UNIVERSIDAD AUTONOMA DE NUEVO LEON

FACULTAD DE INGENIERIA MECANICA Y ELECTRICA
DIVISION DE ESTUDIOS DE POSTGRADO

PROGRAMACION NO LINEAL

POR

LIC. RAMON CANTU CUELLAR

TESIS

EN OPCION AL GRADO DE MAESTRO EN CIENCIAS
DE LA ADMINISTRACION CON ESPECIALIDAD
EN INVESTIGACION DE OPERACIONES

SAN NICOLAS DE LOS GARZA, N. L., MAYO 1996

PROGRAMACION NO LINEAL

IVERSIDAD AUTONOMA DE NUEVO LEON

CULTAD DE INGENIERIA MECANICA Y ELECTRICA

DIVISION DE ESTUDIOS DE POSTGRADO

PROGRAMACION NO LINEAL

POR

LIC. RAMON CANTU CUELLAR

T E S I S

OPCION AL GRADO DE MAESTRO EN CIENCIAS
DE LA ADMINISTRACION CON ESPECIALIDAD
EN INVESTIGACION DE OPERACIONES

SAN NICOLAS DE LOS GARZA, N. L., MAYO 199[illegible]

UNIVERSIDAD AUTONOMA DE NUEVO LEON

FACULTAD DE INGENIERIA MECANICA Y ELECTRICA
DIVISION DE ESTUDIOS DE POSTGRADO

PROGRAMACION NO LINEAL

POR

LIC. RAMON CANTU CUELLAR

T E S I S

EN OPCION AL GRADO DE MAESTRO EN CIENCIAS
DE LA ADMINISTRACION CON ESPECIALIDAD
EN INVESTIGACION DE OPERACIONES

SAN NICOLAS DE LOS GARZA, N. L., MAYO 1996

UNIVERSIDAD AUTONOMA DE NUEVO LEON

FACULTAD DE INGENIERÍA MECANICA Y ELECTRICA

DIVISION DE ESTUDIOS DE POST-GRADO

Los miembros del Comité de tesis recomendamos que la presente tesis realizada por el **Lic. Ramón Cantú Cuéllar** sea aceptada como opción para obtener el grado de Maestro de Ciencias de la Administración con especialidad en **Investigación de Operaciones.**

El Comité de Tesis

Asesor
M.C. Marco A. Méndez Cavazos

Coasesor
M.C. Vicente García Díaz

Coasesor
M.C. Roberto Elizondo Villarreal

División de Estudios de Postgrado
M.C. David Oliva Alvarez

San Nicolás de los Garza, N.L. a Diciembre de 1995.

PROLOGO

La programación no lineal se ocupa del problema de optimizar una función objetivo con la presencia de restricciones tipo de igualdad y/o desigualdad. Si todas las funciones son lineales tenemos un programa lineal de lo contrario, el programa es no lineal y su resolución es el problema de estudio en esta tesis. La popularidad de la programación lineal puede atribuirse a muchos factores, incluyendo su habilidad para modelar problemas grandes y complejos, así como la de su resolución en un intervalo razonable de tiempo mediante el uso del método Simplex, y mas recientemente del método de Karmarkar, y de las computadoras, por parte de los usuarios.

Sin embargo muchos problemas reales no pueden ser adecuadamente representados o aproximados como un programa lineal debido a la naturaleza de la no linealidad de la función objetivo y o la no linealidad de cualquiera de las restricciones.

Los esfuerzos por resolver tales problemas no lineales en forma eficiente provocaron un rapido progreso durante las pasadas tres décadas. Esta tesis presenta estos desarrollos en una forma lógica e independiente.

Asimismo, esta tesis contiene material de consulta para profesionales que requieren aplicar tecnicas de programación no lineal en la resolución de problemas en sus respectivos campos de trabajo como apoyo o referencia en cursos de programacion no lineal o de investigación de operaciones que en sus programas de estudio incluyan la programación no lineal.

INDICE

CAPITULO I

INTRODUCCION

El objetivo de esta tesis esta encaminado a que cualquier, persona, familiarizada con los conceptos de programación lineal pueda de una manera comprensible entender las técnicas de programación no lineal. eliminando las demostraciones matemáticas que ofrecen la mayoría de los textos especializados en el tema. sin perder la continuidad de las ideas, enfocándose mas en la aplicación de tales técnicas.

El contenido de esta tesis es contemplado hacia la investigación y/o docencia referente al área de investigación de operaciones, concretamente a la programación no lineal.

La metodología de esta tesis consiste en desarrollar bases técnicas de los algoritmos y modelos no lineales considerados y ejemplos de su aplicación.

La programación lineal cuenta con los algoritmos Simplex y Karmarkar para la resolución de cualquier problema lineal, siempre y cuando este tenga solución. Desgraciadamente en programacion no lineal no existe un algoritmo matemático que pueda resolver cualquier problema no lineal, sino que dependiendo de la naturaleza del problema, ya sea de la función objetivo o de las restricciones si existen en el problema, se aplica un metodo en particular para la solución de tal problema

La teoría de la convergencia de algoritmos. aunque por si misma constituye un tópico muy importante, no ha sido incluida en esta tesis, enfocada a los aspectos intuitivos en el desarrollo y formulación de dichos algoritmos, para motivar a la posible aplicación de los mismos en las diferentes areas de trabajo.

Un factor que motivo al tema de esta tesis es el no encontrar ninguna elaborada anteriormente que trate en su totalidad el tema de programación no lineal, donde la literatura que existe actualmente del tema, especializada, son obras en ingles en su gran mayoría.

CAPITULO 2

SINTESIS

En esta tesis analizaremos el tema de la programación no lineal, en cinco capítulos y uno final de conclusiones y recomendaciones. Existe dependencia de los capítulos 3 y 4 con los tres siguientes, salvo las aplicaciones del capítulo 3 que pueden ser omitidas sin que se pierda la ilación o continuidad de las ideas. Los capitulos 5, 6 y 7 son independientes completamente, de manera que puede tratarse cada uno como una unidad por separado.

A continuación se presenta un resumen del contenido de los temas a tratar en esta tesis "Programación no lineal".

Capítulo 3

En esta sección se definirá el modelo matematico de programación no lineal, se discutirán las características de este tipo de problemas, se hará una comparación del modelo no lineal con los modelos lineales y se presentara el planteamiento de algunos problemas no lineales.

Capitulo 4

Se analizará en este capítulo conceptos fundamentales en la teoría de programación no lineal, las condiciones de optimalidad para problemas no lineales, así como también la formulación del problema dual y la posible solución de tal problema para conocer la solución del problema primal a partir del dual.

Capítulo 5

El objetivo de este capítulo es analizar algunos de los métodos de búsqueda que existen para resolver problemas de programación no lineal no restringidos, para funciones de una o varias variables, usando derivadas o sin usar derivadas.

Capitulo 6

Esta sección presenta algunos de los métodos existentes que utilizan direcciones factibles, donde, partiendo de un punto factible conocido a otro punto factible mejorado, mas cercano a un punto óptimo. Aquí se consideran problemas no lineales restringidos.

Capítulo 7

Se estudiará en este capítulo algunos modelos especiales de programación no lineal, que pueden ser resueltos con el método Simplex de programación lineal, haciendo ciertos ajustes al modelo matemático.

Analizaremos, por ultimo, un modelo muy interesante llamado programacion geométrica.

Capitulo 8

Se presentan en este capítulo las conclusiones y recomendaciones del presente trabajo.

Como en esta tesis se suponen conocidos algunos conceptos de algebra lineal y analisis matematico, se utilizan estos conceptos para el tema a tratar, se presenta un glosario de terminos para lograr la comprension de tales conceptos que son necesarios en el planteamiento y desarrollo de los algoritmos no lineales.

CAPITULO 3

CONCEPTOS BASICOS DE PROGRAMACION NO LINEAL

Programación no lineal

Se considera como tal al conjunto de métodos utilizados para optimizar una función objetivo, sujeta a una serie de restricciones en los que una o mas de las variables incluidas es no lineal.

Como ejemplo considere el siguiente problema de programación no lineal

Minimizar	$f(x)$	
sujeta a	$g_i(x) \leq 0$	$i = 1, 2, \ldots, m$
	$h_j(x) = 0$	$j = 1, 2, \ldots, l$

donde $f, g_1, g_2, \ldots, g_m, h_1, h_2, \ldots, h_l$ son funciones definidas en E_n, el espacio euclidiano de n dimensiones. X es un subconjunto de E_n, y x es un vector de componentes $x_1, x_2, \ldots, x_n$.

El problema anterior puede ser resuelto para los valores de las variables $x_1, x_2, \ldots, x_n$ que satisfacen las restricciones y que minimicen la función f.

La funcion f es llamada usualmente la función objetivo o la función criterio. Cada una de las restricciones $g_i(x)$ para $i = 1, 2, \ldots, m$ es llamada restricción de desigualdad y cada una de las restricciones $h_j(x)$ para $j = 1, 2, \ldots, l$ es llamada restricción de igualdad. Un vector x que satisface todas las restricciones es llamado una solución factible al problema

La colección de todas las posibles soluciones forman la región factible. El problema de programación no lineal es encontrar un punto factible $\bar{x}$ tal que $f(x) \geq f(\bar{x})$ para cada punto factible x. Un punto tal $\bar{x}$ es llamado una solución óptima o simplemente una solución al problema. Si existe mas de un punto óptimo, estos son referidos como soluciones alternativas óptimas.

Asimismo, un problema de programación no lineal puede expresarse como un problema de maximización y las restricciones de desigualdad pueden estar escritas en la forma $g_i(x) \geq 0$ para $i = 1,2,\ldots,m$. En el caso especial cuando la función objetivo es lineal y cuando todas las restricciones, incluyendo al conjunto X, puede ser representado por desigualdades lineales y/o ecuaciones lineales, el problema anterior es llamado un problema lineal.

Lo anterior indica que la programación lineal es un caso particular de la programación no lineal. Por tal motivo, los problemas no lineales son mucho mas difíciles de resolver que los de programación lineal. Haciendo un analisis de las características de los problemas lineales y no lineales tenemos la siguiente comparación entre ambos problemas.

Programación lineal	Programación no lineal
1. La solución óptima se encuentra en un punto extremo de la región de factibilidad.	1. No siempre la solución óptima se encuentra en un punto extremo de la region de factibilidad.
2. El punto óptimo nunca esta dentro de la región de factibilidad.	2. Hay casos donde el punto optimo esta en el interior de la region factible

3. Sus métodos de optimización generan óptimos absolutos ó globales.

3. Generalmente se encuentra un óptimo local ó relativo, mas no el óptimo global ó absoluto.

4. La región de factibilidad es un conjunto convexo.

4. Se pueden generar regiones de factibilidad que no son necesariamente convexas.

5. Sus funciones objetivo y restricciones son lineales.

5. La función objetivo, las restricciones ó ambas pueden ser no lineales.

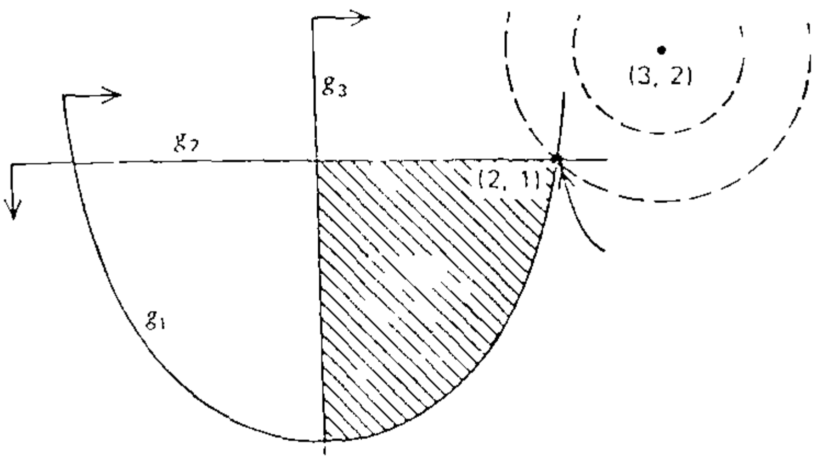
Como ilustración, considere el siguiente problema

Minimizar $(x_1-3)^2+(x_2-2)^2$

sujeta a

$$x_1^2 - x_2 - 3 \le 0$$
$$x_2 - 1 \le 0$$
$$-x_1 \le 0$$

La figura muestra la región factible



El problema es encontrar un punto en la región factible con el valor mas pequeño posible de $(x_1-3)^2+(x_2-2)^2$.

Nótese que los puntos (x_1, x_2) con $(x_1 - 3)^2 + (x_2 - 2)^2$ representan un circulo con radio $\sqrt{c}$ y centro (3,2). Este círculo es llamado curva de nivel de la función objetivo teniendo el valor c.

Puesto que deseamos minimizar c, debemos encontrar el círculo con el radio mas pequeño que intersecte la región factible. Como muestra la figura el tal circulo mas pequeño tiene c=2 e intersecta a la región factible en el punto (2,1). Por lo tanto la solución optima se obtiene en el punto (2,1) y tiene un valor objetivo igual a 2.

El metodo utilizado en el ejemplo anterior consiste en encontrar una solucion óptima determinando la curva de nivel con el mas pequeño valor objetivo que intersecta la region factible. Obviamente este enfoque de resolver geométricamente el problema es solamente conveniente para problemas pequeños y no es practico para problemas con mas de dos variables o aquellos con restricciones y funciones objetivo complicadas o complejas.

Por otra parte existen algunas condiciones o postulados que se cumplen para los dos modelos, los de programación lineal y los de programación no lineal:

1. Al aumentar (disminuir) el lado derecho de una restriccion $\leq$ o $\geq$ se relaja la restricción. Esta no puede contraer y si puede expandir al conjunto factible

2. Aumentar (disminuir) el lado derecho de una restricción estrecha la restriccion. Esto no puede expandir pero si contraer el conjunto no restringido.

3. Una restriccion no puede perjudicar, sino quizá ayudar, al valor optimo de la funcion objetivo.

4. Estrechar una restricción no puede ayudar, sino quiza perjudicar, al valor optimo de la función objetivo.

Planteamiento de problemas

Analicemos ahora el planteamiento de problemas, obteniendo modelos de programación no lineal, para buscar luego métodos de solución a problemas no lineales en los siguientes capítulos.

Ejemplo 1: A una compañía le cuesta c dolares por unidad fabricar un producto. Si la compañía cobra x dólares por unidad del producto, los clientes pedirán $D(x)$ unidades. Para maximizar las ganancias, ¿ Que precio tendrá que poner la compañía ? :

Variable: x

Función objetivo: Maximizar $z = (x - c)\, D(x)$ sin restricciones.

Ejemplo 2: Si se utilizan x_1 unidades de capital y x_2 unidades de trabajo, una compañía puede producir x_1x_2 unidades de un bien manufacturado. Se puede conseguir el capital a N$30 /unidad y el trabajo a N$7/unidad . Se dispone de un total de N$600 para conseguir capital y trabajo. ¿ Como puede la compañía maximizar la cantidad de bienes que se pueden fabricar ?

Variables: x_1 Unidades de capital

x_2 Unidades de trabajo

Función objetivo: Maximizar $z = x_1x_2$

Restricciones $30x_1 + 7x_2 \le 600$

$x_1 \geq 0 \qquad x_2 \geq 0$

Ejemplo 3: Q&H. Company se anuncia en telenovelas y juegos de futbol. Cada comercial en una telenovela cuesta 5000 dolares y cada comercial en un juego de fútbol cuesta 10000 dolares. Si se compran x_1 comerciales en novelas seran vistos por $5\sqrt{x_1}$ hombres y por $20\sqrt{x_1}$ mujeres (los datos vienen en millones de espectadores). Si se compran x_2 comerciales en juegos de futbol serán vistos por $7\sqrt{x_2}$ mujeres y por $17\sqrt{x_2}$ hombres. Por lo menos 40 millones de hombres y por lo menos 60 millones de mujeres quiere Q&H que vean sus comerciales

Variables: x_1 Anuncios en telenovelas

x_2 Anuncios en fútbol

Función objetivo: minimizar $z = 5000x_1 + 1000x_2$

Restricciones: $5\sqrt{x_1} + 17\sqrt{x_2} \geq 40$

$20\sqrt{x_1} + 7\sqrt{x_2} \geq 60$

$x_1 \geq 0$, $x_2 \geq 0$

Ejemplo 4: Una compañía produce un cierto articulo y desea encontrar una demanda conocida, supóngase que el inventario de la producción puede determinarse en un total de k periodos La demanda durante cualquier periodo puede encontrarse a partir del inventario en el principio del periodo y de la producción durante el periodo. La cantidad de producción máxima durante cualquier periodo esta restringida por la capacidad de producción del equipo disponible de modo que este no puede exceder b unidades. Supóngase que se tienen trabajadores suficientes cuando se necesite y despedirlos si son innecesarios, de cualquier manera, no fomentar las fluctuaciones de trabajo pesado incurre un costo provisional a el cuadrado de la diferencia en la fuerza de trabajo en cualquiera de los dos periodos sucesivos. También es incurrido un costo proporcional al inventario suma y sigue de un periodo a otro. Encontrar la fuerza de trabajo de inventariar durante los periodos 1, 2, ... k tal que se satisfaga la demanda y el costo sea minimizado.

Variables: I_k: nivel de inventario

L_k: fuerza de trabajo en el fin del periodo k

U_k: fuerza de trabajo adquirida durante el periodo k

Función objetivo: minimizar $z \quad \sum\left(c_1 u_k + c_2 I_k\right)$

Restricciones: $L_k \quad L \quad + U_k$

$I_k \quad I \quad pL_k - I - d_k$

$0 \cdot I$

$I_k \quad 0 \quad k \quad 1, 2, \ldots\ldots n$

Asimismo:

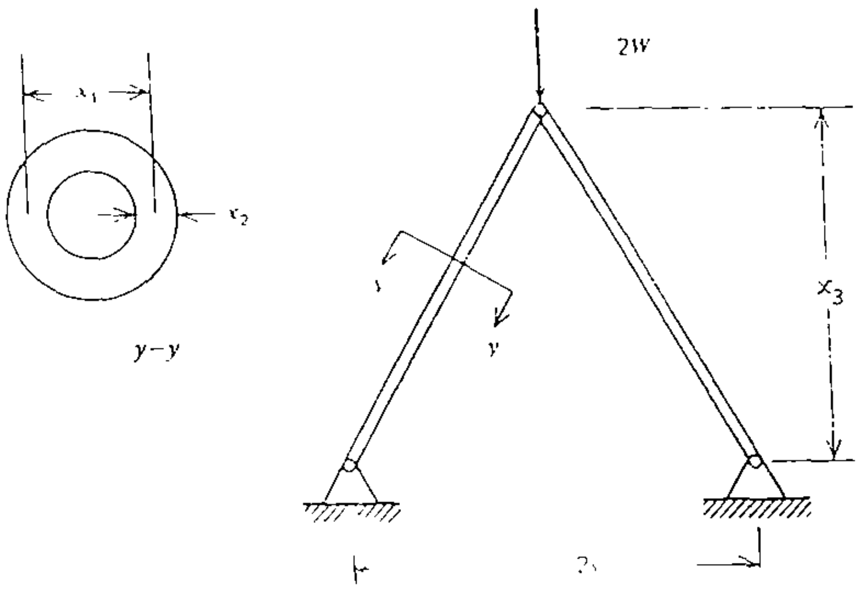
I_0 : Inventario inicial

L_0 : Fuerza de trabajo inicial

d_k : Demanda conocida durante k periodos

p : Numero de unidades producidas por trabajador durante cualquier periodo.

Ejemplo 5: Considere la armazon en un plano mostrada en la figura



La armazón consiste en dos tubos de acero juntos en un extremo y fijos en dos puntos pivotes en el otro extremo. El espacio, esto es, la distancia entre los dos pivotes es fijada en 2s. El problema de diseño es elegir la altura de la armazón, el espesor y el diámetro promedio de la armazón de tubos de acero, tal que esta pueda soportar una carga de 2W y minimizando la altura total de la armazón.

El peso de la estructura es $2\pi\rho x_1 x_2 (s^2 + x_3^2)^{1/2}$ donde ρ es la densidad del tubo de acero

Se presentan las siguientes restricciones adicionales:

1. Debido a limitaciones de espacio, la altura de la estructura no puede exceder b
2. El radio del diámetro del tubo entre el espesor del tubo no debe exceder b2.

3. La tensión de compresión en los tubos de acero no puede exceder la tensión dada del acero, que origina la siguiente restricción donde b es constante

$$w(s^2 + x_3)^{1/2} \le b\, x_1 x_2 x_3$$

4. La altura, diámetro y espesor pueden ser elegidos tal que los tubos no se doblen bajo la carga. Esta restricción puede ser expresada con b_4 como un parámetro conocido

$$w(s^2 + x_3^2) \le b_3 x_1 x_2 x_3$$

Función objetivo: Minimizar $z = x_1 x_2 (s^2 + x_3^2)^{1/2}$

Restricciones:

$$w(s^2 + x_3^2)^{1/2} - b_3 x_1 x_2 x_3 \le 0$$

$$w(s^2 + x_3^2)^{1/2} - b_4 x_1 x_2 x_3 (x_1^2 + x_2^2) < 0$$

$$x_3 \le b_1$$

$$\frac{x_1}{x_2} < b$$

$$x_1, x_2, x_3 \ge 0$$

CAPITULO 4

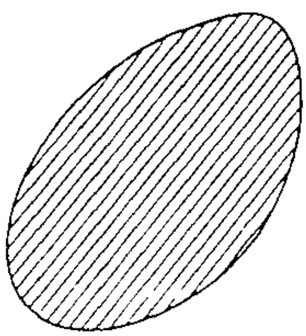
CONDICIONES DE OPTIMALIDAD Y DUALIDAD LAGRANGIANA

Conjuntos convexos

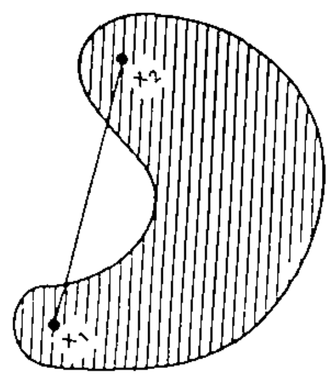
Definición: Un conjunto X en E_n se llama un conjunto convexo si dados dos puntos x_1, x_2 en x entonces $\lambda x_1 + (1 - \lambda)x_2$ pertenece al conjunto X para cada $\lambda \in [0, 1]$

La convexidad de X se puede interpretar geométricamente como sigue: Para cada par de puntos x_1 y x_2 en X, el segmento de recta que los une, debe pertenecer a X.

Ejemplo:



Un conjunto convexo



Un conjunto no convexo

Funciones convexas y cóncavas

Definición: Sea X un convexo no vacío en E_n. La función f sobre X es una función convexa si para cada $x_1, x_2 \in X$ y para cada $\lambda \in [0, 1]$ cumple

$$f(\lambda x_1 + [1 - \lambda]x_2) \leq \lambda f(x_1) + (1 - \lambda) f(x_2)$$

La funcion f es llamada cóncava si $-f$ es convexa, esto es

$$f(\lambda x_1 + [1 - \lambda]x_2) \geq \lambda f(x_1) + (1 - \lambda) f(x_2)$$

Si las desigualdades de esta definición son estrictas, se dice que la función es estrictamente convexa y estrictamente cóncava, respectivamente.

En base a esta definición se tiene que f es convexa si y solo si -f es cóncava y viceversa. Geométricamente en dos dimensiones.

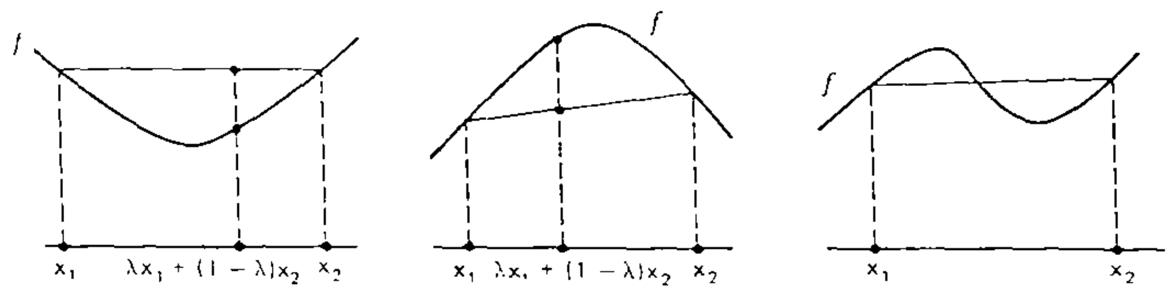


Figura 1 Ejemplos de funciones cóncavas y convexas

La figura 1 muestra que f es convexa si y solo si el segmento de recta que une dos puntos cualesquiera de la curva $y = f(x)$ nunca se encuentra por debajo de la curva $y = f(x)$. De manera similar, $f(x)$ es una función cóncava si y solo si el segmento rectilíneo que une dos puntos cualesquiera de la curva $y = f(x)$, nunca se encuentra por arriba de esta curva.

Trabajando únicamente con funciones cóncavas y convexas, se caería en el campo de la programación convexa, de donde se tienen ciertas propiedades importantes referentes a este tipo de funciones. Veamos los siguientes resultados.

Teorema: Para el problema no lineal

Minimizar $f(x)$

sujeta a $g_i(x) \geq 0$ $\quad$ $i = 1, 2, \ldots, n$

$h_i(x) = 0$ $\quad$ $i = 1, 2, \ldots, n$

si X es un conjunto convexo y si f es convexa sobre X, entonces cualquier mínimo local del problema no lineal anterior es una solución óptima de este problema.

Ahora bien, la manera de saber si f(x) es cóncava o convexa es mediante la segunda derivada de f(x), es decir, analizando el signo de f''(x) como se efectuaba en calculo diferencial para funciones de una sola variable y mediante la matriz hessiana de f para funciones de mas de una variable.

$$H(x) - \begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x\, x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 x_1} & \frac{\partial^2 f(x)}{\partial_2^2} & & \frac{\partial^2 f(x)}{\partial x_2 x_n} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n x_2} & \frac{\partial^2 f(x)}{\partial x_n x_2} & & \frac{\partial^2 f(x)}{\partial x_n^2} \end{pmatrix}$$

Partiendo de este concepto tenemos los dos teoremas siguientes para funciones de una sola variable.

Teorema: Supóngase que f''(x) existe para toda x en X. Entonces f(x) es una función convexa sobre X si y solo si f''(x) ≥ 0 para toda x en X.

Teorema: Supóngase que f''(x) existe para toda x en X. Entonces f(x) es una función concava sobre X si y solo si f''(x) ≤ 0 para toda x en X

Y para funciones de mas de una sola variable se tiene:

Teorema: Supóngase que f(x) tiene derivadas parciales continuas de segundo orden para cada punto x = (x_1, x_2,..., x_n) en X. Entonces f(x) es una función convexa sobre X si y solo si para cada x en X los menores principales de H(x) son no negativos.

Definición: El i-ésimo menor principal de una matriz nxn es el determinante de cualquier matriz ixi que se obtiene al quitar renglones y las n - i columnas correspondiente a la matriz.

Definición: El k-ésimo menor principal dominante de una matriz nxn es el determinante de la matriz kxk que se obtiene al quitar los últimos n - k renglones y columnas de la matriz.

Teorema: Supóngase que f(x) tiene derivadas parciales continuas de segundo orden para cada punto $x = (x_1, x_2, \ldots, x_n)$ en X.

Entonces f(x) es una función cóncava sobre X si y solo si para cada x en X y $k = 1, 2, \ldots, n$, los menores principales diferentes de cero, tienen el mismo signo que $(-1)^k$.

Ejemplo: Determinar si las siguientes funciones son cóncavas, convexas o ninguna de las dos

1. $f(x) = x^2$
2. $f(x) = 3x + 4$
3. $f(x_1, x_2) = x_1^2 - 2x_1x_2 - x_2^2$
4. $f(x_1, x_2) = -x_1 - x_1x_2 - 2x_2^2$
5. $f(x_1, x_2) = x_1 - 3x_1x_2 - 2x_2$

Solución:

1. $f(x) = x^2$

   $f'(x) = 2x$

   $f''(x) = 2 > 0$ Como $f''(x) \geq 0$, f(x) es una función convexa

2. $f(x) = 3x + 4$

   $f'(x) = 3$

f''(x) = 0   f(x) es cóncava y convexa simultáneamente dado que cumple con estas definiciones.

3. Calculando la matriz hessiana para $f(x_1, x_2) = x_1 + 2x_1x_2 + x_2^2$

$$\frac{\partial f(x)}{\partial x_1} = 2x_1 + 2x_2 \qquad \frac{\partial f(x)}{\partial x_2} = 2x_1 + 2x_2$$

$$\frac{\partial^2 f(x)}{\partial x_1^2} = 2 \qquad \frac{\partial^2 f(x)}{\partial x_2^2} = 2$$

$$\frac{\partial^2 f(x)}{\partial x_1 x_2} = 2 \qquad \frac{\partial^2 f(x)}{\partial x_2 x_1} = 2$$

$$H(x) = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$$

Los primeros menores son los elementos de la diagonal principal iguales ambos a $2 \geq 0$

$$\begin{vmatrix} 2 & 2 \\ 2 & 2 \end{vmatrix} = 2(2) \quad 2(2) = 0 \geq 0$$

Como los menores principales de H son no negativos para cualquier punto. $f(x_1,x_2)$ es una función convexa.

4. Siguiendo el mismo procedimiento que en el ejemplo anterior

$$\frac{\partial f(x)}{\partial x_1} = -2x_1 - x_2 \qquad \frac{\partial f(x)}{\partial x_2} = -x_1 - 4x_2$$

$$\frac{\partial^2 f(x)}{\partial x_1^2} = -2 \qquad \frac{\partial^2 f(x)}{\partial x_2^2} = -4$$

$$\frac{\partial^2 f(x)}{\partial x_1 x_2} = -1 \qquad \frac{\partial^2 f(x)}{\partial x_2 x_1} = -1$$

$$H(x) = \begin{pmatrix} -2 & -1 \\ -1 & -4 \end{pmatrix}$$

Los primeros menores principales son los elementos de la diagonal del hessiano, -2 y -4, ambos negativos. Para k = 1

$(-1)^1 = -1$, entonces los dos menores principales tienen el mismo signo, pues son negativos. El segundo menor principal es:

$$\begin{vmatrix} -2 & -1 \\ -1 & -4 \end{vmatrix} = (-2)(-4) - (-1) = 8 - 1 = 7 > 0$$

Y para k = 2 $(-1)^2 = 1$, el segundo menor principal tiene el mismo signo que $(-1)^2$. Por lo tanto f(x) es una función concava.

5. Nuevamente

$$\frac{\partial f(x)}{\partial x_1} = 2x_1 - 3x_2 \qquad \frac{\partial f(x)}{\partial x_2} = -3x_1 + 4x_2$$

$$\frac{\partial^2 f(x)}{\partial x_1^2} = 2 \qquad \frac{\partial^2 f(x)}{\partial x_2^2} = 4$$

$$\frac{\partial^2 f(x)}{\partial x_1 x_2} = -3 \qquad \frac{\partial^2 f(x)}{\partial x_2 x_1} = -3$$

$$H(x) = \begin{pmatrix} 2 & -3 \\ -3 & 4 \end{pmatrix}$$

Los primeros dos menores principales son positivos y para k=1 $(-1)^k = (-1)^k = -1$, lo cual significa que f(x) no es cóncava. Entonces f(x) será convexa si el segundo menor principal, el determinante del hessiano es mayor o igual a cero, entonces tenemos que

$$\begin{vmatrix} 2 & -3 \\ -3 & 4 \end{vmatrix} = 2(4) - (-3)(-3) = 8 - 9 = -1 < 0$$

Por lo tanto como el segundo menor principal es negativo, f(x) tampoco es una función convexa.

Las funciones convexas y cóncavas desempeñan un papel muy importante en el estudio de los problemas de programación no lineal, así como también algunos otros tipos de funciones que son similares a estas, que son generalizaciones de estos conceptos. Analizaremos estos nuevos conceptos, para luego pasar a analizar las condiciones de optimalidad partiendo de estos conceptos fundamentales.

Generalización de funciones cóncavas y convexas

Funciones cuasi-convexas y cuasi-cóncavas

Definición: Una función f es una función cuasi-convexa bajo un conjunto convexo X en E_n si para cada x_1, x_2 en X

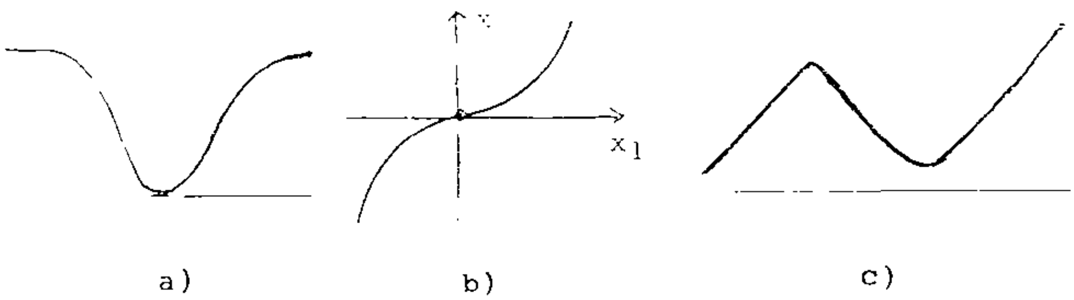
$f(\lambda x_1 + [1-\lambda]x2) \leq$ mínimo $\{f(x_1), f(x_2)\}$

La función f es llamada cuasi-cóncava si -f es cuasi-convexa, es decir si

$f(\lambda x_1 + [1-\lambda]x_2) \geq$ mínimo $\{f(x_1), f(x_2)\}$ para todo, λ tal que $0 \leq \lambda \leq 1$

La figura 2 muestra un ejemplo de una función cuasi-convexa, una función cuasi-cóncava y una que no cumple ninguna de las dos condiciones.

Figura 2



La figura 2a) muestra que f es una función cuasi-convexa si es imposible encontrar tres puntos colineales tal que el punto del centro tiene el mayor valor objetivo al evaluar estos puntos en f, por la definición de función cuasi-convexa. Similarmente la definición de función cuasi-cóncava indica que si f es cuasi-cóncava es imposible encontrar tres puntos colineales tales que el punto del centro tenga el mas pequeño valor objetivo que los otros dos puntos.

La figura 2b) muestra una función que es cuasi-convexa y cuasi-cóncava. dado que cumple estas dos definiciones, al tomar tres puntos en el eje x_1, la imagen del punto central no es ni mayor ni menor que la de los otros dos puntos colineales.

La figura 2c) muestra una función que no cumple con ninguna de estas condiciones.

Una función lineal es simultáneamente cóncava, convexa, cuasi-cóncava y cuasi-convexa.

Una función si es cuasi-convexa no necesariamente es convexa, pero el reciproco no es cierto, porque si f es convexa entonces es también cuasi-convexa. Este mismo criterio puede ser aplicado para funciones cóncavas y cuasi-cóncavas.

Funciones Pseudo-convexas y Pseudo-cóncavas

Antes de definir estos dos nuevos tipos de funciones, es necesario conocer cuando una función convexa es diferenciable.

Definición: Sea X un conjunto no vacío en En. el espacio Euclidiano en n dimensiones y sea f: $X \to E_1$ Entonces f es una función diferenciable en un punto $\bar{x}$ perteneciente a X si existe un vector $\nabla f(\bar{x})$ llamado el vector gradiente. y una funcion $\alpha : En \to E_1$ tal que

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})^t(x - \bar{x}) + \|x - \bar{x}\| \alpha(\bar{x}, x)$$

para cada x en X. donde $\lim_{x \to \bar{x}} \alpha(\bar{x}, x) = 0$, $\|x - \bar{x}\|$ es la norma del vector $x - \bar{x}$, expresada como $\|x - \bar{x}\| = \sqrt{(x_1 - x_2)^2 + \ldots + (x_n - \bar{x}_n)^2}$ y $\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \ldots, \frac{\partial f(x)}{\partial x_n} \right)$

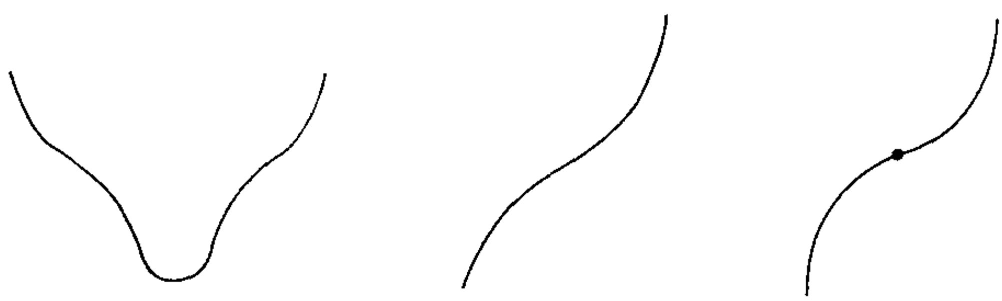
Los componentes del vector gradiente son las derivadas parciales de f con respecto a las variables $x_1, x_2, \ldots, x_n$.

Definición: Sea X un conjunto no vacío en E_n y sea f: $X \to E_1$ diferenciable en X. La función f es llamada pseudo-convexa si para cada x_1, x_2 en X con $\nabla f(x_1)^t (x_2 - x_1) \geq 0$ se tiene que $f(x_2) \geq f(x_1)$ o equivalentemente, si $f(x_2) < f(x_1)$ entonces
$\nabla f(x_1)^t (X_2 - X_1) < 0$

La función f es llamada pseudo-concava si -f es una función pseudo-convexa.

La figura 3 muestra algunos tipos de funciones pseudo-convexas y pseudo-concavas, así como una que no es ninguna de las dos.

Figura 3



a) b) c)

a) Funcion pseudo-convexa, [illegible] uncion pseudo-concava y pseudoconvexa, c [illegible] uncion que no es ni pseudo-convexa ni pseudo-concava.

Estas definiciones [illegible] funciones cuasi-convexas, cuasi-concavas, pseudo-convexas y pseudo-concavas se convierten en estrictas cada una de ellas si la desigualdad correspondiente [illegible] definición de cada tipo de función es estricta

Habiendo definido es[illegible]onceptos fundamentales analicemos las condiciones de optimalidad para funciones [illegible]exas

Condiciones de optimalidad

Teorema: Sea la función f: $E_n \to E_1$, el problema Min f(**x**) sujeto a **x** $\in$ X, donde f es una función convexa y X es un conjunto convexo no vacío en E_n; se tiene que $\bar{x}$ es una solución óptima a este problema si y solo si f tiene un subgradiente ξ en $\bar{x}$ tal que $\xi^t(\bar{x})(x - \bar{x}) \geq 0$ para todo **x** en X.

Asimismo, en base a lo anterior, si el problema es de maximización, $\bar{x}$ es una solución óptima si y solo si $\xi^t(\bar{x})(x - x) \leq 0$ para todo x en X

Corolario: El conjunto de soluciones óptimas alternativas puede ser definido equivalentemente como X^*

$$X^* = \left\{ x \in X: \nabla f(\bar{x})^t (x - \bar{x}) = 0 \quad y \quad \nabla f(x) = (x) \right\}$$

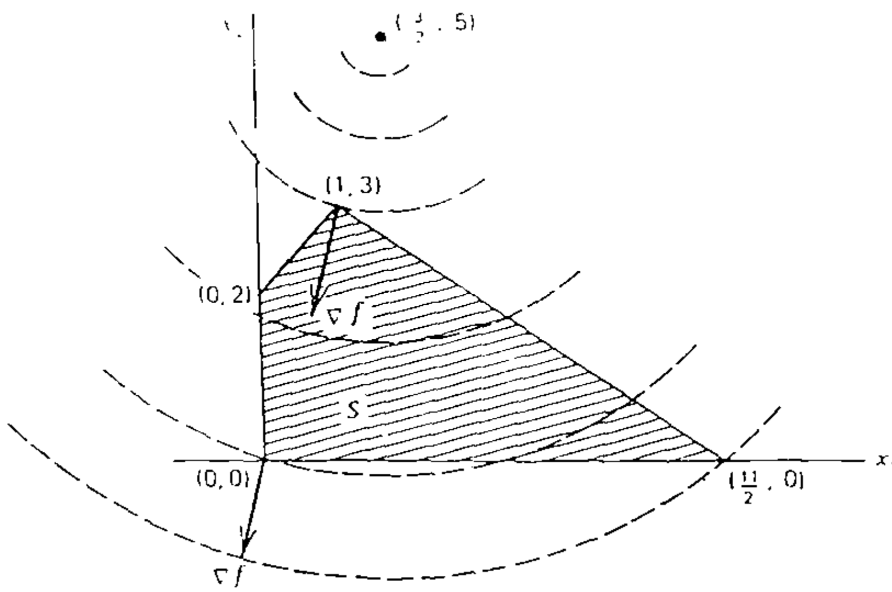
Ejemplo: Para el siguiente problema:

$$\text{Minimizar } f(x) = (x_1 - 3/2)^2 + (x_2 - 5)^2$$

$$\begin{aligned} \text{sujeta a } -x_1 + x_2 &\leq 2 \\ 2x_1 + 3x_2 &\leq 11 \\ -x_1 &\leq 0 \\ -x_2 &\leq 0 \end{aligned}$$

Solución: La función objetivo es una funcion convexa que representa el cuadrado de la distancia del punto (3/2,5). El conjunto poliedrico convexo X es representado en la figura 4

Figura 4.



De la figura, se observa claramente que el punto óptimo es (1,3). El vector gradiente en el punto (1,3) es:

$$\nabla f(x) = (2\,[x_1 - 3/2],\, 2[x_2 - 5])^t$$

con x = (1,3)

$$\nabla f(x) = (2[1 - 3/2],\, 2[3 - 5])^t = (-1\ ,\, -4)^t$$

Se observa geométricamente que el vector (-1, -4) forma un ángulo ≤ 90° con cada vector de la forma $(x_1 - 1,\ x_2 - 3)$, donde $(x_1,\ x_2) \in X$. Entonces la condición de optimalidad establecida anteriormente fue verificada. Sustituyendo en el vector gradiente $\bar{x} = (1, 3)$ tenemos

$$
\begin{aligned}
\nabla f(x) &= \begin{pmatrix} -1 \\ -4 \end{pmatrix}(x_1 - 1,\, x_2 - 3) = -1(x_1 - 1) + (-4)(x_2 - 3) \\
&= -x_1 + 1 - 4x_2 + 12 = -x_1 - 4x_2 + 13 \\
&= 4(3) + 13 = -13 + 13 = 0
\end{aligned}
$$

lo cual indica que (1,3) es el punto optimo.

Condiciones de Optimalidad de Fritz-John

Antes de definir las condiciones de optimalidad de Fritz-John, para problemas de programación no lineal, analicemos el concepto de los multiplicadores de Lagrange. Método de optimización por lagrangianos

El problema a resolver es:

Optimizar $f(x)$

sujeta a $g_i(x) = 0 \quad i = 1, \ldots, m$

donde f, g_1, g_2 ...., g_m son funciones definidas en E_n, x es el vector de n componentes $x_1, x_2, \ldots x_n$ y la palabra optimizar puede significar maximización o minimización.

El método clásico para resolver este problema, cuando las restricciones están igualadas a cero (cuando son desigualdades este método requiere de ciertos ajustes que complican bastante la situación), es utilizando los multiplicadores de Lagrange (λ_i). Esto se realiza definiendo de la siguiente manera una nueva función objetivo, llamada el lagrangiano

$$F(x_1, x_2, \ldots, x_n, \lambda_1, \lambda_2, \ldots, \lambda_m) = f(x) - \lambda_1 g_1(x) - \lambda_2 g_2(x) - \ldots - \lambda_m g_m(x)$$

El lagrangiano F con m+n variables independientes, es equivalente a la función objetivo original f con n variables independientes porque se están restando únicamente ceros, es decir que $\lambda_i g_i(x) = 0$ para $i = 1, \ldots, m$. Como el lagrangiano no tiene restricciones, para resolver este problema en teoría (dado que ocasionalmente es imposible resolverlo por métodos exactos y se requieren métodos de aproximación a analizar en los capítulos siguientes) hay que obtener todas las derivadas parciales de primer orden para cada variable e igualarlas a cero y resolver el sistema de ecuaciones resultantes

$$\begin{array}{ll} \dfrac{\partial F}{\partial x_1} = 0 & \dfrac{\partial F}{\partial \lambda_1} = 0 \\[2ex] \dfrac{\partial F}{\partial x_2} = 0 & \dfrac{\partial F}{\partial \lambda_2} = 0 \\ \vdots & \vdots \\ \dfrac{\partial F}{\partial x_n} = 0 & \dfrac{\partial F}{\partial \lambda_m} = 0 \end{array}$$

Teorema: (Condiciones de Fritz-John) Sea X un conjunto no vacío y sea f: $E_n \rightarrow E_1$ y g_i: $E_n \rightarrow E_1$ para $i = 1, \ldots, m$. Considere el problema P de minimizar $f(\mathbf{x})$ sujeta a $\mathbf{x}$ en X y $g_i(\mathbf{x}) \leq 0$ para $i = 1, \ldots, m$. Sea $\bar{\mathbf{x}}$ una solución factible y sea $I = \{i: g_i(\bar{\mathbf{x}}) = 0\}$

Además, supóngase que f y g_i son diferenciables en $\bar{x}$ y que g_i para $i \in I$ son continuas en $\bar{x}$. Si $\bar{x}$ resuelve localmente el problema, P, entonces existen escalares u_0 y u_i para $i \in I$, tales que

$$u_0 \nabla f(\bar{x}) + \sum_{i \in I} u_i \nabla g_i(\bar{x}) = 0$$

$$u_0,\ u_i \geq 0 \quad \text{para } i \in I$$

$$(u_0,\ u_I) \neq (0,\ 0)$$

donde u_I es el vector cuyas componentes son u_i para $i \in I$. Además, si g_i para $i \notin I$ son también diferenciables en $\bar{x}$, entonces las condiciones anteriormente mencionadas pueden ser escritas en la siguiente forma equivalente:

$$u_0 \nabla f(\bar{x}) + \sum_{i=1}^{m} u_i \nabla g_i(\bar{x}) = 0$$

$$u_i\, g_i(\bar{x}) = 0 \qquad \text{para } i = 1, \ldots, m$$

$$u_0,\ u_i \geq 0 \qquad \text{para } i = 1, \ldots, m$$

$$(u_0,\ u) \neq 0$$

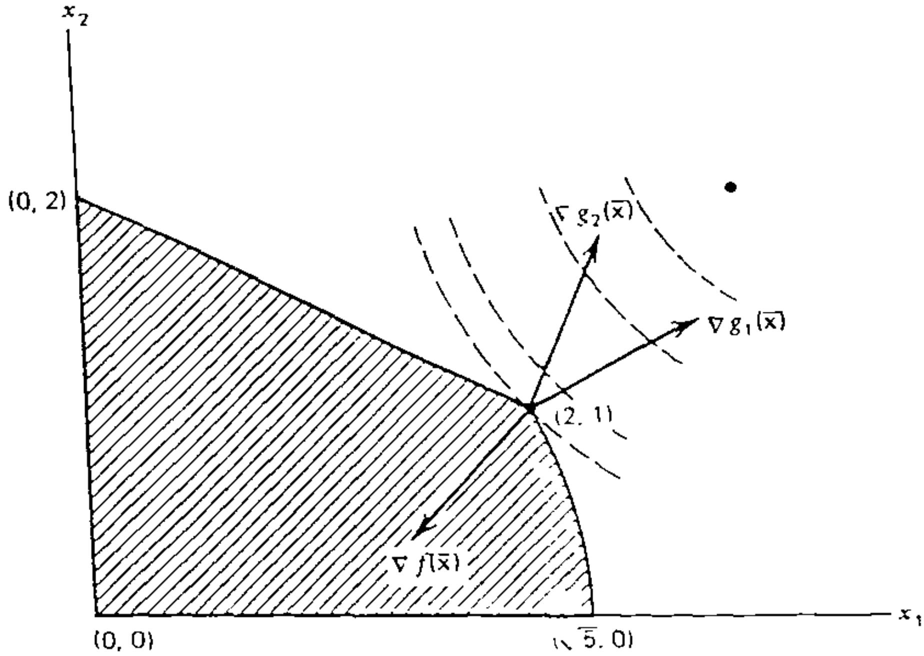
donde u es el vector cuyas componentes son u_i para $i = 1, \ldots, m$.

Ejemplo. Minimizar $(x_1 - 3)^2 + (x_2 - 2)^2$

$$\begin{aligned} \text{sujeta a} \quad & x_1 + x_2 \le 5 \\ & x_1 + 2x_2 \le 4 \\ & -x_1 \le 0 \\ & -x_2 \le 0 \end{aligned}$$

Se tiene la siguiente región factible

Figura 5



Las condiciones de Fritz-John se cumplen en el punto óptimo (2,1). Nótese que el conjunto de restricciones activas, es decir aquellas que funcionan como igualdad en el punto considerado son $I = \{1, 2\}$ para $x = (2, 1)^t$. Calculando los gradientes para la función objetivo y restricciones

$\nabla f(x) = (-2, -2)^t$ $\qquad$ $\nabla g_1(x) = (4, 2)^t$ $\qquad$ $\nabla g_2(x) = (1, 2)^t$

Por lo tanto, para satisfacer las condiciones de Fritz-John, hay que encontrar un vector diferente de cero $(u_0, u_1, u_2) > 0$ que satisfaga

[illegible]

Si $u_1 = u_0 / 3$ y $u_2 = 2u_0 / 3$ para cualquier $u_0 > 0$ se satisfacen las condiciones de Fritz-John dado que se encuentra un vector diferente de cero $(u_0, u_1, u_2) > 0$

Ejemplo: Considere el siguiente problema

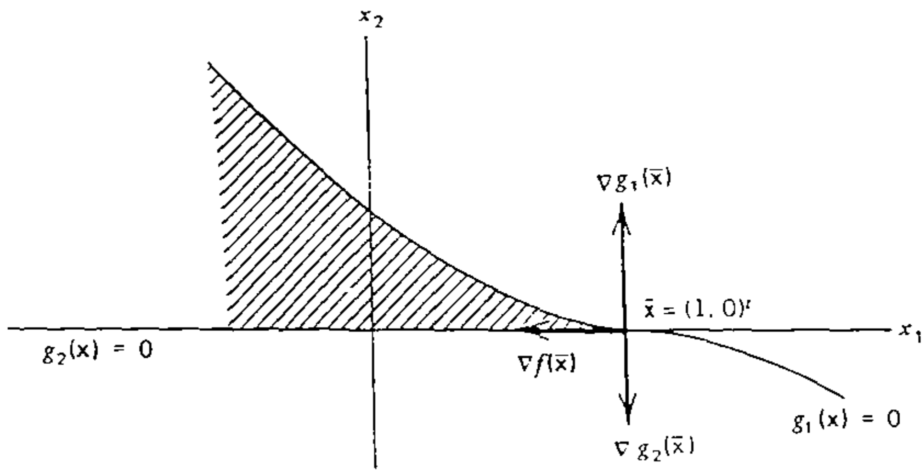
Minimizar $-x_1$

sujeto a $x_2 - (1 - x_1)^3 \leq 0$

$-x_2 \leq 0$

Gráficamente

Figura 6



Las condiciones de Fritz-John se cumplen en el punto $(1,0)^t$. El conjunto le restricciones activas en x es dado por I= {1, 2}. Calculando los gradientes

$\nabla f(x)$ $(1, 0)^t$ $\qquad$ $\nabla g_1(x)$ $(0, 1)$ $\qquad$ $\nabla g_2(x)$ $(0, -1)^t$

de lo cual

$$u\begin{pmatrix} 1 \\ \end{pmatrix} + u\begin{pmatrix} \\ 1 \end{pmatrix} + u_2\begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \begin{pmatrix} \\ \end{pmatrix}$$

Esta condición solo se cumple si u_0 0. Entonces las condiciones de Fritz-John on verdaderas en **x** tomando $u_0 = 0$ y u_1 u α, donde α es un escalar positivo para obtener el vector $(u_0, u_1, u_2) \geq 0$. Analicemos ahora si las condiciones de Fritz-John se cumplen en el punto $(0, 0)^t$. El conjunto de restricciones activas es I {3, 4} y se tiene que u_1 u_4 0

Además

$$\nabla f(x) = (-6, -4)^t \qquad \nabla g_3 = (-1, 0)^t \qquad \nabla g_4 = (0,-1)^t$$

Las condiciones de Fritz-John se cumplen si y solo si

$$u_0\begin{pmatrix}-6\\-4\end{pmatrix} + u_3\begin{pmatrix}-1\\0\end{pmatrix} + u_4\begin{pmatrix}0\\-1\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$$

entonces $u_3=-6u_0$, $u_4=-4u_0$, pero si $u_0>0$ se tiene que u_3, $u_4<0$ contradiciendo las restricciones de no negatividad. Por otra parte si $u_0=0$ entonces $u_3=u_4=0$ lo cual contradice la condición de que el vector (u_0, u_3, u_4) es diferente de cero. Por lo tanto las condiciones de Fritz-John no se cumplen en $x=(0,0)^t$. El origen no es un punto óptimo local.

Las condiciones de Fritz-John no son suficientes para la optimalidad de los problemas de programación lineal.

Condiciones de Optimalidad de Karush-kuhn-Tucker (KKT)

En problemas de programación lineal el método Simplex se puede interpretar como un procedimiento sistemático para alcanzar un punto extremo óptimo de la región de factibilidad que satisface las condiciones de KKT. Para cualquier problema de programación lineal si un punto es óptimo es porque cumple con estas condiciones

Para problemas de programación no lineal las condiciones de Fritz-John y de KKT son necesarias y suficientes para optimalidad cuando la función objetivo, las restricciones de igualdad se encuentran bajo apropiadas suposiciones de convexidad.

Teorema:(Condiciones de Karush-Kuhn-Tucker) Sea X un conjunto no vacío abierto en E_n, y sean $f:E_n \to E_1$ Y $g_i:E_n \to E_1$ para $i=1,\dots,m$. Sea $\bar{x}$ una solución factible y denótese $I=\{i: g_i(\bar{x})=0\}$. Si f y g_i para cada $i \in I$ son diferenciables $\bar{x}$. Si $\bar{x}$ resuelve localmente el problema

Minimizar f(x) $\quad x \in X$

sujeta a $\quad g_i(x) \leq 0 \quad i=1,...,m$

deben existir escalares ui para $i \in I$ tal que

$$\nabla f(x) + \sum_{i=1}^{m} u_i \nabla g_i\left(x\right) = 0$$

$$u_i g_i(x) = 0 \qquad \text{para } i = 1, ..., m$$

$$u_i \geq 0 \qquad \text{para } i = 1, ..., m$$

Estas mismas condiciones también son validas para que x sea punto óptimo para el problema

Maximizar f(x)

sujeta a $g_i(x) \geq 0 \quad i = 1, ..., m$

donde f y g_i tienen las mismas características que el problema de minimizacion.

Ejemplo: Para el problema

Minimizar $10x_1 - 2x_1^2 - x_1^3 + 8x_2 - x_2$

sujeta a $x_1 + x_2 < 0$

$-x_1 < 0$

$x_2 \leq 0$

Verificaremos si se cumplen las condiciones de KKT en el punto (1,1) para determinar si es o no óptimo. El conjunto de restricciones activas $I = \{1\}$. Calculando los gradientes

$$\nabla f = \begin{pmatrix} 10-4x_1-3x_1 \\ 8-2x_2 \end{pmatrix} \qquad \nabla f(x) = (3, 6)^t \qquad \nabla g_1(x) = (1, 1)^t$$

Como solo la primera restricción es activa, tenemos que $u_2 = u_3 = 0$. De la primera condición

$$\nabla f(x) + \sum_{i}^{m} u_i g_i(x) = 0$$

$$\binom{1}{0} + u_1\binom{1}{1} = \binom{0}{0}$$

Como debe ser $u_1 \geq 0$ se describe un sistema incompatible. Por lo tanto el punto $x = (1, 1)^t$ no es óptimo.

Ejemplo: Minimizar $x_1^4 + x_2^4 + 12x_1^2 + 6x_2^2 - x_1x_2 - x_1 - x_2$

sujeta a

$$\begin{aligned} -x_1 - x_2 &\leq -6 \\ -2x_1 + x_2 &\leq -3 \\ x_1 &\geq 0 \\ x_2 &\geq 0 \end{aligned}$$

Para el punto (3,3) verifiquemos las condiciones de KKT. Para

$I = \{1, 2\}$ $\quad u_3 = u_4 = 0$

$$\nabla f = \begin{pmatrix} 4x_1^3 + 24x_1 - x_2 - 1 \\ 4x_2^3 + 12x_2 - x_1 - 1 \end{pmatrix} \qquad \nabla f(x) = (176, 140)^t$$

$$\nabla g(x) = (1, 1)^t \qquad \nabla f_2(x) = (2, -1)^t$$

Resolviendo el sistema resultante

$$\begin{aligned} -u_1 - 2u_2 &= -176 \\ -u_1 + u_2 &= -140 \end{aligned}$$

Se tiene que $u_1 = 152$ y $u_2 = 12$, dando a estos parámetros estos valores se cumplen las condiciones de KKT en el punto (3,3).

Dualidad Lagrangiana

Dado un problema de programación no lineal llamado primal, existe un problema que es estrechamente asociado con este llamado dual lagrangiano

Estos dos problemas son dados a continuación

Problema Primal P: Minimizar f(x)

$$\text{sujeta a } g_i(x) \leq 0 \qquad i=1,\dots,m$$

$$h_i(x) = 0 \qquad i=1,\dots,l$$

donde f, g_i, h_i: $E_n \to E_1$ y X es un conjunto no vacío en E_n. Sean g y h las m y l funciones vectoriales con componentes i-ésimos g_i y h_i, respectivamente.

Problema Dual lagrangiano D: Maximizar $\theta(u, v)$

sujeta a $u \geq 0$

donde $\theta(u, v) = \inf \left\{ f(x) + \sum_{i=1}^{m} u_i g_i(x) + \sum_{i=1}^{l} v_i h_i(x) : x \in X \right\}$

Los vectores u y v pertenecen a E_m y E_l, respectivamente. En este problema, las restricciones $g_i(x) \leq 0$ y $h_i(x)$ pueden ser incorporadas en la función objetivo usando los multiplicadores de Lagrange u_i y v_i. El multiplicador u_i asociado con la restricción de desigualdad $g_i(x) \leq 0$ es no negativo, en tanto que el multiplicador v_i asociado con la restricción $h_i(x) = 0$ es no restringido en signo.

Entonces el problema dual consiste de maximizar el Infimum (maxima cota inferior) de la funcion

$$f(x) + \sum_{i=1}^{m} u_i g_i(x) + \sum_{i=1}^{l} v_i h_i(x)$$

Asimismo, dado un problema de programación no lineal, algunos problemas duales lagrangianos pueden concebirse dependiendo en cuales restricciones son tomadas como

$g_i(x) \leq 0$ y $h_i(x) = 0$

y cuales restricciones son tomadas por el conjunto X. Esta elección puede afectar el valor óptimo de D (en un problema no convexo) y complicar el procedimiento que consiste en evaluar y actualizar la función dual e durante el curso de la resolución del problema dual.

Mediante esta técnica pueden resolverse problemas no lineales convexos o no convexos. También en problemas de optimización discreta donde todas o algunas de las variables pueden ser restringidas a ser números enteros.

Ejemplo: Para el siguiente problema primal

Minimizar $x_1^2 + x_2^2$

sujeta a $-x_1 - x_2 + 4 \leq 0$

$x_1, x_2 \geq 0$

Para este problema, la solución óptima ocurre en el punto (2, 2) con valor objetivo igual a 8.

Sea $g(x) = -x_1 - x_2 + 4$ y $X = \{(x_1, x_2) : x_1, x_2 \geq 0\}$. La función dual es dada por

$$\theta(u) = \inf \{x_1^2 + x_2^2 + u(-x_1 - x_2 + 4) : x_1, x_2 \geq 0\}$$

$$= \inf \{x_1^2 - ux_1 : x_1 \geq 0\} + \inf \{x_2^2 - ux_2 : x_2 \geq 0\} + 4u$$

La máxima cota inferior se obtiene para $\theta(u)$ de la siguiente forma: Para x_1 y x_2 derivando de la función anterior

$2x_1 - u = 0$ $\qquad$ $2x_2 - u = 0$

$x_1 = u/2$ $\qquad$ $x_2 = u/2$

lo anterior es válido si $u \geq 0$. Si $u \leq 0$, entonces $x_1 = x_2 = 0$, de lo cual si $u \geq 0$

$$\theta(u) = \inf \left\{ \left(\frac{u}{2}\right)^2 + \left(\frac{u}{2}\right)^2 + u\left(-\frac{u}{2} - \frac{u}{2} + 4\right) : x_1, x_2 \geq 0 \right\}$$

$\theta(u) = \inf \left\{\frac{1}{2} u^2 + 4u\right\} = \frac{1}{2} u^2 + 4u$

Si $u < 0$

$\theta(u) = \inf \{0^2 + 0^2 + u(-0 \quad 0 + 4)\} = \inf\{4u\} = 4u$

$$\theta(u)=\begin{cases} 4u & \text{si } u \; [illegible] \; 0 \\ \frac{1}{2}u \; [illegible] +4u & \text{si } u \geq 0 \end{cases}$$

El valor máximo de θ (u) se obtiene para u = 4 , entonces $\theta(u) = 8$. Los valores óptimos de los problemas primal y dual son ambos iguales a 8.

Bajo ciertas suposiciones de convexidad y teniendo como requisito apropiadas restricciones, los problemas primal y dual tienen el mismo valor objetivo óptimo y por lo tanto es posible resolver el problema primal indirectamente resolviendo el problema dual, como en el ejemplo anterior, pero esto, en la mayoría de los casos no es valido.

Aquí surge otra diferencia entre los problemas de programación lineal y no lineal, que mientras en programacion lineal los problemas primal y dual tienen los mismos valores objetivos optimos, en programación no lineal, en general el valor objetivo óptimo del primal es mayor o igual que el valor óptimo del problema dual.

Ejemplo: Para el problema de minimizacion

Minimizar $f(x)$ [illegible] $-2x$ [illegible]

sujeta a $\quad x_1 + x_2 - 3 = 0$

$(x_1, x$ [illegible] X

donde X [illegible] {(0, 0), (0, 4), (4, 4), (4, 0), (1, 2), (2, 1)}

De los seis puntos solo cumplen la restricción de igualdad (1, 2) y (2, 1) de los cuales el óptimo es (2, 1). El valor objetivo mínimo óptimo es -3

La función objetivo dual es dada por

$$\theta(v) = \inf\{(-2x_1 + x_2) + v(x_1 + x_2 - 3) : (x_1, x_2) \in X\}$$
$$\qquad = \inf\{-2x_1 + vx_1 : x_1 \in X_1\} + \inf\{x_2 + vx_2 : x_2 \in X\} - 3v$$
con $X_1 = X_2 = \{0, 1, 2, 4\}$

Derivando la primera y segunda ecuaciones con respecto a x_1 y x_2, respectivamente, e igualando a cero se tiene

$$-2 + v = 0 \qquad\qquad 1 + v = 0$$
$$v = 2 \qquad\qquad v = -1$$

con $v < -1$, para minimizar, tomando el punto (4, 4) y sustituyendo en la ecuacion

$\theta(v) = \inf\{-2(4) + 4 + v(4 + 4 - 3)\} = \inf\{-4 + 5v\} = -4 + 5v$

Para $-1 \leq v < 2$, tomando el punto (4,0) se tiene

$\theta(v) = \inf\{-2(4) + 0 + v[4 + 0 - 3]\} = \inf\{-8 + v\} = -8 - v$

Si v 2, tomando para minimizar $\theta(v)$ el punto (0,0)

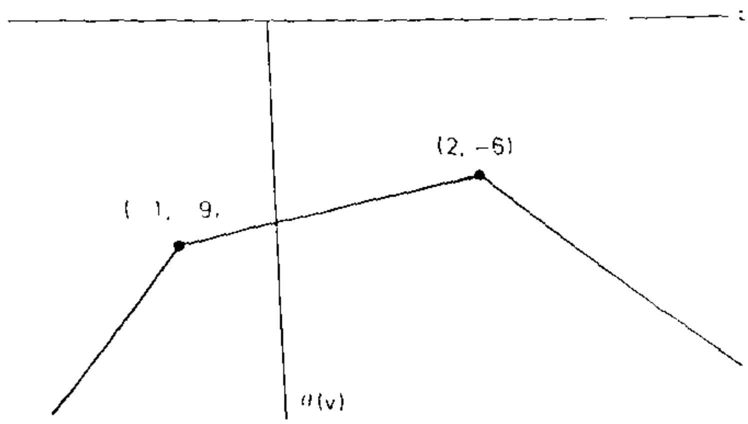
$\theta(v) = \inf\{-2(0) + 0 + v(0 + 0 - 3)\} = -3v$

Entonces

$$\theta(v) = \begin{cases} -4 + 5v & \text{si } v < -1 \\ -8 - v & \text{si } -1 < v \leq 2 \\ -3v & \text{si } v \geq 2 \end{cases}$$

La gráfica de la función dual es mostrada en la figura 7 y la solución optima es v = 2 con valor objetivo -6

Figura 7



Ejemplo: Para el problema

Minimizar $f(x) = 3x_1 + 7x_2 + 10x_3$

sujeta a $\quad -x_1 - 3x_2 - 5x_3 + 7 \leq 0$

$x_1, x_2, x_3 \quad 0 \text{ ó } 1$

La función lagrangiana dual obtenida a partir del problema anterior es :

$\theta(u) = \inf\{3x_1 + 7x_2 + 10x_3 + u(-x_1 - 3x_2 - 5x_3 + 7): x_1, x_2, x_3 \quad 0 \text{ ó } 1\}$

$\inf\{(3 - u)x_1 + (7 - 3u) + (10 - 5u)x_3 + 7u\}$

$x_1, x_2, x_3 = 0 \text{ ó } 1$

El valor mínimo es calculado haciendo $x_j \quad 0$ si este coeficiente de costo reducido es positivo y haciendo $x_j = 1$ si el coeficiente de costo reducido es negativo. El resultado es la familia de soluciones

	x_1	x_2	x_3
$0 \leq u \leq 2$	0	0	0
$2 \leq u \leq 7/3$	0	0	1
$7/3 \leq u \leq 3$	0	1	1
$3 \leq u$	1	1	1

de lo cual tenemos que

$$\theta(u) \begin{cases} 7u & \text{si } 0 \leq u \leq 2 \\ 2u + 10 & \text{si } 2 \leq u \leq 7/3 \\ -u + 17 & \text{si } 7/3 \leq u \leq 3 \\ -2u + 20 & \text{si } 3 \leq u \end{cases}$$

para la función dual el valor óptimo es $u = 7/3$ y su valor objetivo y su valor objetivo es igual a $44/3$

CAPITULO 5

OPTIMIZACION NO LINEAL SIN RESTRICCIONES

Metodos de busqueda

Uno de los principales campos de la programación no lineal es el de la optimizacion no restringida u optimización libre, que trata el problema de minimizar o maximizar una función en ausencia de cualquier restricción. Existen métodos de busqueda del valor óptimo de una función f(x) que pueden ser aplicados para funciones de una ó varias variables, usando derivadas o sin usar derivadas.

A continuación analizaremos algunos de los métodos de busqueda que existen para resolver problemas de programación no lineal no restringidos, despues de presentar algunos conceptos fundamentales.

Metodos de optimizacion de funciones unimodales de una sola variable en problemas no restringidos

Definición: Se dice que una funcion de una sola variable es unimodal, cuando tiene un solo punto minimo o maximo

Si f(x) es estrictamente unimodal entonces (para el caso de maximización):

para $x_1 < x_2 < x^*$ entonces $f(x_1) < f(x_2) < f(x^*)$

y $x^* < x_2 < x_1$ entonces $f(x^*) > f(x_2) > f(x_1)$

siendo x^* el valor que optimiza la funcion f(x), es decir $f(x) \le f(x^*)$ para toda x

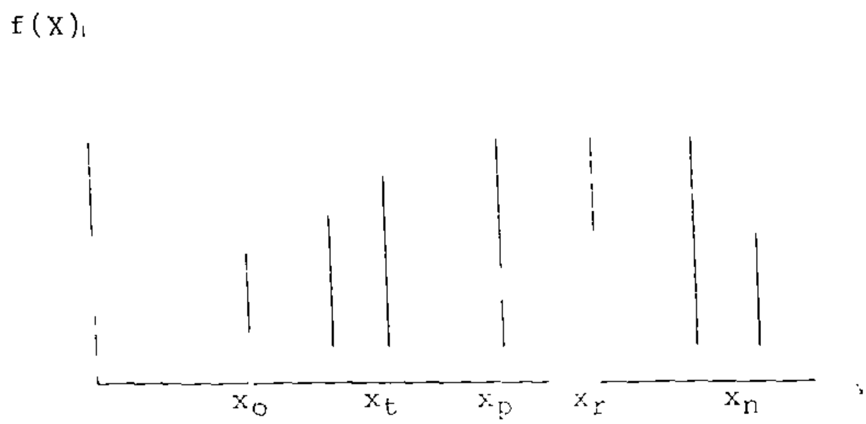
Dada una función f(x) estrictamente unimodal y dos puntos x_1 y x_2 con $x_1 < x_2$, existen tres posibilidades a analizar para efectuar la eliminación. Por ejemplo para caso de maximización

a) Si $f(x_1) > f(x_2)$ entonces $x^* < x_2$

b) Si $f(x_1) < f(x_2)$ entonces $x_1 < x^*$

c) Si $f(x_1) = f(x_2)$ entonces $x_1 < x^* < x_2$

Suponiendo que se requieren hacer k evaluaciones de una función f(x), que es estrictamente unimodal, sean x_0 y x_n los extremos izquierdo y derecho del intervalo de búsqueda, respectivamente, sea x_p el mejor punto obtenido y sean x_t y x_r los puntos a la izquierda y derecha de x_p, respectivamente. Obviamente el punto óptimo x^* debe encontrarse en el intervalo $x_t < x^* < x_r$. La distancia $x_r - x_t$ se denomina intervalo de incertidumbre



Una medida de la efectividad de los métodos de búsqueda de puntos óptimos en funciones unimodales de una sola variable es la siguiente

$$F = \frac{x_r - x_t}{x_n - x_0}$$

Búsqueda en un intervalo con derivadas

Metodo de busqueda de Fibonacci

El método de Fibonacci es un método de búsqueda en un intervalo para minimizar una función estrictamente cuasi-convexa

Este método es bastante eficiente para aproximar, bajo cierta tolerancia de error, un punto mínimo o máximo en funciones unimodales de una sola variable. Este método tiene la ventaja de que, como no utiliza derivadas, puede calcular puntos óptimos en funciones no diferenciables. Este método supone que se conoce un rango inicial de búsqueda.

Procedimiento

1.- Con el porcentaje de aproximación requerido, se determina el valor de F_n, que deberá de utilizarse.

2.- Con el intervalo de búsqueda dado inicial, $a < x^* < b$, se define el valor inicial L .

$Lo = b - a$

3.- Con la expresion

$$\Delta = L_{i-1}\frac{F_{(i+1)}}{F_{(i-1)}}$$

se calcula Δ para $i = 1$ $\quad \Delta_1 = Lo\frac{F_{n-2}}{F_n}$

4.- Se procede a calcular los valores de x_1, x_2, $x_1 = a + \Delta_1$, $x_2 = b - \Delta_1$ y se procede a evaluar $f(x_1)$ y $f(x_2)$.

5.- Se efectua la eliminación de la forma siguiente (para el caso de minimizacion).

Si $f(x_1) < f(x_2)$, entonces

- Eliminar el intervalo $x_2 < x < b$
- Hacer $b = x_2$
- Hacer L_i [illegible]
- Hacer i [illegible] y calcular Δ_1
- Hacer $x_3 = a + \Delta_1$ y evaluar $f(x_3)$
- Si $x_1 < x_3$, hacer $x_2 = x_3$ y $f(x_2) = f(x_3)$ en caso contrario, hacer $x_2 = x_1$, $f(x_2) = f(x_1)$ y $x_1 = x_3$, $f(x_1) = f(x_3)$
- Se continua con el paso 6

Si $f(x_1) > f(x_2)$, entonces

- Eliminar el intervalo $a < x < x_1$

- Hacer $a = x_1$
- Hacer $L_1 = b - a$
- Hacer $i = i + 1$ y calcular Δi
- Hacer $x_3 = b - \Delta_i$ y evaluar $f(x_3)$
- Si $x_2 < x_3$ hacer $x_1 = x_2$, $f(x_1) = f(x_2)$ y $x_2 = x_3$, $f(x_2) = f(x_3)$ en caso contrario, hacer $x_1 = x_3$ y $f(x_1) = f(x_3)$
- Se continua con el paso 6

6.- Se continua con el paso 5 hasta que $i = n - 1$.

Ejemplo: Encontrar el punto mínimo de la función unimodal

$f(x) = x^2 - 6x + 2$ con un 3% de error o menos, en el rango $0 < x < 10$

n	Fn	1/Fn	
0	1	1.0000	
1	1	1.0000	
2	2	0.5000	
3	3	0.3333	
4	5	0.2000	
5	8	0.1250	
6	13	0.0769	
7	21	0.0476	
8	34	0.0294	(margen de error)
9	55	0.0182	

La solucion x^* esta en el intervalo $2.9412 \leq x \leq 3.2353$

$$\text{Aproximación} = \frac{3.2353 - 2.942}{10 - 0} = 0.0294$$

El valor mínimo encontrado para la funcion es -6.996

Los cálculos para obtener esta solucion se presentan en la siguiente tabla

Método de búsqueda de Fibonacci

a	b	Li	i	Δi	x_1	x_2	$f(x_1)$	$f(x_2)$	Rango a eliminar	Nuevo Rango	x_3	$f(x_3)$
0	10	10	1	3.8235	3.8235	6.1765	-6.3218	3.0900	6.1765<x<10.00	0<x<6.1765		
0	6.1765	6.1765	2	2.3529	2.3529	3.8235	-6.5813	-6.3218	3.8235<x<6.1765	0<x<3.8235	2.3529	-6.5813
0	3.8235	3.8235	3	1.4706	1.4706	2.3529	-4.6609	-6.5813	0.0000<x<1.4706	1.4706<x<3.8235	1.4706	-4.6609
1.4706	3.8235	2.3529	4	0.8824	2.3529	2.9412	-6.5813	-6.9965	1.4706<x<2.3529	2.3529<x<3.8235	2.9412	-6.9965
2.3529	3.8235	1.4706	5	0.5882	2.9412	3.2353	-6.9965	-6.9446	3.2353<x<3.8235	2.3529<x<3.2353	3.2353	-6.9446
2.3529	3.2353	0.8824	6	0.2941	2.6471	2.9412	-6.8754	-6.9965	2.3529<x<2.6471	2.6471<x<3.2353	2.6471	-6.8754
2.6471	3.2353	0.5882	7	0.2941	2.9412	2.9412	-6.9965	-6.9965	2.6471<x<2.9412	2.9412<x<3.2353	2.9412	-6.9965

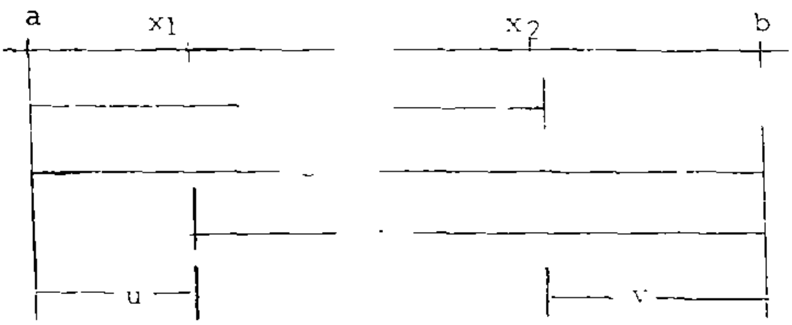
Método de búsqueda de la Sección de Oro

Este método también requiere que la función sea de una sola variable y unimodal que sea estrictamente cuasi-convexa. Si el rango de incertidumbre original es: $a \leq x \leq b$, el proceso reduce este intervalo en cada evaluación, en un valor constante T (representa la letra griega tao)

Este valor constante llamado "Sección de Oro" se calcula de la siguiente manera:

Supóngase que se esta examinando el intervalo de búsqueda $a \leq x \leq b$ con dos valores x_1 y x_2 conocidos, tales que a [illegible] b

Gráficamente



Sea $T =$ [illegible] donde r [illegible] t y entonce [illegible] s, si al hacer la búsqueda se eliminara el rang [illegible] $x_2 \leq x \leq b$ lo que se requiere es que [illegible] siguiente evaluacion se mantenga el valor de T para encontrar el nuevo valor de [illegible]

T = u/r = [illegible] [illegible] [illegible] = 1/T − 1 [illegible]

de donde se obtiene la ecuacion [illegible] resolviendo se tiene que T [illegible] 0.618

Procedimiento

1.- Establecer el valor de la aprox[illegible] que se requiere (e)

2.- Con el intervalo de búsqueda [illegible] inicial, a [illegible] x^* [illegible] b y el valor de T [illegible] 0.618, se calcula [illegible] b [illegible] a y los valore[illegible]

x_1 [illegible] (1 [illegible])(b - a) + a [illegible] + a y se procede a evaluar $f(x_1)$ y $f(x_2)$

3.- Se efectúa la eliminación de la forma siguiente (Para el caso de minimización)

Si $f(x_1) > f(x_2)$, entonces

- Hacer $a = x_1$
- Hacer $x_1 = x_2$ y $f(x_1) = f(x_2)$
- Calcular $x_2 = T (b - a) + a$
- Evaluar $f(x_2)$
- Continuar con el paso 4

Si $f(x_1) < f(x_2)$, entonces

- Hacer $b = x_2$
- Hacer $x_2 = x_1$ y $f(x_2) = f(x_1)$
- Calcular $x_1 = (1 - T)(b - a) + a$
- Evaluar $f(x_1)$
- Continuar con el paso 4

4.- Si $\frac{b-a}{Lo} < e$ se ha cumplido la tolerancia, en caso contrario continuar con el paso 3.

Ejemplo: Encontrar el punto mínimo de la función unimodal

$f(x) = 5x^2 + 8x - 2$ con un 5% de error o menos, en el intervalo $-1 \le x \le 5$

$Lo = b - a = 5 -(-1) = 6$

Los cálculos e iteraciones del metodo se presentan en la siguiente tabla. Partiendo de ello, la solución esta en el intervalo $-0.8722 \le x \le -0.6655$

$$\text{Aproximación} = \frac{-0.6655 - 0.8722}{5 - 1.0} = 0.0344$$

Método de la sección de oro

a	b	x1	x2	f(x1)	f(x2)	Rango a Eliminar	Nuevo Rango	$\frac{b-a}{Lo}$
-1	5	1.2920	2.7080	16.6823	56.3303	2.7080<x<5.0000	-1<x<2.7080	0.6180
-1	2.7080	0.4165	1.2920	2.19882	16.6823	1.2920<x<2.7080	-1<x<1.2920	0.3820
-1	1.2920	-01245	0.4165	-2.9182	2.19882	0.4165<x<1.2920	-1<x<0.4165	0.2361
-1	0.4165	0.4589	-0.1245	-4.6183	-2.9182	-0.1245<x<0.4165	-1<x<-0.1245	0.1459
1	-0.1245	[illegible]	4589	-5.1096	-4.6183	-0.4589<x<-0.1245	-1<x<-0.4589	0.0902
-	-0.4589		-0.6655	-5.1997	-5.1096	-0.6655<x<-0.4589	-1<x<-0.6655	0.0557
-1	-0.6655	-0.8722	-1.7933	-5.1739	-5.1987	-1.0000<x<-0.8722	-0.8722<x<-0.6655	0.0344

Método de Interpolación Cuadrada

Los métodos de Fibonacci y de la sección de oro son utilizados para funciones unimodales, pero si una función no es de este tipo, no pueden aplicarse los métodos anteriores, como en las funciones multimodales.

Definición: Una función de una sola variable es multimodal si tiene mas de un valor óptimo, ya sea máximo ó mínimo.

Para funciones multimodales no diferenciables en los puntos de discontinuidad, se utiliza para resolver estos problemas de programación no lineal el método de interpolación cuadrada.

Procedimiento

1.- Dados valores del punto arbitrario x y una dirección arbitraria s, evalúe

$$g(\alpha) = f(x + \alpha s)$$

El punto x es el punto de partida de la búsqueda del óptimo local, mientras que s es la dirección de búsqueda.

2.- Evalue $g(\alpha)$ para i= 0, 1, 2, 4, 8, 16, ... a, b, c, donde c es el primer valor de α para el cual $g(\alpha)$ se ha incrementado. El valor óptimo de α, $a < \alpha^* < c$, se encuentra en el rango $a < \alpha^* < c$.

3.- Una vez que se conocen los valores de g(a), g(b) y g(c) se ajusta un polinomio cuadrado a f(x) y se calcula el mínimo local del polinomio α_c, dado por:

$$\alpha_c = \frac{1}{2}\,\frac{g(a)(c^2-b^2) + g(b)(a^2-c^2) + g(c)(b^2-a^2)}{g(a)(c-b) + g(b)(a-c) + g(c)(b-a)}$$

Si $g(\alpha_c) < g(b)$ entonces $\alpha^* = \alpha_c$. De otra manera, si $g(\alpha_c) > g(b)$ se toma $\alpha^* = b$.

Ejemplo: Minimizar la función $f(x) = x^4 - 10x^3$ [illegible] en el punto x = [illegible] y la direccion s = 1 como valores iniciales.

Fase 1

$$g(\alpha) = f(x + \alpha s)$$

$$= f(3 + \alpha)$$

$$= (3 + \alpha)^4 - 10(3 + \alpha)^3 + 35(3 + \alpha)^2 - 50(3 + \alpha)$$

Fase 2

α	$g(\alpha)$
0	-24
1	-24
2	0

En $\alpha = 2$ el valor de $g(\alpha)$ se incremento. Por lo tanto $c = 2$, $b = 1$ y $a = 0$. Se tiene que $0 < \alpha^* < 2$.

Fase 3

El mínimo α_c del polinomio cuadrado se calcula con la formula anterior. Sustituyendo los valores citados de a, b y c

$$\alpha_c = \frac{1}{2}\,\frac{g(a)(c^2-b^2) + g(b)(a^2-c^2) + g(c)(b^2-a^2)}{g(a)(c - b) + g(b)(a - c) + g(c)(b - a)}$$

$$= \frac{1}{2}\,\frac{(-24)(2^2 - 1^2) + (-24)(0^2 - 2^2) + 0(1^2 - 0^2)}{(-24)(2 - 1) + (-24)(0 - 2) + 0(1 - 0)} = 0.50$$

$$g(\alpha_c) = g(0\text{-}50) = -24.94$$

Como $g(\alpha_c) < g(b)$, se tiene que $\alpha^* = \alpha_c = 0.50$

Pero dado que

$$g(\alpha) = f(x + \alpha s) = f(3 + \alpha)$$

Por lo tanto el mínimo aproximado de la funcion original es

$$x^* = 3 + \alpha^* = 3 + 0.5 = 3.5$$

y

$$f(x^*) = f(3.5) = -24.94.$$

Búsqueda en un intervalo con derivadas

Método de bisecciones sucesivas.

Supóngase que se desea minimizar una función f bajo un intervalo cerrado y acotado. Además supóngase que f(x) es pseudo-convexa y diferenciable. En la k-esima iteración, sea el intervalo de incertidumbre [a_k, b_k].Supóngase que la derivada $f'(x_k)$ es conocida y considere los siguientes tres casos posibles:

1.- Si $f'(x_k) = 0$, entonces por pseudo-convexidad de la función f, x_k es un punto mínimo.

2.- Si $f'(x_k)>0$, entonces para $x > x_k$, se tiene que $f'(x_k)(x - x_k) > 0$ y por pseudo-convexidad de f tenemos que $f(x) \geq f(x_k)$. En otras palabras, el mínimo se encuentra a la izquierda de x_k, tal que el nuevo intervalo de incertidumbre $[a_k + 1, b_k - 1]$ es dado por $[a_k, b_k]$.

3.- Si $f'(x_k) < 0$, entonces para $x < x_k$, $f'(x_k)(x - x_k)>0$, tal que $f(x) > f(x_k)$. Entonces el mínimo se encuentra a la derecha de x_k, tal que el nuevo intervalo de incertidumbre $[a_k+1, b_k+1]$ es dado por $[x_k, b_k]$.

La posición de x_k en el intervalo $[a_k, b_k]$ puede ser elegida tal que la máxima longitud posible del nuevo intervalo de incertidumbre es minimizado. Esto es, x_k puede ser elegida tal que minimice el máximo de $x_k - a_k$ y $b_k - x_k$. Obviamente la localización óptima de x_k es el punto medio $1/2(a_k - b_k)$.

Procedimiento

1.- Sea $[a_1, b_1]$ el intervalo inicial de incertidumbre y sea g el intervalo final de incertidumbre. Sea n el menor entero positivo tal que $(1/2)n \cdot \ell\, b - a_1$. Sea k 1 y continúe con el paso 2.

2.- Sea f_k $1/2(a_k - b_k)$ y evalúe $f'(x_k)$. Si $f'(x_k)$ 0 parar; x_k es una solución óptima. De lo contrario continuar con el paso 3. si $f'(x_k) > 0$ y continuar con el paso 4 si $f'(x_k) < 0$.

3- Sea a_k+1 a_k y b_k+1 x_k. Continuar con el paso 5.

4.- Sea a_k+1 x_k y b_k+1 b_k- Continuar con el paso 5.

5 - Si k n parar, el mínimo se ubica en el intervalo $[a_k+1, b_k+1]$ de lo contrario reemplazar k por k + 1 y repetir el paso 2

Ejemplo: Minimizar $f(x) = x^2 + 2x$

sujeta a $-3 \leq x \leq 6$

Reducir el intervalo de incertidumbre a un intervalo cuya longitud £ es menor o igual a 0.2.

El numero de observaciones n satisface

$(1/2)n \leq £/b_1-a_1$ 0.2 6 - (-3) ~ 0.2/9 0.0222

para n = 6.

La siguiente tabla resume los cálculos usando el método de bisecciones sucesivas.

Interacción k	a_k	b_k	x_k	$f'(x_k)$
1	-3.000	6.0000	1.500	5.0000
2	-3.000	1.5000	-0.750	0.5000
3	-3.000	-0.7500	-1.8750	-1.7500
4	-1.875	-0.7500	-1.3125	-0.6250
5	-1.3125	-0.7500	-1.0313	-0.6250
6	-1.0313	-0.7500	-0.8907	0.2186
7	-1.0313	-0.8907		

El intervalo final de incertidumbre es [-1.0313, -0.8907], tal que el minimo puede ser tomado como el punto medio del intervalo -0.961

Método de Interpolación Cubica

Este método a diferencia del de interpolacion cuadratica requiere que la funcion sea diferenciable. El método consiste en encontrar un valor óptimo de una variable α, denominada α^*, tal que la función

$$g(\alpha) = f(x + \alpha s)$$

obtenga un mínimo local, donde X y s son valores iniciales arbitrarios. El punto X es el punto de partida de la busqueda de los óptimos, mientras que s es la direccion de búsqueda. Este mínimo local se obtiene en tres pasos o fases.

Procedimiento

1.- Dados valores arbitrarios de x y s evalúe $g(\alpha) = f(x + \alpha s)$

2.- Evalúe $g(\alpha)$ y $g'(\alpha)$ para valores de α igual a 0,1,2,4,8,16, ...,a, b, donde b es el primer valor para el cual $g'(\alpha)$ es no negativo ó $g(\alpha)$ no ha decrecido. El valor óptimo se encuentra en el rango $a < \alpha^* \leq b$.

3.- Se ajusta un polinomio cubico tomando en consideración los valores g(a), g(b), g'(a) y g'(b). El valor mínimo α^* se representa en esta iteración por α_e, donde

$$\alpha_e = b - \frac{g'(b) + W - Z}{g'(b) - g'(a) + 2W}(b-a)$$

$$Z = 3\left(\frac{g(a) - g(b)}{b-a}\right) + g'(a) + g'(b)$$

$$W = (Z^2 - g'(a)g'(b))^{1/2}$$

Si g(a) o g(b) no son menores a $g(\alpha_e)$, entonces se acepta que $\alpha^* = \alpha_e$, manejando un valor aproximado. Si no, es decir, si $g(a) < g(\alpha_e)$ o $g(b) < g(\alpha_e)$, entonces:

a) si $g'(\alpha_e) > 0$, se repite el mismo procedimiento en el intervalo $a < \alpha^* \leq b$ donde $a = a$ y $b = \alpha_e$. Regresar al paso 3

b) si $g'(\alpha_e) < 0$, se repite el mismo procedimiento en el intervalo $a \leq \alpha^* < b$, donde $a = \alpha_e$ y $b = b$. Regresar al paso 3.

Ejemplo: Encontrar un mínimo local de la función unimodal de una sola variable $f(x) = x^4 - 10x^3 + 35x^2 - 50x$, tomando como valores iniciales $x = 1$ y dirección inicial $s = 1$.

Fase 1

$$g(\alpha) = f(x + \alpha s)$$
$$= f(1 + \alpha(1))$$
$$= f(1 + \alpha)$$
$$= (1+\alpha)^4 - 10(1+\alpha)^3 + 35(1+\alpha)^2 - 50(1+\alpha)$$
$$g'(\alpha) = 4(1+\alpha)^3 - 30(1+\alpha)^2 + 70(1+\alpha) - 50$$

$$g'(\alpha) = 4(1+\alpha)^3 - 30(1+\alpha)^2 + 70(1+\alpha) - 50$$

Fase 2

α	$g(\alpha)$	$g'(\alpha)$
0	-24	-6
1	-24	2

Para $\alpha = 1$, se tiene el primer valor donde $g'(\alpha) > 0$. Por lo tanto $b = 1$ y $a = 0$. Se tiene que $0 \leq \alpha^* \leq 1$.

Fase 3

$$Z = 3\left(\frac{g(a) - g(b)}{b - a}\right) + g'(a) + g'(b)$$

$$= 3 \frac{-24 + 24}{1 - 0} - 6 + 2 = -4$$

$$W = (Z^2 - g'(a)g'(b))^{1/2}$$

$$= ((-4)^2 - (-6)(2))^{1/2}$$

$$= (16 + 12)^{1/2} = 5.29$$

$$\alpha_e = b - \frac{g'(b) + W - Z}{g'(b) - g'(a) + 2W}(b - a)$$

$$\alpha = 1 - \left(\frac{2 + 5.29 + 4}{2 + 6 + 2(5.29)}\right)(1 - 0)$$

$$= 0.39$$

$$g(\alpha_e) = g\ 0.39) = -25$$

Como $g(\alpha_e) < g(a)$ y $g(\alpha_e) < g(b)$ entonces el mínimo local α^* se acepta que sea α_e, es decir,

$$\alpha^* = \alpha_e = 0.39$$

Por lo tanto, como $g(\alpha^*) = f(1 + \alpha^*)$, el óptimo relativo es:

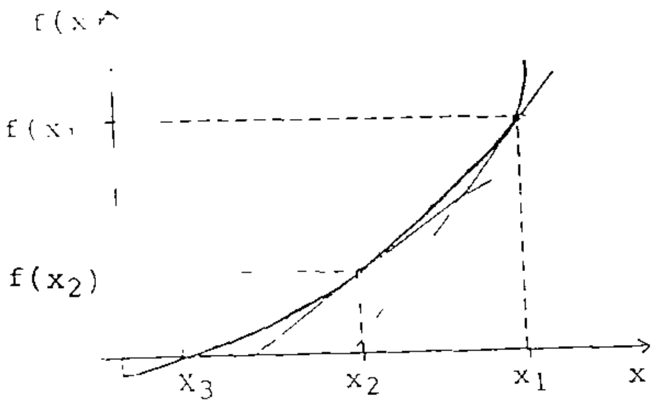
$$x^* = 1 + \alpha^* = 1 + 0.39 = 1.39$$

$$f(x^*) = f(1.39) = -25.$$

Valor mínimo $x = 1.39$.

Método de Newton-Raphson

Este método es también llamado "Método de Newton de Primer Orden" o "Método de las Tangentes" y se utiliza para encontrar la raíz de una ecuación. Para su aplicación se requiere conocer un valor de "x" cercano a la solución y que para este punto el valor de la primera derivada sea diferente de cero, tal y como se ilustra en la figura 1



Para obtener la ecuación que representa el algoritmo de solución, se procede de la siguiente forma:

1. Para encontrar el valor x_2 (ver figura) considere que el punto $(x_1, f(x_1))$ esta razonablemente cerca de la solución exacta y que la pendiente en dicho punto es diferente de cero. Trazando una tangente en el punto $(x_1, f(x_1))$ de tal forma que cruce el eje X, el valor de la pendiente de dicha tangente esta dada por:

$$\text{Pendiente de la tangente} = \frac{f(x_1) - y}{x_1 - x}$$

Como para x_2 el valor de $y = 0$, y la pendiente de la tangente en el punto $(x_1, f(x_1))$ esta dada por la primera derivada de la funcion evaluada en x_1, entonces:

$$\text{Pendiente de la tangente} = \frac{f(x_1)}{x_1 - x_2} = f'(x_1)$$

Por lo tanto, el nuevo valor x_2 mas cercano a la raiz puede ser calculado si se conoce el valor de la primera derivada en el punto donde se trazo la tangente, dicho valor esta dado por la siguiente ecuación:

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}$$

2.- En forma similar se tiene que, trazando una tangente en el punto (x_2, $f(x_2)$) se obtiene:

$$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)}$$

3.- Generalizando el procedimiento anterior, se tiene que la formula del algoritmo para este método esta dada por:

$$x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}$$

Aunque el método de Newton-Raphson es utilizado para encontrar la raíz de una ecuación, también puede ser utilizada para determinar el punto en el que se optimiza una función no lineal.

Supóngase que se tiene una función no lineal de una sola variable que puede ser multimodal, para la cual interesa determinar alguno de los valores que optimizan la funcion, en un problema no restringido.

Entonces, podemos afirmar que en el punto en el que la funcion toma un valor óptimo, ya sea mínimo o máximo, la pendiente de la función g(x) es cero. De tal forma que si graficamos la funcion $f(x) = g'(x)$ y aplicamos el metodo de Newton-Raphson para encontrar la raíz o una de las raices de f(x), esto equivaldria a que para el valor de "x" para el cual f(x) es igual a cero, la pendiente de la función g(x) también es cero ($g'(x) = 0$) y si para ese valor de "x" la pendiente de g(x) es cero, entonces se trata de un valor optimo de la función g(x)

Si a partir de un valor conocido de "x" se desea encontrar un valor que optimice la función dada g(x), entonces aplicando el método de Newton-Raphson, se procederá a hacer:

$$f(x) = g'(x)$$

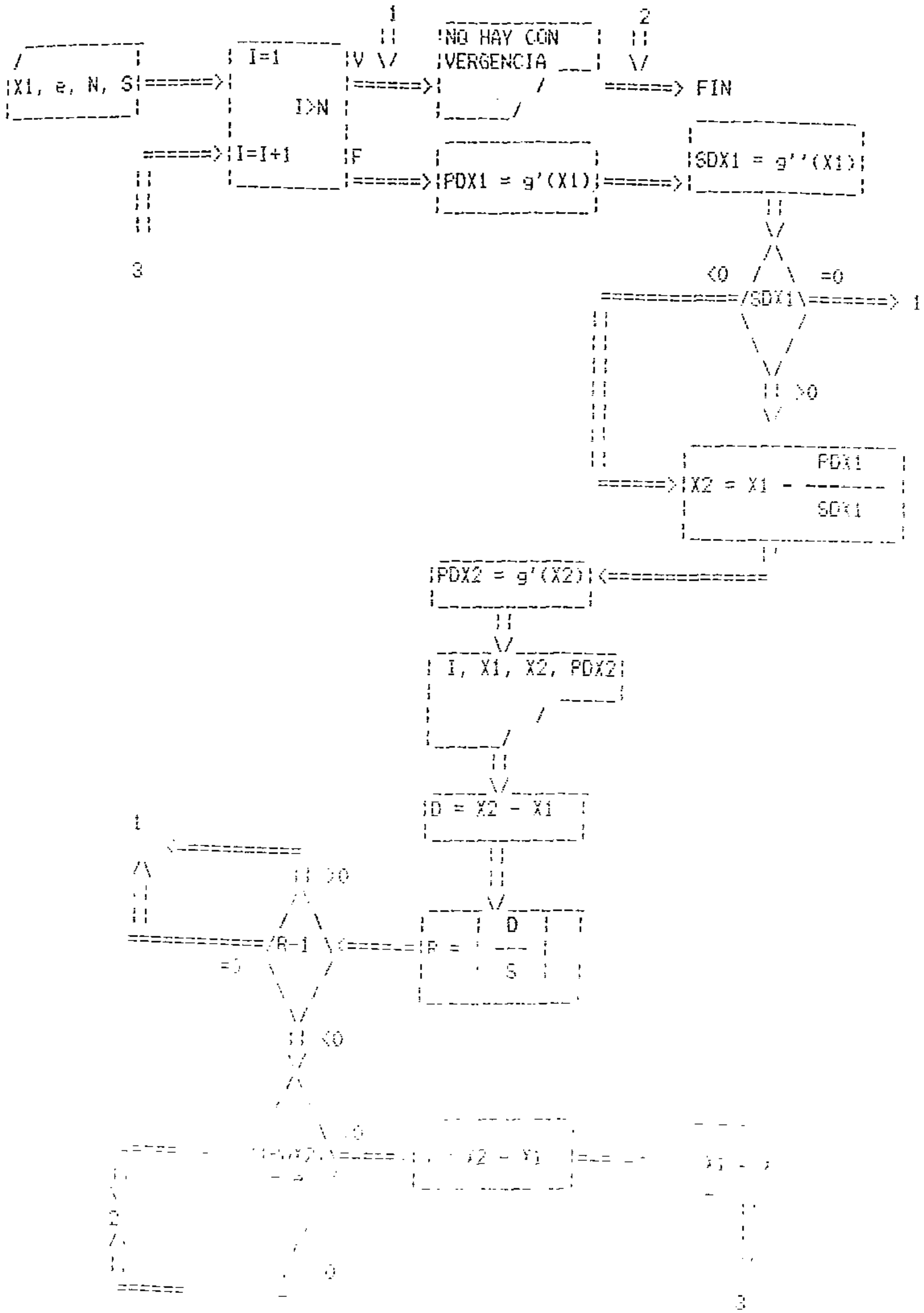
y aplicar la ecuacion general:

$$x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}$$

o bien, a partir de la función g(x) esta ecuación puede expresarse como:

$$x_n = x_{n-1} - \frac{g(x_{n-1})}{g'(x_{n-1})}$$

El diagrama de flujo del método de Newton-Raphson se presenta a continuación.


1
2
NO HAY CON VERGENCIA
X1, e, N, S
I=1
I>N
V
F
I=I+1
FIN
3
PDX1 = g'(X1)
SDX1 = g''(X1)
<0
=0
SDX1
>0
X2 = X1 - PDX1 / SDX1
PDX2 = g'(X2)
I, X1, X2, PDX2
D = X2 - X1
R = D / S
R-1
<0
X2 - X1

Ejemplo: Encontrar un valor óptimo de la función $g(x) = (80/x) + 20x + 20$, a partir de los siguientes datos: $x_1 = 1$, $E = 0.00001$, $N = 15$ y $s = 100$

$$f(x) = g'(x) = \frac{-80}{x^2} + 20 \qquad f'(x) = g''(x) = \frac{160}{x^3}$$

Los resultados se resumen en la siguiente tabla:

I	X_1	PDX_1	SDX_1	X_2	PDX_2	D	R	R-1	PDX_2-E	S	$g(X_2)$
1	1.0000	-60.00000	160.00	1.3750	-22.31405	0.37500	0.00375	-0.99625	22.31404	0.3750	105.68182
2	1.3750	-22.31405	61.55	1.7375	-6.49817	0.36255	0.96680	-0.03320	6.49816	0.3625	100.79285
3	1.7375	-6.49817	30.50	1.9506	-1.02587	0.21305	0.58765	-0.41235	1.02586	0.2130	100.02502
4	1.9506	-1.02587	21.56	1.9982	-0.03635	0.04759	0.22335	0.77665	0.03634	0.0475	100.00003
5	1.9982	0.03635	20.05	2.0000	-0.00005	0.00181	0.03810	-0.96190	0.00004	0.0018	100.00000
6	2.0000	-0.00005	20.00	2.0000	[illegible]	0.00000	0.00136	[illegible]	-0.00001		100.0000

para $x = 2$, el valor óptimo de $g(x)$ sera 100.

Ejemplo: Encontrar un valor óptimo de la funcion $g(x) = 2x^4 - 7x^3 + 4x - 2$, a partir de los siguientes datos: $x_1 = -1$, $E = 0.00001$, $N = 15$ y $S = 100$

$$f(x) = g'(x) = 8x^3 - 21x^2 + 4 \qquad f'(x) = 24x^2 - 42x$$

Resumiendo las iteraciones y resultados, tenemos la siguiente tabla a continuacion:

I	X_1	PDX_1	SDX_1	X_2	PDX_2	D	R	R-1	PDX_2|-E	S	$g(X_2)$
1	-1.0000	-25.00000	66.00	-0.6212	-6.02182	0.37879	0.00379	-0.99621	6.02181	0.37879	-2.50890
2	-0.6212	-6.02182	35.35	-0.4509	-1.00234	0.17034	0.00170	-0.99829	1.00233	0.17034	-3.07249
3	-0.4509	-1.00234	23.82	-0.4088	-0.05577	0.04209	0.00042	-0.99957	0.05576	0.04209	-3.10112
4	-0.4088	-0.05577	21.18	-0.4062	-0.00021	0.00263	0.00003	-0.99997	0.00020	0.00263	-3.10119
5	-0.4062	-0.00021	21.02	-0.4061	0.00000	0.00001	0.00000	-0.99999	-0.00001		-3.10119

Para x = -0.4061, el valor óptimo de g(x) sera: -3.10119

Metodos de busqueda multidimensional

Optimización de funciones sin usar derivadas

Una de las grandes desventajas de los metodos que emplean derivadas para funciones de varias variables es que en estos es necesario evaluar el gradiente y en varios casos, tambien el Hessiano.

Otro problema en consideración, en el caso de que la funcion sea diferenciable, es el evaluar analíticamente en cada iteración las funciones contenidas en el gradiente y el Hessiano. Esta tarea de evaluacion analítica, puede ser muy laboriosa y en ocasiones imposible, sobre todo cuando se trate de funciones complejas con muchas variables independientes. Hay métodos numéricos que permiten expresar estas derivadas por procesos que no son analíticos, y por lo tanto se puede eliminar este problema. La gran desventaja con los metodos numericos es el error de redondeo que introducen y, que acumulado en un gran número de operaciones e iteraciones, puede generar resultados bastante alejados de los óptimos reales

Por este motivo, se han desarrollado una serie de metodos iterativos que, sin ser tan eficientes en cuanto al número de iteraciones necesarias para alcanzar un óptimo como los metodos basados en el gradiente (que veremos posteriormente), tienen las siguientes ventajas.

a) No requieren de gradientes y Hessianos, por lo que sirven tanto para funciones diferenciables como no diferenciables

b) No requieren de métodos numéricos y por lo tanto su error de redondeo se reduce considerablemente.

Analizaremos ahora algunos de estos métodos.

Metodo de Coordenadas Cíclicas

Este método usa los ejes coordenados como direcciones de búsqueda. Mas específicamente, el método busca a lo largo de las direcciones $d_1,...,d_n$, donde d_j es un vector de ceros, excepto por un 1 en la coordenada j. Asi, a lo largo de la dirección de búsqueda d_j, la variable x_j es cambiada mientras todas las otras variables quedan fijas.

Si la función es diferenciable, entonces el metodo converge a un punto estacionario.

Procedimiento

Paso 0: Elegir un escalar $\varepsilon > 0$ para ser usado para saber cuando terminar el algoritmo como margen de incertidumbre y sean $d_1,..., d_n$ las direcciones coordenadas. Elegir un punto inicial x_1, sea y_1 x_1, k j 1 y continuar con el paso 1

Paso 1: Sea λ una solución óptima a el problema de minimizar la funcion $f(y_j + \lambda d_j)$ sujeta a $\lambda \in E_1$ y sea $y_j+1 = y_j + \lambda_j d_j$. Si $j<m$, reemplazar j por j+l y repetir el paso 1. De otra manera, si j n, continuar con el paso 2.

Paso 2 : Sea x_{k+1} y_{n+1}. Si $\|x_{k+1} - x_k\| < \varepsilon$, parar. De lo contrario sea y_1 x_{k+1}, sea $j = 1$, reemplazar k por k + 1 y repetir el paso 1.

Ejemplo: Minimizar $(x_1 - 2)^2 + (x_1 - 2x_2)^2$ partiendo del punto inicial (0,3).

La tabla siguiente muestra un resumen de los calculos para el metodo de coordenadas ciclicas. En cada iteración los vectores y_1 y y_2 son obtenidos mediante la representacion de una linea de búsqueda en las direcciones (1, 0) y (0, 1), respectivamente. Para el punto (2.22, 1.11) la funcion objetivo tiene un valor de 0.0023, muy cercano del valor óptimo 0, con punto óptimo (0, 0). De hecho, podemos hacer que el valor óptimo sea mas aproximado a 0, haciendo ε mas pequeño, pero utilizando un numero mayor de iteraciones.

Iteracion k	x_i $f(x_i)$	j	d_j	y_i	λ_j	y_{i+1}
1	(0.00, 3.00)	1	(1.0, 0.0)	(0.00, 3.00)	3.13	(3.13, 3.00)
	52.00	2	(0.0, 1.0)	(3.13, 3.00)	-1.44	(3.13, 1.56)
2	(3.13, 1.56)	1	(1.0, 0.0)	(3.13, 1.56)	-0.50	(2.63, 1.56)
	1.63	2	(0.0, 1.0)	(2.63, 1.56)	-0.25	(2.63,1.31)
3	(2.63, 1.31)	1	(1.0, 0.0)	(2.63, 1.31)	-0.19	(2.44, 1.31)
	0.16	2	(0.0, 1.0)	(2.44, 1.31)	-0.09	(2.44, 1.22)
4	(2.44, 1.22)	1	(1.0, 0.0)	(2.44, 1.22)	-0.09	(2.35, 1.22)
	0.04	2	(0.0, 1.0)	(2.35, 1.22)	-0.05	(2.35, 1.17)
5	(2.35, 1.17)	1	(1.0, 0.0)	(2.35, 1.17)	-0.06	(2.29, 1.17)
	0.015	2	(0.0, 1.0)	(2.29, 1.17)	-0.03	(2.29, 1.14)
6	(2.29, 1.14)	1	(1.0, 0.0)	(2.29, 1.14)	-0.04	(2.25, 1.14)
	0.007	2	(0.0, 1.0)	(2.25, 1.14)	-0.02	(2.25, 1.12)
7	(2.25, 1.12)	1	(1.0, 0.0)	(2.25, 1.12)	-0.03	(2.22, 1.12)
	0.004	2	(0.0, 1.0)	(2.22, 1.12)	-0.01	(2.22, 1.11)

La ilustración de este método, para este ejemplo, se presenta en la siguiente figura. (ver figura 2)

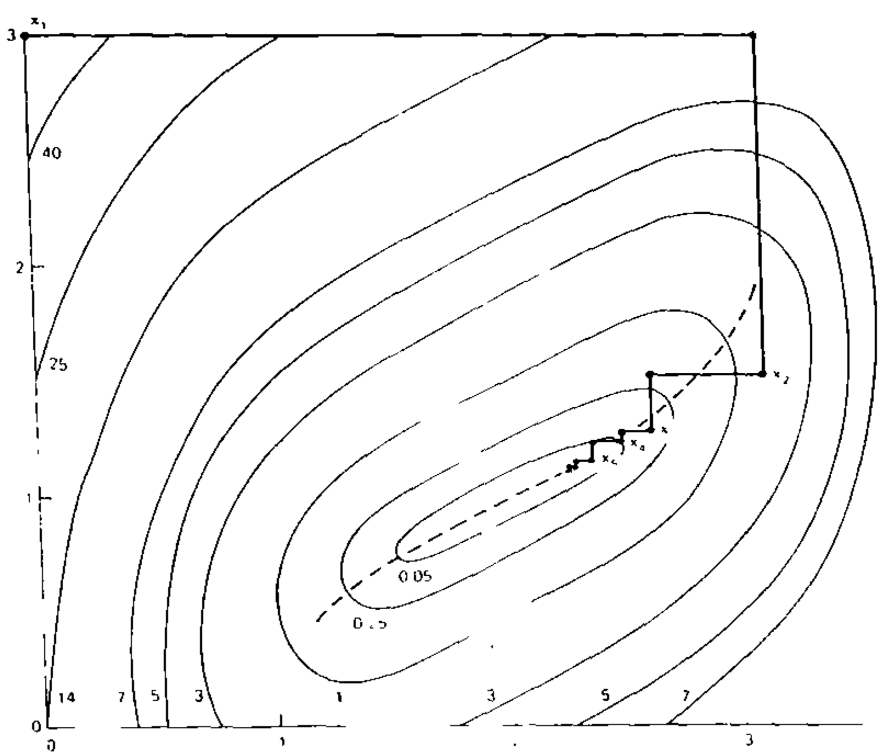


Figura 2 Ilustracion del metodo de coordenadas ciclicas.

Metodo de Hooke y Jeeves

El metodo de Hooke y Jeeves representa dos tipos de busqueda exploratoria y patron de busqueda. Dado x_1, una busqueda exploratoria a lo largo de de las direcciones coordenadas produce el punto $\mathbf{x}_2$. Un patron de busqueda a lo largo de la direccion $\mathbf{x}_2$ $\mathbf{x}_1$ conduce al punto $\mathbf{y}$. Otra busqueda exploratoria a partir del punto $\mathbf{y}$ conduce al punto x_3. El siguiente patron de busqueda es a lo largo de la direccion $\mathbf{x}_3$ $\mathbf{x}_2$, obteniendose $\mathbf{y}'$. El proceso es repetido. La figura 3 muestra las primeras dos iteraciones del metodo. Las lineas que van de $\mathbf{x}_1$ a $\mathbf{x}_2$ y las que van de y a x, representan la busqueda exploratoria a lo largo de los ejes coordenados y la linea que une a $\mathbf{x}_2$ y $\mathbf{x}_3$ y la que une a $\mathbf{x}_2$ y $\mathbf{y}$ representan el patron de busqueda.

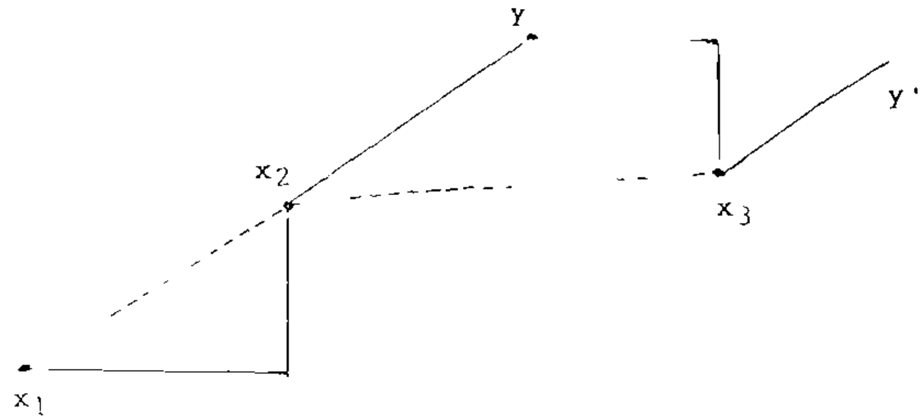


Figura 3 . patron de busqueda

El metodo de Hooke y Jeeves maneja direcciones de linea a lo largo de las direcciones coordenadas $d_1,\ldots,d_n$ y un patron de busqueda en su desarrollo.

Procedimiento

Paso 0. Elegir un escalar $\varepsilon > 0$ como margen de incertidumbre. Elegir un punto de partida x_1, sea $y_1 = x_1$, sea $k=j=1$ y seguir con el paso 1.

Paso 1: Sea λ_j una solucion optima a el problema de minimizar la funcion $f(y_j + \lambda d_j)$ sujeta a $\lambda \in E_1$ y sea $y_{j+1} = y_j + \lambda_j d_j$. Si $j<n$, reemplazar j por $j+1$, y repetir el paso 1. De lo contrario, si $j=n$, sea $x_{k+1} = y_{n+1}$. Si $\| x_{k+1} - x_k \| < \varepsilon$, parar, de lo contrario, continuar con el paso 2.

Paso 2: Sea $d = x_{k+1} - x_k$ y sea λ una solucion optima al problema de minimizar $f(x_{k+1} + \lambda d)$ sujeto a $\lambda \in E_1$. Sea $y_1 = x_{k+1} + \lambda d$, sea $j = 1$, remplazar k por $k+1$, y repetir el paso 1.

Ejemplo: Minimizar $(x_1 - 2)^4 + (x_1 - 2x_2)^2$ partiendo del punto (0,3).

La siguiente tabla resume los calculos del metodo de Hooke y Jeeves.

En cada iteracion, una busqueda exploratoria a lo largo de las direcciones coordenadas dados los puntos y_2 y y_3, y un patron de busqueda a lo largo de la direccion $d = x_{k+1} - x_k$ dado el punto y_1, excepto en la iteracion k = 1, donde $y_1 = x_1$. En cuatro iteraciones se llega del punto inicial a la solucion optima(2.00, 1.00), con valor objetivo igual a cero. En este punto, $\|x_5 - x_4\| = 0.045$ y el procedimiento es terminado. La figura 4 ilustra el progreso del metodo.

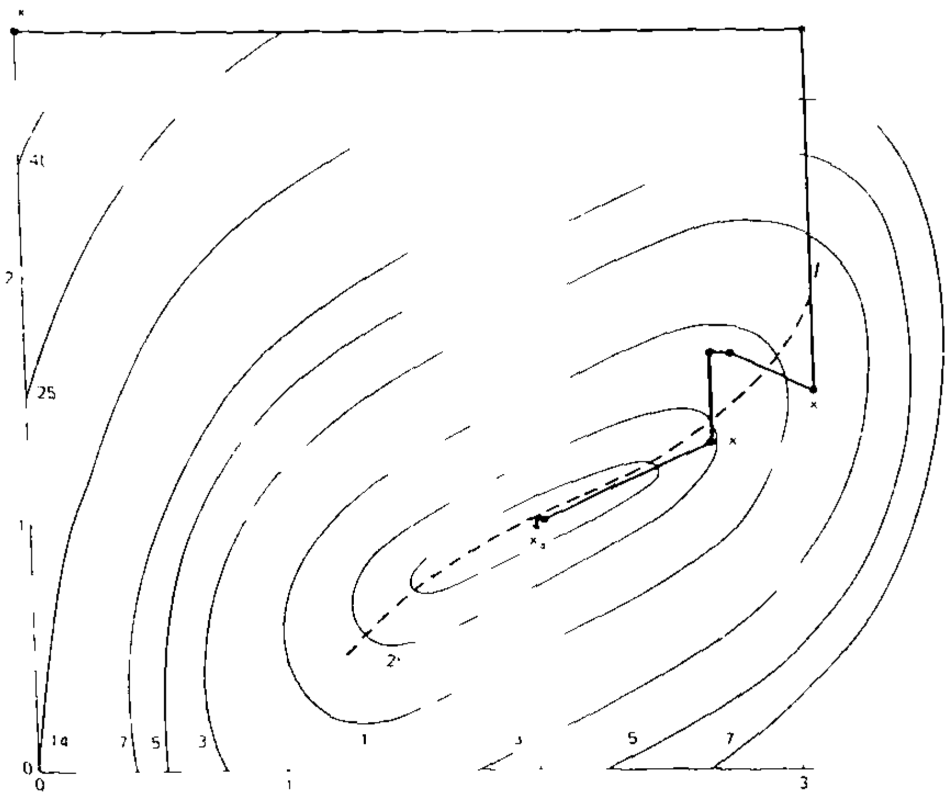


Figura 4 : Ilustracion del metodo de Hooke y Jeeves

Metodo de busqueda de Hooke y Jeeves

Iteracion k	x_i f(x_i)	j	y_j	d	λ	y_{j+1}	d	λ	$y_3 + \lambda d$
1	(0.00, 3.00)	1	(0.00, 3.00)	(1.0, 0.0)	3.13	(3.13, 3.00)		—	—
	52.00	2	(3.13, 3.00)	(0.0, 1.0)	-1.44	(3.13, 1.56)	(3.13, -1.44)	-0.10	(2.82, 1.70)
2	(3.13, 1.56)	1	(2.82, 1.70)	(1.0, 0.0)	-0.12	(2.70, 1.70)		—	
	1.63	2	(2.70, 1.70)	(0.0, 1.0)	-0.35	(2.70, 1.35)	(-0.43, -0.21)	1.50	(2.06, 1.04)
3	(2.70, 1.35)	1	(2.06, 1.04)	(1.0, 0.0)	-0.02	(2.04, 1.04)		—	
	0.24	2	(2.04, 1.04)	(0.0, 1.0)	-0.02	([illegible]4, 1.02)	(-0.66, [illegible])	0.06	(2.[illegible], 1.[illegible])
4	(2.04, 1.02)	1	(2.00, 1.00)	(1.0, 0.0)	0.0[illegible]	[illegible]			
	0.000003	2	(2.00, 1.00)	(0.0, 1.0)	[illegible].00	([illegible], 1.0[illegible])			
5	(2.00, 1.00)								
	0.00								

Metodo de Rosenbrock

El metodo de Rosenbrock utiliza lineas de busqueda para minimizar una funcion f de varias variables. Si f es diferenciable entonces el metodo converge a un punto con gradiente cero.

Procedimiento:

Paso 0: Sea $\varepsilon > 0$, elegido como terminacion escalar. Elegir $d_1,\ldots,d_n$ como las direcciones coordenadas. Elegir un punto de partida x_1, sea $y_1 = x_1$, $k = j = 1$, y continuar con el paso 1.

Paso 1: Sea λ_j una solucion optima a el problema de minimizar $f(y_j + \lambda d_j)$ sujeto a $\lambda \in E_1$, y sea $y_{j+1} = y_j + \lambda_j d_j$. Si $j<n$, reemplazar j por j+1 y repetir el paso 1. De lo contrario, seguir con el paso 2.

Paso 2: Sea $x_{k+1} = y_{n+1}$. Si $\|x_{k+1} - x_k\| < \varepsilon$, parar, de lo contrario sea $y_1 = x_{k+1}$, reemplazar k por k+1, sea $j = 1$ y continuar con el paso 3.

Paso 3: Formar un nuevo conjunto de direcciones de busqueda linealmente independientes. (las direcciones $d_1,\ldots,d_n$ son linealmente independientes si $\sum_{j=1}^{n} \lambda_j d_j = 0$ implica que $\lambda_j = 0$ para $j = 1,\ldots,n$) cada uno con norma igual a uno. Ademas, supongase que estos vectores son mutuamente ortogonales, esto es, $d_i^t d_j = 0$ para $i \neq j$. Partiendo del vector comun x_k, la funcion objetivo f es minimizada a lo largo de las direcciones iterativamente, conduciendo a el punto x_{k+1}. En particular $x_{k+1} - x_k = \sum_{j=1}^{n} \lambda_j d_j$, donde λ_j es la distancia recorrida a lo largo de d_j. La nueva coleccion de direcciones $d_1,\ldots,d_n$ son obtenidas a partir del procedimiento de Gram-Schmidt como sigue

$$a_j = \begin{cases} d_j & \text{si } j = 0 \\ \sum_{i=j}^{n} \lambda_i d_i & \text{si } j \neq 0 \end{cases}$$

$$a \qquad\qquad j \;\; 1$$

$$b \quad a \quad \sum (a_j\, d\,)d \qquad j > 0$$

$$d \quad \frac{b}{b}$$

Ejemplo: Minimizar $(x_1 - 2)^4 + (x_1 - 2x_2)^2$ usando el punto inicial (0, 3).

Para $y_1 = (0, 3)$, el punto y_2 es obtenido optimizando la funcion a lo largo de y_1, y y_3 es obtenido optimizando la funcion a lo largo de la direccion d_2 partiendo de y_2. Despues de la primera iteracion tenemos que $\lambda_1 = 3.13$ y $\lambda_2 = -1.44$. Usando el metodo de Gram-Schmidt, las nuevas direcciones de busqueda son (0.91, -0.42) y (-0.42, -0.91). Despues de cuatro iteraciones, el punto (2.21, 1.1) es alcanzado, y el valor de la funcion objetivo es 0.002. A partir de esto se tiene que $\|x_4 - x_3\| = 0.15$ y el procedimiento termina. En la figura 5 se observa el progreso del metodo de Rosenbrock., y la tabla siguiente condensa los calculos de cuatro iteraciones del metodo hasta alcanzar el punto optimo.

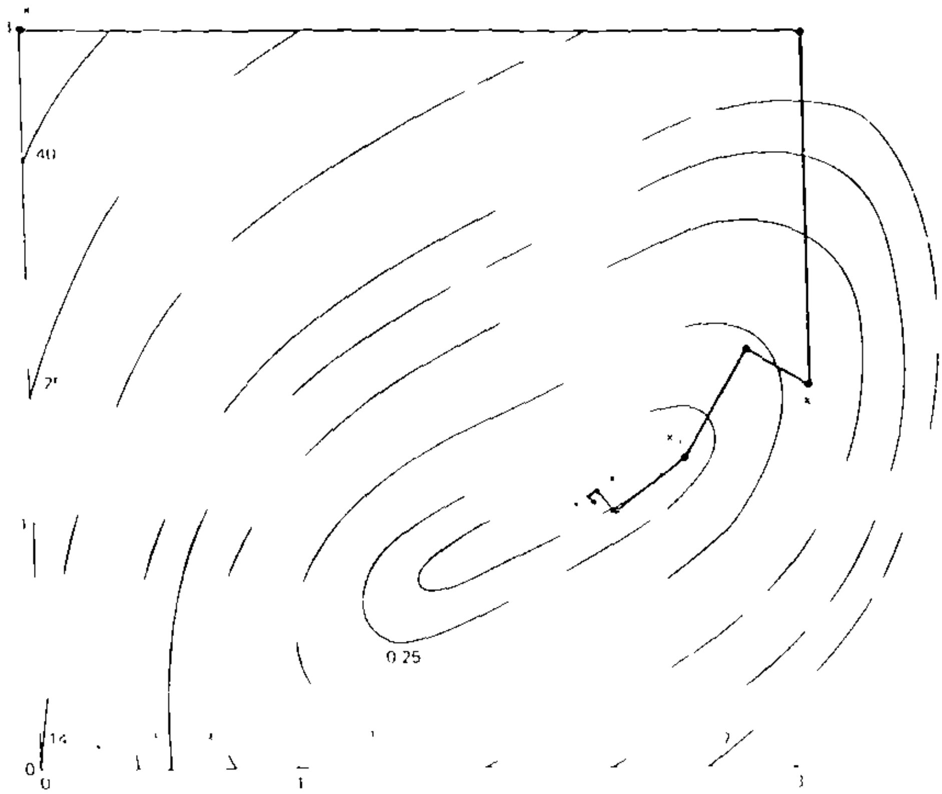


Figura 5 : Ilustracion del metodo de Rosenbrock

Resumen de los calculos del metodo de Rosenbrock

Iteracion k	x_k $f(x_k)$	j	y_j $f(y_j)$	d_j	λ	y_{j+1} $f(y_{j+1})$
1	(0.00, 3.00) 52.00	1	(0.00, 3.00) 52.00	(1.0, 0.0)	3.13	(3.13, 3.00) 9.87
		2	(3.13, 3.00) 9.87	(0.0, 1.0)	-1.44	(3.13, 1.56) 1.63
2	(3.13, 1.56) 1.63	1	(3.13, 1.56) 1.63	(0.91, -0.42)	-0.34	(2.82, 1.70) 0.79
		2	(2.82, 1.70) 0.79	(-0.42, -0.19)	0.51	(2.16, 1.24) 0.16
3	(2.61, 1.24) 0.16	1	(2.61, 1.24) 0.16	(-0.85, 0.52)	0.38	(2.29, 1.04) 0.05
		2	(2.29, 1.04) 0.05	(0.52, -0.85)	-0.10	(2.24, 1.13) 0.004
4	(2.24, 1.13) 0.004	1	(2.44, 1.13) 0.004	(-0.96, -0.28)	0.04	(2.20, 1.12) 0.003
		2	(2.20, 1.12) 0.003	(0.28, -0.96)	0.02	(2.21, 1.10) 0.002

Metodos de busqueda multidimensionales usando derivadas

La metodologia de las tecnicas de optimizacion no lineal para funciones sin usar derivadas consistia en establecer un metodo que cambie el valor de la funcion objetivo hacia la direccion del maximo o minimo, para los metodos que utilizan derivadas las cuestiones basicas de metodologia son las mismas:

1.- ¿ Hacia que direccion moverse en R^n ?
2.- ¿ Cuando parar las iteraciones ?

Obviamente la respuesta a la pregunta (2) es facil, parar las iteraciones cuando f se empiece a incrementar o disminuir

Una eleccion natural direccional es determinada mediante calculo: hacer que f crezca o decrezca tan rapido como sea posible.

Una manera de lograr esto es dada por el gradiente de f(x).

Los metodos de gradientes determinan la mejor direccion y la mejor longitud de recorrido en esa direccion, para poder alcanzar el maximo o minimo, en el menor tiempo posible (menor numero de iteraciones).

Metodo del gradiente

Ejemplo: Encontrar el valor optimo de la funcion

$f(x_1, x_2) = 7x_1 + 4x_2 + x_1x_2 - x_1^2 - x_2^2$ usando el metodo del gradiente con una aproximacion del 1% y a partir de los valores $x_1 = 8$ y $x_2 = 2$

Calculando el vector gradiente

$\nabla f(x) = (7 + x_2 - 2x_1, 4 + x_1 - 2x_2)^t$

$\nabla f(x_0) = (-7, 8)^t$

Partiendo del punto inicial x_0 (8, 2), al moverse una unidad en la direccion x, decrece el nivel de f en 7 unidades. En cambio, un movimiento unitario en la direccion x, hace crecer el nivel de f en 8 unidades. Entonces, del punto (8, 2) hay que moverse a (8, 2+8k) donde k es la longitud de paso. Para calcularla se tiene x (8, 2+8k)

$$f(x) = 7(8) + 4(2+8k) + 8(2+8k) - (8) - (2+8k)^2$$

$$= -64k^2 + 6k + 12$$

Derivando la funcion $f(x_1)$ e igualandola a cero:

$$-128k + 64 = 0$$

$$k = 0.05$$

Entonces hay que moverse al punto (8, 2+8(0.05 (8, 6) Se rep el procedimiento para x_1 (8, 6), hasta que en 6 iteraciones se llega a x_6 (6.03 5.625). El metodo convierte al punto optimo (6, 5) que genera un valor maximo de La tabla siguiente resume estos calculos.

Metodo del Gradiente

I	X1	X2	f(X1, X2)	Ganancia	Gradientes de		Vect[illegible]				f([illegible])			k
					X1	X2	X1	k	[illegible]2	k	k[illegible]	k	k0	par[illegible] (d. d, [illegible])
0	0.000	2.0000	12.0000		-7.0000	8.0000	8.000	0.0000	2[illegible]000	[illegible]000	-64.0000	64.000[illegible]	12.00	[illegible]
1	8.000	6.0000	28.0000	16.0000	-3.0000	0.0000	8.000	-3.0000	6.0000	0.0000	-9.0000	9.0000	28.00	0.5[illegible]
2	6.500	6.0000	30.2500	2.2500	0.0000	-1.5000	6.500	0.0000	6.0000	-1.5000	-2.2500	2.2500	30.25	0.50
3	6.500	5.2500	30.8125	0.5625	-0.7500	0.0000	6.500	-0.7500	5.2500	0.0000	-0.5625	0.5625	30.81	0.50
4	6.125	5.2500	30.9531	0.1406	0.0000	-0.3750	6.125	0.0000	5.2500	-0.3750	-0.140[illegible]	0.1406	30.95	0.5[illegible]
5	6.125	5.0625	30.9983	0.0352	-0.1875	0.0000	6.125	-0.1875	5.0625	0.0000	[illegible]52	[illegible]352	3[illegible]99	[illegible]
6	6.031	5.0625	30.9971	0.0088	0.0000	-0.0938	6.0[illegible]1	0.0000	[illegible]625	[illegible]0938	[illegible]088	0.[illegible]088	31.00	0.50

Metodo de Ascenso/Descenso acelerado

Sea $f(x)$ una funcion diferenciable. Dado un punto inicial de partida $x = (x_1, x_2, \ldots, x_n)$, la direccion de ascenso acelerado, en el punto x_0 esta dada por $\nabla f(x_0)$ y la longitud de recorrido estara dada por un parametro $k_i > 0$, tal que el punto generado en la iteracion i, para i = 0,1,2,.. es

$$x_{i+1} = x_i + k_i \nabla f(x_i) \qquad k_i > 0 \text{ para } i = 0,1,2,\ldots$$

Para encontrar el valor de k_i, se maximiza la funcion de una sola variable

$$g(k_i) = f(x_i + k_i \nabla f(x_i))$$

por algun metodo conocido.

El vector $x_i + k\nabla f(x_i)$ es tangente al contorno $g(k_i)$ y como $\nabla f(x_{i+1})$ es perpendicular a este contorno, los movimientos sucesivos son en angulos rectos.

Un criterio para terminar el algoritmo es comparar el valor de la funcion en dos iteraciones sucesivas y parar cuando esa diferencia es menor a una tolerancia establecida, es decir, cuando

$$| f(x_{i+1}) - f(x_i) | < e \qquad \text{donde } e > 0$$

Este metodo encontrara un minimo o maximo local, generalmente el que esta cercano al punto de inicializacion x_0.

Ejemplo: Encontrar el valor maximo de la funcion

$$f(x) = 8x_1 + 2x_2 + 2x_1x_2 - 2x_1^2 - x_2^2$$

con una aproximacion del 1% y a partir de los valores $x_1 = 2$ y $x_2 = 10$.

Calculando el vector gradiente

$$\nabla f(x) = (8 - 2x_2 - 4x_1 , 2 + 2x_1 - 2x_2)^t$$

$$\nabla f(x_0) = (20, -14)$$

Se parte del punto inicial $x_0 = (2+20k, 10-14k)^t$

$$f(x_1) = 8(2+20k) + 2(10-14k) + 2(2+20k)(10-14k) - 2(2+20k)^2-(10-14k)$$

simplificando se tiene

$$g(k) = -1556k^2 + 596k + 52$$

Derivando la funcion anterior e igualando su derivada a cero se tiene

$$-3112k + 596 = 0$$

$$k = 0.1915$$

Entonces $x_1 = (5.83, 7.319)$, calculando el valor del gradiente en

x_1, $\nabla f(x_1) = (-0.6813, -0.9769)$ se tiene que

$$x = (5.83-0.6838k, 7.319-0.9769k)^t$$

evaluando el punto en la funcion objetivo se llega a

$$g(k) = -0.55k^2 + 1.42k + 25.07$$

derivando la ecuacion e igualando a cero, se tiene que

$k = 1.2845$ y $x_2 = (4.952, 6.064)$

Se sigue el mismo procedimiento hasta llegar al punto (5, 6) que es el valor maximo en 5 iteraciones.

La tabla siguiente resume los calculos del metodo de ascenso descenso acelerado para este ejemplo.

Metodo de ascenso / descenso acelerado

					Gradientes de		Vector X^{i+1}				$f(x^{i+1})$			k
I	X1	X2	f(X1, X2)	*Ganancia*	X1	X2	X1	k	X2	k	k2	k	k0	para $(d_f/d_k)=0$
0	2.000	10.000	-32.0000		20.0000	14.0000	2.000	20.000	10.000	-14.000	-1556.00	596.00	-32.00	0.1915
1	5.830	7.319	25.0720	57.0720	-0.6838	-0.9763	5.830	-0.684	7.319	-0.977	-0.55	1.42	25.07	1.2845
2	4.952	6.064	25.9852	0.9132	0.3200	-0.2240	4.952	0.320	6.064	-0.224	-0.40	0.15	25.99	0.1915
3	5.013	6.021	25.9998	0.0146	-0.0109	-0.0156	5.013	-0.011	6.021	0.016	-0.001	0.0004	26.00	1.2845
4	4.999	6.001	25.999996	0.0002	0.0051	-0.0036	4.999	0.005	6.001	0.004	0.0001	0.00004	26.00	0.1915
5	5.000	6.000	26.0000	0.000004										

Metodo de Newton

El metodo de Newton es muy parecido al metodo de ascenso descenso acelerado, y esta basado en la siguiente regla iterativa

$$x_{k+1} = x_k - H^{-1}(x_k)\nabla f(x_k) \qquad \text{(para minimizacion)}$$

$$x_{k+1} = x_k + H^{-1}(x_k)\nabla f(x_k) \qquad \text{(para maximizacion)}$$

donde el punto x_0 en E_n es arbitrario, $f(x)$ es diferenciable, $\nabla f(x_k)$ es la transpuesta del gradiente evaluado en el punto x_k de la k-esima iteracion y $H^{-1}(x_k)$ es la matriz inversa del Hessiano evaluado en el punto x_k de la k-esima iteracion.

Ejemplo: Minimizar $(x_1 - 2) + (x_1 - 2x_2)^2$ partiendo del punto inicial (0, 3)

La siguiente tabla muestra un resumen de los calculos. En cada iteracion, $x_{k+1} = x_k - H(x_k)^{-1} \nabla f(x_k)$. Despues de 6 iteraciones se obtiene el punto $x_7 = (1.83, 0.91)$.

En este punto $\|\nabla f(x_7)\| = 0.04$ y el procedimiento es terminado. La tabla siguiente resume los calculos y la figura 6 ilustra el progreso del metodo.

Metodo de Newton

Iteracion k	x_k $f(x_k)$	$\nabla f(x_k)$	$H(x_k)$	$H(x_k)^{-1}$	$H(x_k)^{-1}\nabla f(x_k)$	x_{k+1}
1	(0.00, 3.00) 52.00	([illegible]4.0, 24.0)	$\begin{bmatrix} 50.0 & 4.0 \\ 4.0 & 8.0 \end{bmatrix}$	[illegible]	([illegible] 2.67)	([illegible])
2	(0.67, 0.33) 3.13	(−9.39, [illegible]0.04)	$\begin{bmatrix} 23.23 & 4.0 \\ 4.0 & 8.0 \end{bmatrix}$	$\frac{1}{169.84}\begin{bmatrix} 8.0 & [illegible] \\ [illegible] & 23.23 \end{bmatrix}$	([illegible]4[illegible], [illegible]0.23)	(1.11, [illegible])
3	(1.11, 0.56) 0.63	([illegible]2.84, −0.04)	$\begin{bmatrix} 11.50 & 4.0 \\ 4.0 & 8.0 \end{bmatrix}$	$\frac{1}{76}\begin{bmatrix} 8.0 & [illegible]4.0 \\ [illegible]4.0 & 11.50 \end{bmatrix}$	(0.30, 0.14)	(1.41, 0.70)
4	(1.41, 0.70) 0.12	(−0.80, [illegible]0.04)	$\begin{bmatrix} 6.18 & 4.0 \\ 4.0 & 8.0 \end{bmatrix}$	$\frac{1}{33.[illegible]}\begin{bmatrix} [illegible] & [illegible] \\ [illegible] & [illegible] \end{bmatrix}$	([illegible], [illegible]10)	(1.61, [illegible])
5	(1.61, 0.80) 0.02	([illegible]0.22, [illegible]0.04)	$\begin{bmatrix} 3.83 & 4.0 \\ 4.0 & 8.0 \end{bmatrix}$	$\frac{1}{[illegible]}\begin{bmatrix} [illegible] & [illegible] \\ [illegible] & [illegible] \end{bmatrix}$	([illegible])	(1.74, [illegible]87)
6	(1.74, 0.87) 0.005	(−0.07, 0.00)	$\begin{bmatrix} 2.81 & 4.0 \\ 4.0 & 8.0 \end{bmatrix}$	$\frac{1}{6.[illegible]}\begin{bmatrix} 8.0 & [illegible] \\ [illegible]4.0 & [illegible]81 \end{bmatrix}$	(0.09, 0.04)	(1.83, 0.91)
7	(1.83, 0.91) 0.0009	(0.0003, [illegible]0.04)				

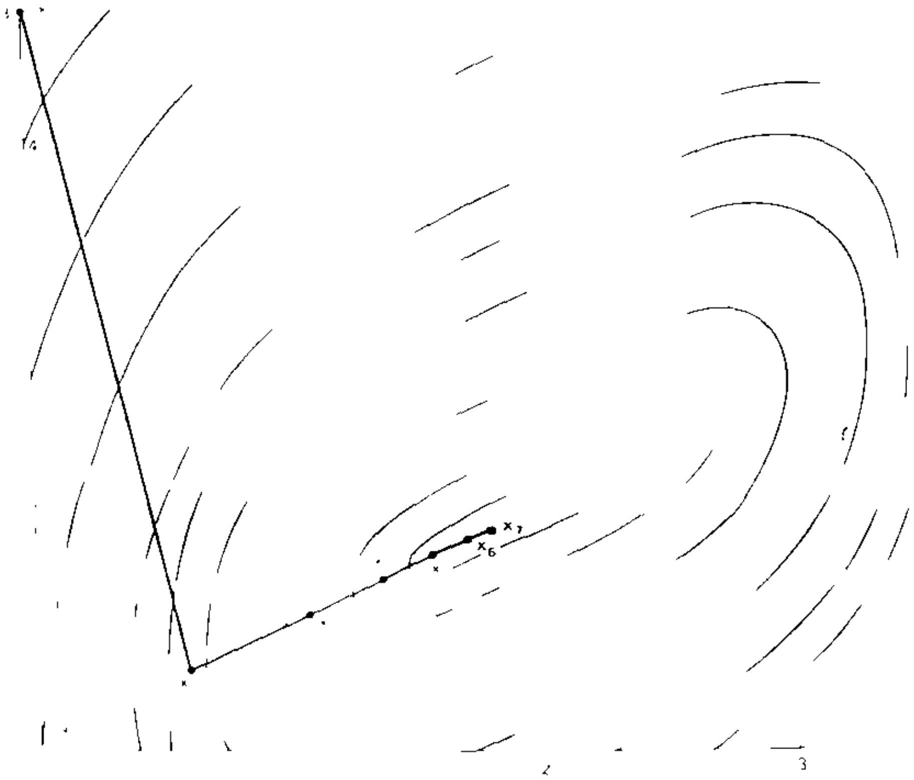


Figura 6: Ilustracion del metodo de Newton

Metodo de Davidon-Fletcher-Powell

Los tres metodos anteriores, el de ascenso-descenso acelerado, del gradiente y el de Newton convergen muy lentamente a la solucion buscada, sobre todo cuando el gradiente adquiere un valor cercano a cero. Esta situacion puede eliminarse utilizando direcciones conjugadas, en vez de direcciones perpendiculares (forman angulos de 90°) como las que manejan los metodos anteriormente citados.

El metodo de Davidon-Fletcher-Powell se basa en direcciones conjugadas , y es considerado como uno de los algoritmos mas eficientes para calcular puntos optimos locales en funciones diferenciables de varias variables.

Definicion: Sea H una matriz simetrica. Los vectores $d_1,\ldots,d_k$ son llamados conjugados si son linealmente independientes.

Procedimiento:

Paso 0: Sea $\varepsilon > 0$ el escalar que señala la terminacion del metodo. Elegir un punto inicial x_1 y una matriz simetrica positiva definida inicial D_1. Sea $y_1 = x_1$, sea $k = j = 1$, y continuar con el paso 1.

Paso 1: Si $\|\nabla f(x_i)\| < \varepsilon$, parar. De lo contrario, sea $d_j = -D_j \nabla f(y_j)$ sujeta a $\lambda \geq 0$. Sea $y_{j+1} = y_j + \lambda_j d_j$. Si $j < n$, continuar con el paso 2. Si $j = n$, sea $y_1 = x_{k+1} = y_{n+1}$, reemplazar k por $k + 1$, sea $j = 1$, y repetir el paso 1.

Paso 2: Construir D_{j+1} como sigue:

$$D_{j+1} = D_j + \frac{p_j p_j^t}{p_j^t q_j} - \frac{D_j q_j q_j^t D_j}{q_j^t D_j q_j}$$

donde

$p_j = \lambda_j d_j = y_{j+1} - y_j$

$q_j = \nabla f(y_{j+1}) - \nabla f(y_j)$

Reemplazar j por $j + 1$ y repetir el paso 1.

Ejemplo: Minimizar $(x_1 - 2)^4 + (x_1 - 2x_2)^2$

En la primera iteracion

$$D_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Para cada iteracion, para $j = 1,2$, d_j es dado por $-D_j \nabla f(y_j)$, donde D_1 es la matriz identidad y D_2 se obtiene con las formulas de D_{j+1}, p_j y q_j. En la iteracion 1, tenemos que $p_1 = (2.7, -1.49)^t$ y $q_1 = (44.33, -22.72)^t$ en la formula de D_{j+1}. En la iteracion 2, tenemos que $p_1 = (-0.1, 0.05)^t$ y $q_1 = (-0.7, 0.8)^t$, y finalmente en en la interacion 3 tenemos $p_1 = (-0.02, 0.02)^t$ y $q_1 = (-0.1, 0.24)^t$. El punto y_{j+1} es calculado optimizando a lo largo de la direccion d_j partiendo de y_j para $j = 1,2$.

El procedimiento es terminado en el punto $y_2 = (2.115, 1.058)$ en la cuarta iteracion, ya que $\nabla f(y_2) = 0.006$ es bastante pequeño. El. resumen de los calculos se expresa en la siguiente tabla, y la trayectoria tomada por el metodo se ilustra en La figura 7.

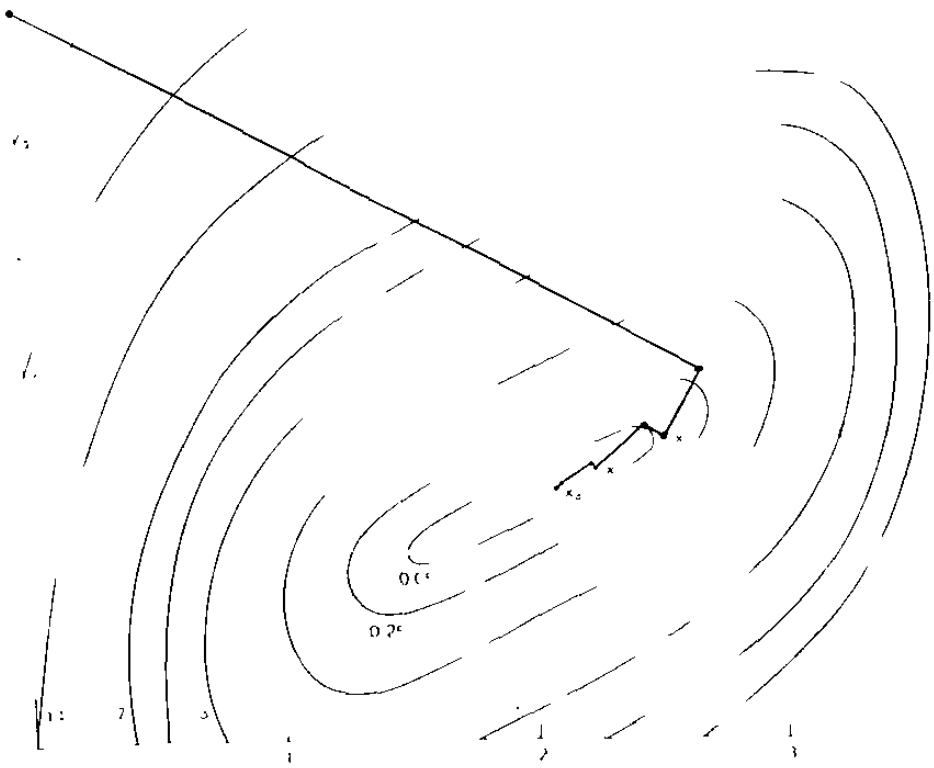


Figura 7: Ilustracion del metodo de Davidon-Fletcher-Powell

Método de Davidon-Fletcher-Powell

Iteración k	x_k $f(x_k)$	j	y_j $f(y_j)$	$\nabla f(y_j)$	$\lVert\nabla f(y_j)\rVert$	D	d	λ	y_{j+1}
1	(0.00, 3.00) (52.00)	1	(0.00, 3.00) (52.00)	(−44.00, 24.00)	50.12	$\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$	(44.00, −24.00)	0.062	(2.70, 1.51)
		2	(2.70, 1.51) (0.34)	(0.73, 1.28)	1.47	$\begin{bmatrix}0.25 & 0.38\\0.38 & 0.81\end{bmatrix}$	(−0.67, −1.31)	0.22	(2.55, 1.22)
2	(2.55, 1.22) (0.1036)	1	(2.55, 1.22) (0.1036)	(0.89, −0.44)	[illegible]	[illegible]	(−0.89, 0.44)	0.11	(2.45, 1.27)
		2	(2.45, 1.27) (0.0490)	(0.18, −0.36)	0.40	$\begin{bmatrix}0.65 & [illegible]\\0.45 & 0.46\end{bmatrix}$	(0.28, [illegible])	0.64	(2.27, 1.11)
3	(2.27, 1.11) (0.008)	1	(2.27, 1.11) (0.008)	(0.18, −0.20)	0.27	$\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$	(−0.18, 0.20)	0.10	(2.25, 1.13)
		2	(2.25, 1.13) [illegible]	(0.04, −0.04)	0.06	[illegible]	(0.05, −0.03)	2.64	(2.12, 1.05)
4	(2.12, 1.05) (0.0005)	1	(2.12, 1.05) (0.0005)	[illegible]	[illegible]		[illegible]	[illegible]	[illegible]
		2	(2.115, 1.058) (0.0002)	(0.004, 0.004)	0.006				

Metodo de Fletcher-Reeves

Este metodo tambien esta basado en la logica de las direcciones conjugadas. No es tan eficiente como el metodo de Davidon-Fletcher Powell, pero es mucho mas sencillo en calculos manuales.

Procedimiento

Paso 0. Elegir un escalar de terminacion $\varepsilon > 0$ y un punto inicial x_1. Sea $y_1 = x_1$, $d_1 = -\nabla f(y_1)$, $k = j = 1$ y continuar con el paso 1.

Paso 1. Si $\|\nabla f(y_j)\| < \varepsilon$, parar. De lo contrario, sea λ_j una solucion optima al problema de minimizar $f(y_j + \lambda d_j)$ sujeta a $\lambda > 0$, y sea $y_{j+1} = y_j + \lambda_j d_j$. Si $j < n$, continuar con el paso 2. De otra manera, seguir con el paso 3.

Paso 2. Sea $d_{j+1} = -\nabla f(y_{j+1}) + \alpha_j d_j$,

$$\alpha = \frac{\|\nabla f(y_{j+1})\|^2}{\|\nabla f(y_j)\|^2}$$

Reemplazar j por j+1 y continuar con el paso 3

Paso 3: Sea $y_1 = x_{k+1} = y_{n+1}$ y sea $d_1 = -\nabla f(y_1)$. Sea $j = 1$, reemplazar k por k+1 y continuar con el paso 1.

Ejemplo: Minimizar $(x_1 - 2)^4 + (x_1 - 2x_2)^2$

En cada iteracion d_1 es dado por $-\nabla f(y_1)$ y d_2 es dado por $-\nabla f(y_2) + \alpha_1 d_1$, donde α_1 se obtiene con la formula anterior. Ademas, y_{j+1} es obtenido optimizando a lo largo de d_j, partiendo de y_j. En la iteracion 4, el punto $y_2 = (2.185, 1.094)$, el cual es muy cercano al optimo (2.00, 1.00), es alcanzado. Como la norma del gradiente es igual a 0.02, la cual es muy pequeña, se paran las iteraciones en el paso 4.

La tabla siguiente es un compendio de las iteraciones para el problema con el metodo de Fletcher-Reeves, y el progreso del algoritmo se muestra en la figura 8.

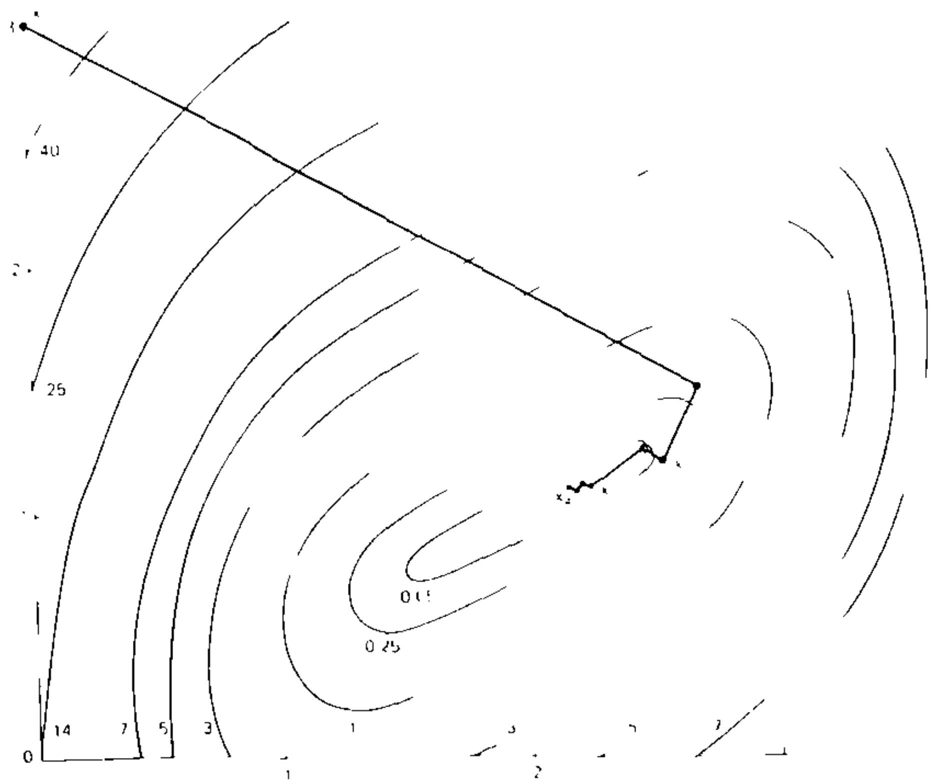


Figura 8 · Ilustracion del metodo de Fletcher-Reeves

Método de Fletcher-Reeves

Iteración k	x_k $f(x_k)$	j	y_j $f(y_j)$	$\nabla f(y_j)$	$\lVert\nabla f(y_j)\rVert$	α	d	λ	y_{j+1}
1	(0.00, 3.00) 52.00	1	(0.00, 3.00) 52.00	(44.00, 24.00)	50.12		([illegible]00, 24.00)	0.062	(2.70, 1.51)
		2	(2.70, 1.51) [illegible]	(0.73, 1.28)	1.47	[illegible]	69. [illegible]	[illegible]23	(2.54, 1.21)
[illegible]	([illegible]54, 1.[illegible]) [illegible]10	1	(2.54, [illegible]) 0.10	[illegible] 0.45	[illegible]		57. [illegible]	[illegible]	[illegible]
		2	2.44, 1.26 0.04	(0.18, 0.32)	0.37	1.[illegible]	30. [illegible]	[illegible]	[illegible]25 [illegible]
3	(2.25, 1.10) 0.008	1	(2.25, 1.10) 0.008	(0.16, 0.20)	0.32		(0.16, 0.20)	0.10	(2.23, 1.12)
		2	(2.23, 1.12) 0.003	(0.03, 0.04)	0.05	0.0[illegible]	0.036, [illegible]	1.02	(2.19, 1.09)
4	(2.19, 1.09) 0.0017	1	(2.19, 1.09) 0.0017	[illegible]	[illegible]		05. [illegible]		[illegible]
		2	(2.185, 1.094) 0.0012	(0.02, 0.01)	0.02				

CAPITULO 6

METODOS DE DIRECCIONES FACTIBLES

Esta clase de metodos resuelven problemas de programacion no lineal moviendose en cada iteracion de un punto factible a un punto factible mejorado, mas cercano a un punto optimo. La metodologia de los metodos de direcciones factibles es: Dado un punto factible x_k, se determina una direccion d_k tal que para $\lambda>0$ suficientemente pequeno, las siguientes dos propiedades son verdaderas: (1) $x_k + \lambda d_k$ es factible, y (2) el valor objetivo en $x_k + \lambda d_k$ es mejor que el valor objetivo en x_k. Despues que tal direccion es determinada, un problema de optimizacion uni-dimensional que debe ser resuelto es determinar hasta donde continuar a lo largo de d_k. Esto conduce a un nuevo punto x_{k+1} y el proceso es repetido. Ya que la factibilidad primal es mantenida durante el proceso de optimizacion, estos procedimientos son con frecuencia referidos como "metodos primales".

Metodos de este tipo consisten en como converger a soluciones que cumplen las condiciones de Karush-Kuhn-Tucker o algunas veces a puntos que cumplen con las condiciones de Fritz-John.

Metodo de Zoutendijk

Definicion: Considere el problema de minimizar f(x) sujeto a $x \in s$ donde $f:E_n \rightarrow E_1$ y S es un conjunto no vacio en E_n. Un vector d diferente de cero es llamado una direccion factible en $x \in s$ si existe un $\delta>0$ tal que $x-\lambda d \in s$ para todo $\lambda \in (0, \delta)$.

De otra manera, d es llamada una direccion mejorada en $x \in s$ si existe un $\delta>0$ tal que $f(x+\lambda d)<f(x)$ y $x + \lambda d \in s$ para todo $\lambda \in (0, \delta)$.

El metodo de Zoutendijk es aplicable para resolver el problema de optimizacion

Minimizar $f(x)$ A: matriz m x n E: matriz l x n

sujeta a Ax b: m-vector e: l-vector

Ex e

El metodo encuentra en cada iteracion una direccion factible y un tamaño de avance (longitud de paso) en esa direccion, tal como en programacion no restringida, la cual tiene dos propiedades

a) Factibilidad: Consiste en que un avance en esa direccion no viola ninguna restriccion

b) Utilidad. Un avance en esa direccion mejora el valor de la funcion objetivo.

Procedimiento

Paso 0 Encontrar una solucion factible inicial x con $Ax_1 \leq b$ y Ex_1 e. Sea k 1 y seguir con el paso 1

Paso 1 Dado x_k, supongase que A^t y b^t son descompuestas en (A_1^t, A_2) y (b^t, b_2^t) tal que A_1x_k b y A_2x_k b_2. Sea d_k una solucion optima a el siguiente problema

Minimizar $\nabla f(x)^t d$

sujeta a $A_1d \leq 0$

Ed 0

$-1 < d_i < 1$ para j 1,...,n

Si $\nabla f(x_k)$ dk 0, parar; x_k es un punto que cumple con las condiciones de Karush-Kuhn-Tucker, de otra manera, continuar con el paso 2.

Paso 2 Sea λk una solucion optima a el siguiente problema de busqueda de linea.

Minimizar $f(x_k + \lambda d_k)$

sujeto a $0 \leq \lambda < \lambda max$

donde

$$\lambda max = \begin{cases} minimo\{b \quad d_i : d_i > 0\} & \text{si d } < 0 \\ \infty & \text{si d} \leq 0 \end{cases}$$

con b [illegible] x_k

d [illegible]

sea x_{k+1} [illegible] $x_k + \lambda d_k$, identificar el nuevo conjunto de restricciones activas en x_{k+1} y obtener [illegible] nuevos valores de A_1 y A_2.

Reemplazar k por k + 1 y repetir el paso 1.

Ejemplo: Minimizar $f(x) = 2X_1^2 + 2x_2^2 - 2x_1x_2 - 4x_1 - 6x_2$

$$
\begin{aligned}
\text{sujeta a} \quad & x_1 + x_2 \leq 2 \\
& x_1 + 5x_2 \leq 5 \\
& -x_1 \leq 0 \\
& -x_2 \leq 0
\end{aligned}
$$

usando el metodo de Zoutendijk, partiendo del punto inicial

$x_1 = (0, 0)^t$

El vector gradiente $\nabla f(x)$ es:

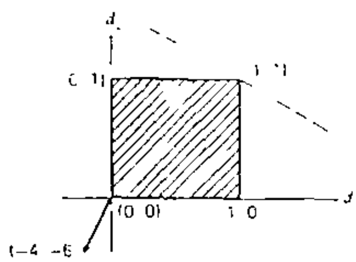
$\nabla f(x) = (4x_1 - 2x_2 - 4, 4x_2 - 2x_1 - 6)^t$

Primera iteracion

En el punto x_1, solamente las restricciones de no negatividad son activas, que se cumplen como restricciones de igualdad, por lo cual I = {3, 4}. El problema para encontrar una direccion factible es:

$$
\begin{aligned}
\text{Minimizar } & -4d_1 - 6d_2 \\
\text{sujeto a } & -d_1 \leq 0 \\
& -d_2 \leq 0 \\
& -1 \leq d_1 \leq 1 \\
& -1 \leq d_2 \leq 1
\end{aligned}
$$

El problema puede resolverse con el metodo simplex o con el metodo grafico, la solucion optima es $d_1 = (1\ ,\ 1)$ y el valor optimo para el problema de direccion factible es igual a 10. La siguiente figura ilustra la iteracion 1.



Ahora necesitamos encontrar el punto factible a lo largo de la direccion (1 , 1), partiendo del punto (0 , 0) con un valor minimo de $f(x) = 2x_1^2 + 2x_2^2 - 2x_1x_2 - 4x_1 - 6x_2$. Cualquier punto a lo largo de esta direccion puede ser escrito como $x_1 + \lambda d_1 = (\lambda, \lambda)$. Entonces $f(x_1 + \lambda d_1) = -10\lambda + 2\lambda^2$

El valor maximo de λ para el cual $x_1 + \lambda d_1$ es factible se obtiene como sigue

$$\hat{b} = \begin{pmatrix} 2 \\ 5 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 1 & 5 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \end{pmatrix}$$

$$\hat{d} = \begin{pmatrix} 1 & 1 \\ 1 & 5 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \end{pmatrix}$$

$$max = \left\{ \frac{2}{2}, \frac{5}{6} \right\} = \frac{5}{6}$$

Entonces, para el nuevo punto $x_1 + \lambda d_1$, el valor de λ_1 se obtiene resolviendo el siguiente problema de busqueda uni-dimensional

Minimizar $-10\lambda + 2\lambda^2$

sujeta a $0 \leq \lambda \leq {}^5/_6$

Como la funcion objetivo es convexa y el minimo no restringido es ${}^5/_2$, de lo cual

$$x_2 = x_1 + \lambda d_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \frac{5}{6}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{5}{6} \\ \frac{5}{6} \end{pmatrix}$$

Segunda iteracion

En el punto $({}^5/_6, {}^5/_6)^t$ se tiene que $\nabla f(x_2) = (-{}^7/_3, -{}^{13}/_3)$. El conjunto de restricciones activas en el punto x_2 es $I = \{2\}$, tal que la direccion factible es obtenida resolviendo el problema

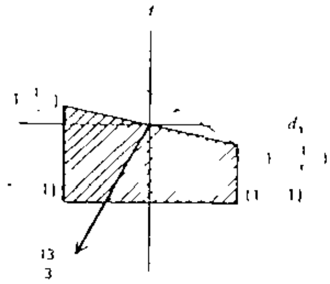
Minimizar $-{}^7/_3\, d_1 - {}^{13}/_3\, d_2$

sujeta a $d_1 + 5d_2 \leq 0$

$-1 \leq d_1 \leq 1$

$-1 \leq d_2 \leq 1$

Usando el metodo grafico se llega a la solucion optima al prol a lineal anterior que es d_2 $(1, -^1{}_5)^t$ y el valor de la funcion objetivo es $-^{22}/_{15}$. La si nte figura ilustra esta iteracion.



Partiendo del punto x_2, cualquier punto en la direccion d_2 puede ser escrito como x_2 + λd_2 $(^5/_{\ } - \lambda, ^5 \ , \ ^1 \ \lambda)$, de lo cual $f(x + \lambda d_2)$ $-^{125}/_8 - ^{22}/_1 \lambda$ λ^2

$$b - \begin{pmatrix} 2 & 1 & 1 & ^5{}_6 \\ 5 & 1 & 5 & \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$d = \begin{pmatrix} 1 & 1 \\ 1 & 5 \end{pmatrix} \begin{pmatrix} 1 \\ ^1{}_5 \end{pmatrix} - \begin{pmatrix} ^4 \\ 0 \end{pmatrix}$$

como d_2 0, se tiene que λmax = min $\{^1{}_3/^4{}_3\}$ $_{12}$

Por lo tanto $\lambda 2$ es la solucion optima al siguiente problema

Minimizar $-^{125}/_8 - ^{22}/_{15}\lambda + ^{62}/_{25}\lambda^2$

sujeta a $0 < \lambda <$ $^5{}_{12}$

La solucion optima es λ_2 $_{186}$, tal que x_3 $x_2 - \lambda d_2$ $(^{55}{}_{31}, {}_{31})^t$.

Tercera iteracion

En $x_3 = (^{55}{}_{31}, ^{24}/_{31})$, tenemos que $\nabla f(x_3)$ $(-^{32}{}_{31}, -^{160}{}_{31})^t$.

El conjunto de restricciones activas en el punto x_3 es dada por I {2}, tal que la direccion factible se obtiene resolviendo:

Minimizar $-^{32}{}_{31}d_1 - ^{160}{}_{31}d_2$

sujeta a $d_1 + 5d_2 \leq 0$

$1 \leq d_1 \leq 1$

$-1 \leq d_2 \leq 1$

Usando nuevamente el metodo grafico, se tiene que $d_3 = (1, -^1{}_5)$, lo cual hace que $f(x_3 + d_3) = -^{125}/_8 - ^{22}/_{15}\lambda + ^{62}/_{25}\lambda^2$ pues $d_2 = d^3$.

Asi mismo

$$/ \quad \begin{pmatrix} 2 \\ 5 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 \\ 1 & 5 \end{pmatrix}\begin{pmatrix} \frac{35}{3} \\ -\frac{4}{1} \end{pmatrix} \quad \begin{pmatrix} \frac{3}{31} \\ 0 \end{pmatrix}$$

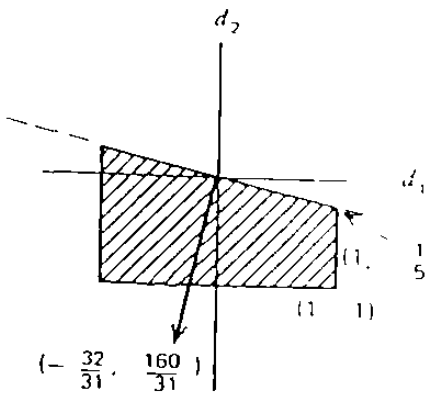
$$d = \begin{pmatrix} 1 & 1 \\ 1 & 5 \end{pmatrix}\begin{pmatrix} 1 \\ \frac{1}{5} \end{pmatrix} - \begin{matrix} 4 \\ 0 \end{matrix}$$

como d_3 0, λ minimo $\{^3{}_{31}/^4/_5\}$ $^{15}{}_{124}$, se tiene ahora que λ_3 es la solucion del problema

$-^{125}{}_8 - ^{12}/_{25}\lambda + ^{62}\ \lambda^2\ 0$

sujeta a 0 λ < ^{1} $_4$

que no tiene solucion, por lo tanto el valor optimo es x_2 $(^{35}{}_{31}, ^{-1}{}_{31})^t$. La figura siguiente ilustra este proceso final.



Metodo de Topkis-Veinott

El metodo de Zoutendijk tiene un variante, con el cual se puede resolver el problema lineal

Minimizar f(x)

sujeta a $g_i(x) \leq 0$ $\quad$ i = 1,...,m

donde todas o algunas de las restricciones son no lineales, pero el punto optimo que se obtiene con el metodo puede no ser optimo sino aproximado, sin cumplir las condiciones de Karush-Kuhn-Tucker.

El metodo de Topkis-Veinott garantiza la convergencia del punto optimo para este problema, obteniendo un punto que cumple con las condiciones de Fritz-John de optimalidad.

Procedimiento

Paso 0: Elegir un punto x_1 tal que $g_i(x_1) \leq 0$ para $i = 1,...,m$. Sea $k = 1$ y seguir con el paso 1

Paso 1: Sea (z_k, d_k) una solucion optima a el siguiente problema de programacion lineal

Minimizar z

sujeta a $\nabla f(x_k)^t d - z \leq 0$

$\nabla g_i(x_k)^t d - z \leq -g_i(x_k)$ para $i = 1,..., m$

$-1 \leq d_j \leq 1$ para $j = 1,..., n$

Si $z_k = 0$, parar, x_k es un punto Fritz-John. De lo contrario, si $z_k < 0$ continuar con el paso 2

Paso 2: Sea λ_k una solucion optima al problema de busqueda unidimensional

Minimizar $f(x_k + \lambda d_k)$

sujeta a $0 \leq \lambda \leq \lambda max$

donde $\lambda max = \sup \{\lambda : g_i(x_k + \lambda d_k) \leq 0$ para $i = 1,...,m\}$. Sea $x_{k+1} = x_k + \lambda_k d_k$, reemplazar k por k + 1, y volver al paso 1

Ejemplo: Minimizar $2x_1^2 + 2x_2^2 - 2x_1x_2 - 4x_1 - 6x_2$

sujeta a $x_1 + 5x_2 \leq 5$

$2x_1^2 - x_2 \leq 0$

$-x_1 \leq 0$

$-x_2 \leq 0$

partiendo del punto $x_1 = (0.00, 0.75)^t$

El gradiente de la funcion objetivo es

$\nabla f(x) = (4x_1 - 2x_2 - 4, 4x_2 - 2x_1 - 6)^t$

Los gradientes de las funciones restricciones son $(1, 5)$, $(4x_1, -1)^t$ $(-1, 0)^t$ y $(0, -1)^t$, los cuales son utilizados para determinar la direccion factible del problema en cada iteracion.

Iteracion 1

En $x_1 = (0.0, 0.75)$ tenemos que $\nabla f(x_1) = (-5.5, -3.0)$. Por lo tanto el problema de direccion de busqueda es:

$$\begin{aligned} &\text{Minimizar } z \\ &\text{sujeta a } -5.5d_1 - 3d_2 - z \leq 0 \\ &\qquad d_1 + 5d_2 - z \leq 1.25 \\ &\qquad -d_2 - z \leq 0.75 \\ &\qquad -d_1 - z \leq 0 \\ &\qquad -d_2 - z \leq 0 \\ &\qquad -1 \leq d_j \leq 1 \qquad \text{para } j = 1, 2 \end{aligned}$$

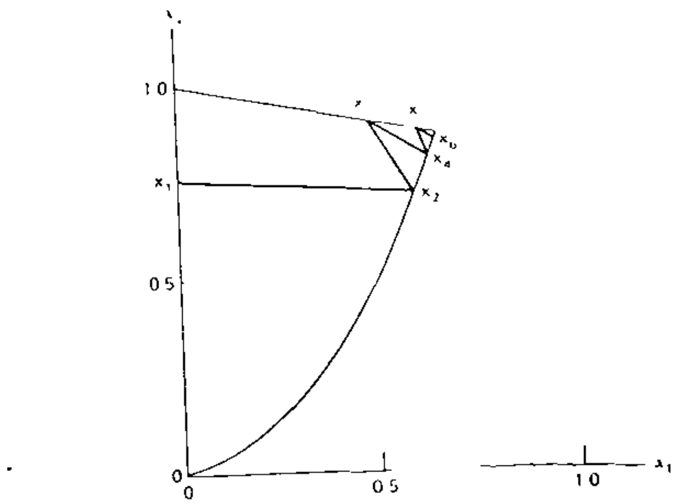
La solucion optima al problema anterior es $d_1 = (0.7143, -0.3571)^t$ y $z = -0.7143$, usando el metodo simplex.

Evaluando la funcion en el punto $x_1 + \lambda d_1$ se tiene que

$$f(x_1 + \lambda d_1) = 0.972\lambda^2 - 4.036\lambda - 3.375$$

El valor minimo de la funcion es $\lambda_1 = 0.84$, entonces $\lambda max = 0.84$ y este valor resuelve el problema de minimizar $f(x_1 + d_1)$ sujeta a $0 \leq \lambda \leq 0.84$, con lo cual $x_2 = x_1 + \lambda_1 d_1 = (0.60, 0.72)$.

El proceso es repetido. La tabla siguiente resume los calculos para 5 iteraciones obteniendose el punto (0.6548, 0.8575) con valor objetivo de -6.5590, muy cercano al optimo (0.658872, 0.868226) obtenido con el paquete de computadora GINO, con valor en la funcion objetivo -6.613086. El progreso del algoritmo se muestra en la siguiente figura



Método de Topkis-Veinott

Iteración			Dirección de búsqueda			Dirección de línea		
k	x_k	$f(x_k)$	$\nabla f(x_k)$	d_k	z_k	λ_{max}	λ_k	x_{k+1}
1	(0.0000, 0.7500)	3.3750	(5.50, 3.00)	(0.7143, 0.03571)	0.7143	0.84	0.84	(0.6000, 0.7200)
2	(0.6000, 0.7200)	5.8272	(3.04, 4.32)	(0.07123, 0.1[illegible]67)	0.2877	1.561676	1.561676	(0.4888, 0.9022)
3	(0.4888, 0.9022)	6.1446	(3.8492, 3.3688	09574 0 [illegible]47	181[illegible]	1.56[illegible]	[illegible]56395	(0.6[illegible]5, 0.8154
4	(0.6385, 0.8154	6.3425	5.6[illegible], 4.0154	0.[illegible]1595, [illegible]29	084	[illegible]	1.418[illegible]	0.61[illegible], [illegible]8768
5	(0.6159, 0.8768)	6.[illegible]82	(3.29 [illegible]246	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Metodo de proyeccion de gradiente de Rosen

El metodo de proyeccion del gradiente esta motivado por el metodo de ascenso-descenso acelerado para problemas sin restricciones. El gradiente negativo se proyecta sobre la superficie de trabajo para definir la direccion de movimiento.

Con este metodo se resuelve el problema de la forma:

Minimizar $f(x)$

sujeta a $Ax \leq b$

$Ex = e$

Procedimiento

Paso 0. Elegir un punto x_1 con $Ax_1 \leq b$, $Ex = e$. Supongase que A^t y b^t son descompuestos en (A_1^t, A_2^t) y (b_1^t, b_2^t) tal que $A_1x_1 = b_1$ y $A_2x_1 < b_2$. Sea $k = 1$ y seguir con el paso 1.

Paso 1: Sea $M^t = (A^t_1, E^t)$. Si M es vacio, parar si $\nabla f(x_k) = 0$. Sea $d_k = -\nabla f(x_k)$, y continuar con el paso 2. De lo contrario sea $P = I - M^t(MM^t)^{-1}M$ y $d_k = -P\nabla f(x_k)$. Si $d_k \neq 0$, seguir con el paso 2. Si $d_k = 0$, calcular $w = -(MM^t)^{-1}M\nabla f(x_k)$ y sea $w^t = (u^t, v^t)$. Si $u \geq 0$, parar. x_k es un punto que cumple con las condiciones de Karush-Kuhn-Tucker. Si $u \ngeq 0$, elegir un componente negativo de u, sea, u_i. Actualizando A_1 eliminando la columna, correspondiente a u_i y repetir el paso 1.

Paso 2: Sea x_k una solucion optima al siguiente problema de busqueda en un intervalo

Minimizar $f(x_k + \lambda d_k)$

sujeta a $0 < \lambda \leq \lambda max$

donde λmax se calcula igual que en el metodo de Zoutendijk.

Sea $x_{k+1} = x_k + \lambda_k d_k$, y suponer que A^t y b^t son descompuestos en (A_1^t, A_2^t) y (b_1^t, b_2^t) tal que $A_1x_{k+1} = b_1$ y $A_2x_{k+1} < b_2$.

Reemplazar k por k+1 y repetir el paso 1

Ejemplo: Resolver el siguiente problema con el metodo de proyeccion de gradiente de Rosen, partiendo del punto (0, 0)

$$\text{Minimizar } 2x_1^2 + 2x_2^2 - 2x_1x_2 - 4x_1 - 6x_2$$

$$\text{sujeta a } x_1 + x_2 \le 2$$

$$x_1 + 5x_2 \le 5$$

$$-x_1 \le 0$$

$$-x_2 \le 0$$

Primera iteracion: En $x_1 = (0, 0)$, el gradiente

$$\nabla f(x_1) = (4x_1 - 2x_2 - 4,\ 4x_2 - 2x_1 - 6)^t$$

$$= (-4,-6)^t$$

Solamente las restricciones de no negatividad son activas en x_1, tal que, $I = \{3, 4\}$, con

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad A_2 = \begin{pmatrix} 1 & 1 \\ 1 & 5 \end{pmatrix}$$

Entonces si esto es verdadero, se tiene que

$$P = I - A_1^t (A_1A_1^t)^{-1} A_1 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

y $d_1 = -P\nabla f(x_1) = (0, 0)$. Calculando

$$w = u = (A_1A_1^t)A_1\nabla f(x_1) = (-4, -6)$$

Eligiendo $u_4 = -6$ y suprimiendo el gradiente de la cuarta restriccion, la matriz A_1 es modificada dado $A_1 = (-1, 0)$. La matriz de proyeccion modificada P que se obtiene es

$$P = I - A_1^t (A_1A_1^t)^{-1} A_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

y la direccion de movimiento d_1 es dada por

$$d_1 = -P\nabla f(x_1) = -\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} -4 \\ -6 \end{pmatrix} = \begin{pmatrix} 0 \\ 6 \end{pmatrix}$$

Por otra parte, cualquier punto x_2 en la direccion d_1 partiendo del punto x_1 puede ser escrito como $x_2 = x_1 + \lambda d = (0, 6\lambda)$ y el correspondiente valor objetivo es $f(x_2) = 72\lambda^2 - 36\lambda$. El valor maximo de λ para el cual $x_1 + \lambda d$ es factible es obtenido como sigue

$$b \quad b \quad 1\,x \quad \begin{matrix} 2 \\ 5 \end{matrix} \quad \begin{matrix} 1 & 1 \\ 1 & 5 \end{matrix} \quad \begin{matrix} 0 \\ 0 \end{matrix} \quad \begin{matrix} 2 \\ 5 \end{matrix}$$

$$d \quad 1\,d \quad \begin{pmatrix} 1 & 1 \\ 1 & 5 \end{pmatrix}\begin{pmatrix} 0 \\ 6 \end{pmatrix} \quad \begin{matrix} 6 \\ 30 \end{matrix}$$

$\lambda max = \text{minimo}\,\{\frac{2}{6}, \frac{5}{30}\} = \frac{1}{6}$

entonces λ es una solucion optima del problema

Minimizar $72\lambda^2 - 36\lambda$

sujeta a $0 < \lambda < \frac{1}{6}$

La solucion optima es $\lambda_1 = \frac{1}{6}$, tal que

$x_2 = x_1 + \lambda d_1 = (0, 1)^t$

En el final de la tercera iteracion

$u = -(A_1 A_1^t)^{-1} A_1 \nabla f(x_3) = \frac{2}{3} > 0$

Por lo tanto x_3 es el punto optimo para $u_2 = \frac{2}{3}$, $\nabla f(x_3) + u_2 \nabla g_2(x_3) = 0$

tal que x_3 es un punto que cumple con las condiciones de Karush-Kuhn-Tucker.

Los calculos para las tres iteraciones se resumen en la siguiente tabla.

Método de Proyección de Gradiente de Rosen

Iteración k	x_k	$f(x_k)$	$\nabla f(x_k)$	I	A	P	d_k	u	λ	λ_{max}	λ
					Dirección de búsqueda				Dirección de línea		
			[illegible]	[illegible]							
					[illegible]		0, [illegible]				
2	(0, 1 [illegible]	4 00	(6, 2)	2, 3}	$\begin{bmatrix} 1 & 5 \\ 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 \end{bmatrix}$ [illegible]	0, 0	([illegible])			
				{2	[1, 5]	[illegible]	[illegible]		[illegible]		([illegible])
3	([illegible], [illegible])	−7 16	([illegible], [illegible])	2}	[1, [illegible]	[illegible]	0, [illegible]	[illegible]			

Metodo del gradiente reducido de Wolfe

Este metodo resuelve problemas de programacion no lineal, sujetos a restricciones lineales, en forma matematica

Minimizar $f(x)$

sujeta a $Ax = b$

$x \geq 0$

El algoritmo converge a un punto que cumple con las condiciones de Karush-Kuhn-Tucker

Procedimiento:

Paso 0: Elegir un punto x_1 que satisfaga $Ax_1 = b$, $x_1 \geq 0$. Sea $k = 1$ y seguir con el paso 1.

Paso 1 Sea $d_k = (d_B, d_N)$. Si $d_k = 0$, parar: x_k es un punto que cumple con las condiciones de Karush Kuhn-Tucker. De lo contrario seguir con el paso 2

I_k: Conjunto de indices de los m componentes mas grandes de x_k

$$B = \{a_j : j \in I_k\} \qquad N = \{a_j : j \notin I_k\} \qquad (1)$$

$$r = \nabla f(x_k)^t - \nabla_B f(x_k)^t B^{-1} A \qquad (2)$$

$$d_j = \begin{cases} -r_j & \text{si } j \notin I_k \text{ y } r_j \leq 0 \\ -x_j r_j & \text{si } j \notin I_k \text{ y } r_j > 0 \end{cases} \qquad (3)$$

$$d_B = -B^{-1} N d_N \qquad (4)$$

Paso 2 :Resolver el siguiente problema lineal

Minimizar $f(x_k + \lambda d_k)$

sujeta a $0 < \lambda < \max$

donde

$$\lambda\max = \begin{cases} \underset{1 \leq j \leq n}{\text{minimo}} \left\{ \dfrac{-x_{jk}}{d_{jk}} : d_j < 0 \right\} & \text{si } d_k > 0 \\ \infty & \text{si } d_k \geq 0 \end{cases}$$

y x_{jk}, d_{jk} son los j-esimos componentes de x_k y d_k, respectivamente.

Sea λ_k una solucion optima y sea $x_{k+1} = x_k + \lambda_k d_k$.

Reemplazar k por k + 1 y repetir el paso 1

Ejemplo: Resolver el siguiente problema usando el metodo del gradiente reducido de Wolfe, partiendo del punto $x_1 = (0, 0, 2, 5)$

$$\text{Minimizar } 2x_1^2 + 2x_2^2 - 2x_1x_2 - 4x_1 - 6x_2$$

$$\begin{array}{ll} \text{sujeta a} & x_1 + x_2 + x_3 = 2 \\ & x_1 + 5x_2 + x_4 = 5 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{array}$$

Calculando el vector gradiente

$$\nabla f(x_1) = (4x_1 - 2x_2 - 4,\ 4x_2 - 2x_1 - 6,\ 0,\ 0)$$

Primera iteracion: En el punto $x_1 = (0, 0, 2, 5)$ tenemos que $\nabla f(x_1) = (-4, -6, 0, 0)^t$. I = {3,4}, tal que B = [3, 4]. El gradiente reducido es dado por la ecuacion (2)

$$r = (-4, -6, 0, 0) - (0, 0) \begin{matrix} 1\,1\,1\,0 \\ 1\,5\,0\,1 \end{matrix} = (-4, -6, 0, 0)$$

Y $r_j = 0$ para $j \in I$. La informacion en este punto se resume en la siguiente tabla

			x_1	x_2	x_3	x_4
solucion x_1			0	0	2	5
$\nabla f(x_1)$			-4	-6	0	0
$\nabla_B f(x_1)$	0	x_3	1	1	1	0
	0	x_4	1	5	0	1
r			-4	-6	0	0

Por la ecuacion (3) se tiene que $d_N = (d_1, d_2)^t = (4, 6)^t$. Calculando d_B con la formula (4)

$$d_B = (d_3, d_4) = -B^{-1} N d_N = -\begin{pmatrix} 1 & 1 \\ 1 & 5 \end{pmatrix} \begin{pmatrix} 4 \\ 6 \end{pmatrix} = (-10, -34)^t$$

$B^{-1}N$ es expresando bajo las variables correspondientes a N, a saber x_1 y x_2.

[illegible]nces el vector direccion es $d_1 = (4, 6, -10, -34)^t$ Ahora, partiendo de x_1 [illegible] $(0, 0, 2, 5)^t$, [illegible] desea minimizar la funcion objetivo a lo largo de la direccion d_1 [illegible] $(4, 6, -10, -34)^t$. [illegible] que calcular el maximo valor de λ tal que $x_1 + \lambda d_1$ es factible

λmax [illegible] minimo $\{ ^2/_{10}, ^5/_{34} \}$ [illegible] $^5/_{34}$

$f(x_1 + \lambda d_1)$ [illegible] $56\lambda^2 - 52\lambda$ tal que λ_1 es la solucion a el siguiente problema

Minimizar 56λ [illegible] 52λ

sujeta a $0 < \lambda \leq$ [illegible]

[illegible]ramente λ_1 [illegible] $^5/_{34}$, tal que x_2 [illegible] $x_1 + \lambda d$ [illegible] $(^{10}/_{17}, ^{15}/_{17}, ^{9}/_{17}, 0)^t$

Siguiendo este mismo proceso, en la iteracion 3, de 3), d_N [illegible] (d_1, d_4) [illegible] y de 4) d_B [illegible] $(d_1, d_2)^t$ [illegible] $(0,0)^t$. Por lo tanto d [illegible] 0 y la solucion x_3 es optima. La tabla [illegible]guiente resume [illegible]s calculos.

Iteración			Dirección de búsqueda		Dirección de línea	
k	x_k	$f(x_k)$	r_k	d_k	λ_k	x_{k+1}
1	(0, 0, 2, 5)	0.0	([illegible] 4, −6, 0, 0)	(4, 6, [illegible] 10, −34)	[illegible]	[illegible]
[illegible]	$(\frac{10}{17}, \frac{15}{17}, \frac{9}{17}, 0)$	6.436	(0, 0, [illegible])	([illegible], [illegible], $\frac{5}{[illegible]}$, 0)	[illegible]	[illegible]
[illegible]	([illegible], $\frac{[illegible]}{31}$, $\frac{3}{31}$, 0)	[illegible]16	(0, 0, [illegible])	(0, 0, 0, 0)		

CAPITULO 7

MODELOS ESPECIALES DE PROGRAMACION NO LINEAL

Programacion Cuadratica

El problema de programacion cuadratica se define como la maximizacion o minimizacion de una funcion cuadratica como funcion objetivo sujeta a restricciones lineales. Para el problema de maximizacion la funcion objetivo debe ser concava y debe ser convexa para el problema de minimizacion. La representacion matematica de dicho problema es

$$
\begin{aligned}
&\text{Maximizar (o minimizar) } Z = CX + X^t DX \\
&\text{Sujeta a} \quad AX \leq P_0 \\
&\qquad\qquad X > 0
\end{aligned}
$$

donde

$X = (x_1, x_2, \ldots\ldots, x_n)^t$

$C = (c_1, c_2, \ldots\ldots, c_n)$ $\qquad A = \begin{pmatrix} a_1 & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_m & \ldots & a_{nn} \end{pmatrix}$

$P_0 = (b_1, b_2, \ldots\ldots, b_m)^t$

$$D = \begin{pmatrix} d_{11} & \ldots & d \\ \vdots & & \vdots \\ d_{n1} & \ldots & d_{11} \end{pmatrix}$$

Como las restricciones se suponen lineales en este caso, garantiza que la region factible sea un conjunto convexo.

Cualquier punto que satisfaga las condiciones de Karush-Kuhn-Tucker sera una solucion de este problema de programacion cuadratica. En base a esto la solucion a este problema se asegura por la aplicacion directa de las condiciones de K.K.T. El problema puede escribirse para el caso de maximizacion como

Maximizar $z = CX + X^t DX$

$$\text{Sujeta a } g(x) = \begin{pmatrix} A \\ -I \end{pmatrix} X - \begin{pmatrix} P \\ 0 \end{pmatrix} \leq 0$$

Las condiciones de K.K.T. pueden combinarse como

$$\begin{pmatrix} 2D & A^t & -I & 0 \\ A & 0 & 0 & I \end{pmatrix} \begin{pmatrix} X \\ \lambda \\ U \\ S \end{pmatrix} = \begin{pmatrix} -C^t \\ P \end{pmatrix}$$

$\mu_j x_j = 0 = \lambda_i S_i$ para toda i y toda j

$\lambda, U, X, S \geq 0$

El problema es equivalente a resolver un conjunto de ecuaciones lineales, mientras se satisfacen las condiciones adicionales $\mu_j x_j = 0 = \lambda_i S_i$. La solucion al sistema anterior se obtiene utilizando la fase 1 del metodo de las dos fases. La unica restriccion aqui es que debe mantenerse siempre la condicion $\lambda_i S_i = 0 = \lambda_j x_j$. Esto quiere decir que λ_i y S_i no pueden ser positivas al mismo tiempo asi como tambien λ_j y x_j no pueden ser positivas simultaneamente. La fase 1 terminara en la forma usual con la suma de las variables artificiales igual a cero unicamente si el problema tiene un espacio factible. Esta tecnica descrita se conoce como el metodo de Wolfe para programacion cuadratica.

Ejemplo: Maximizar $4x_1 + 6x_2 - 2x_1^2 - 2x_1x_2 - 2x_2^2$

Sujeta a $\quad x_1 + 2x_2 \leq 2$

$x_1, x_2 \geq 0$

Escribiendo el problema en forma matricial

$$\text{Maximizar } Z = (4,6)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + (x_1, x_2)\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\text{Sujeta a} \quad (1,2)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leq 2$$

$$x_1, x_2 \geq 0$$

Entonces las condiciones de K. K. T. son :

$$\begin{pmatrix} 4 & 2 & 1 & -1 & 0 & 0 \\ 2 & 4 & 2 & 0 & -1 & 0 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \lambda_1 \\ \mu_1 \\ \mu_2 \\ S_1 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \\ 2 \end{pmatrix}$$

$$4x_1 + 2x_2 + \lambda_1 - \mu_1 = 4$$
$$2x_1 + 4x_2 + 2\lambda_1 - \mu_2 = 6$$
$$x_1 + 2x_2 + S_1 = 2$$

Introduciendo las variables artificiales R_1 y R_2 se tiene :

$$4x_1 + 2x_2 + \lambda_1 - \mu_1 + R_1 = 4$$
$$2x_1 + 4x_2 + 2\lambda_1 - \mu_2 + R_2 = 6$$
$$x_1 + 2x_2 + S_1 = 2$$

Resolviendo este problema de programacion lineal restringido. formandose la siguiente tabla :

$$W\min = R_1 + R_2$$

Cj de las Var Bas	Variable Basica	x_1	x_2	λ_1	μ_1	μ_2	R_1	R_2	S_1	A	A_i/c_i
1	R_1	4	2	1	-1	0	1	0	0	4	1---Sale
1	R_2	2	4	2	0	-1	0	1	0	6	3
0	S_1	1	2	0	0	0	0	0	1	2	2
	Cj	0	0	0	0	0	1	1	0	Como $\mu_1 = 0$, puede	
	Sol. Bas. Fac	0	0	0	0	0	4	6	2	entrar x_1 a la base, y	
	Δ_i	-6	-6	-3	1	1	0	0	0	se cumple $\mu_1 x_1 = 0$	

└Entra

Cj de las Var Bas	Variable Basica	x_1	x_2	λ_1	μ_1	μ_2	R_1	R_2	S_1	A	A/c
0	x_1	1	-	1/4	1/4	0	[illegible]	0	0	1	?
1	R_2	0	3	3/2	1/2	1	1/2	1	0	1	4
0	S_1	0	3/2	-1/4	[illegible]	0	-1/4	0		1	[illegible] --- Sale
Cj		0	0	0	0	0	0	1	0	Como [illegible] 0, entrar	
Sol. Bas. Fac.		1	0	0	0	0	0	4	1	x_2 a la base y	
Δ_j		0	-3	-3/2	1/2		3/2	0		$\mu_1 x_1$ 0	

└Entra

Cj de las Var Bas	Variable Basica	x_1	x_2	λ_1	μ_1	μ_2	R_1	R_2	S_1	A	A_i/c_i
1	x_1	1	0	1/3	-1/3	0	1/3	0	-1/3		?
1	R_2	0	0	2	0	1	0	1	-2		1 -- Sale
0	x_2	0	1	-1/6	-1/6	0	-1/6	0	2/3	2/3	4
Cj		0	0	0	0	0	1	1	0	Como S_1 0,	
Sol. Bas. Fac.		2/3	2/3	0	0	0	0	2	0	entrar λ_1 a la base	
Δ_j		0	0	-2	0	1	1	0	2	$\lambda_1 S_1$ 0	

└Entra

Cj de las Var Bas	Variable Basica	x_1	x_2	λ_1	μ_1	μ_2	R_1	R_2	S_1	A	A_i/c_{ie}
0	x_1	1	0	0	-1/3	1/6	1/3	-1/6	0	1/3	
0	λ_1	0	0	1	0	-1/2	0	[illegible]	-1	1	
0	x_2	0	1	0	1/6	-1/12	-1/6	1/12	1/2	5/6	
Cj		0	0	0	0	0	0	1	1	0	
Sol. Bas. Fac.		1/3	5/6	1	0	0	0	0	0	0	
Δ_j		0	0	0	0	0	0	1	1	0	

Solucion optima

$x_1 = 1/3$ $\quad x_2 = 5/6$ $\quad \lambda_1 = 1$ $\quad \mu_1 = 0$

$\mu_2 = 0$ $\quad R_1 = 0$ $\quad R_2 = 0$ $\quad S_1 = 0$

Programacion Separable

Definicion: Un problema de programacion no lineal es separable si puede ser expresado de la siguiente forma

$$\text{Maximizar (minimizar) } f(x) = \sum_{j=1}^{n} f_j(x_j)$$

$$\text{Sujeta a} \quad \sum_{j=1}^{n} g_{ij}(x_j) \le b_i \qquad \text{para } i = 1, \ldots, m$$

$$x_j \ge 0 \qquad \text{para } j = 1, \ldots, n$$

En programacion separable la funcion objetivo y las restricciones deben expresarse como sumas de funciones de una sola variable cada una.

Ejemplo: La funcion lineal

$$f(x) = a_1x_1 + a_2x_2 + \ldots + a_nx_n$$

donde $a_1, a_2, \ldots, a_n$ son constantes separables

Ejemplo: La funcion

$$h(x) = x_1^2 + x_1 sen(x_1 + x_2) + x_2 e^{x_3}$$

no es separable pues aparecen dos funciones (segundo y tercer sumando) en terminos de mas de una variable.

Solucion aproximada al problema separable

Una solucion aproximada para cualquier problema separable puede obtenerse con el metodo Simplex de programacion lineal.

Supongase que f(x) ha de ser aproximada sobre el intervalo [a,b]. Definase a_k, siendo k=1,...,K como el k-esimo punto de separacion en el eje X tal que $a_1 < a_2 < \ldots < a_k$. Los puntos a_1 y a_k coinciden con los puntos terminales a y b del intervalo en estudio. Por lo tanto, f(x) se aproxima como sigue

$$f(x) = \sum_{k=1}^{n} f(a_k) t_k \qquad x = \sum_{k=1}^{n} a_k t_k$$

donde t_k es un peso no negativo asociado al punto k-esimo de separacion tal que

$$\sum_{k=1}^{n} t_k = 1 \qquad \text{Esta aproximacion es valida si:}$$

i) A lo mas dos t_k son positivas
ii) Si t_k es positiva, unicamente se permite que una t_k adyacente (t^*_{k+1} o bien t_k^*-1) sea positiva.

En el formato del metodo Simplex, la base restringida o conjunto de variables basicas especifica que no mas de dos t_i positivas pueden aparecer en la base. Ademas dos t^k pueden ser positivas unicamente si son adyacentes. La condicion de optimalidad estricta del metodo Simplex se utiliza para seleccionar la variable de entrada t_i unicamente si satisface las condiciones anteriores. El procedimiento se repite hasta que la condicion de optimalidad se satisface o hasta que es imposible introducir nuevas t_i sin violar la condicion de base restringida, lo que ocurra primero. En este punto, la ultima tabla da la solucion optima aproximada al problema.

Ejemplo: Maximizar $z = x_1 + x_2^4$

$$\text{Sujeta a } 3x_1 + 2x_2 \le 9$$

$$x_1, x_2 \ge 0$$

Considerando las funciones separables

$$f_1(x_1) = x_1 \qquad g_1^1 = 3x_1$$

$$f_2(x_2) = x_2^4 \qquad g_1^2 = 2x_2$$

Las funciones $f_1(x_1)$ y $g_1^1(x_1)$ quedan en su forma presente, dado que son lineales. En este caso se trata x_1 como una de las variables. Considerando $f_2(x_2)$ y $g_1^2(x_2)$, supongase que existen cuatro puntos de separacion ($K_2 = 4$). Ya que el valor de x_2 no puede exceder de 3 (por el intervalo (0,3)), se deduce que

k	a_2^k	$f_2(a_2^k)$	$g_1^2(a_2^k)$
1	0	0	0
2	1	1	2
3	2	16	8
4	3	81	18

De lo cual

$f_2(x_2) = t^1_2 f_2(a^1_2) + t^2_2 f_2(a^2_2) + t^3_2 f_2(a^3_2) + t^4_2 f_2(a^4_2)$

$= 0(t^1_2) + 1(t^2_2) + 16(t^3_2) + 81(t^4_2)$

$= t^2_2 + 16t^3_2 + 81t^4_2$

$g(x_2) = t^1_2 g(a_2) + t^2_2 g(a^2_2) + t^3_2 g^2_1(a_2) + t^4_2 g(a^4_2)$

$= 2t^2_2 + 8t^3_2 + 18t^4_2$

El problema de aproximacion es:

Maximizar $z = x_1 + t^2_2 + 16t^3_2 + 81t^4_2$

sujeta a $3x_1 + 2t^2_2 + 8t^3_2 + 18t^4_2 < 9$

$t^1_2 + t^2_2 + t^3_2 + t^4_2 = 1$

$t^k_2 \geq 0 \qquad k = 1,2,3,4$

$x_1 \geq 0$

Añadiendo una variable de holgura a la primera restriccion

$3x_1 + 2x_2 + 8t_2 + 18t_2 + S_1 = 9$

La tabla Simplex inicial, con las columnas reordenadas para dar una solucion de inicio es la siguiente

Cj de las Var Bas	Variable Basica	x_1	t^2_2	t^3_2	t^4_2	S_1	t^1_2	A_i	A_i/c_{ie}
0	S	3	2	8	18	1	0	9	9/8
0	t^1_2	0	1	1	1	0	1	1	1--Sale
Cj		1	1	16	81	0	1		
Sol Bas Fac		0	0	0	0	9	0		
–		1	1	16	81	0	1		

└Entra

Si t^4_2 es la variable de entrada, la variable de salida es S_1, estando en la base t^1_2 y t^4_2, lo que viola la condicion de factibilidad. Entonces la variable que debe salir es S_1 y no puede entrar a la base t^4_2. Considerando t_2 como variable de entrada, sale de la base t_2 y se cumplen las condiciones de factibilidad.

La tabla de la primera iteracion del metodo Simplex aparece a continuacion.

Cj de las Var Bas	Variable Basica	x	t	t	t	S_1	t	A	A	
0	S_1		[illegible]	[illegible]	10	1	8		1 1[illegible]	--Sale
16	t	0	[illegible]	1	1	0	1	1	1	
Cj		1		[illegible]	81	0	0			
Sol. Bas. Fac.		0	0		0	1	0			
[illegible]		1	1[illegible]	0	6[illegible]	0	-16			

└Entra

La variable de entrada es t^4 y como esta en la base t , entra en lugar de S_1.

Cj de las Var Bas	Variable Basica	x_1	t	t_2	t_2	S	t	A_1
81	t^4	3 10	-6 10	0	1	1 10	-8 10	1/10
16	t	[illegible] 0	1[illegible]	1	0	1	18 0	9 10
Cj				16	81	0	1	
Sol. Bas. Fac.		0	[illegible]	9 10	1 10	0	0	
[illegible]		-37 2	24	0	0	-13 2	36	

La tabla muestra que t^1_2 y t^3_2 son candidatos a ser variable de entrada. Como t^1_2 no es un punto adyacente a t^3_2 y t^4_2 basicas, no puede admitirse. t^2_2 no puede entrar a la base, dado que si entra, sale de la base t^3_2 y quedan dos valores no adyacentes en la base. El proceso termina en este punto.

Solucion

Optima x_1 0 $\quad t^2_2$ 0 $\quad t^3_2$ 9 10

t^1_2 1 10 $\quad S_1$ 0 $\quad t^4_2$ 1 10

Zmax = 0 + 0 + 16(9 10) + 81(1 10)

= 22.5

Programacion Fraccional lineal

Ahora consideraremos un problema en el cual la funcion objetivo es el cociente de dos funciones lineales y las restricciones lineales. Tales problemas son llamados "problemas de programacion lineal fraccional" y pueden ser establecidos como sigue

$$\text{Minimizar } \frac{p^t x + \alpha}{q^t x + \beta}$$

$$\text{Sujeta a } Ax = b$$
$$x \geq 0$$

donde p y q son vectores con n componentes, b es un vector con m componentes, A es una matriz m x n y α y β son escalares

Lema: Sea $f(x) = (p^t x + \alpha)/(q^t x + \beta)$ y sea S un conjunto convexo tal que $q^t x + \beta \neq 0$ bajo S. Entonces, f es tanto pseudoconvexa y pseudoconcava bajo S

Del lema anterior surgen algunas implicaciones para problemas de programacion lineal fraccional

1. Puesto que la funcion objetivo es pseudoconvexa y pseudoconcava bajo S entonces tambien es cuasi-convexa, cuasi-concava, estrictamente cuasi-convexa y cuasi concava
2. Puesto que la funcion objetivo es pseudo-convexa y pseudoconcava, entonces, un punto que satisface las condiciones de K.K.T. para un problema de minimizacion es tambien un minimo global bajo la region factible. Igualmente, un punto que satisface las condiciones de K.K.T. para un problema de maximizacion es tambien un maximo global bajo la region factible.
3. Como la funcion objetivo es estrictamente cuasi-convexa y estrictamente cuasi-concava, entonces un minimo local es tambien un minimo global bajo la region factible. Igualmente, un maximo local es tambien un maximo global bajo la region factible.
4. Como la funcion objetivo es cuasi-concava y cuasi-convexa, si la region factible es acotada,(delimitada completamente su grafica) entonces, la funcion objetivo tiene un minimo en un punto extremo de la region factible y tambien tiene un maximo en un punto de la region factible

Metodo de Charnes y Cooper

Este metodo consiste en convertir el problema de programacion lineal fraccional en un problema de programacion lineal que pueda ser resuelto con el metodo Simplex.

Considere el siguiente problema

$$\text{Minimizar } \frac{p^t x + \alpha}{q^t x + \beta}$$

$$\text{sujeta a } Ax \leq b$$

$$x \geq 0$$

Supongase que el conjunto $S = \{x: Ax \leq b \text{ y } x \geq 0\}$ es compacto (acotado y cerrado, que contiene sus puntos limites o puntos frontera) y supongase que $q^t x + \beta > 0$ para cada $x \in S$. Haciendo $z = 1/(q^t x + \beta)$ y $y = zx$, el problema anterior puede reducirse al siguiente programa lineal

$$\text{Minimizar } -p^t y - \alpha z$$

$$\text{Sujeta a } Ay - bz \leq 0$$

$$-q^t y - z = 1$$

$$y \geq 0$$

$$z \geq 0$$

Ahora, si $q^t x + \beta < 0$ para todo $x \in S$, haciendo $-z = 1/(q^t x + \beta)$ y $y = zx$ se tiene que el nuevo problema es

$$\text{Minimizar } -p^t y - \alpha z$$

$$\text{Sujeta a } Ay - bz \leq 0$$

$$-q^t y - z = 1$$

$$y \geq 0$$

$$z \geq 0$$

Finalmente, si existen x_1 y $x_2 \in S$ tales que $q^t x_1 + \beta > 0$ y $q^t x_2 + \beta < 0$, entonces la solucion optima al problema fraccional es no acotada

Ejemplo: Minimizar $\dfrac{-2x_1 + x_2 + 2}{x_1 + 3x_2 + 4}$

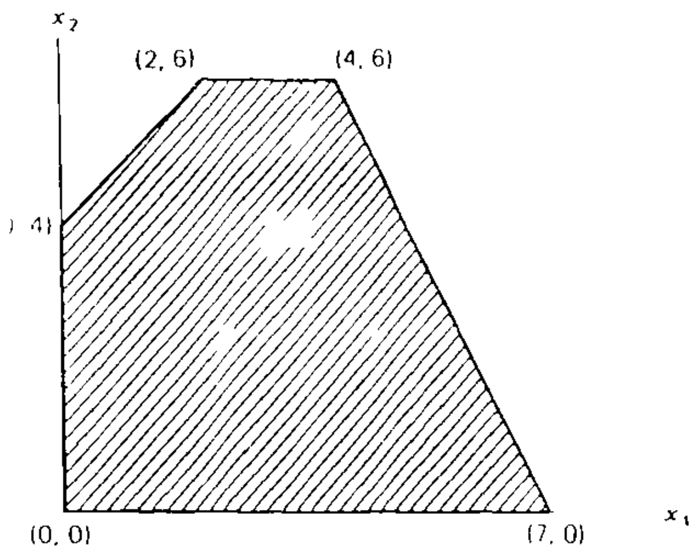
$$\text{Sujeta a } -x_1 + x_2 \leq 4$$

$$2x_1 + x_2 \leq 14$$

$$x_2 \leq 6$$

$$x_1, x_2 \geq 0$$

La siguiente figura muestra la region factible.



Como el punto (0.0) es factible, y en este punto $-x_1 + 3x_2 + 4 > 0$. Por lo tanto, el denominador es positivo bajo la region factible entera, se obtiene el siguiente problema lineal

$$\text{Minimizar } -2y_1 + y_2 + 2z$$

$$\text{Sujeta a } -y_1 + y_2 - 4z \leq 0$$

$$2y_1 + y_2 - 14z \leq 0$$

$$y_2 - 6z \leq 0$$

$$y_1 + 3y_2 + 4z = 1$$

$$y_1, y_2, z \geq 0$$

Resolviendo este problema con el metodo Simplex, se tiene que la solucion optima es $y_1 = 7/11$, $y_2 = 0$ y $z = 1/11$. La solucion optima al problema original es $x_1 = y_1/z_1 = 7$ y $x_2 = y_2/z_2 = 0$. El correspondiente valor objetivo es -1.09.

Programacion Geometrica

Una de las mas recientes tecnicas de optimizacion matematica es la programacion geometrica, que es manejada para funciones y restricciones no lineale descubierta por R. Duffin y C. Zenner.

Desigualdad de la media aritmetica-media geometrica

Si $x_1,\ldots,x_n$ son cualesquier numeros no negativos y $\lambda_1, \ldots, \lambda_n$ tienen la propiedad

$$\sum_{j=1}^{n} \lambda_j = 1$$

y $\lambda_j > 0$ para todo j (j = 1,2,...,n), entonces

$$\sum_{j=1}^{n} \lambda_j x_j \ge \prod_{j=1}^{n} x_j^{\lambda_j} \qquad (1)$$

Existe una forma especial de esta desigualdad que puede ser usada en ciertos problemas de minimizacion, supongase que se tiene una funcion f de la forma

$$f = f(x_1, \ldots, x_n) = y_1 + y_2 + \ldots + y_n$$

donde

$$y_j = y_j(x_1, \ldots, x_2) \ge 0, \qquad j = 1, \ldots, m$$

Esto es, cada termino y_j esta en funcion de $x_1,\ldots,x_n$ y f es la suma de tales terminos. Por otra parte si en la ecuacion (1) si $\lambda = 1/n$ se tiene que

$$\frac{1}{n}\sum_{j=1}^{n} \lambda_j x_j \ge \prod_{j=1}^{n} x_j^{\frac{1}{n}} \qquad (2)$$

Ahora supongase que

$$\prod_{j=1}^{m} y_j^{1/m} = c \qquad (3)$$

donde c es constante. De la ecuacion (2) se tiene que

$$\sum_{j=1}^{m} \frac{1}{m} y_j \ge \prod_{j=1}^{m} y_j^{1/m}$$

Partiendo de esta definicion, considerando el posinomio

$$f(x_1,\ldots,x_n) = \sum_{i=1}^{m} c_i x_1^{a_{i1}} \cdots x_n^{a_{in}} = \sum_{i=1}^{m} c_i p_i$$

aplicando la desigualdad (5) se obtiene

$$\sum_{i=1}^{m} c_i p_i \geq \prod_{i=1}^{m} \left(\frac{c_i p_i}{\lambda_i} \right)^{\lambda_i}$$

o tambien

$$\sum_{i=1}^{m} c_i x_1^{a_{i1}} \ldots x_n^{a_{in}} \geq \prod_{i=1}^{m} \left(\frac{c_i x_1^{a_{i1}} \ldots x_n^{a_{in}}}{\lambda_i} \right)^{\lambda_i}$$

Se busca encontrar los valores de $\lambda_1, \ldots\ldots, \lambda_m$ que satisfagan las condiciones

$$\begin{aligned}
a_{11}\lambda_1 + a_{21}\lambda_2 + \ldots + a_{m1}\lambda_m &= 0 \\
a_{12}\lambda_1 + a_{22}\lambda_2 + \ldots + a_{m2}\lambda_m &= 0 \\
\vdots \qquad\qquad & \quad \vdots \\
a_{1n}\lambda_1 + a_{2n}\lambda_2 + \ldots + a_{mn}\lambda_m &= 0 \\
\lambda_1 + \lambda_2 + \ldots\ldots + \lambda_m &= 1 \\
\lambda_1, \lambda_2, \ldots\ldots \lambda_m &> 0
\end{aligned}$$

y asimismo, que pueda obtenerse el valor de la funcion objetivo f.

Ejemplo: Minimizar

$$f(x_1, x_2, x_3) = \frac{4}{x_1 x_2 x_3} + 8x_1x_2 + 4x_2x_3 + 4x_1x_3$$

Examinando

$$(4x_1^{-1} x_2^{-1} x_3^{-1})^{\lambda 1} (8x_1x_2)^{\lambda 2} (4x_2x_3)^{\lambda 3} (4x_1x_3)^{\lambda 4}$$

Para lograr que se eliminen las variables al buscar el valor de la funcion objetivo, tomando

Por la ecuacion (3) la desigualdad se expresa como

$$\sum^{r} y_i \quad f(x_1,\ldots,x_n) \cdot mc \qquad (4)$$

para todos los vectores $(x_1,\ldots,x_n)$

Ejemplo: Para $f(x)$ $1x^3 + 2x$ $4x$, x 0 con m 3 y y_1 $1x^3$, y_2 $2x^2$, y_3 $4x$

Usando la relacion (3)

$$\prod^{n} y_i^{1/m} \quad y^{1/3} y2^{1/3} y3^{1/3}$$

$$(1x^3)^{1/3}(2x^2)^{1/3}(4x)$$

$$(8)^{1/3} \quad 2$$

Entonces, por (4) se tiene que

$f(x) \geq mc$ $3(2)$ 6, para $x > 0$.

Resolucion de problemas de programacion geometrica no restringidos

Sean $y_1, y_2, \ldots, y_m$ m numeros no negativos, entonces la desigualdad (1) puede ser escrita como

$$\lambda_1 y_1 \quad + \lambda_m y_m \geq y_1^{\lambda_1} \ldots y_m^{\lambda_m}$$

$$\sum_{i}^{m} \lambda_i \quad 1, \lambda_i > 0 \text{ para todo } i$$

Si $y_i = \lambda_i y_i$, entonces

$$y_1 + y_2 + \ldots + y_m \geq \left(\frac{y_1}{\lambda_1}\right)^{\lambda_1} \left(\frac{y_2}{\lambda_2}\right)^{\lambda_2} \ldots\ldots \left(\frac{y_m}{\lambda_m}\right)^{\lambda_m} \ldots\ldots (5)$$

La funcion general f puede considerarse compuesta de m terminos

y_j (j 1,....m) tal que

$$y_j \quad c_j x_1^{aj1} x_2^{aj2} \ldots x_n^{ajn}$$

donde $c_j > 0$ y para $k = 1,2,\ldots,n$ a_{jk} es un numero real y x_k es una variable que solo toma valores positivos, es llamada posinomio.

$\lambda \quad \lambda \quad \lambda_4 \quad ^1{}_2\,\lambda_1$

y con $\sum \lambda \quad 1\ \lambda \quad > 0$

se tiene que

$\lambda_1 \quad 2\ 5\ \lambda_3 \quad \lambda_3 \quad \lambda_4 \quad 1\ 5$

De lo cual

$$f(x_1x_2x_3) \quad 2\ 5 \quad 2\ 5\left(\frac{4}{x\ x\ x_3}\right) + 1/5\,(5\cdot 8x_1x_2) + 1/5\,(5\cdot 4x_2x_3) + 1\ 5\,(5\cdot 4x_1x_3)$$

$$\geq 5/4\left(\frac{4}{x_1x_2x_3}\right)^{\cdot\,\cdot}(5\cdot 8x_1x_2) + (5\cdot 4x_2x_3)^{1\ 5}(5\cdot 4x_1x_3)^{1\ 5}$$

$$\geq (10)\ \ (40)\ \ (20)^{1\ 5}(20)^{1\ 5}$$

$$\geq (1600000)^{1\ 5}$$

Los valores de las variables x_1, x_2, x_3 se obtienen resolviendo las siguientes ecuaciones

$$\frac{4}{x\ x_2 x_3} \quad 10^{-5}$$

$$8x_1x_2 \quad 40^{1\ 5}$$

$$4x_2x_3 = 20^{1\ 5}$$

$$4x_1x \quad 20^{1\ 5}$$

Solucion optima:

$x_1 \quad 3.50 \qquad x_2 = 3.50 \qquad x_3 \quad 0.13$

Valor de la funcion objetivo: 104.5

En el ejemplo anterior se observa una caracteristica importante de la programacion geometrica, que permite conocer acerca del valor de la funcion objetivo antes que la combinacion optima de las variables.

Programacion geometrica restringida

Estos problemas de programacion geometrica restringida se resuelven mediante su problema dual.

El problema de programacion geometrica restringida

$$\min f(x) = \sum_{j=1}^{n} y_j(x)$$

$$\text{sujeta a} \quad h_i(x) = \sum_{j} y_j(x) \le 1, \qquad i = 1, \ldots, k$$

donde $y_j(x) = c_j x_1^{a_{1j}} x_2^{a_{2j}} \ldots x_n^{a_{nj}}$ $\quad j = 1, \ldots, p$

y $\quad c_j > 0, \qquad j = 1, \ldots, p$

con $\quad p_1 = m, \qquad p_k = p, \qquad p_i > p_{i-1}$

El problema dual se expresa como

$\max g(\lambda_1, \lambda_2, \ldots, \lambda_p) = \max g(\lambda)$

$$= \prod_{i=1}^{p} \left(\frac{c_i}{\lambda_i}\right)^{\lambda_i} \prod_{r=1}^{k} (\lambda_{p_{r-1}+1} + \ldots + \lambda_{p_r})^{(\lambda_{p_{r-1}+1} + \ldots + \lambda_{p_r})}$$

sujeta a $\lambda_1, \ldots, \lambda_p \ge 0$

$$\sum_{i=1}^{p} \lambda_i = 1$$

$$\sum_{i=1}^{p} a_{ij} \lambda_i = 0 \qquad j = 1, \ldots, n$$

Ejemplo: Minimizar $f(x_1, x_2, x_3) = 3x_2x_3 + 1/9\, x_1$

sujeta a $1/3\, x_1^{1/2} x_2^{-2} x_3^{1/2} + 3x_1^{-1} x_2 x_3^{-2} \le 1$

El problema dual es

$$\text{Max } g(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = \left(\frac{3}{\lambda_1}\right)^{\lambda_1} \left(\frac{1/9}{\lambda_2}\right)^{\lambda_2} \left(\frac{1/3}{\lambda_3}\right)^{\lambda_3} \left(\frac{3}{\lambda_4}\right)^{\lambda_4} (\lambda_3 + \lambda_4)^{\lambda_3 + \lambda_4}$$

$$\begin{aligned}
\text{sujeta a} \quad & \lambda_1 + \lambda_2 = 1 \\
& \lambda_2 + 1/2\, \lambda_3 + \lambda_4 = 0 \\
& \lambda_1 - 2\lambda_3 + \lambda_4 = 0 \\
& \lambda_1 + 1/2\lambda_3 - 2\lambda_4 = 0 \\
& \lambda_1, \lambda_2, \lambda_3, \lambda_4 \ge 0
\end{aligned}$$

La primera restriccion se refiere a la suma de los λ_i i 1,2, dado que son dos terminos en la funcion objetivo. Las otras tres restricciones son referentes a los terminos que tienen las variables x_1, x , x_3, respectivamente con coeficientes la potencia de cada termino El arreglo de coeficientes

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1\ 2 & -1 \\ 1 & 0 & 2 & 1 \\ 1 & 0 & 1\ 2 & 2 \end{pmatrix}$$

Para hallar los valores de λ_2, λ_3, λ_4, calculando la matriz inversa del arreglo anterior, se tiene

$$\begin{matrix} 7\ 9 & 7\ 9 & 1\ 9 & 1\ 3 \\ 2\ 9 & 7\ 9 & 1\ 9 & 1\ 3 \\ 2\ 3 & 2\ 3 & -2\ 3 & 0 \\ 5\ 9 & 5\ 9 & 2\ 9 & 1\ 3 \end{matrix}$$

de lo cual

λ_1 7 9 λ_2 2 9 λ_3 2 3 λ_4 5 9

El valor optimo de los problemas dual y primal es:

$$g(\lambda^*) \quad \left(\frac{27}{7}\right)^{7\ 9}\left(\frac{1}{2}\right)^{2\ 9} \ \frac{1}{2} \quad \left(\frac{27}{5}\right)^{5\ 9}\left(\frac{11}{9}\right)^{11\ 9}$$

$$5.032$$

Entonces, para minimizar $x_1.x_2.x_3$, se obtiene la unica solucion resolviendo el siguiente sistema de ecuaciones

$$3x_1x_2 \quad 7\ 9\ g(\lambda^*)$$

$$1\ 9\ x_1 - 2\ 9\ g(\lambda^*)$$

$$1\ 3\ x_1^{1\ 2}x_2^{\ 2}x_3^{1\ 2} \quad 6\ 11$$

$$3x_1^{-1}x_2x_3^{-2} - 5\ 11$$

Los terminos del lado derecho de las dos ultimas igualdades se obtienen dividiendo el λ_i correspondiente $i = 3,4$ entre la suma de $\lambda_3 + \lambda_4$

Llegando finalmente a la solucion optima

$$x_1 = 10.065 \qquad x_2 = 4.678 \qquad x_3 = 0.278$$

habiendo aplicado el siguiente teorema al ejemplo anterior

Teorema: Si el problema primal tiene solucion, entonces para las soluciones optimas respectivas x^*, λ^*, se tiene que $f(x^*) = g(\lambda^*)$.

El concepto de metodos de direccion factible es una ampliacion directa y logica de los metodos usados para problemas sin restricciones, pero da lugar a algunas dificultades sutiles, como la falta de convergencia global en el valor optimo del problema no lineal. Los problemas con restricciones de desigualdad se pueden tratar con una estrategia de conjunto activo. Con este enfoque, ciertas restricciones se tratan como activas, y las restantes como inactivas. Se determina el conjunto correcto de restricciones activas durante el proceso de busqueda anadiendo y eliminando restricciones del conjunto de trabajo.

Los metodos de direccion factible mas practicos son el metodo de proyeccion de gradiente de Rosen y el metodo de gradiente reducido de Wolfe. Estos metodos basicos pueden considerarse como el metodo de ascenso-descenso acelerado aplicado en la superficie definida por las restricciones activas. De los dos metodos, el del gradiente reducido de Wolfe es recomendado, pues converge en un menor numero de iteraciones para la mayoria de los problemas que el metodo de proyeccion de gradiente de Rosen.

Por ultimo, los metodos de solucion de programacion no lineal se pueden clasificar en terminos generales como procedimientos directos o indirectos. Ejemplos de los metodos directos son los algoritmos de gradiente, donde el maximo(minimo) de un problema se busca siguiendo la tasa de incremento(disminucion) mas rapida de la funcion objetivo en un punto. En los metodos indirectos, el problema original se transforma primero en un problema auxiliar del cual se determina el optimo. Algunos ejemplos de estas situaciones son la programacion cuadratica, la programacion separable, la programacion fraccional y la programacion geometrica.

Observemos que los problemas auxiliares en estos casos pueden producir una solucion exacta o aproximada del problema original. Por ejemplo, el uso d[illegible]ndiciones de Karush-Kuhn-Tucker con la programacion cuadratica produce una solucion exacta, en tanto que la programacion separable genera solo una solucion aproximada.

BIBLIOGRAFIA

1. Armitano, O., Edelman, J. y Garcia, E. "Programacion no lineal", Limusa, 1985.
2. Bazaraa, M., Jarvis, J. y Sherali, H. "Linear Programming and Network Flows", John Willey and Sons, 1990.
3. Bazaraa, M., Sherali, H. y Shetty, C. "Nonlinear Programming" John Willey and Sons, 1993.
4. Luenberger, D. "Programacion lineal y no lineal", Addison Wesley Iberoamericana, 1989.
5. Prawda, J., "Metodos y Modelos de Investigacion de Operaciones", Limusa, 1986.
6. REA (Staff of Research and Education Association), "The Operations Research Problem Solver", REA, 1989.
7. Sivaslian, B., Stanfel, L., "Optimization Tecnics in Operations Research", Prentice Hall, 1975.
8. Strang, G. "Karmarkar's Algoritm and Its Place in Applied Mathematics", The Mathematical Intelligencer, Vol. 9, Springer Verlag, 1987.
9. Taha, H. "Investigacion de Operaciones", Alfaomega, 1991.
10. Winston, W. "Investigacion de operaciones", Grupo Editorial Iberoamerica, 1994

APENDICE

GLOSARIO

Base: En el metodo Simplex, conjunto de variables basicas.

Biseccion: Tecnica que consiste en dividir repetitivamente a la mitad a los subintervalos de [a,b], y en cada paso, localizar la mitad que contiene a λ_k, punto optimo.

Conjunto acotado: Un conjunto es acotado si existe un numero k tal que $|x| \leq k$ para cada x en el conjunto.

Conjunto cerrado: Un conjunto que contiene sus puntos frontera o puntos limite.

Conjunto compacto: Un conjunto acotado y cerrado.

Convergencia: El hecho que una sucesion de pasos o iteraciones tenga un limite.

Convexidad: Propiedad de los conjuntos convexos, aquellos conjuntos que para cualesquier par de puntos en el conjunto, el segmento que los une, debe pertenecer al conjunto

Direccion factible: Un vector cuyo movimiento a lo largo del mismo no viola ninguna de las restricciones del problema.

Direcciones conjugadas: Las direcciones $d_0, d_1, \ldots, d_{n-1}$ son conjugadas de la matriz Hessiana si cumple $d_i^t H d_j = 0$ $i \neq j$, $i,j = 0,1,\ldots,n-1$ con $d_i, d_j \neq 0$ para toda i y j.

Espacio euclidiano: Conjunto de todos los vectores de dimension n definidos sobre los numeros reales.

Funcion bimodal: Aquella funcion que tiene dos optimos locales o relativos.

Funcion multimodal: Una funcion que tiene varios optimos locales o relativos.

Funcion unimodal: Una funcion que tiene solo un punto minimo o maximo global o absoluto.

Gradiente : Vector que tiene como componentes las derivadas parciales de f con respecto a las variables $x_1, x_2, \ldots, x_n$.

Interpolacion: Tecnica que consiste en encontrar un valor optimo de una funcion objetivo partiendo de una nueva funcion.

Iteraccion : Un desarrollo de una serie de etapas de las que consta un algoritmo.

Longitud de paso : Magnitud de avance de un punto factible a otro mas cercano a la solucion optima en una iteracion

Matriz Hessiana : La matriz compuesta por las derivadas parciales de segundo orden $\partial_2 f(x)\, \partial x_i x_j \; f_{ij}(x)$ para $i \; 1, \ldots, n$, $j \; 1, \ldots, n$

Matriz simetrica : Una matriz cuadrada que es igual a su transpuesta.

Medida de eficiencia : El numero de iteraciones de los metodos de busqueda, entre menos iteraciones se efectuen es mejor la medida de eficiencia de un metodo.

Metodos primales : Aquellos metodos en los que la factibilidad del problema primal es mantenida durante el proceso de optimizacion.

Multiplicadores de Lagrange : Son las constantes que aparecen en el lagrangiano λ_i $i \; 1, \ldots, m$.

Problema lineal : Es un problema de minimizar o maximizar una funcion lineal en la presencia de restricciones lineales del tipo de desigualdad, igualdad o ambas.

Restricciones activas : Son aquellas restricciones de desigualdad que son validas como igualdad.

Restricciones inactivas : Aquellas restricciones de desigualdad que no son validas como igualdad Son lo contrario a las restricciones activas.

Solucion Factible : Un vector perteneciente a un subconjunto del espacio euclidiano E_n que satisface las restricciones.

Solucion optima : Un punto x tal que $f(x) \leq f(x)$ para el caso de maximizacion. Para el caso de minimizacion $f(x) \geq f(x)$.

Variable basica : En programacion lineal, en un sistema de m ecuaciones con n variables ($n > m$) una solucion basica es aquella que se obtiene al fijar n - m variables del sistema iguales a cero y resolviendo el sistema en funcion de las m restantes, llamadas variables basicas.

Vector: Arreglo de n numeros escritos en una fila o una columna.

Vectores linealmente independientes: Aquellos vectores $a_1, a_2, \ldots, a_k$ de dimension n que

$\sum \lambda_i a_i = 0$ implica que $\lambda i = 0$ para $j = 1, \ldots, k$.

AUTOBIOGRAFIA

El licenciado Ramon Cantu Cuellar nacio el 31 de agosto de 1967 en Monterrey, Nuevo Leon. Sus padres son el Prof. Ruben Cantu Cañamar y Argelia Cuellar de Cantu. Estudio en la facultad de Ciencias Fisico Matematicas de la U.A.N.L. de 1984 a 1988, obteniendo el titulo de Licenciado en Matematicas en 1989. Desde agosto de 1990 se desempeña como catedratico en Matematicas y Estadistica en la Facultad de Ingenieria Mecanica y Electrica de la U.A.N.L. Busca obtener el grado de Maestro en Ciencias de la Administracion con especialidad en Investigacion de Operaciones con la tesis: Programacion no lineal