}^m y X_{min}^m son constantes para $m = k-1, k-2, ..., m'$, entonces el algoritmo se puede implementar por el trazo de líneas en el plano t-x de t^m a t^{m-1}. Si alguna de estas líneas cruza los límites de almacenamiento, entonces se termina el procedimiento.

Una vez que el procedimiento se ha repetido para $k = K, K-1, ..., 1$, las pendientes de todas las funciones $v^{k-1}(^*)$ son conocidas para toda $m = 0, 1, ..., K$. La u^m óptima se podrá encontrar por la fase de avance de la Programación Dinámica. Sin embargo, dados los supuestos del problema (seccionalmente lineal y convexa), también se pueden obtener la u^m óptima con el siguiente procedimiento.

Algoritmo: Cálculo de u^k y x^k óptimas.

Caso 1. Si el nivel de almacenamiento x^{k-1} se ubica en una línea correspondiente al caso 1, caracterizado por i^* y j^* el costo de flujo óptimo $u^k = U_{j^*}^k$. El nivel de almacenamiento en el periodo t^k es por lo tanto $x^k = X_{i^*}^k$. Luego, la trayectoria de almacenamiento está a lo largo de la frontera entre las dos regiones.

Caso 2. Si el nivel de almacenamiento x^{k-1} se ubica en una región correspondiente al caso 2, caracterizado por i^* y j^* (i.e., la región delimitada con un tope (t^k, $X_{i^*}^k$)) el nivel de almacenamiento óptimo en el tiempo t^k es $x^k = X_{i^*}^k$, así que la tasa de flujo óptima es:

$$u^k = (x^{k-1} - X_{i^*}^k) / \Delta t^k$$

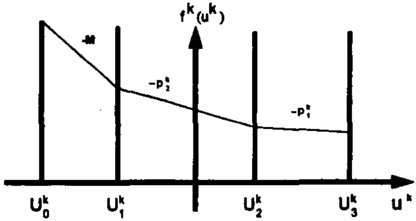
Caso 3. Si el nivel de almacenamiento x^{k-1} se ubica en una región correspondiente al caso 3, caracterizado por i^* y j^* (i.e., la región delimitada por las líneas paralelas con pendientes $- U_{j^*}^k$) entonces la tasa de flujo óptima es $u^k = U_{j^*}^k$. El nivel de almacenamiento óptimo en t^k es por tanto $x^k = x^{k-1} - U_{j^*}^k \Delta t^k$. La trayectoria óptima de almacenamiento es paralela a los límites de la región.

Caso 4. Si el nivel x^{k-1} se ubica en una región correspondiente al caso 4, caracterizada por i^* y j^*, cualquier tasa de flujo u^k que consiga un nivel de almacenamiento x^k en el periodo t^k en el rango $X_{i^*-1}^k \leq x^k \leq X_{i^*}^k$ es óptimo. Esto es:

$$\max \{ U_{j^*-1}^k , (x^{k-1} - X_{i^*}^k) / \Delta t^k \} \leq u^k \leq \min \{ U_{j^*}^k , (x^{k-1} - X_{i^*-1}^k) / \Delta t \}$$

Los valores óptimos ρ^k se pueden encontrar de las regiones entre las cuales pasa la trayectoria óptima. Si la trayectoria se ubica en los límites entre las dos regiones (i.e., el caso 1), entonces se puede usar (5.3.23) para encontrar ρ^k.

Enseguida se presenta la solución de un ejemplo con este procedimiento. Del ejemplo del capítulo 3 se tiene que la representación gráfica de las unidades de producción, mediante la reparametrización hecha con u^k es:



Gráfica 8. Representación de $f^k(u^k)$ del ejemplo

donde:

$$U_0^k = -(c_2^k + c_1^k)$$

$$U_1^k = s^k - (c_2^k + c_1^k)$$

$$U_2^k = s^k - c_1^k$$

$U_3^k = s^k$

para $k = 5, 4, 3, 2, 1$. Además, $\Delta t^k = 1$.

Luego, los resultados que se obtienen con este procedimiento son:

Como el nivel de almacenamiento inicial es $x^0 = 0$

$k = 1$	$u^1 = -35$
	$x^1 = 35$
$k = 2$	$u^2 = -65$
	$x^2 = 100$
$k = 3$	$u^3 \varepsilon [5,45]$
	tómese por ejemplo: $u^3 = 40$
	$x^3 = 60$
$k = 4$	$u^4 = 55$
	$x^4 = 5$
$K = 5$	$u^5 = 5$
	$x^5 = 0$

Lo que se puede interpretar como:

i) el primer mes producir con las máximas capacidades de las unidades de producción, así se cumple con satisfacer a toda la demanda del mes ($\eta^1 = 90$) y; además, almacenar 35 unidades del producto; por tanto, el nivel del almacén cuenta con 35 unidades al terminar el mes;

ii) el segundo mes con las máximas capacidades de las unidades de producción, así se cumple con satisfacer a toda la demanda del mes ($\eta^2 = 50$) y; además, almacenar 65 unidades adicionales del producto, $x^2 = 100$;

iii) el tercer mes tiene la particularidad de que para satisfacer la demanda del periodo se sacan 40 unidades del almacén y las restantes se hacen en la unidad 1 de producción, el nivel del almacén se reduce a $x^3 = 60$ unidades del producto;

iv) el cuarto mes, al igual que en el mes anterior, se extraen 55 unidades de producto del almacén y para satisfacer completamente la demanda (s^4 = 100) se producen las restantes 45 unidades del producto en la unidad 1 de producción; el nivel del almacén es $x^4 = 5$, por último;

v) el quinto mes, dado que el almacén se debe vaciar para su mantenimiento, se sacan las últimas 5 unidades del producto, las restantes 90 se producen con el uso de las dos unidades de producción a su máxima capacidad.

El valor marginal de un incremento adicional del contenido de almacén en el tiempo t^{k-1} es $\rho^k = 0.12$, $k = 1, 2, 3, 4, 5$.

Conclusiones y Recomendaciones

A través de las páginas de este documento se observa la trasendencia de la formación y el conocimiento de cursos como Álgebra Lineal, Cálculo Diferencial y Programación Lineal y Dinámica. Esta formación y conocimiento de las materias ya mencionadas y algunas otras que llevan implícita la matemática son requisitos que debe cubrir un alumno de la carrera de Actuaría. De aquí que sea posible señalar alactuario como un profesionista con la capacidad de aportar métodos alternativos para la optimización de recursos.

Para desarrollar este método ha sido necesario conocer y comprender un poco de la teoría de las Programaciones Lineal y Dinámica, así como de precisar las características del problema a optimizar.

Se expone un método de optimización para una variedad de situaciones prácticas, en donde se permite la aproximación deterministica. Este método resuelve problemas que constan de un almacén y un sistema lineal de producción de comodidad, en el cual cada etapa no se afecta por operaciones en otras etapas.

Se desarrolla, en el presente documento, un método alternativo para este tipo de problemas puesto que el almacenamiento complica considerablemente la optimización del sistema, ya que encadena las operaciones entre los periodos, así que la solución del problema no se puede hacer en forma independientementeen cada periodo y; si se intenta resolver este problema con la Programación Lineal el número de restricciones es proporcional al número de periodos (K - k). Dado que el tiempo para solucionar programas lineales son aproximadamente igual al cubo del número de restricciones, hay evidentemente severas restricciones en el número de periodos que se pueden acomodar si este sistema lineal se resuelve directamente. Mientras que, utilizando Programación Dinámica existe la posibilidad que hubiera dificultad en los cálculos por el problema que llamó Bellman como "plaga de la dimensionalidad".

Entonces, se propone reformular el programa lineal como uno de Programación Dinámica; este uso combinado de las dos programaciones se hacen con el fin de aprovechar las ventajas de cada una de ellas. En donde, se utilizan las características de programas dinámicos de una sola etapa y se resuelve así una sucesión de pequeños problemas a optimizar con Programación Lineal.

La combinación de la Programacion Lineal y Dinámica se utiliza para derivar un algoritmo, el cual genera la función de costo total $v^{k-1}(*)$ dada una función de costo total $v^k(*)$ y una función de costo de producción $f^k(*)$ ambas seccionalmente lineales y convexas. Este procedimiento puede ser aplicado recurrentemente para $k = K, K-1, \ldots, 1$ tal que cada función $v^k(*)$ se caracteriza. Las tasas de flujo y niveles de almacenamiento óptimos se pueden encontrar utilizando una fase de avance, de forma similar al utilizado en uno de Programación Dinámica.

El algoritmo solución se divide en dos etapas:

1.- Un procedimiento de etiquetado para $v^{k-1}(*)$, el cual puede observarse gráficamente después de que todas las pendientes de todas las funciones $v^{k-1}(*)$ se conocen, y

2.- Dado la estructura del problema se calculan los valores óptimos para u^k y x^k.

El procedimiento da resultados exactos, porque satisface sistemática y explícitamente las condiciones necesarias y suficientes Kuhn-Tucker para la solución del programa lineal.

También se da en el método una interpretación gráfica de los resultados. Además, este método puede resolver problemas con sistemas complejos de producción y de larga duración.

BIBLIOGRAFÍA

- Apostol, T.M., (1989)
  "Análisis Matemático"
  Reverté
  España

- Bertsekas, D.P., (1982)
  "Constrained Optimization and Multiplier Methods"
  Academic Press, Inc.
  London

- Bertsekas, D.P., (1987)
  "Dynamic Programming"
  Prentice Hall
  New Jersey

- Courant, R., F. John, (1978)
  "Introducción al Cálculo y al Análisis Matemático" Vol. 2
  Noriega
  México

- Dechamps, C., R. Nuytten y S. T. Lee, (1980)
  "Optimal Operation of Storage Devices in a Power System"
  IEEE PES Winter Meeting
  New York

- Denardo, E.V., (1982)
  "Dynamic Programming"
  Prentice Hall
  New Jersey

- Hadley, G., (1964)
  "Nonlinear and Dynamic Programming"
  Addison Wesley
  USA

- Hoffman, K., R. Kunze, (1991)
  "Álgebra Lineal"
  Prentice Hall
  México

- Intrilligator M.D., (1972)
  "Mathematical Optimization and Economic Theory"
  Prentice Hall
  New Jersey

- Kaufmann A., G. Desbazeille,(1965)
  "La Méthode du Chemin Critique"
  Dunod
  Paris

- Kaufmann, A., (1967)
  "La Programación Dinámica"
  Continental
  México, D.F.

- Kirk, D. E., (1970)
  "Optimal Control Theory -An Intoduction"
  Prentice Hall
  Englewood Cilffs, N. J.

- Kuhn, H.W., (1970)
  "Proceedings of the Princeton Symposium on Mathematical Programming"
  Princeton University Press
  Princeton, New Jersey

- Luenberger, D.G., (1989)
  "Introducción a la Programación Lineal y no Lineal"
  Adisson Wessley
  México, D.F.

- Nemhauser, G.L., (1966)
  "Dynamic Programming, Introduction to"
  John Wiley and Sons, Inc
  USA

- Prawda J., (1990)
  "Métodos y Modelos de Investigación de Operaciones" Vol. I
  Limusa
  México D. F.

- Taha H.A., (1989)
  "Investigación de Operaciones"
  Alfaomega
  México D. F.