



UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

9
2y

ESCUELA NACIONAL DE ESTUDIOS PROFESIONALES "ACATLAN"

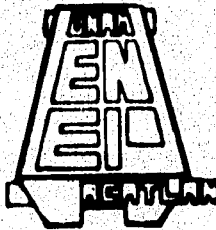
LA PROGRAMACION NO LINEAL EN LA SOLUCION DEL PROBLEMA DE LOCALIZACION

T E S I S

QUE PARA OBTENER EL TITULO DE LICENCIADA EN MATEMATICAS APLICADAS Y COMPUTACION

PRESENTA: ADELIA GUADALUPE COPAS OSIO

ASESOR: DR. RICARDO



ACATLAN, EDO. DE MEXICO



1996

TESIS CON FALLA DE ORIGEN

TESIS CON FALLA DE ORIGEN



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



UNIVERSIDAD NACIONAL
AVENIDA DE
MEXICO

ESCUELA NACIONAL DE ESTUDIOS PROFESIONALES "ACATLAN"

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

SRITA. ADELIA GUADALUPE COPAS OSIO
Alumna de la carrera de M.A.C.
P r e s e n t e .

De acuerdo a su solicitud presentada con fecha 11 de mayo de 1995, me complace notificarle que esta Jefatura tuvo a bien asignarle el siguiente tema de tesis: "LA PROGRAMACION NO LINEAL EN LA SOLUCION DEL PROBLEMA DE LOCALIZACION", el cual se desarrollará como sigue:

INTRODUCCION.

- CAP. I Introducción a la Investigación de Operaciones.
 - CAP. II Programación Matemática.
 - CAP. III Aplicación.
 - CAP. IV Solución.
 - CAP. V Conclusiones y Extensiones.
- BIBLIOGRAFIA.

Asimismo, fué designado como Asesor de Tesis: M. en I. RICARDO - ACEVES GARCIA.

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

E.N.E.P. ACATLAN

ATENTAMENTE
"POR MI RAZA HABLARA EL ESPAÑOL"
Acatlán, Edo. Mex. agosto 2



ACT. LAURA MA. RIVERA BECERRA
Jefe del Programa de Actuaria y MATEMATICAS
y M.A.C. APLICADAS Y COMPUTACION

cg'

*"Existen muchos métodos de
Optimización para llegar al óptimo,
pero los matemáticos en ese punto
siempre son los mismos".*

Varela.

GRACIAS A DIOS

A MIS PADRES

Por la gran confianza que siempre han
depositado en mí.

A MIS HERMANOS

Por el apoyo brindado.

A MIS LUEGOS . . .

Y a todas aquellas personas que
directamente o indirectamente han
contribuido al logro de este objetivo.

CONTENIDO

Introducción	i
I Introducción a la Investigación de Operaciones	1
I.1 ¿Qué es Investigación de Operaciones?	1
I.2 Antecedentes	2
I.3 Modelos	5
I.4 Tipos de problemas más comunes	12
II Programación Matemática	20
II.1 Programación lineal	21
II.2 Optimización clásica	29
II.3 Programación no lineal	34
II.3.1 Métodos iterativos para la solución de problemas no lineales	39
II.3.1.1 Métodos de optimización sin restricciones	39
II.3.1.2 Métodos de optimización con restricciones	72
III Aplicación	98
III.1 Problema de Localización de servicios	98
III.2 Conceptualizar del problema de localización de servicios	99
III.3 Localización de un solo servicio	101
III.4 Localización de varios servicios	110
III.5 Problema de aplicación	118

IV Solución	121
IV.1 Comparación de métodos no lineales para resolver el problema de localización de servicios.	121
V Conclusiones y Extensiones	124
Apéndice A	127
- Teorema de Holgura complementaria	127
- Multiplicadores de Lagrange	128
Apéndice B	130
- Método de Aproximación Hiperbólico	130
Bibliografía	133

INTRODUCCIÓN

En años recientes, la Investigación de Operaciones ha tenido un impacto creciente en la Administración y Tomas de Decisiones, tanto en los sectores público como privado, debido en parte al desarrollo tecnológico que han tenido las computadoras, las cuales han servido como una herramienta fundamental de apoyo, para los modelos y técnicas de solución que se emplean.

La Investigación de Operaciones contempla áreas como: Programación Lineal, Programación Dinámica, Programación Entera, Teoría de Inventarios Programación no Lineal, Teoría de Localización etc.; en las dos últimas áreas se enfoca este trabajo.

La programación no lineal tiene gran capacidad para manejar restricciones de desigualdad, el alcance de la discusión de este trabajo sobre Optimización será ampliada considerablemente. Además de su aplicabilidad obvia a problemas prácticos de la dirección y administración de empresas, también facilita a los economistas estudiar la teoría del consumo, producción y asignación de recursos desde una nueva perspectiva, cabe mencionar que la aplicación de la programación no lineal es muy amplia, dado que muchos problemas reales son no lineales.

La teoría de localización de servicios es una técnica importante pero poco conocida, la cual ha ofrecido y ofrece un gran potencial para la solución de problemas donde se tenga que localizar uno o más servicios, que a su vez deben atender o satisfacer a un conjunto de usuarios o centros de demanda. Entendiéndose por servicio: un hospital, una estación de bomberos o de policía, un distrito político, etc.

Uno de los problemas que enfrenta la teoría de localización es cuando la localización para el nuevo servicio coincide con la localización de algún centro de demanda; esto significa que no existe alguna garantía de que la localización óptima del nuevo servicio nunca será igual a la localización de algún centro de

demanda, entonces, por las condiciones del problema se decidió aplicar algunos métodos de programación no lineal que son completamente iterativos.

El principal propósito de este trabajo es comprender y analizar la programación no lineal y los diferentes métodos que se desprende de ella y su aplicación en el problema de localización de servicios, y a su vez eliminar la dificultad que presenta la localización de servicio con métodos que no requieren que la función sea diferenciable. Específicamente, deseamos diseñar e implementar un sistema de cómputo que resuelva de forma óptima el problema de localización de servicios, que evite utilizar el método de aproximación hiperbólico.

Este trabajo se desarrolla como sigue: En el primer capítulo, se presenta un panorama general de la Investigación de Operaciones (I. de O.), por ejemplo, el surgimiento y desarrollo de la I. de O., así como, la piedra angular de la mismas que es el modelo y se explican los diferentes modelos que utiliza la I. de O. y también se explican los problemas más comunes que aborda la I. de O.

En el segundo capítulo, se presenta una introducción de programación lineal, enunciando principalmente las ventajas y desventajas que presenta su utilización. Posteriormente, se enuncia la optimización clásica y de igual forma, como en el caso anterior se enuncia las ventajas y desventajas que presenta en la solución de problemas. Esto permitirá dar un marco de referencia para el análisis y formulación del problema de programación no lineal. Se analiza los principales métodos que utilizan para resolver problemas de optimización sin restricciones y con restricciones.

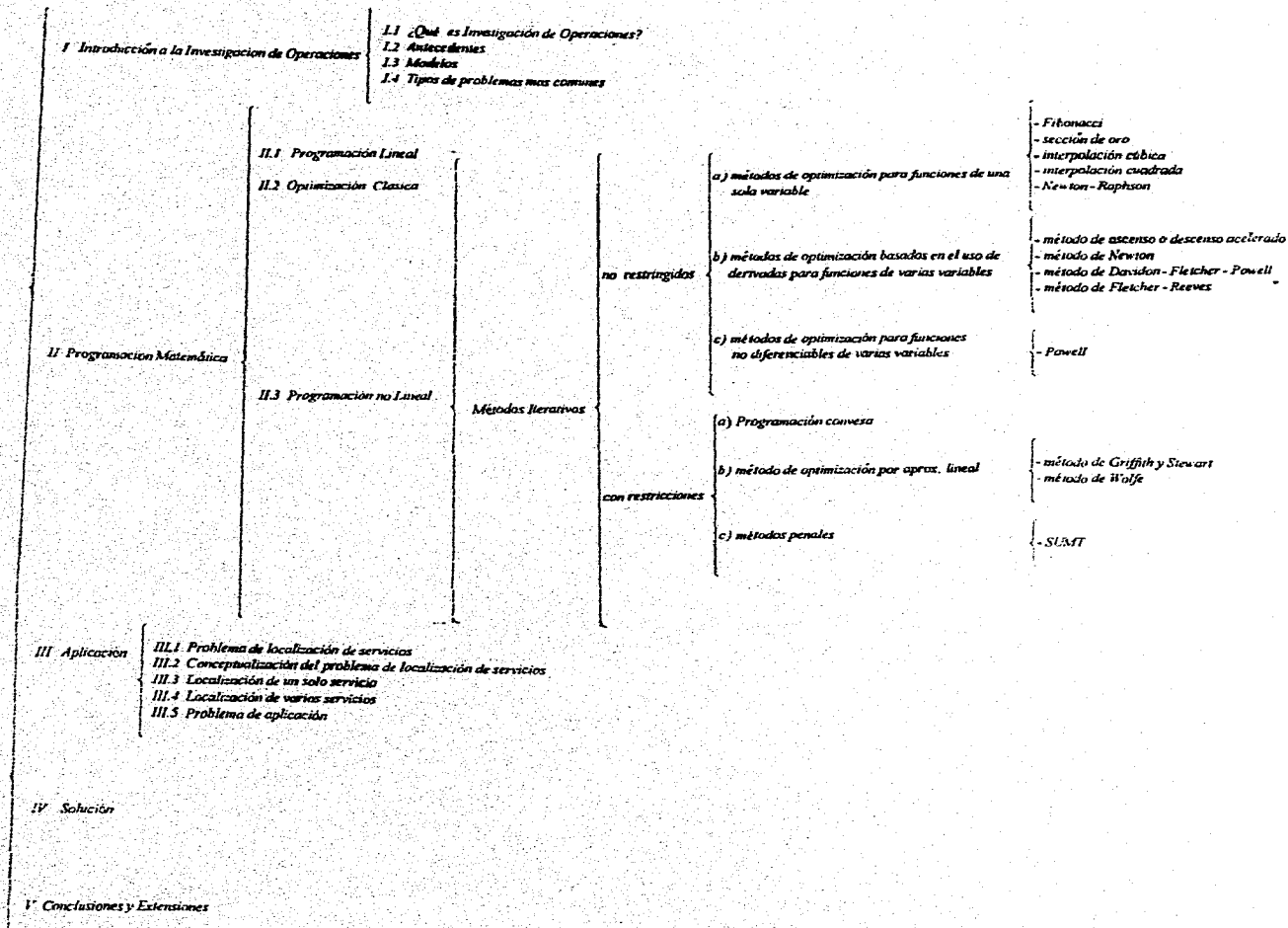
En el tercer capítulo, se presenta los elementos analíticos necesarios para la formulación del problema de localización en el análisis y resolución del problema de localización de servicios en el espacio de soluciones continuo, es decir, se describe el problema de localización de servicios, la cual se divide en dos partes; la primera parte es para un solo servicio, en este caso se trabaja con la norma rectangular o Manhattan, considera que la distancia entre dos puntos no es la línea recta sino el mínimo número de calles que debe recorrer, el otro tipo de problema es la llamada norma euclidiana, considera que la más corta entre dos puntos es la línea recta que los une. En la segunda parte, es la localización de varios nuevos servicios, con respecto a múltiples centros de demanda. En este punto el problema de localización de un solo servicio puede ser considerado como un caso particular del problema de multiservicios, este tipo de

localizaciones ocurre en el mismo contexto que se presentó en la sección de localización de un solo servicio.

En el cuarto capítulo, se presenta un análisis y comparación de los principales algoritmos de solución no lineales, para resolver problemas de localización de servicios. Los métodos que se programaron son: Fibonacci, sección de oro y Powell, estos métodos cumplieron satisfactoriamente las condiciones del problema, es decir, el problema requiere que los métodos por la cual se va a resolver debe ser directo, que no se base en el gradiente o matriz Hessiana. Los métodos de programación no lineal se compara con el método de aproximación hiperbólico; con esta comparación se deduce que los métodos de programación no lineal proporciona mejor solución que el método planteado por localización de servicios.

En el quinto capítulo, se establecen las condiciones del trabajo y se presenta algunas sugerencias de acuerdo a los resultados obtenidos en los capítulos anteriores.

Lo anterior se resume en el siguiente cuadro sinóptico.



INTRODUCCIÓN A LA INVESTIGACIÓN DE OPERACIONES

El objetivo de este capítulo es presentar al lector lo que es la Investigación de Operaciones, mostrando la manera de como se le usa y sus ventajas.

Se explica en este capítulo las condiciones que deben existir para que la Investigación de Operaciones se pueda aplicar en la solución de problemas reales. Además, de explicar las diferentes etapas por las que necesariamente debe pasar todo estudio o proyecto de Investigación de Operaciones.

1.1 ¿Qué es Investigación de Operaciones?

No es fácil definir qué es Investigación de Operaciones (I. de O.). Existen diversas definiciones en los textos, pero se podría establecer que la Investigación de Operaciones es un enfoque científico interdisciplinario para la solución de problemas, que envuelve la interacción compleja, dinámica y subjetiva de recursos métodos y sistemas, a los cuales, en algunos casos, no se les puede proporcionar una solución exacta por medio de los procedimientos matemáticos o por medio de técnicas de ensayo y error. La Investigación de Operaciones utiliza modelos matemáticos como un recurso primario, está diseñada para cuantificar y acotar estos problemas dentro de un marco de restricciones específicas, medidas objetivas y variables, de tal forma que se busquen controles óptimos de operación, decisiones, niveles y soluciones. O como la define (Churchman, Ackoff y Arnoff, 1957), la Investigación de Operaciones es la aplicación, por grupos interdisciplinarios, del método científico a problemas con el control de las organizaciones o sistemas (hombre-máquina) a fin de que se produzcan soluciones que mejor sirvan a los objetos de toda la organización.

1.2 Antecedentes

La Investigación de Operaciones (I. de O.) se remontan a los años 1759 cuando el economista Quesnay empieza a utilizar modelos primitivos de programación matemática. Más tarde, otro economista de nombre Walras, en 1874, hace uso de técnicas similares. Los modelos lineales de Investigación de Operaciones tienen como precursores a Jordán en 1873, Miskowsky en 1896 y a Farkas en 1903. Los modelos dinámicos probabilísticos tienen su origen en Markov a fines del siglo pasado. El desarrollo de los modelos de inventarios, así como el de tiempos y movimientos, se lleva a cabo por la década de los veinte de este siglo, mientras que los modelos de líneas de espera se origina con los estudios de Erlang, a principios del siglo XX. Los problemas de asignación se estudian con métodos matemáticos por los húngaros König y Egervary en la segunda y tercera década de este siglo. Los problemas de distribución se estudia por el ruso Kantorovich en 1939. Von Neuman cimienta, en 1937 lo que años más tarde culminaría como la Teoría de Juegos y la Teoría de Preferencias (esta última desarrollada en conjunto con Morgenstern). Los modelos matemáticos de la I. de O. que utilizaron estos precursores, estaban basados en el Cálculo Diferencial e Integral (Newton, Lagrange, Laplace, Lebesgue, Leibnitz, Reimman, Stieltjes, por mencionar algunos), en la probabilidad y la estadística (Bernoulli, Poisson, Gauss, Bayes, Gosset, Snedecor, etc.)

En la Segunda Guerra Mundial fue cuando la I. de O. empezó a tomar auge, primero se le utilizó en la logística estratégica para vencer al enemigo (teoría de Juegos) y, más tarde al finalizar la guerra, en la logística de distribución de los recursos militares de los aliados dispersos por todo el mundo. Precisamente en este problema, que la fuerza aérea norteamericana, a través de su Centro de Investigación Rand Corporation, comisionó a un grupo de matemáticos para que resolviera este problema que estaba consumiendo tantos recursos humanos, financieros y materiales. Fue el doctor Dantzing el que en 1947, resumiendo el trabajo de muchos de sus precursores inventara el método simplex, con lo cual dio

inicio a la programación lineal. Con el avance de las computadoras digitales se empezó a extender la Investigación de Operaciones durante la década de los cincuenta en las áreas de programación dinámica (Bellman), programación no lineal (Kuhn Tucker), programación entera (Gomory), redes de optimización (Ford y Fulkerson), simulación (Markowitz), inventarios (Raiffa) y procesos markovianos de decisión (Howard).

Actualmente la Investigación de Operaciones no sólo se aplica en el sector privado (Industrias, sistemas de comercialización, Sistemas financieros, transportes sistemas de salud etc.) sino también en el sector de los servicios públicos, tanto en los países desarrollados como en los países del tercer mundo.

En la actualidad se han diseñado programas profesionales y de posgrado (maestría y doctorado) en la especialidad de Investigación de Operaciones.

Resumiendo, la Investigación de Operaciones se aplica a todas las áreas del conocimiento, al igual que la filosofía, sólo que esta maneja conceptos e ideas teóricas, mientras que la Investigación de Operaciones es la herramienta práctica para la toma de decisiones, como se observa en la siguiente figura.

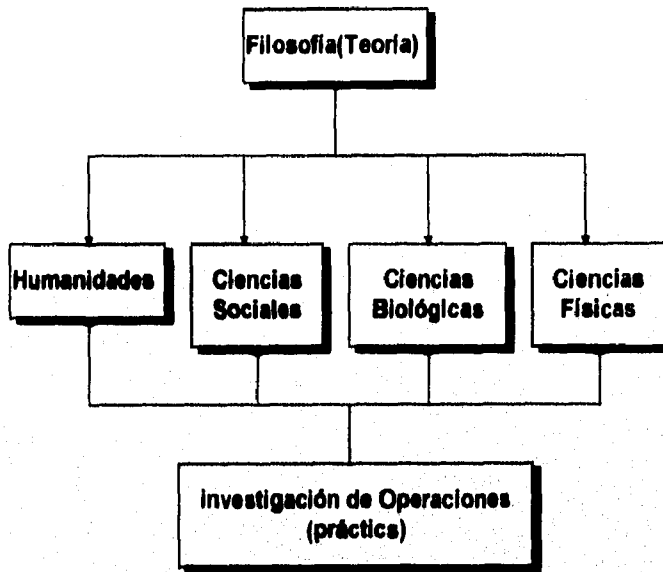


Figura 1.2.1.

La Investigación de Operaciones todavía se encuentra en una edad incipiente y hay mucho por hacer en el desarrollo de este campo, tanto en su teoría como en su aplicación.

La piedra angular de la Investigación de Operaciones es el modelo que debe representar al fenómeno en estudio y del cual se obtendrá la solución al problema. Si dicha solución satisface o resuelve al problema existente en la realidad, la representación del modelo habrá sido satisfactoria, pero si por el contrario, al llevarse a la práctica los resultados no resuelve el problema, el modelo no corresponde a la realidad.

1.3 Modelo

Definiremos un modelo como la representación simbólica, simulada o física de la realidad mediante la abstracción, donde por medio de una teoría adecuada se lleva a cabo una investigación y solución al problema de la realidad, como se observa en la siguiente figura.

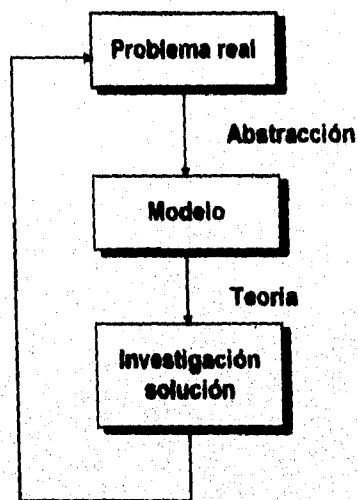


Figura 1.3.2.

Existen diferentes clases de modelos los más comunes son: Icónicos, Analógicos o descriptivos, Simbólicos o explicativos.

1.- Modelos Icónicos: Representa la realidad en proporcionalidad, es decir, las propiedades relevantes del fenómeno real se representa de acuerdo con las mismas propiedades, normalmente con un cambio de escala. De aquí que, en general, estos modelos se ven como lo que representan, pero diferente en tamaño, son imágenes. Algunos ejemplos comunes son las fotografías, las

maquetas, dibujos y modelos a escala de barcos, etc. En general, los modelos icónicos son específicos, concretos, y difíciles de manejar para fines experimentales.

2.- Modelos Analógicos o Descriptivos: Los modelos analógicos pueden representar situaciones dinámicas y se usan más que los icónicos, porque pueden mostrar las características del acontecimiento que se estudia. Por ejemplo, las curvas de demanda, las curvas de distribución de frecuencia en las estadísticas, los diagramas de flujo, etc. A menudo un modelo analógico es muy adecuado para representar relaciones cuantitativas entre las propiedades de los objetos de varias clases. Al transformar las propiedades en propiedades analógicas, con frecuencia podemos incrementar nuestra capacidad de hacer cambios. Otra ventaja de los modelos analógicos sobre los icónicos es que ordinariamente puede hacerse que los primeros represente muchos procesos distintos del mismo tipo lo que se hace evidentemente en el flujo de trabajo en proceso y de productos terminados de una fábrica. No podría usarse eficientemente un modelo icónico para estudiar los efectos de ciertos cambios en el control de calidad. Un diagrama de flujo es un modelo analógico muy sencillo y eficaz en esas situaciones.

3. Modelos Simbólicos o Explicativos: Estos modelos son conceptualizaciones abstractas del problema real, se base en el uso de letras, números, variables y ecuaciones, este tipo de modelos son fáciles de manipular y se puede hacer con ellos un gran número de experimentos. Los modelos simbólicos toman la forma de relaciones matemáticas (casi siempre ecuaciones o desigualdades "inecuaciones") que refleja la estructura de lo que representa.

En Investigación de Operaciones siempre que sea posible, se buscan modelos simbólicos, no sólo porque son más fáciles de manipular sino porque normalmente producen resultados más precisos, en comparación con los modelos icónicos o análogos.

Metodología de la Investigación de Operaciones

En este punto se va a explicar y desarrollar las diferentes etapas que son necesarias para el estudio de un proyecto de Investigación de Operaciones.

Recopilación y organización de los datos: Para encontrar la solución de un problema, primero debemos ser capaces de encontrar el problema y formularlo de manera que sea factible someterlo a una investigación. Por lo regular, el análisis de la investigación se puede entender mejor en la parte analógica. Al igual que un médico con los síntomas y no con el diagnóstico, continuamente deben buscar síntomas adicionales, antes de hacer un diagnóstico correcto. Para encontrar y formular correctamente un problema. Es decir, cualquier estudio de Investigación de Operaciones es esencial que el problema en consideración esté definido claramente. Por que es imposible obtener la respuesta correcta a partir de un modelo mal formulado o planteado.

En la formulación del problema deben estar bien establecidos los objetivos, los cursos alternativos de acción, las restricciones y los efectos del sistema en estudio sobre sistemas relacionados. Debe haber completo acuerdo en estos puntos entre las personas que inicia el estudio de la Investigación de Operaciones y las personas que lo realizan. Además deben haberse convenido entre las partes involucradas es una medida de efectividad. Esta medida de efectividad debe estar en armonía con los objetivos de la organización total.

Construcción del modelo: Una vez formulado el problema del tomador de decisiones, la siguiente etapa consiste en formularlo de manera conveniente para su análisis. La forma convencional como la Investigación de operaciones realiza esto, es construyendo un modelo matemático que represente la esencia del problema. Un modelo matemático es un conjunto de ecuaciones, que describen un sistema o problema. El modelo matemático generalmente contiene dos clases de ecuaciones: (1) La función de efectividad y (2) las restricciones. La función de

efectividad, frecuentemente denominada ecuación objetivo o función objetivo, es una expresión matemática de objetivo del estudio; Por ejemplo, la expresión matemática de la ganancia o costo de operación particular. Ejemplos de restricciones, las expresiones matemáticas de las limitaciones sobre una operación o sistema.

La función objetivo y de restricción son funciones de dos tipos de variables, variables controlables (de decisión) y variables incontrolables. Una variable controlable es aquella que puede ser directamente controlada por quien toma las decisiones. Los valores de esas variables deben determinarse. Las variables incontrolables son aquellas que no están bajo el control de quien toma las decisiones.

Debe recordarse que un modelo es una aproximación de un sistema real. Por consiguiente, todas las variables pueden no estar incluidas en el modelo. Esto a veces es mal interpretado por personas no familiarizadas con el enfoque de la Investigación de Operaciones. Ningún especialista de Investigación de Operaciones reclamará que su modelo incluye todas las posibles variables o que las repuestas obtenidas a partir del modelo son infalibles cuando se aplican a un sistema real en estudio. Cualquier procedimiento está sujeto a algún error. Lo que se pretende es hacer el error tan pequeño como sea posible.

En un modelo se debe aclarar qué variables son importantes y qué datos son necesarios para el análisis de un sistema.

Una vez que se plantea el modelo correcto, el problema consiste en optimizar la función objetivo. La función objetivo o función económica, es aquella parte del modelo que representa, por ejemplo, el tiempo que tarda el proyecto, el costo de realización de algún bien o servicio, las utilidades de una empresa. Esta función objetivo o función económica es una función de varias variables. La optimización de esta función consiste en maximizar o minimizar dicha función según el caso.

Anteproyecto del problema: Una vez establecido el modelo matemático, el siguiente paso es obtener una solución al problema a partir del modelo. Esto se lleva a cabo determinando la solución óptima del modelo y luego aplicando esta solución al problema real. Ya que un modelo es una aproximación de un sistema real o problema, la solución óptima del modelo no garantiza una solución óptima del problema real.

Prueba de modelo: Una de las reglas de investigación de operaciones dice que por lo general, no es suficiente confiar sólo en la propia intuición. Debe tomarse esta precaución no sólo al obtener la solución de un problema, sino también al evaluar el modelo que se formuló para representarlo. El criterio apropiado para juzgar la validez de un modelo, es su capacidad de predecir los efectos relativos de los cursos de acción alternativos con suficiente exactitud, para que permita tomar decisiones adecuadas. No importa cuán plausible puede parecer el modelo, no debe aceptarse bajo la creencia de que esta condición quedará satisfecha. Dada la dificultad para comunicar y entender todos los aspectos y sutilezas de un problema operacional complejo, existe la rara posibilidad de que el equipo de investigación de operaciones, no haya sido informado de todos los hechos de la situación o no los haya interpretado correctamente. Por ejemplo, puede ser que no haya incorporado al modelo un factor o una interrelación importante, o tal vez no haya estimado con exactitud algún parámetro de entrada.

Antes de emprender pruebas más elaboradas, es bueno comenzar por verificar los errores obvios o lo que se pasó por alto en el modelo. Al examinar de nuevo la formulación del problema pueden descubrirse las equivocaciones de este tipo. Otra prueba útil, es la de asegurarse de que todas las expresiones matemáticas, son consistentes en las dimensiones de las unidades que emplean. Además, puede obtenerse un mejor conocimiento de la validez del modelo, al variar los parámetros de entrada y/o salida, de las variables de decisión y comprobando que los resultados del modelo se comportan de manera factible. Con frecuencia, esto es especialmente revelador cuando se asignan a los

parámetros o a las variables, valores extremos cercanos a su máximo o a su mínimo.

Un enfoque más sistemático para la prueba del modelo es emplear una prueba retrospectiva, la cual hace uso de datos históricos y reconstruye el pasado, para determinar si el modelo y la solución resultante hubieran tenido un buen desempeño, de haberse usado. La comparación de la efectividad de este desempeño hipotético con lo que en realidad ocurrió, indica si el uso del modelo da mejoras significativas sobre la práctica actual. Pueden indicarse áreas en las que el modelo tiene fallas y requiere modificaciones, al emplear las alternativas de solución y determinar sus desempeños históricos hipotéticos, y se pueden reunir evidencias en cuanto a lo bien que el modelo predice los efectos relativos de los diferentes cursos de acción.

Por otra parte la prueba retrospectiva tiene la desventaja de que usa los mismos datos que sirvieron para formular el modelo. Entonces, la pregunta crucial es si el pasado en realidad representa el futuro. Si no es así, el modelo puede tener un desempeño distinto en el futuro del que hubiere tenido en el pasado.

Para salvar esta desventaja, a veces es útil continuar con las cosas como están por una temporada. Esto proporcionará datos con los que no contaba cuando se construyó el modelo.

Si la solución final es usada repetidas veces, es importante continuar verificando el modelo y su solución, después de la implantación inicial, para asegurarse de que siguen siendo válidos.

Aplicación o Ejecución: El éxito de la fase de implantación depende en gran parte, del apoyo que proporcionen tanto la alta administración como el equipo de investigación de operaciones. En consecuencia los investigadores de operaciones deben mostrar la participación activa de la gerencia, al formular el problema y

evaluar la solución. La guía de la gerencia es valiosa en sí, para identificar las consideraciones especiales relevantes y evitar con estas fallas potenciales durante estas etapas, sin embargo, hacer que la gerencia se convierta en parte integrante del estudio sirve también, para comprometer su apoyo activo al llevarla a la práctica.

En esta fase el grupo de Investigación de Operaciones, explica la solución a la administración responsable del sistema en estudio. Es importante que la explicación de la solución, se haga en función de los procedimientos usados en el sistema real. Una vez que haya acuerdo sobre la solución, es responsabilidad de ambas partes, traducir la solución en un procedimiento de operación de fácil comprensión. Después de aplicar la solución al sistema, el grupo de Investigación de Operaciones debe observar la respuesta del sistema a los cambios realizados.

El éxito de un estudio de Investigación de Operaciones, es parte del apoyo recibido de la administración.

Control: Una vez que un modelo y su solución se consideran aceptables, deben colocarse controles sobre la solución. Estos controles se establecen para determinar cualquier cambio significativo de las condiciones en las cuales se basa el modelo.

Aunque dichas etapas de un proyecto de Investigación de Operaciones se inician en el orden enumerado, por lo general no terminan en ese mismo orden. De hecho, cada fase procede normalmente hasta que se termine el proyecto, e interacciona en forma continua con las otras. Aunque debemos examinar por separado estas fases de la Investigación de Operaciones, debe tenerse en mente que es probable que, en el transcurso del tiempo, las mismas se superpongan e interaccionen.

En el siguiente esquema, se resume todo el proceso que se sigue para aplicar la investigación de operaciones.

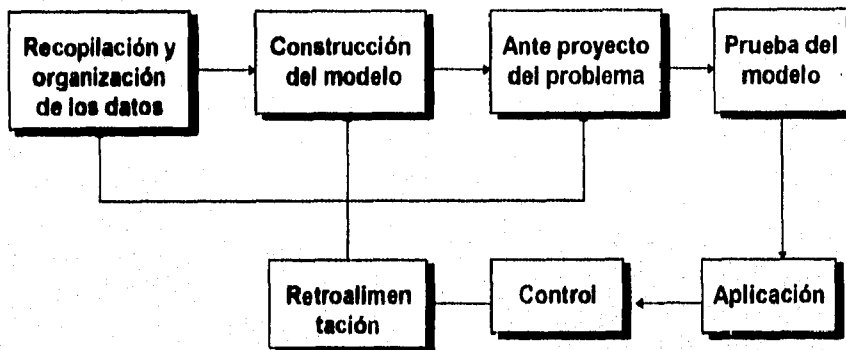


Figura 1.3.3.

En el siguiente punto se contempla los problemas típicos de la Investigación de Operaciones.

1.4 Tipos de problemas más comunes

Desde sus inicios, la Investigación de Operaciones se ha aplicado a una gran variedad de problemas. Sin embargo, la mayoría de éstos, ha sido de naturaleza táctica más que estratégica. La diferencia entre problemas tácticos y estratégicos no es sencilla, debido a que se basa cuando menos, en tres características, cada una de las cuales implica una cuestión de magnitud.

1.- Un problema es más táctico que otro si el efecto de su solución es de menor duración, es decir, si su solución puede modificarse o anularse fácilmente. Es decir, cuanto mayor es la duración del efecto de la solución en el problema, éste es más estratégico. Por lo tanto, un problema táctico consiste en establecer lo que producirá mañana y uno estratégico, es referente al lugar dónde se deberá construir una planta adicional. La Investigación de Operaciones se ha aplicado con más frecuencia a problemas de corto plazo, que a largo plazo. Refiriéndonos a esta característica del problema, como su rango.

2.- Un problema es más estratégico, cuando mayor sea la parte de la organización afectada directamente por su solución. Por lo tanto, un problema que asocie la selección de un convenio contable, es aparentemente más táctico que, por ejemplo, el presupuesto de la empresa. Esta característica del problema puede referirse como su alcance.

3.- Un problema es más estratégico en cuanto más implica la determinación de fines, meta u objetivos. Todos los problemas comprenden la selección de medios para alcanzar los resultados deseados, pero muchos consideran los resultados deseados como dados o proporcionados. En la medida en que lo hacen, son tácticos. De aquí que la planeación de la empresa, que debe establecer metas y objetivos en la organización es más estratégico, que un problema que trate de minimizar los costos de transporte, en el que dicha minimización se considera como resultado conveniente. Esta característica del problema puede referirse como su orientación afines.

No hay puntos de separación definidos en la escala, que representan estas características que distinguen los problemas tácticos de los estratégicos. Por esta razón, lo más que podemos establecer de un problema, es que es más o menos estratégico que otro, en función de esas características.

Una importante consecuencia de la aplicación de la Investigación de Operaciones a una amplia variedad de problemas tácticos, derivada de la práctica, es que se han identificado un grupo de problemas con ciertas características comunes y que además se requieren modelar. Debido al frecuente ocurrencia de éstos, se han desarrollado técnicas para modelarlos y obtener soluciones de los mismos. Estos problemas prototipos son los siguientes: Inventario, Reemplazo, Líneas de espera (Teoría de colas), Secuencia y coordinación, Transporte, Competencia (Teoría de juegos), Localización etc., estos problemas se describen a continuación:

Modelos de asignación: Los problemas de asignación se presentan, cuando los recursos disponibles no son suficientes para permitir que cada trabajo se efectúe de la manera más eficiente. Por lo tanto, el objetivo es asignar los recursos a los trabajos de manera que se minimice el costo total o se maximice el rendimiento total.

La mayoría de los problemas de asignación pueden representarse mediante una matriz como en la que se presenta en la tabla 1. Los valores en las casillas, C_{ij} representan el costo o el rendimiento que resulta de asignar una unidad de recurso R_i al trabajo J_j . Las C_{ij} pueden ser independientes o interdependientes. Por ejemplo, el costo de asignar un camión a una ruta de entrega particular, no depende del modo en que los otros camiones se asignen a las rutas.

Tabla 1. Problemas de asignación típica

Lo que se debe hacer							
Recursos	J_1	J_2	...	J_j	...	J_n	Cantidad de recursos disponibles
R_1	C_{11}	C_{12}	...	C_{1j}	...	C_{1n}	b_1
R_2	C_{21}	C_{22}	...	C_{2j}	...	C_{2n}	b_2
\vdots	\vdots	\vdots	...	\vdots	...	\vdots	\vdots
R_i	C_{i1}	C_{i2}	...	C_{ij}	...	C_{in}	b_i
\vdots	\vdots	\vdots	...	\vdots	...	\vdots	\vdots
R_m	C_{m1}	C_{m2}	...	C_{mj}	...	C_{mn}	b_m
Cant. de recursos requeridos							a_1 a_2 ... a_j ... a_n

Modelos de inventarios: El problema de inventarios, es aquel en donde se requiere almacenar un bien o artículo, con el propósito de satisfacer demandas sobre un horizonte de tiempo, ya sea finito o infinito, hablaremos de la existencia de un problema de inventario. Generalmente un inventario consiste de recursos utilizables, que en grandes volúmenes pueden estar ociosos. Estos recursos pueden ser de cualquier tipo por ejemplo: hombres, materiales, máquinas o dinero. Cuando dicho recurso es materia, o artículos en cualquier etapa de su acabado, el inventario generalmente se menciona como "existencia en almacén".

Normalmente, el objetivo es minimizar el costo total (real o esperado). Sin embargo, si el inventario afecta la demanda (el volumen solicitado por los clientes o usuarios), el objetivo puede ser maximizar las utilidades (reales o esperadas).

Modelos de reemplazo: Generalmente los problemas de reemplazo se tienen cuando los artículos se deterioran con el tiempo, y/o el uso. Y para alargar su vida útil, se recurre a dos diferentes tipos de mantenimiento preventivo (antes de que ocurra una falla) y correctivo (cuando ya ocurrió). Sin embargo, el deterioro puede

llegar a ser tal que el mantenimiento resulte muy costoso y se haga necesario el reemplazo del recurso. En ocasiones los recursos no demasiado deteriorados se reemplazan por otros más modernos y eficaces, para el desempeño de una función similar.

Modelos de líneas de espera: La teoría de líneas de espera, llamada a veces teoría de colas se ocupa del análisis matemático de los fenómenos de las líneas de espera o colas. Las colas se presentan con frecuencia cuando se solicita un servicio como los clientes son de tipo probabilístico. La teoría de colas no pretende en ningún momento resolver directamente el problema de la espera en cola sino más bien describe la situación que presenta una cola a través del tiempo y extrae lo de que éstas se ha dado en llamar, las características operacionales de la cola. Algunas de éstas son; el número promedio de clientes en la cola, su tiempo promedio de espera en la cola, el porcentaje de tiempo que al despachador o despachadores están ocupados, etc.

Modelos de secuenciación: Estos modelos comprenden la determinación de una secuencia óptima, para una serie de tareas o eventos, o la mejor selección de un orden adecuado para atender a los clientes que están esperando. En situaciones que corresponde, a proyectos o trabajos, que consisten de actividades que se deben realizarse en una secuencia específica. Estos modelos implican determinar cuánto esfuerzo se debe hacer para ejecutar cada actividad y cuándo programarla, de modo que se optimase una parte de la ejecución integral del proyecto.

Modelos de Transporte: El problema clásico de transporte contempla una serie de fuentes de operación, fábricas o bodegas, y unas terminales, destinos o agencias donde se demanda un bien o producto, además se tiene un costo unitario de transporte entre cada fuente y cada terminal.

El planteamiento matemático en términos de programación lineal es el siguiente:

$$\text{Min. } z = \sum_{j=1}^m \sum_{i=1}^n c_{ij} x_{ij}$$

S.a:

$$\sum_{j=1}^m x_{ij} = a_i \quad i = 1, 2, 3, \dots, n$$

$$\sum_{i=1}^n x_{ij} = b_j \quad j = 1, 2, 3, \dots, m$$

$$\sum_{i=1}^n a_i = \sum_{j=1}^m b_j$$

x_{ij} es la cantidad asignada desde el origen i hasta el destino j .

c_{ij} es el costo o ganancia de asignar 1 unidad desde el origen i hasta el destino j .

a_i son las cantidades disponibles en cada origen.

b_j son las cantidades requeridas en cada destino.

Modelos de Competencia: Los problemas de competencia también se conocen como teoría de juegos. Un juego es una situación competitiva entre n personas o grupos denominados jugadores, que se realiza bajo un conjunto de reglas previamente establecidas con consecuencias conocidas. Las reglas definen las actividades elementales o movimientos de juego. Puede permitirse diferentes movimientos para los distintos jugadores, pero cada jugador cuenta con un solo movimiento de que disponen los otros jugadores.

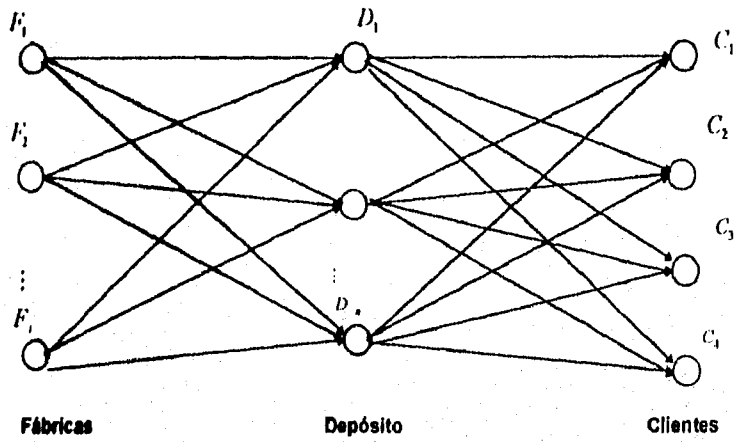
La teoría de juegos se ha desarrollado básicamente de acuerdo con el juego suma-cero.

Modelos de Localización: Una técnica importante pero poco conocida, es la denominada teoría de localización de servicios, la cual ha ofrecido y ofrece un gran potencial para la solución de problemas, donde se tenga que localizar uno o más servicios, que su vez deban atender o satisfacer a un conjunto de usuarios o centros de demanda. Entendiéndose por servicio: un hospital, una estación de bomberos o de policía, un distrito político, una fábrica o bodega, etc.

El problema de localización de servicios es muy antiguo en la literatura matemática, Cavalieri en 1647 consideró el problema de determinar un punto, cuya suma de sus distancias a tres puntos dados, sea mínima, demostró que cada lado deberá tener un ángulo menor a 120 grados, que el punto dado. Fagnano en 1775 demostró, que el punto para el cual la suma de las distancias a los vértices de un cuadrilátero, es mínima, esta dada por la intersección de las diagonales. Tedenat en 1810 encontró para el caso de n puntos, la siguiente condición necesaria: la suma de los cosenos de los ángulos entre alguna línea arbitraria en el plano y el conjunto de líneas que une los n puntos, con el punto mínimo, debe ser igual a cero. Finalmente Steiner probó en 1837, que la condición necesaria y suficiente, es que la suma de los cosenos y senos de los ángulos anteriormente mencionados debe ser igual a cero.

En 1929 Alfred Weber en su estudio clásico, presenta una caracterización de la localización triangular y la discusión del concepto de punto mínimo. En 1956 Walter Isard en su estudio de localización industrial menciona el problema de una multiplicidad de productores, pero lo restringe al caso de varias áreas de mercado, considerando a cada una de ellas con un sólo productor. Lo que reduce el problema de localización, a uno de una sola fuente para cada área de demanda, evitando el problema de determinar la localización de varias fuentes simultáneamente. Sin embargo, fue hasta los trabajos de Kuhn en 1963, que al problema se le considera completamente tratado y resuelto. En la actualidad el problema de localización se ha vinculado fuertemente con las técnicas de optimización, debido a los diferentes contextos en los que surge, pudiéndose utilizar en su análisis y solución, a la programación lineal, la programación dinámica, la programación no lineal y la programación entera.

Los problemas de localización se dirigen a investigar la decisión de dónde localizar una o unas facilidades centrales, que a su vez satisfacen unos puntos de demanda o clientes; por lo general se hacen mención de sistemas de distribución o sistemas logísticos. Por ejemplo, un sistema en el cual se tienen fábricas (F) que distribuyen su producto, y se deben localizar unos depósitos (D) para que finalmente se haga entrega del producto a los clientes asignados C .



Un sistema de distribución

II PROGRAMACIÓN MATEMÁTICA

En los últimos años, ha habido un creciente interés en utilizar técnicas matemáticas de optimización, para resolver problemas de programación que no es posible solucionar con los métodos clásicos del cálculo diferencial y del cálculo de variaciones. Recientemente se han desarrollado nuevas herramientas matemáticas aplicables a los más diversos problemas y a campos tan amplios como: la economía, ingeniería y administración de empresas. Ya que todas estas áreas, se tienen un gran número de actividades interrelacionadas, que pueden ser expresadas por una estructura formulada por símbolos y ecuaciones, que nos definen a un modelo matemático en donde, el objetivo primordial es encontrar el mejor punto que optimice el modelo matemático (económico).

La formulación matemática general que tienen el modelo de programación matemática, se pueden representar como:

$$\begin{aligned} &\text{Optimizar}^1 f(X), \\ &\text{Sujeto a:} \\ &g_j(X) \leq 0 \quad j = 1, 2, \dots, m \\ &X \geq 0. \end{aligned}$$

donde $f(X)$ es de n variables o incógnitas, llamada función objetivo o económica y está sujeta a un conjunto de m restricciones o limitaciones expresada por la función de $g_j(X)$.

Si la función objetivo y las restricciones del problema están formadas por expresiones de primer grado (funciones lineales), y las variables son no negativas, entonces se tiene caso el de programación lineal. Si la función objetivo y/o las restricciones son no lineales, se hablará de programación no lineal.

¹ Optimizar es sinónimo de buscar lo mejor, también alcanzar la ganancia máxima o tener la pérdida mínima.

A continuación se presentarán algunos métodos analíticos para obtener la solución de los problemas planteados con la formulación matemática anterior, como son los de programación geométrica, programación lineal, programación dinámica (discreta), programación no lineal, técnicas de búsqueda, principio del máximo (discreto), programación cuadrática, programación separable, programación convexa, programación entera, programación combinatoria, programación heurística, etc.

Este capítulo se desarrollará en la siguiente forma; primero se tratará de dar un panorama general relacionado a problemas de programación lineal, y a su formulación matemática para con ello proporcionar una idea intuitiva de lo que es un modelo de programación lineal. En segundo lugar se abordará el tema de optimización clásica, para dar una idea general de la dificultad de solucionar problemas que tienen muchas variables. Esta complejidad del problema refleja principalmente en tratar de resolver las derivadas parciales y por último se tratará a la programación no lineal describiendo su formulación general y algunos métodos iterativos, para la solución de problemas no lineales sin restricciones y con restricciones; considerando para cada uno de éstos, las ventajas y desventajas que tienen.

II.1 Programación Lineal

El hombre a lo largo de su historia ha intentado siempre proyectarse hacia la cumbre o alcanzar el éxito en sus actividades, sean éstas empresariales, científicas o políticas. En todas ellas las técnicas de optimización han formalizado y cuantificado, mediante procedimientos matemáticos, la forma de alcanzar lo mejor en una circunstancia o problema bien definido.

Los científicos, en especial los matemáticos, por lo general se han ocupado de fenómenos que intentan hallar los puntos extremos, máximos y mínimos en un problema de optimización. Euclides planteó el problema de encontrar la distancia más corta entre dos puntos y una recta. Herón de Alejandria es considerado como uno de los primeros en estudiar un problema de optimización al asegurar

que la luz viajaba entre dos puntos por ruta más corta, sin embargo 18 siglos más tarde Fermat descubrió el principio más general asegurando que la luz viajaba entre dos puntos en un tiempo mínimo.

A finales de los años 30 aparece el interés en el desarrollo del cálculo de variaciones, sin embargo, el verdadero esfuerzo en optimización se origina con la Segunda Guerra Mundial y con la aparición de la computadora digital. En los 40 Danzig reconoce la estructura matemática de muchos problemas de logística militar y desarrolla el método simplex de programación lineal. Por el rápido desarrollo de la computadora digital. La programación lineal se transforma de un interesante tópico matemático en un importante y ampliamente aplicado procedimiento de optimización. Y La posibilidad con esta máquina para realizar complejas manipulaciones matemáticas sobre un gran conjunto de ecuaciones lineales, ha permitido la solución de grandes problemas industriales.

Finalmente en los años 50 la optimización recibe otro impulso con el advenimiento de la era espacial. La trayectoria óptima de los proyectiles, es uno de los numerosos problemas para los cuales se desarrollaron los métodos de programación dinámica y el principio del máximo, extendiéndose su uso rápidamente a las áreas de economía e ingeniería.

Uno de los avances más importantes dentro del campo de las matemáticas en este siglo es la aparición y desarrollo de la programación lineal (PL). En la actualidad, una herramienta común en multitud de problemas derivados de la necesidad de optimizar los recursos, destinados a diversas actividades en las organizaciones productivas.

El problema más común que aborda la programación lineal, es la asignación de recursos limitados entre actividades competitivas de la mejor manera posible. Este problema de asignación puede surgir, cuando debe elegirse el nivel de ciertas actividades que compite por recursos escasos necesarios para realizarlas. La variedad de situaciones a las que se puede aplicar esta descripción, es sin duda muy grande y va desde la asignación de productos, hasta la asignación de recursos nacionales a las necesidades de un país; desde la selección de una cartera de inversiones, hasta la selección de los patrones de

envió; desde la planeación agrícola, hasta el diseño de un terapia de radiación, etc. No obstante, el ingrediente común de todas estas situaciones es la necesidad de asignar recursos a las actividades.

La programación lineal utiliza un modelo matemático para describir el problema. El adjetivo lineal significa que todas las funciones matemáticas del modelo deben ser lineales. En este caso, la palabra programación no se refiere a programación en computadoras; en esencia es sinónimo de planeación. Así, la programación lineal trata la planeación de las actividades para obtener un resultado óptimo, esto es, el resultado que mejor alcance la meta especificada, entre todas las alternativas de solución.

Formulación matemática general

Definición 1: Se entiende por programación lineal aquel que optimiza

$$\text{Opt. } Z = cX \quad (1)$$

Sujeta a:

$$AX \leq b \quad (2)$$

$$X \geq 0. \quad (3)$$

donde la función lineal (1) se llama función objetivo; las desigualdades (2) por lo general, se hace referencia a las limitaciones y se llaman restricciones, estas desigualdades también pueden ser mayor o igual que b y a las restricciones (3) se llaman condiciones de no negatividad. La palabra optimizar puede significar maximizar o minimizar.

En el problema lineal definido anteriormente se tiene que x es un vector columna con n componentes. A este vector se le denomina el vector de actividades y sus n componentes son variables de decisión. Sea entonces

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad (4)$$

Al vector renglón c , también con n componentes (c_1, c_2, \dots, c_n) se le denomina el vector de precios o costos unitarios. El vector columna b , con m componentes, se le denomina el vector de disponibilidad de recursos. El vector 0 es un vector columna de n ceros. Por último la matriz A , con m renglones y n columnas se le denomina matriz de coeficientes tecnológicos. Cada elemento a_{ij} en la matriz A , con $i=1,2,\dots,m$ y $j=1,2,\dots,n$ representa la cantidad de recursos j que se necesita por unidad de la actividad i .

Matricialmente se reescribe al programa lineal como:

$$\text{Opt. } Z = (c_1, c_2, \dots, c_n) * \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad (5)$$

Sujeto a:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} * \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \quad (6)$$

y

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \geq \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (7)$$

Otra forma de reescribirlo es:

$$\text{Opt. } Z = c_1x_1 + c_2x_2 + \dots + c_nx_n \quad (8)$$

sujeto a las restricciones:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &\leq b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &\leq b_2 \\ \vdots & \quad \quad \quad \vdots \quad \quad \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &\leq b_m \end{aligned} \quad (9)$$

$$x_1, x_2, \dots, x_n \geq 0. \quad (10)$$

Por último, también se puede escribir como:

$$\text{Opt. } Z = \sum_{i=1}^n c_i X_i \quad (11)$$

Sujeto a:

$$\sum_{i=1}^n a_{ij} X_i \begin{cases} \leq b_j \\ > b_j \end{cases} \quad j = 1, \dots, m \quad (12)$$

$$X_i \geq 0, \quad i = 1, \dots, n \quad (13)$$

Todas las formas son equivalentes. Las componentes de c , b y a son números reales, es decir, pueden ser negativos, ceros o positivos.

Terminología para las soluciones del modelo

En programación lineal, cualquier conjunto de valores específicos para las variables de decisión (x_1, x_2, \dots, x_n) se llama solución, sin importar si es una posibilidad deseable o ni siquiera permitida. A continuación se identifican los diferentes tipos de soluciones usando un adjetivo apropiado.

Una solución factible es aquella para la que todas las restricciones se satisfacen.

La región factible: es la colección de todas las soluciones factibles. Si existe soluciones factibles, la meta de la programación lineal es encontrar la mejor, medida según el valor de la función objetivo en el modelo.

Una solución óptima: es una solución factible que lleva al valor más favorable de la función objetivo. El valor más favorable significa el valor más grande o más pequeño, dependiendo de si el objetivo es maximizar o minimizar. Entonces, una solución óptima maximiza/minimiza la función objetivo sobre toda la región factible.

La muchos problemas tendrán nada más una solución óptima. Sin embargo, también es posible tener más de una solución. Otra posibilidad es que el problema no tenga soluciones óptimas. Esto ocurre sólo si: 1) no tiene soluciones factibles o 2) las restricciones no impiden que el valor de la función objetivo (Z) crezca indefinidamente en la dirección favorable (positiva o negativa).

Suposiciones de la programación lineal

Todas las suposiciones de la programación lineal están implícitas en la formulación del modelo, sin embargo, vale la pena hacer hincapié en ellas para que sea más sencillo de evaluar si esta técnica es adecuada en una situación dada.

Proporcionalidad: la proporcionalidad es una suposición sobre las actividades individuales que se consideran independientes unas de las otras. Por lo tanto, considérese el caso de que sólo una de n actividades se realiza, llámese a ésta la actividad K , de manera que $x_j = 0$ para toda $j = 1, 2, \dots, m$ excepto para $j = K$.

La suposición dice que: 1) la medida global de efectividad de Z es igual a $c_1 x_1$ y 2) el consumo de cada recurso i es igual a $a_{i1} x_1$; o bien, ambas cantidades son directamente proporcionales en el nivel donde se lleva a cabo cada actividad K ($K = 1, 2, \dots, n$).

En particular, esto significa que no hay cargos extras debidos al inicio de la actividad (costos fijos) y que la proporcionalidad se cumple en todo en el rango de niveles de la actividad.

Aditividad: la condición de proporcionalidad no es suficiente para garantizar que la función objetivo y las restricciones sean lineales. Si existe interacción entre algunas actividades que puedan cambiar la medida total de efectividad o el consumo total de algún recurso, pueden surgir términos de producto cruzado. La aditividad supone que no existen interacciones de este tipo, entre ninguna de las actividades, de manera que no habrá términos de productos cruzados en el modelo.

De manera más específica, la suposición de aditividad (al igual que la proporcionalidad) se aplica tanto a la función objetivo como a las restricciones. Este último tipo de función representa la utilización total de algún recurso. Para ambos tipos de funciones, la suposición concierne a la comparación entre el valor total de la función que se obtiene al realizar conjuntamente las actividades en sus respectivos niveles (x_1, x_2, \dots, x_n) y las contribuciones individuales al valor de la función, al realizar cada actividad por separado (estableciendo todas las variables en cero). En programación lineal, estas contribuciones individuales son: $c_j x_j$ para la función objetivo y $a_{ij} x_j$ para las restricciones. Es decir, la suposición de aditividad es para cada función y el valor total ésta se puede obtener sumando las contribuciones individuales de las actividades respectivas.

En algunos problemas, si la aditividad no es una suposición razonable, de forma que algunas o todas las funciones matemáticas del modelo necesariamente son no lineales (debido a términos de producto cruzado), resulta definitivamente la entrada en el ámbito de la programación no lineal.

Divisibilidad: Algunas variables de decisión sólo tienen significado físico cuando adquieren valores enteros. La solución óptima que se obtiene en programación lineal con mucha frecuencia no es entera. Por esto, la suposición de divisibilidad se refiere a que las unidades de cada actividad se puedan dividir en cualquier

nivel fraccional, para que se permitan valores no enteros de las variables de decisión.

Certidumbre: La suposición de certidumbre dice que todos los parámetros del modelo (los valores a_{ij} , b_i y c_j) son constantes conocidas. En los problemas reales, muy pocas veces se satisface por completo esta suposición. Casi siempre se formula un modelo de programación lineal para elegir un curso de acción futuro. Entonces, los parámetros que se emplean están basados en una predicción de las condiciones futuras, lo que inevitablemente introduce un cierto grado de incertidumbre.

Por esta razón siempre es importante realizar un análisis de sensibilidad después de encontrar una solución óptima para los valores supuestos de los parámetros. El propósito general del análisis de sensibilidad es identificar los parámetros sensibles (es decir, aquellos que no pueden cambiar mucho sin cambiar la solución óptima), para tratar de estimarlos con mayor exactitud, y después elegir una solución que sea buena en toda la gama de valores posibles para estos parámetros sensibles.

La programación lineal es una técnica poderosa para tratar problemas de asignación de recursos entre actividades que compiten, al igual que otros problemas cuya formulación matemática es parecida. Se ha convertido en una herramienta estándar de gran importancia para muchas organizaciones industriales y de negocios. Aún más casi cualquier organización social tiene el problema de asignar recursos en algún contexto y cada vez es mayor el reconocimiento de la aplicabilidad tan amplia de esta técnica.

Sin embargo, no todos los problemas de asignación de recursos limitados se pueden formular de manera que se ajusten a un modelo de programación lineal, ni siquiera como una aproximación razonable. Cuando no se cumplen una o más de las suposiciones de programación lineal, en este caso es posible aplicar los de modelos matemáticos, o de programación no lineal. La falta de linealidad puede surgir por diversos motivos. Por ejemplo, en un problema de producción en programación lineal, el beneficio bruto de cada producto es constante. Puede ser una función decreciente del nivel de producción, porque una gran producción

tiende a disminuir el precio de mercado (ingreso medio), o bien porque una producción creciente tiende a aumentar el costo variable medio del producto. Si esto es así, la función objetivo lineal se debe reemplazar por una versión no lineal y de igual forma sus restricciones. El problema de transporte con descuentos por volumen en los precios de embarque, un problema de transporte es determinar un plan óptimo para mandar bienes desde varias orígenes hasta varios destinos, dadas las restricciones de recursos y demanda, con el fin de minimizar el costo total de transporte.

Hoy en día una práctica común entre los administradores de grandes carteras de inversión usan como una guía, modelos computarizados basados en programación no lineal. Puesto que los inversionistas se preocupan tanto, por el rendimiento esperado (ganancia) como por el riesgo asociado con su inversión, la programación no lineal se usa para determinar una cartera que con ciertas suposiciones proporcione un balance óptimo entres estos dos factores. Siempre que se presente algún problema de este tipo, será más apropiado una formulación no lineal que una lineal.

II.2 Optimización clásica

La teoría de la optimización clásica considera el uso del cálculo diferencial para determinar puntos de máximos y mínimos(extremos) para funciones restringidas y no restringidas. Los métodos expuestos pueden no ser adecuados para cálculos numéricos eficientes. Sin embargo, la teoría subyacente proporciona la base para visualizar la mayoría de los algoritmos de programación no lineal.

El problema básico, es determinar la utilidad máxima o costo mínimo de un sistema usando la teoría clásica, se transforma en el problema de ubicar todos los puntos máximos y mínimos locales, y luego comparar los valores individuales de la función en esos puntos, con el objeto de determinar el máximo o mínimo absoluto. Para ello, es necesario examinar:

- 1.- Los puntos estacionarios. (Los puntos estacionarios son aquéllos donde la primera derivada es cero.)
- 2.- Los puntos a lo largo del contorno.
- 3.- Aquellos puntos donde la primera derivada es discontinua.

Un punto extremo de una función $f(x)$ define un máximo o un mínimo de la función. Matemáticamente, un punto $X^0 = (x_1, \dots, x_j, \dots, x_n)$ es un máximo si

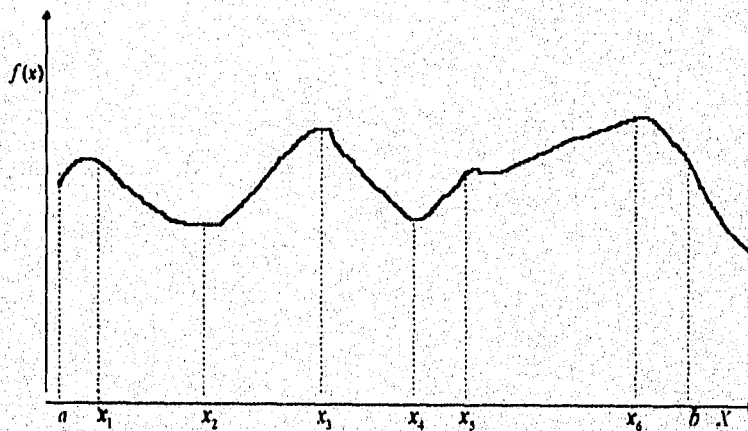
$$f(X^0 + h) \leq f(X^0) \quad (14)$$

para toda $h = (h_1, \dots, h_j, \dots, h_n)$ tal que $|h_j|$ es suficientemente pequeña para toda j .

En otras palabras, X^0 no excede a $f(X^0)$. En forma similar, X^0 es un mínimo si para h , tal como se definió anteriormente,

$$f(X^0 + h) \geq f(X^0) \quad (15)$$

En la siguiente figura se ilustra el máximo y el mínimo de una función $f(x)$ de una sola variable sobre el intervalo $[a, b]$.



Gráfica 2.2.1.

Los puntos x_1, x_2, x_3, x_4 y x_6 son extremos de $f(x)$. Esto incluye a x_1, x_3 y x_6 como máximo, y x_2 y x_4 como mínimo. Ya que

$$f(x_6) = \max\{f(x_1), f(x_3), f(x_6)\}$$

donde $f(x_6)$ se conoce como máximo global o absoluto, mientras que $f(x_1)$ y $f(x_3)$ son máximos locales o relativos. De igual manera, $f(x_4)$ es un mínimo local y $f(x_2)$ es un mínimo global.

Condiciones necesarias y suficientes para extremos

Para establecer las condiciones necesarias y suficientes para que una función $f(X)$ de n variables, tenga puntos extremos superiores, se supone que las primeras y segundas derivadas de $f(X)$ sean continuas en cada X .

Teorema 1: Una condición necesaria para que X^0 sea un punto extremo de $f(X)$ es que

$$\nabla f(X^0) = 0 \quad (16)$$

El siguiente teorema establece las condiciones de suficiencia para que X^0 sea un punto extremo.

Teorema 2: Una condición suficiente para que un punto estacionario sea extremo es que la matriz hessiana (o de hesse) H evaluada en X^0 sea

- a) positiva definida cuando X^0 es un punto mínimo, y
- b) negativa definida cuando X^0 es un punto máximo

Para encontrar un criterio que permita establecer si se tiene un máximo o un mínimo local, se resolverá un desarrollo en series de Taylor alrededor del

punto estacionario. Entonces para el vector $X = (x_1, x_2, \dots, x_n)'$, el desarrollo de Taylor queda como:

$$f(x) = f(X^0) = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \Big|_{x^0} (x_i - x_i^0) + \frac{1}{2} \sum_{j=1}^n \sum_{i=1}^n \frac{\partial^2 f}{\partial x_j \partial x_i} \Big|_{x^0} (x_j - x_j^0)(x_i - x_i^0) + \dots \quad (17)$$

O en forma vectorial como:

$$f(x) = f(X^0) + \nabla f(X^0)(X - X^0) + 1/2(X - X^0)' H^0(X - X^0) + \dots, \quad (18)$$

donde H^0 es la matriz hessiana de las derivadas parciales evaluada en X^0

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial X_1 \partial X_1} & \frac{\partial^2 f}{\partial X_1 \partial X_2} & \dots & \frac{\partial^2 f}{\partial X_1 \partial X_n} \\ \frac{\partial^2 f}{\partial X_2 \partial X_1} & \frac{\partial^2 f}{\partial X_2 \partial X_2} & \dots & \frac{\partial^2 f}{\partial X_2 \partial X_n} \\ \vdots & \vdots & \dots & \vdots \\ \frac{\partial^2 f}{\partial X_n \partial X_1} & \frac{\partial^2 f}{\partial X_n \partial X_2} & \dots & \frac{\partial^2 f}{\partial X_n \partial X_n} \end{bmatrix} \quad (19)$$

En general, si H es indefinida X^0 debe ser un punto de silla. Sin embargo, para el caso donde esto no es concluyente, X^0 puede ser o no punto extremo y el desarrollo de una condición de suficiencia llega a ser muy difícil, ya que sería necesario considerar términos de orden superior en el desarrollo de Taylor. Sin embargo, en algunos casos tales procedimientos complicados pueden no ser necesario, ya que la diagonalización de H puede llevar a información más concluyente.

La condición de suficiencia establecida por el teorema 2 se deduce fácilmente para funciones de una sola variable. Dado que X^0 es un punto estacionario, entonces,

- i) $f''(X^0) < 0$ es una condición suficiente para que X^0 sea un máximo.
- ii) $f''(X^0) > 0$ es una condición suficiente para que X^0 sea un mínimo.

Estas condiciones se determinan directamente considerando la matriz Hessiana con un elemento. Si en el caso de una sola variable $f''(X^0)$ se anula, las derivadas de orden superior deben investigarse, como se muestra en el siguiente teorema:

Teorema 3: Si un punto estacionario X^0 de $f(X)$ las primeras $(n-1)$ derivadas se anulan y $f''(X) \neq 0$, entonces en $X = X^0$ ocurre que $f(X)$ tiene:

- i) Un punto de inflexión si n es impar, y
- ii) Un punto extremo si n es par. Este punto extremo será un máximo si $f^{(n)}(X^0) < 0$ y un mínimo si $f^{(n)}(X^0) > 0$.

Como se puede apreciar, el cálculo sirve para resolver problemas simples, pero en problemas más grandes y restringidos, es difícil resolverlos por los métodos clásicos de cálculo. Por ejemplo, en un problema con restricciones, al igualar el gradiente $\nabla f(X)$ a cero, que es la condición necesaria para que un punto sea óptimo local o relativo, se generan n ecuaciones no lineales que hay que resolver. Esto en general es muy difícil, si no es que imposible. Por otro lado, existen ocasiones en que $\nabla f(X) = 0$ no puede expresarse en forma explícita, si no más bien implícitamente, sin embargo, cuando aún se pudiera resolver el sistema de ecuaciones generado por $\nabla f(X) = 0$, no será posible garantizar que ese punto es un óptimo local, pues puede darse el caso de tratarse un punto de silla.

Una limitación más del enfoque del cálculo, es su insuficiencia para cubrir el caso de las restricciones de igualdad. Por esta razón, la restricción presupuestaria en el modelo de maximización de la utilidad, por ejemplo, se enuncia de la forma, que el gasto total sea exactamente igual a (y no menos o igual que) una suma específica. En otras palabras, la limitación de la aproximación por el cálculo niega al consumidor la opción de ahorrar parte de sus fondos disponibles.

De ahí es la necesidad de introducir la formulación de la programación no lineal, y sus diferentes métodos para solucionar este tipo de problemas con

restricciones o sin restricciones, estos métodos de optimización se basan principalmente en el teorema 1 y 2 descritos anteriormente.

11.3 Programación no lineal

Con la teoría de la programación no lineal, se establecen las óptimas, condiciones necesarias y suficientes que determinan una solución óptima, más que cuestiones de computación. Esta teoría incluye principalmente el estudio de los multiplicadores de Lagrange, incluyendo el teorema de Kunh-Tucker y sus extensiones. Esto mejora considerablemente la comprensión de la filosofía de la optimización con restricciones y proporciona fundamentos básicos satisfactorios para otras disciplinas, como la teoría de la empresa, la economía del consumidor y la teoría del consumidor.

Una suposición importante de la programación lineal es que todas sus funciones (función objetivo y funciones de restricción) son lineales. Aunque, en esencia, esta suposición se cumple para muchos problemas prácticos, es frecuente que no sea así. De hecho, muchos economistas han encontrado que cierto grado de no linealidad es la regla y no la excepción, en los problemas de planeación económica, por lo cual, muchas veces es necesario manejar métodos de programación no lineal.

De manera general, el problema de programación no lineal consiste en encontrar $X = (x_1, x_2, \dots, x_n)$ para optimizar los recursos, como se muestra en el siguiente modelo:

$$\begin{aligned} \text{Opt. } & f(X) && (20) \\ \text{Sujeto a:} & && \\ & h_j(X) = 0, && j = 1, \dots, m \\ & g_j(X) \geq 0, && j = m+1, \dots, p \\ & X \in E^n && \end{aligned}$$

Es decir sea $f(X)$ una función continua, que denota a la función objetivo, $h_1(X), h_2(X), \dots, h_m(X)$ funciones continuas que denotan restricciones de igualdad, $g_{m+1}(X), g_{m+2}(X), \dots, g_p(X)$, funciones continuas que denotan restricciones de desigualdad y $X = (X_1, X_2, \dots, X_n)^T$ un vector en el espacio euclidiano de n -dimensiones, E^n .

Otra forma equivalentemente de representar un problema de optimización no lineal es el siguiente:

$$\text{Opt. } \{f(X) / X \in R\}, \quad (21)$$

donde el conjunto R se definen como:

$$R = \{ X / h_j(X) = 0, g_k(X) \geq 0, X \in E^n, \text{ para toda } j, k \}.$$

donde $f(X), h_j(X), g_k(X)$, son como se describió anteriormente.

Los problemas no lineales pueden ser de diferentes tipos:

- a) **Restringidos:** cuando se tienen restricciones (lineales o no lineales).
- b) **No restringidos:** cuando no se tienen restricciones y sólo se optimiza la función objetivo, que desde luego, no es lineal,
- c) **Continuos:** cuando alguna de las variables y funciones son continuas.
- d) **Discretos:** cuando alguna de las variables y/o funciones es discreta
- e) **Diferenciables:** cuando todas las funciones del problema son doblemente diferenciables (que exista el límite).
- f) **Con restricciones de igualdad y/o desigualdad**
- g) **Convexos, cuadráticos, separables.**
- h) **Con una sola variable independiente o con varias variables independientes.**

Como se observa existen muchos tipos de problemas no lineales, aunque, para este trabajo, se enfocará únicamente a los problemas restringidos y no restringidos, así también se analizarán algunos métodos iterativos sin

Es decir sea $f(X)$ una función continua, que denota a la función objetivo, $h_1(X), h_2(X), \dots, h_m(X)$ funciones continuas que denotan restricciones de igualdad, $g_{m+1}(X), g_{m+2}(X), \dots, g_p(X)$, funciones continuas que denotan restricciones de desigualdad y $X = (X_1, X_2, \dots, X_n)^T$ un vector en el espacio euclidiano de n -dimensiones, E^n .

Otra forma equivalentemente de representar un problema de optimización no lineal es el siguiente:

$$\text{Opt. } \{f(X) \mid X \in R\}, \quad (21)$$

donde el conjunto R se definen como:

$$R = \{ X \mid h_j(X) = 0, g_k(X) \geq 0, X \in E^n, \text{ para toda } j, k \}.$$

donde $f(X), h_j(X), g_k(X)$, son como se describió anteriormente.

Los problemas no lineales pueden ser de diferentes tipos:

- a) **Restringidos:** cuando se tienen restricciones (lineales o no lineales).
- b) **No restringidos:** cuando no se tienen restricciones y sólo se optimiza la función objetivo, que desde luego, no es lineal.
- c) **Continuos:** cuando alguna de las variables y funciones son continuas.
- d) **Discretos:** cuando alguna de las variables y/o funciones es discreta.
- e) **Diferenciables:** cuando todas las funciones del problema son doblemente diferenciables (que exista el límite).
- f) **Con restricciones de igualdad y/o desigualdad**
- g) **Convexos, cuadráticos, separables.**
- h) **Con una sola variable independiente o con varias variables independientes.**

Como se observa existen muchos tipos de problemas no lineales, aunque, para este trabajo, se enfocará únicamente a los problemas restringidos y no restringidos, así también se analizarán algunos métodos iterativos sin

restricciones y con restricciones. A continuación se mencionan algunos ejemplos de aplicación.

Ejemplos de aplicación

El problema de la mezcla con elasticidad de productos en los precios.

El problema de la mezcla de productos, la meta es determinar la mezcla óptima de niveles de producción para los productos de una empresa, dada las limitaciones sobre los recursos necesarios para producirlos, con el fin de maximizar la ganancia total de la empresa. En algunos casos existe una ganancia unitaria fija asociada a cada producto, con lo que la función objetivo que se tiene es lineal. Sin embargo, en muchos problemas de mezcla de productos, ciertos factores introducen no linealidad en la función objetivo. Por ejemplo: Un fabricante puede encontrar precios elásticos mediante las cuales la cantidad que se puede vender de un producto va en relación inversa con el precio que se cobra. Así, la curva precio-demanda puede parecerse a la que se muestra en la figura (2.3.1), en donde $f(X)$ es el precio que se necesita para poder vender X unidades. Si el costo unitario de producir el artículo está fijo en c (línea punteada de la figura), entonces la ganancia de la empresa por producir y vender X unidades está dada por una función no lineal,

$$f(X) = Xf(X) - CX$$

Como se puede observar en la figura (2.3.2). Si cada uno de los n productos de la empresa tiene una función parecida para la ganancia, por ejemplo $f_i(X_i)$ por producir y vender X_i unidades del producto i ($i = 1, \dots, n$), entonces es la función objetivo global es

$$f(X) = \sum_{j=1}^n f_j(X_j)$$

$f(X)$ es una suma de funciones no lineales.

El problema de transporte con descuentos por volumen en los precios de embarque

Un problema de transporte es determinar un plan óptimo para mandar bienes desde varios orígenes hasta varios destinos, dadas las restricciones de recursos y demanda, con el fin de minimizar el costo total de transporte. Los costos pueden no ser fijos, a veces se dispone de descuentos por volumen para cantidades grandes, con lo que el costo marginal de mandar una unidad más puede seguir un patrón como el que se muestra en la gráfica (2.3.3). El costo que resulta al embarcar x unidades está dado entonces por una función no lineal $C(x)$ que es una función lineal por partes con pendiente igual al costo marginal, como la que se muestra en la gráfica (2.3.4). En consecuencia, si cada combinación de origen y destino tiene una función de costo similar, es decir, si el costo de mandar x_{ij} unidades del origen i ($i = 1, 2, \dots, m$) al destino j ($j = 1, 2, \dots, n$) está dado por una función no lineal $C_{ij}(x_{ij})$ entonces la función objetivo global que se va a minimizar es el siguiente:

$$f(X) = \sum_{i=1}^m \sum_{j=1}^n C_{ij}(x_{ij})$$

Aun con esta función objetivo no lineal, es normal que las restricciones sean del tipo de restricciones lineales especiales que se ajustan al modelo del modelo de transporte.

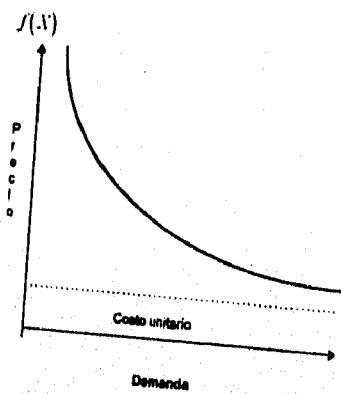


Figura 2.3.1. curva precio demanda.

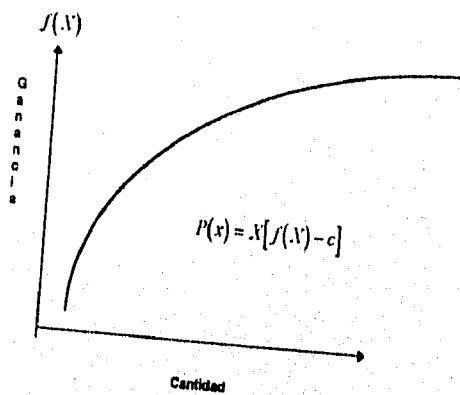


Figura 2.3.2. función de ganancia.

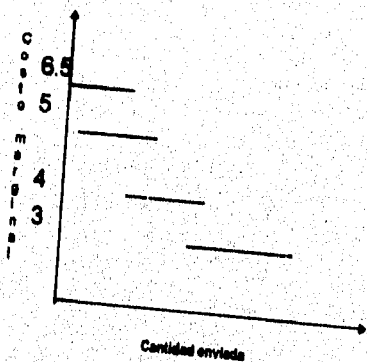


Figura 2.3.3. Costo marginal de transporte.

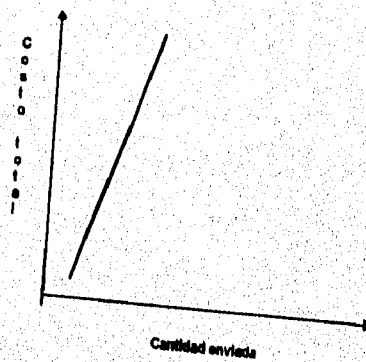


Figura 2.3.4. Función de costo de transporte.

II.3.1 Métodos iterativos para la solución de problemas no lineales

En el área de programación no lineal, existen muchos métodos para resolver problemas, una de las razones es que no existe un algoritmo general que resuelva la variedad de problemas no lineales que existe en el mundo real. Se han hecho grandes logros en lo que se refiere a algunos importantes casos especiales de este problema, haciendo algunas suposiciones sobre las funciones, aunque la investigación sigue activa.

En forma breve se plantearán los métodos, aunque en este trabajo es de sumo interés los métodos que no requieren que la función sea diferenciable, porque de acuerdo a la aplicación se necesitan los métodos que no usen derivadas.

Como ya se mencionó anteriormente, en esta etapa se va a dividir los problemas no lineales en dos ramas, los problemas no restringidos y los problemas restringidos, por lo tanto, el tema se va a analizar de la siguiente forma:

Primero se va a analizar los problemas de optimización no lineales y no restringidos junto con los métodos iterativos que resuelven dicho problema y en segundo lugar, se analizarán los problemas no lineales con restricciones, con sus respectivos métodos iterativos.

II.3.1.1 Métodos de optimización sin restricciones

En esta sección se analizan los problemas de optimización sin restricciones con la siguiente formulación matemática

minimizar $f(x)$

sujeto a:

$x \in E^n$

donde $f(x)$ es una función real y E^n es el espacio euclidiano de n -dimensiones.

Los métodos que se contemplan para esta sección son los siguientes:

- a) Métodos de optimización para funciones de una sola variable, en esta sección se explicarán los diferentes métodos iterativos que se tienen, son los siguientes: Fibonacci, Sección de oro, Interpolación cúbica, Interpolación cuadrada y Newton Raphson.
- b) Métodos basados en el uso de gradiente, para funciones de varias variables. Entre estos métodos se explicara los de gradiente, métodos basados en segundas derivadas, y métodos de direcciones conjugadas.
- c) Métodos de búsqueda para funciones de varias variables, no necesariamente diferenciables, por ejemplo método de Powell.

Los algoritmos que se tratarán son importantes para las aplicaciones prácticas, ya que suelen ofrecer posibilidades más sencillas y directas para la obtención de soluciones; pero quizá su mayor importancia radica en que establecen ciertos niveles de referencia respecto a la dificultad de aplicación y a la rapidez de convergencia.

En el caso de los métodos que usan gradientes para resolver funciones de varias variables, existe una estructura general para casi todos los algoritmos de descenso que se analizan. Se empieza en un punto inicial, se determina, de acuerdo con una regla fija, una dirección de movimiento y después se sigue esa misma dirección hacia un mínimo (relativo) de la función objetivo de esa recta. En el nuevo punto se determina una nueva dirección, y se repite el proceso. Las diferencias entre algoritmos radican en la regla mediante la cual se seleccionan las diferencias sucesivas del movimiento. Una vez que se ha hecho la selección

todos los algoritmos exigen movimiento hacia el punto mínimo de la recta correspondiente.

El proceso que determina el punto mínimo en una recta dada se denomina búsqueda lineal. De hecho, para funciones no lineales generales que no se puedan minimizar analíticamente, este proceso se completa buscando detenidamente en la recta el punto mínimo. Estas técnicas de búsqueda lineal, que no son otra cosa que procedimientos de resolución de problemas de minimización unidimensionales, constituyen la base fundamental de los algoritmos de programación no lineal, dado que los problemas de dimensiones superiores se resuelven realizando una secuencia de búsquedas lineales sucesiva.

Ejemplos de problemas sin restricciones

Los problemas de optimización sin restricciones se presentan en diversos contextos, pero son más frecuentes cuando la formulación del problema es sencilla. A menudo, las formulaciones más complejas implican restricciones explícitas. Sin embargo, muchos problemas con restricciones suelen convertirse en problemas sin restricciones utilizando las restricciones para establecer relaciones entre variables, reduciendo así, el número efectivo de variables.

Ejemplo 1 (Producción). Determinar la manera de combinar varias entradas para producir cierto bien, es un problema común en teoría económica. Hay una función de producción producida $f(x_1, x_2, \dots, x_n)$, que representa la cantidad de bienes producidos $j = 1, 2, \dots, n$. El precio por unidad de bien producido es q , y los precios unitarios de las entradas son p_1, p_2, \dots, p_n . El producto que desee maximizar la ganancia debe resolver el problema

$$qf(x_1, x_2, \dots, x_n) - p_1x_1 - p_2x_2 - \dots - p_nx_n$$

obteniendo la derivada parcial, la cual conduce directamente a las n ecuaciones

$$q \frac{\partial \gamma}{\partial x_i}(x_1, x_2, \dots, x_n) = p_i \quad i = 1, 2, \dots, n$$

Estas ecuaciones se pueden interpretar como confirmación, en la solución, de que el valor marginal debido a un incremento pequeño en la i -ésima entrada debe ser igual al precio p_i .

Ejemplo 2 (Aproximación). Un uso común de la optimización es en la aproximación de funciones. Suponga que durante un experimento se observa el valor de la función g en m puntos, x_1, x_2, \dots, x_m con lo que se conocen los valores $g(x_1), g(x_2), \dots, g(x_m)$ y se desea aproximar la función por un polinomio de grado n (o menor), donde $n < m$.

$$h(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_0$$

En relación con cualquier elección del polinomio de aproximación, habrá un conjunto de errores $\varepsilon_k = g(x_k) - h(x_k)$. Se define la mejor aproximación como el polinomio que minimiza la suma de los cuadrados de estos errores, es decir, se minimiza

$$\sum_{k=1}^m (\varepsilon_k)^2$$

Lo cual, a su vez, significa que se minimiza

$$f(a) = \sum_{k=1}^m \left[g(x_k) - (a_n x_k^n + a_{n-1} x_k^{n-1} + \dots + a_0) \right]^2$$

respecto a $a = (a_0, a_1, \dots, a_n)$ para hallar los mejores coeficientes. Esta es una expresión cuadrática en los coeficientes a . Para hallar una representación compacta de este objetivo se define

$$a_j = \sum_{k=1}^m (x_k)^{j+1}, \quad b_j = \sum_{k=1}^m g(x_k) (x_k)^j \quad \text{y} \quad c = \sum_{k=1}^m g(x_k)^2$$

Entonces, recurriendo al álgebra se puede mostrar lo siguiente:

$$f(a) = a^T Q a - 2b^T a + c$$

donde

$$Q = \{q_{ij}\}, \quad b = (b_1, b_2, \dots, b_{n+1})$$

Esto deduce directamente al sistema de $n+1$ ecuaciones $Qa = b$ que se pueden resolver para determinar a .

Ejemplo 3 (Control). Los problemas dinámicos, donde las variables corresponden a acciones tomadas en una sucesión de momentos de tiempo, a menudo se puede formular como problemas de optimización sin restricciones, por ejemplo, supóngase que la posición de un objeto grande se controla por una serie de fuerzas de control correctoras. El error de posición (la distancia a la posición deseada) está controlada por la siguiente ecuación:

$$x_{k+1} = x_k + u_k$$

donde x_k es el error en el tiempo k , y u_k es la fuerza efectiva aplicada al tiempo u_k . El valor x_0 está dado. La sucesión $u_0, u_1, u_2, \dots, u_n$ deberá seleccionarse para minimizar el objetivo.

$$J = \sum_{k=0}^n \{x_k^2 + u_k^2\}$$

esto representa un compromiso entre el deseo de igualar x_k a 0 y el reconocimiento de que la acción de control u_k es costosa.

El problema se puede convertir en un problema sin restricciones eliminando las variables x_k , $k = 1, 2, \dots, n$, del objetivo. Se ve en seguida la siguiente ecuación:

$$x_k = x_0 + u_0 + u_1 + u_2 + \dots + u_{k-1}$$

Así pues, el objetivo se puede volver a expresar como:

$$J = \sum_{k=0}^n \{(x_0 + u_0 + u_1 + \dots + u_{k-1})^2 + u_k^2\}$$

que es una función cuadrática de las incógnitas u_i , y tiene la misma estructura que la del **ejemplo 2**, por la cual se puede tratar de forma similar.

a) Métodos de optimización para funciones de una sola variable

Definición 2: Una función unimodal es cuando se tiene solamente un máximo o un mínimo.

Si una función unimodal decrece (aumenta) en forma monótona y después tiende a crecer (decrecer) también en forma monótona esto implica que es una función estrictamente unimodal, esto quiere decir que si $f(x)$ es estrictamente unimodal, entonces

$$X_1 < X_2 < X^* \Rightarrow f(X_1) < f(X_2) < f(X^*);$$

$$X^* < X_3 < X_4 \Rightarrow f(X^*) > f(X_3) > f(X_4);$$

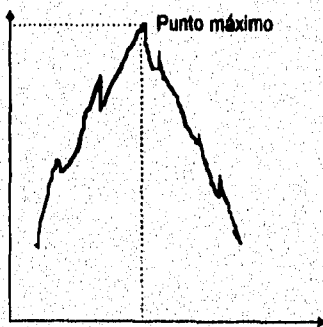


Figura 2.3.5.

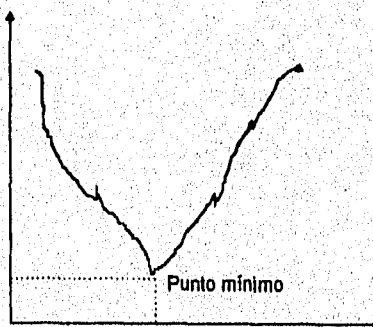


Figura 2.3.6.

De acuerdo a la dos figuras (2.3.5) y (2.3.6) anteriores 1 y 2 se deduce lo siguiente: toda función cóncava o convexa es unimodal, pero no toda función unimodal es cóncava o convexa.

Si se necesitan hacer k evaluaciones en una $f(x)$; y es estrictamente unimodal, entonces sea X_0 y X_{k+1} respectivamente los extremos izquierdo y derecho de intervalo de investigación. Sea p_k el mejor punto obtenido y sea $t_k < X^* < r_k$. La distancia $r_k - t_k$ se llama el intervalo de incertidumbre. Una medida de la efectividad de los métodos de búsqueda de puntos óptimos de una sola variable es la siguiente:

$$E = \frac{r_k - t_k}{X_{k+1} - X_0} \quad (22)$$

Comparación de métodos

Método	Efectividad
Búsqueda uniforme	$\frac{2}{(n+1)}$
Búsqueda dicotómica uniforme	$\frac{1}{\binom{n}{2} + 1}$
Búsqueda dicotómica secuencial	$\frac{1}{2^n}$
Búsqueda de Fibonacci	$\frac{1}{F_n}$
Búsqueda de sección de oro	$(0.618)^{1-n}$

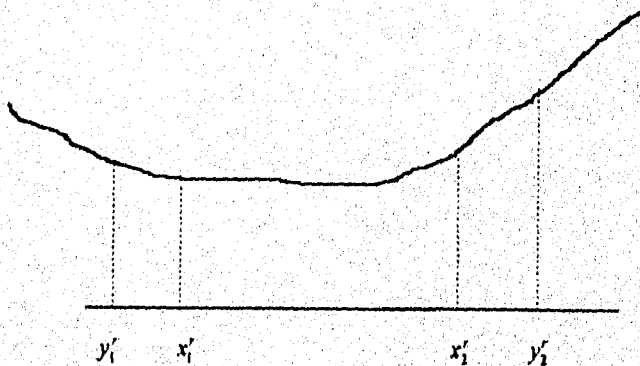
Método de Fibonacci

Este método es bastante eficiente para aproximar bajo cierta tolerancia de error, un punto mínimo o máximo en funciones unimodales¹ de una sola variable. Tiene la ventaja de que no utiliza derivada, por lo cual puede calcularse el punto óptimo en funciones no diferenciables. El método supone que se conoce un rango inicial de búsqueda en cada evaluación, el objetivo del método es acorralar al punto óptimo.

La búsqueda del mínimo de una función unimodal, es probablemente el criterio más simple y elegante de esta familia de procedimientos. Supóngase que se tiene una función $f(X)$ unimodal definida en un intervalo $[y_1', y_2']$, entonces, tomando puntos de prueba x_1' y x_2' sabemos que se puede suceder al menos una de las siguientes afirmaciones:

- si $f(x_1') < f(x_2')$ el mínimo se encuentra entre y_1' y x_2'
- si $f(x_1') > f(x_2')$ el mínimo se encuentra entre x_1' y y_2'
- si $f(x_1') = f(x_2')$ el mínimo se encuentra entre x_1' y x_2'

Gráficamente la solución es como sigue:



¹ Es una función que tiene solo un máximo o un mínimo.

La mejor estrategia para elegir la ubicación de los puntos de prueba se basa en calcular primero, o si se quiere en definir a priori, el número de iteraciones n que se van a realizar. Después de $r-1$ iteraciones, los puntos de prueba óptimos están dados por las siguientes relaciones:

$$x_1^r = \frac{F_{n-r-1}}{F_{n-r+1}}(y_2^r - y_1^r) + y_1^r; \quad r = 1, 2, \dots, n-2 \quad (23)$$

$$x_2^r = \frac{F_{n-r}}{F_{n-r+1}}(y_2^r - y_1^r) + y_1^r; \quad (24)$$

donde $(y_2^r - y_1^r)$ es la longitud del intervalo al cabo de la $r-1$ -ésima prueba y F_i son los números de Fibonacci que se obtienen mediante la siguiente ecuación iterativa,

$$F_0 = F_1 = 1$$

$$F_{n+1} = F_n + F_{n-1} \quad n = 1, 2, \dots \quad (25)$$

Si se observa las ecuaciones (23) y (24) se puede ver que para $r=n-1$ se tendrá la siguiente igualdad:

$$x_1^{n-1} = x_2^{n-1} \quad (26)$$

A fin de evitar esta situación se recurre a la siguiente expresión modificada,

$$x_1^{n-1} = \frac{1}{2}(y_2^{n-1} - y_1^{n-1}) \quad (27)$$

$$x_2^{n-1} = \left(\frac{1}{2} + \varepsilon\right)(y_2^{n-1} - y_1^{n-1}) + y_1^{n-1} \quad (29)$$

donde ε es tan pequeño como se quiera, para determinar en que intervalo se encuentra el mínimo. Se puede mostrar con el último intervalo después de n iteraciones tendrá una magnitud dada por:

$$y_1^n - y^n = \frac{1}{2F_n}(y_1^0 - y_1^0) + E \quad (30)$$

donde el superíndice 0 representa el valor original. En consecuencia, para una determinada precisión, debido a las propiedades de los número de Fibonacci, sólo se requiere de una evaluación de la función por iteración, después de las dos pruebas iniciales. Explícitamente, supóngase que la s-ésima iteración, el mínimo queda localizado entre y_1^s y x_2^s , entonces en la siguiente iteración se tendrá que:

$$x_2^{s+1} = \frac{F_n - (s+1)}{F_n + (s+1)}(x_2^s - y_1^s) + y_1^s \quad (31)$$

y sustituyendo este valor en la expresión (24) se tiene lo siguiente:

$$x_2^{s+1} = \frac{F_n - s - 1}{F_n + 1 - s}(y_2^s - y_1^s) + y_1^s = x_1^s$$

el mismo argumento se puede emplear en el caso de que el mínimo se localice en el otro semi-intervalo.

El número de iteraciones se puede determinar con una tabla de números de Fibonacci. Para ejemplificar, a continuación se ilustra una tabla.

Tabla de efectividad en función al número de iteraciones

Número de Iteraciones	Serie Fibonacci	Efectividad
n	F_n	$\frac{1}{F_n}$
1	1	.1
2	2	.5
3	3	.33
4	5	.2
5	8	.125

F_n es el número de Fibonacci para n evaluaciones y se define como en la ecuación (25).

La desventaja del método de Fibonacci es que, únicamente sirve para funciones de una sola variable. Si la función de una variable es bimodal² o multimodal³ este método sólo localizará un óptimo local o relativo. Su gran ventaja es que se le utiliza como una subrutina de búsqueda en los métodos de optimización de problemas no restringidos en funciones de varias variables.

Método de la sección de oro

Este método resulta, al hacer tender a infinito el número n de puntos de medición permitidos en una búsqueda de Fibonacci, el cual, se puede argumentar, basándose en la propiedad optimal del método finito de Fibonacci. La versión infinita correspondiente, genera una sucesión de intervalos de incertidumbre, las cuales intervalo tienden a cero más rápido de lo que se conseguiría por otros métodos; este método se utiliza cuando no se conoce el número total de iteraciones a realizar. Y el algoritmo requiere que la función sea de una sola variable y unimodal.

Si el rango de incertidumbre original es $a^1 \leq X^* \leq b^1$, al cual se le denotará por $[a^1, b^1]$, el proceso reduce el intervalo en cada evaluación, a un intervalo constante k .

Etapas del algoritmo

Paso 1. Se dan los intervalos a^1 y b^1 del rango de incertidumbre, $a^1 \leq X^* \leq b^1$. se calcula X_1^1 y X_2^1 de la siguiente manera:

² Bimodal significa que la función tiene dos óptimos locales o relativos.

³ Multimodal significa que la función tiene varios óptimos locales o relativos.

$$X_1^t = (1 - \tau)(b^t - a^t) + a^t$$

$$X_2^t = \tau(b^t - a^t) + a^t$$

donde

$$\tau = \frac{\sqrt{5} - 1}{2} \cong 0.618$$

Se hace $k=1$ y se va al paso 3, calculando $F(X_1^1)$ y $F(X_2^1)$.

Paso 2. $X_1^k = (1 - \tau)(b^k - a^k) + a^k$

$$X_2^k = \tau(b^k - a^k) + a^k$$

Calcule $F(X_1^k)$ y $F(X_2^k)$. Si $X_2^k - X_1^k < \epsilon$, donde $\epsilon > 0$ es una tolerancia prefijada, pare. De otro manera ir al paso 3.

Paso 3. Si $F(X_1^k) > F(X_2^k)$, hágase

$$a^{k+1} = X_1^k$$

$$X_1^{k+1} = X_2^k$$

$$b^{k+1} = b^k$$

$$k = k + 1$$

y regrese al paso 2 calculando únicamente X_2^k .

Si $F(X_1^k) < F(X_2^k)$,

hágase

$$b^{k+1} = X_2^k$$

$$X_2^{k+1} = X_1^k$$

$$a^{k+1} = a^k$$

$$k = k + 1$$

y se regresa al paso 2 calculando únicamente X_1^k .

Interpolación cúbica

Este método fue desarrollado por Davidon y es necesario que la función 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 búsqueda de los óptimos, mientras que s es la dirección de la búsqueda. Este mínimo local se obtiene en tres fases.

Etapas del algoritmo

Sea $g'(\alpha) = \frac{dg}{d\alpha}$

Paso 1. Dados valores arbitrarios para X y s , evaluar $g(\alpha) = f(X + \alpha s)$.

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

Paso 3. Se ajusta un polinomio cúbico tomando en consideración los valores $g(a)$, $g(b)$, $g'(a)$, $g'(b)$. El valor mínimo α^* se representa en esta iteración por α_* , donde:

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

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

y

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

Si $g(a)$ o $g(b)$ no son menores a $g(\alpha_*)$, entonces se acepta que $\alpha^* = \alpha_*$ (valor aproximado).

Si no, es decir, si $g(a) < g(\alpha_*)$ ó $g(b) < g(\alpha_*)$, entonces:

- a) Si $g'(a) > 0$, se repite el mismo procedimiento en el intervalo $a < \alpha^* \leq b$, donde $a = a$ y $b = \alpha_*$. Regrese a la paso 3.
- b) Si $g'(a) < 0$, se repite el mismo procedimiento en el intervalo $a < \alpha^* \leq b$, donde $a = \alpha_*$, y $b = b$. Regrese al paso 3.

Interpolación cuadrada

Cuando una función multimodal de una variable es discreta, la interpolación cúbica no sirve (puesto que la función no es diferenciable en los punto de discontinuidad, y los métodos de Fibonacci y sección de oro tampoco sirven, puesto que la función, aún siendo discreta, no es unimodal. En estos casos el método de interpolación cuadrada es el apropiado.

Este método fue diseñado por Powell y mejorado más tarde por Zangwill. En este caso presentamos el original (de Powell), que consiste en tres pasos.

Etapas del algoritmo

Paso 1. Dados los valores de punto arbitrario X y una dirección arbitraria s , evaluar

$$g(\alpha) = f(X + \alpha s).$$

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

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

Paso 3. Una vez que se conocen los valores $g(a)$, $g(b)$ y $g(c)$ se ajusta a un polinomio cuadrado a $f(x)$ y se calcula el mínimo local del polinomio, α_* , dado por la siguiente expresión:

$$\alpha_* = \frac{g(a)(c^2 - b^2) + g(b)(a^2 - c^2) + g(c)(b^2 - a^2)}{g(a)(c - b) + g(b)(a - c) + g(c)(b - a)}$$

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

Método de Newton-Raphson

Se supone una función $f(X)$ de una sola variable, que es continua y diferenciable. Este método localiza el punto donde la función cruza el eje de la variable independiente. A ese punto se le llama el cero de la función.

Una condición necesaria para que un punto X^* sea un óptimo local, es que $f'(X^*) = 0$, entonces este método localiza el o los ceros de la función $f'(X)$

Si se supone que se proporciona un punto inicial de búsqueda (X^1), entonces en la K -ésima iteración se tiene lo siguiente:

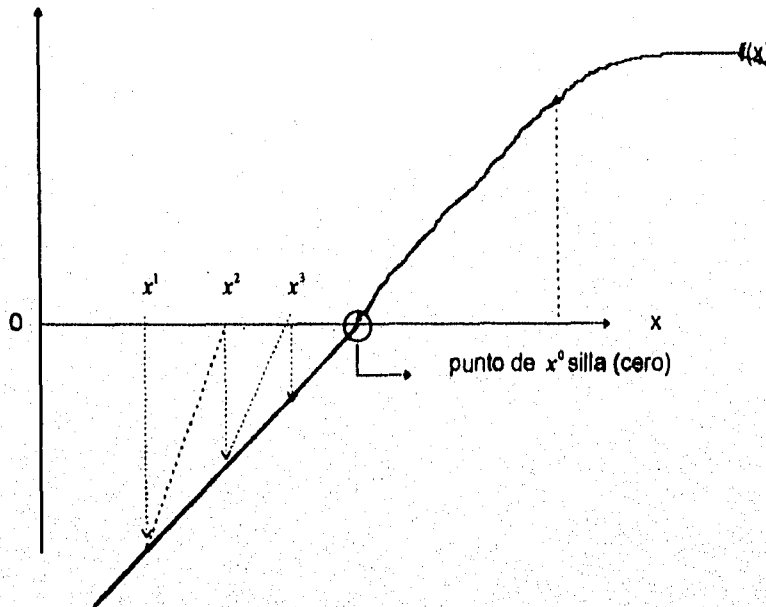
$$X^{k+1} = X^k - \frac{f(X^k)}{f'(X^k)} \quad f''(X^k) \neq 0,$$

donde

$$f'(X^k) = \left. \frac{df(X)}{dX} \right|_{X=X^k}$$

$$f''(X) = \frac{df'(X)}{dX} \Big|_{X=x^k}$$

Este procedimiento converge al cero de la función en un número finito de iteraciones en cualquier función cuadrática, tal como se muestra a continuación.



Sin embargo, existe el peligro que para ciertas funciones no cuadráticas, el método puede divergir. Es decir, este método localiza un cero a la vez, si la función tiene varios ceros (esto quiere decir que la función antiderivada tiene varios puntos extremos), entonces este método localizará esos ceros en función del valor inicial de búsqueda. Por tanto se debe emplear varias veces este método, cada vez, con un punto inicial de búsqueda diferente.

A continuación se plantean algunos métodos de optimización que utilizan derivadas para funciones de varias variables en problemas no restringidos. Método de ascenso y descenso acelerado, método de Newton, método de

Davidon-Fletcher-Powell y método de Fletcher Reeves. Estos métodos tienen como común denominador el uso del gradiente.

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

Los métodos se presentan en este orden porque los primeros dos ilustran de manera sencilla la filosofía del método, aunque no sea del todo práctico para aplicaciones comerciales, y los otros dos porque resulta ser de los más eficientes y rápidos para resolver problemas no restringidos (Davidon-Fletcher-Powell y Fletcher Reeves).

Un ejemplo empírico, para entender el funcionamiento de los métodos de gradientes el siguiente: supóngase que un alpinista desea alcanzar el pico de una montaña (máximo local), este necesita de tres elementos:

- a) Un punto conocido de partida.
- b) Una dirección de caminata, y
- c) Una longitud de caminata por esa dirección.

Una vez que el alpinista haya recorrido la distancia establecida, en la dirección sugerida, desde el punto inicial conocido, se encontrará en un segundo punto, en el cual fijará una nueva dirección y determinará una nueva longitud de recorrido. Así en iteraciones sucesivas de direcciones determinadas de movimiento y de longitud de recorrido de los mismos, el alpinista podrá llegar al pico de una montaña o la profundidad de un barranco.

Los métodos de gradiente hacen exactamente eso, únicamente que determinan la mejor dirección y la mejor longitud de recorrido en esa dirección, para poder alcanzar el máximo o el mínimo local, en el menor tiempo posible (menor número de iteraciones).

Método de ascenso o descenso acelerado

Sea $f(X)$ una función diferenciable en E^n . Dado un punto inicial de partida $X^0 = (x_1^0, x_2^0, \dots, x_n^0)^T$, la dirección de descenso acelerado en ese punto X^0 está dada por $-\nabla f(X)$, y la longitud de recorrido por un parámetro, $\alpha_j > 0$, tal que el punto generado en la iteración $i, i = 1, 2, \dots$, es:

$$X^{i+1} = X^i - \alpha_j \cdot \nabla f(X^i), \quad \text{Minimizar} \quad j = 1, 2, \dots \quad i = 0, 1, 2, \dots$$

$$X^{i+1} = X^i + \alpha_j \cdot \nabla f(X^i), \quad \text{Maximizar} \quad j = 1, 2, \dots \quad i = 0, 1, 2, \dots$$

Para encontrar el valor de α_j , se minimiza la función de una sola variable

$$g(\alpha_j) = f(X^i - \alpha_j \cdot \nabla f(X^i))$$

por alguno de los métodos discutidos en la sección anterior.

El vector $X^i - \alpha_j \cdot \nabla f(X^i)$, es tangente al contorno $g(\alpha_j)$ y como $\nabla f(X^{i+1})$, es perpendicular a este contorno, los movimientos sucesivos son ángulos rectos de 90° .

Existe dos criterios para terminar el algoritmo. 1) Consiste en parar en cuando $\nabla f(X^*) = 0$ en el punto X^* , pues se sabe que este debe ser una condición necesaria (más no suficiente) que debe cumplir un punto máximo o mínimo. Entonces, que el gradiente en el punto X^* se aproxima al valor cero, lo cual, es equivalente a:

$$\sum_{i=1}^n \left[\frac{\partial f(X^*)}{\partial x_i} \right]^2 < \epsilon$$

donde $\frac{\partial f(X^k)}{\partial X_i}$ son los componentes del gradiente, evaluadas en el punto X^k y

$\varepsilon > 0$ es tolerancia arbitraria. Se utiliza el cuadrado del valor de las componentes, con objeto de evaluar la suma de los valores absolutos. 2) Consiste en comparar el valor de la función en dos iteraciones sucesivas y para el algoritmo cuando esa diferencia es menor a una tolerancia arbitraria, es decir, cuando

$$|f(X^{k+1}) - f(X^k)| < \varepsilon, \quad \text{donde } \varepsilon > 0$$

Este método encontrará un mínimo o máximo local, generalmente el que está cercano al punto de inicialización X^0 . Para encontrar todos los mínimos o máximos locales hay que repetir el método con diferentes puntos de inicialización X^0 .

La desventaja de este método es que su convergencia al punto óptimo local buscado es muy lenta, debido precisamente a los movimientos perpendiculares. Esto puede ocasionar un -zigzag- ineficiente en funciones cuyos contornos son excéntricos como el que se muestra en la siguiente figura.

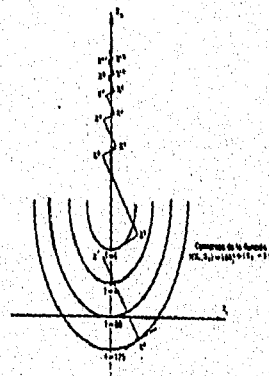


Figura.2.3.7.

Si la función tiene contornos circulares, el óptimo local se encuentra en una sola iteración.

Desgraciadamente, en la práctica, las funciones de varias variables tienen un componente muy irregular, ya que sus contornos son excéntricos o bien no circulares. En estos casos los métodos de ascenso y descenso acelerado son muy ineficientes, por su convergencia lenta.

Método de Newton

El método de Newton también se conoce con el nombre de Newton-Raphson de varias variables; este método es muy parecido al anterior, y está basado en la siguiente regla iterativa:

$$X^{k+1} = X^k - H^{-1}(X^k) \nabla f(X^k) \quad (\text{para minimización})$$

y

$$X^{k+1} = X^k + H^{-1}(X^k) \nabla f(X^k) \quad (\text{para maximización})$$

Dado un punto inicial $X^0 = (x_1^0, x_2^0, \dots, x_n^0)^T$,

donde X^0 en E^n es arbitrario, $f(X)$ es diferenciable, $\nabla f(X^k)$ es la transpuesta de gradiente evaluado en el punto X^k de la k -ésima iteración y $H^{-1}(X^k)$ es la matriz inversa del Hessiano, evaluado en el punto X^k de la k -ésima iteración.

Criterio para parar el método

$$|f(X^{k+1}) - f(X^k)| < \varepsilon, \quad \text{donde } \varepsilon > 0$$

Direcciones conjugadas

Los dos métodos anteriores, (ascenso-descenso acelerado y el de Newton) converge muy lentamente a la solución buscada, sobre todo cuando el gradiente adquiere un valor cercano a cero (en los contornos topográficos tipo valles).

Recientemente se ha descubierto una serie de técnicas que eliminan esta desventaja. La lógica detrás de estos métodos es similar a la de los dos anteriores y sólo difiere en que, en vez de tomar direcciones sucesivas que son perpendiculares (forman ángulos de 90°), se calculan las llamadas direcciones conjugadas.

Dada una función diferenciable de varias variables, la serie de expansión de Taylor en torno al óptimo local X^* , dice que

$$f(X) \cong f(X^*) + (X - X^*)^T \nabla f(X^*) + \frac{1}{2} (X - X^*)^T H[f(X^*)] (X - X^*),$$

pero por la condición necesaria que debe satisfacer el óptimo local X^* , que es $\nabla f(X^*) = 0$, se tiene

$$f(X) \cong f(X^*) + \frac{1}{2} (X - X^*)^T H[f(X^*)] (X - X^*) \quad (32)$$

Se supone que el Hessiano $H[f(X^*)]$ es positivo definido. bajo tales suposiciones, $f(X)$ se comporta como una función cuadrática en el contorno del punto X^* . De aquí se infiere, que cualquier método que se comporta eficientemente con una forma cuadrática, lo hará también con una función de tipo general diferenciable, ya que por la serie de expansión de Taylor cualquier función diferenciable de tipo general, puede aproximarse por la forma cuadrática (32) que a su vez se puede representar en forma matricial como:

$$f(X) = a + b^T X + \frac{1}{2} X^T H X,$$

Se supone que $X \in E^n$, $f(X)$ es diferenciable, b es un vector con n componentes conocidas a es un escalar y H es el Hessiano, al cual se supone definido y simétrico. El Hessiano debe evaluarse en cada punto X^k , generado en la iteración k , $k = 1, \dots, n$. se tiene lo siguiente:

$$\nabla f(X) = b + HX.$$

Si X^* es un mínimo local de $f(X)$, entonces es necesario que $\nabla f(X^*) = 0$, por lo que

$$b + HX^* = 0. \quad (33)$$

Dado un punto inicial $X^0 = (x_1^0, x_2^0, \dots, x_n^0)^T$ y un conjunto de n direcciones independientes $(s_0, s_1, \dots, s_{n-1})^T$, es posible calcular un conjunto de constantes β_i tal que

$$X^* = X^0 + \sum_{i=0}^{n-1} \beta_i s_i \quad (34)$$

Cada iteración $s_i, i = 0, \dots, n-1$, es un vector con n componentes. Se dice que las direcciones $(s_0, s_1, \dots, s_{n-1})^T$, son conjugadas de la matriz H , si es que cumple con la siguiente condición:

$$s_i^T H s_j = 0, \quad i \neq j, \quad i, j = 0, 1, \dots, n-1 \quad (35)$$

con $s_i, s_j \neq 0$ para toda i, j . Se puede probar que cualquier conjunto de direcciones que satisface la definición de ser conjugadas, es un conjunto linealmente independiente. Por lo tanto, se calculan los coeficientes $\beta_i, i = 0, \dots, n-1$ de la siguiente manera:

$$X^* = X^0 + \sum_{i=0}^{n-1} \beta_i s_i$$

Multiplicado por $s_j^T H$,

$$\begin{aligned} s_j^T H X^* &= s_j^T H X^0 + \sum_{i=0}^{n-1} s_j^T H \beta_i s_i \\ &= s_j^T H X^0 + \sum_{i=0}^{n-1} \beta_i s_j^T H s_i \end{aligned} \quad (36)$$

De la definición (36) se tiene que

$$s_j^T H s_i = 0 \quad \text{para } i \neq j,$$

por ser s_i, s_j direcciones conjugadas, por lo que (37) se convierte en

$$s_j^T H X^* = s_j^T H X^0 + \beta_j s_j^T H s_j \quad (37)$$

De (33) se tiene que

$$H X^* = -b$$

Por (37) se puede escribir

$$\begin{aligned} s_j^T (-b) &= s_j^T H X^0 + \beta_j s_j^T H s_j \\ \beta_j s_j^T H s_j &= s_j^T (-b) - s_j^T H X^0 \\ \beta_j &= -(b + H(X^0)X^0)^T \frac{s_j}{s_j^T H(X^0)s_j} \end{aligned} \quad (38)$$

Esta última ecuación establece la lógica iterativa. Empezando en un punto inicial X^0 y un dirección arbitraria s_0 , se calcula un nuevo punto X^1 de la siguiente manera:

$$X^1 = X^0 + \beta_0 s_0$$

donde β_0 se obtiene de la optimización de la función de una sola variable

$$g(\beta_0) = f(X^0 + \beta_0 s_0)$$

Pero esta optimización genera precisamente la expresión (38) en efecto, el gradiente de $g(\beta_0)$ es

$$s_0^T \nabla F(X^0 + \beta_0 s_0)$$

Una condición necesaria para un óptimo local es que esta gradiente sea igual a cero. Por lo tanto

$$s_0^T \nabla F(X^0 + \beta_0 s_0) = 0$$

Pero esto implica de (33) que

$$s_0^T (b + H[X^0 + \beta_0 s_0]) = 0$$

$$s_0^T b + s_0^T H X^0 + s_0^T H s_0 \beta_0 = 0,$$

O sea que

$$\beta_0 = \frac{-s_0^T - s_0^T H(X^0) X^0}{H(X^0) s_0}$$

$$\beta_0 = -(b + H(X^0) X^0)^T \frac{s_0}{s_0^T H(X^0) s_0}$$

Pero esta expresión es precisamente (38). Una vez obtenida β_0 y por consiguiente el nuevo punto X^1 , hay que generar una dirección s_1 , que sea conjugada a s_0 , es decir, que satisfaga (35) y así poder calcular el punto X^2 , donde:

$$X^2 = X^1 + \beta_1 s_1$$

y β_1 se calcula de (38) es decir,

$$\beta_1 = -(b + H(X^1) X^1)^T \frac{s_1}{s_1^T H(X^1) s_1}$$

Para una función el óptimo local X^* , se genera en n iteraciones sucesivas.

Los siguientes dos métodos están basados en direcciones conjugadas, el primero, llamado el método de Davidon-Fletcher-Powell, es considerado como uno de los mejores en cuanto a eficiencia entre todos los métodos conocidos de optimización no restringidas, desde el punto de vista de su convergencia

acelerada en la solución de problemas prácticos, y el segundo método es de Fletcher-Reeves.

Método de Davidon-Fletcher-Powell

Este método, basado en la lógica de direcciones conjugadas es una combinación de los trabajos de Davidon y Fletcher-Powell. Este método es considerado por los expertos de este campo (programación matemática), como uno de los algoritmos más eficientes para calcular puntos óptimos locales en funciones diferenciables de varias variables.

Se considera una función $f(X)$ de n variables, que es diferenciable. El punto principal de este método es una matriz D_k , simétrica, positiva definida, que se evalúa en cada iteración k , $k = 1, 2, \dots, n$.

Esta matriz se actualiza en cada iteración y con ella se calculan las direcciones conjugadas s_k y las longitudes β_k de la iteraciones k , $k = 1, 2, \dots, n$.

Siendo D_0 una matriz arbitraria, simétrica, positiva definida en n por n , y X^0 n punto inicial arbitrario, entonces la primera dirección $s_0 \in E^n$ se obtiene de

$$s_0 = -D_0 \nabla f(X^0).$$

La longitud β_0 se obtiene de la minimización de la función de una sola variable

$$g(\beta_0) = f((X^0)^T + \beta_0 (s_0^T)).$$

Se define un parámetro σ_0 , donde $\sigma_0 = \beta_0 s_0$
por lo que

$$X^1 = X^0 + \sigma_0.$$

La actualización de matriz D se calcula de $D_1 = D_0 + A_0 + B_0$

donde

$$A_0 = \frac{\sigma_0 \sigma_0^T}{\sigma_0^T \Delta_0}$$

$$\Delta_0 = \nabla f(X^1) - \nabla f(X^0)$$

Y

$$\beta_0 = \frac{-(D_0 \Delta_0 \Delta_0^T D_0)}{\Delta_0^T D_0 \Delta_0}$$

En forma general, se tienen conocidos D_k y X_k en la iteración $k, k = 1, 2, \dots, n$ el nuevo punto X^{k+1} se obtiene de

$$X^{k+1} = X^k + \sigma_k$$

donde

$$\sigma_k = \beta_k s_k$$

La dirección $s_k \in \mathbb{R}^n$ se obtiene de

$$s_k = -D_k \nabla f(X^k),$$

y β_k se obtiene de la minimización

$$g(\beta_k) = f((X^k)^T + \beta_k (s_k)^T)$$

La actualización de D_{k+1} se calcula de

$$D_{k+1} = D_k + A_k + B_k,$$

donde

$$A_k = \frac{\sigma_k \sigma_k^T}{\sigma_k^T \Delta_k},$$

$$\Delta_k = \nabla f(X^{k+1}) - \nabla f(X^k),$$

Y

$$\beta_k = \frac{-(D_k \Delta_k \Delta_k^T D_k)}{\Delta_k^T D_k \Delta_k}$$

Cabe mencionar que el numerador son matrices, mientras que los denominadores son escalares.

Los autores de este método definen lo siguiente:

a) La matriz D_k es positiva, definida y simétrica en cada iteración k , $k = 1, 2, \dots, n$. por lo tanto, el método converge a una solución óptima local, porque

$$\frac{dg(\beta_k)}{d\beta_k} = \frac{df((X^k)^T + \beta_k (s_k)^T)}{d\beta_k} = \nabla f((X^k)^T + \beta_k (s_k)^T) s_k.$$

Para $\beta_k = 0$ se tiene que

$$\frac{dg(\beta_k)}{d\beta_k} = \nabla f^T(X^k) s_k$$

Pero $s_k = -D_k \nabla f^T(X^k)$, por lo que

$$\frac{dg(\beta_k)}{d\beta_k} = -\nabla f^T(X^k) D_k \nabla f(X_k) < 0.$$

Por que D_k es positiva definida. Es decir, la desigualdad anterior significa que la función $f(X)$ inicialmente disminuye o decrece a lo largo de la dirección s_k , y por lo tanto se le puede hacer más pequeña en cada iteración.

b) Cuando este método se le aplica a una función cuadrática de la forma:

$$f(X) = \alpha + b^T X + \frac{1}{2} X^T A X$$

entonces las direcciones $s_k \in E^n$ son conjugadas de la matriz A , y por lo tanto el método conduce a un mínimo en n iteraciones. También D_k converge a la inversa de la matriz A a medida que se continua iterando, es decir, $\lim_{k \rightarrow \infty} D_k = A^{-1}$.

Método de Fletcher-Reeves

Como ya se mencionó anteriormente este método está basado en la lógica de las direcciones conjugadas y es un trabajo de Fletcher y Reeves. Este método no es tan eficiente como el método de Davidon-Fletcher-Powell, pero es mucho más sencillo de usar para cálculos manuales, además de tener otras ventajas que se describen posteriormente, aunque este método no es tan potente como el anterior sigue siendo más eficiente que los métodos de ascenso-descenso acelerado y el de Newton.

Este método supone que $f(x)$ es una función diferenciable con seE^n . Dado un punto arbitrario X^0 , el método calcula una dirección inicial

$$s_0 = \begin{bmatrix} s_1^0 \\ s_2^0 \\ \dots \\ s_n^0 \end{bmatrix}$$

donde

$$s_0 = -\nabla f(X^0)$$

Se escoge una longitud β_0 minimizando

$$g(\beta_0) = f(X^0 + \beta_0 s_0)$$

En el punto X^1 es $X^1 = X^0 + \beta_0 s_0$.

La nueva dirección es

$$s_1 = \begin{bmatrix} s_1^1 \\ s_2^1 \\ \dots \\ s_n^1 \end{bmatrix}$$

se obtiene de

$$s_1 = -\nabla f(X^1) + \alpha_0 s_0.$$

donde

$$\alpha_0 = \frac{\nabla f^T(X^1) \nabla f(X^1)}{\nabla f^T(X^0) \nabla f(X^0)}.$$

En forma general, se tiene que al terminar k , $k = 1, 2, \dots, m$, el punto X^{k+1} se calcula de

$$X^{k+1} = X^k + \beta_k s_k,$$

donde β_k se obtiene de la minimización de

$$g(\beta_k) = f((X^k)^T + \beta_k (s_k)^T)$$

$$s_{k+1} = -\nabla f((X^{k+1})^T) + \alpha_k (s_k)^T$$

donde

$$\alpha_k = \frac{\nabla f^T(X^{k+1}) \nabla f(X^{k+1})}{\nabla f^T(X^k) \nabla f(X^k)}$$

Ventajas del este método es que aun siendo menos eficiente que el de Davidon-Fletcher-Powell, requiere mucho menos memoria para su procesamiento en una computadora. Esto es muy significativo cuando el valor de n es muy grande.

c) Métodos de optimización para funciones no diferenciables de varias variables

Método de Powell

Una de las grandes desventajas de los métodos analizados anteriormente es que suponen que la función debe ser diferenciable, con objeto de poder evaluar el gradiente y en varios casos, también el Hessiano.

Otro de los problemas es en el caso de que la función sea diferenciable, es el evaluar analíticamente en cada iteración las funciones contenidas en el gradiente y en el Hessiano.

Por tal motivo, se han desarrollado una serie de métodos iterativos que, sin ser tan eficientes en cuanto al número de iteraciones necesarias para alcanzar un óptimo como los métodos basados en el gradiente (Davidon-Fletcher-Powell, Fletcher-Reeves), tiene las siguientes ventajas:

- a) No requiere de gradiente y Hessianos, por lo que sirve tanto para funciones diferenciables como no diferenciables,
- b) No requiere de métodos numéricos y por lo tanto su error de redondeo se reduce considerablemente.

El método de Powell requiere en cada iteración, de n minimizaciones de funciones de una variable, por métodos que no usan derivadas, como son la interpolación cuadrada para funciones multimodales de una sola variable y los métodos de Fibonacci o sección de Oro, para funciones unimodales de una sola variable. Estas minimizaciones se hace en la dirección de n vectores independientes, s_1, s_2, \dots, s_n , cada vector con n componentes. Como resultado de estos n procesos de minimización, se define en cada iteración una nueva dirección. Si esta dirección s aprueba un cierto criterio, entonces reemplaza algunas de las n direcciones antiguas s_1, s_2, \dots, s_n .

El proceso se inicializa con n direcciones arbitrarias s_1, s_2, \dots, s_n y un punto arbitrario $X^0 = (X_1^0, X_2^0, \dots, X_n^0)^T$.

Powell recomienda que estas direcciones arbitrarias s_1, s_2, \dots, s_n se definen de la siguiente manera:

$$\begin{aligned} s_1 &= (X_1^0, 0, 0, \dots, 0) \\ s_2 &= (0, X_2^0, 0, \dots, 0) \\ &\vdots \quad \quad \quad \vdots \\ s_n &= (0, 0, 0, \dots, X_n^0) \end{aligned}$$

Dado X^0 , y las n direcciones arbitrarias s_1, s_2, \dots, s_n , el método de Powell procede de la forma ilustrada a manera de un diagrama de flujo, que se ilustra en la página 70 y 71.

Se utiliza la siguiente notación: en la iteración k , generan $n+2$ puntos, $X^{k,0}, X^{k,1}, \dots, X^{k,n+1}$, cada punto con n componentes. Por ejemplo, el punto $X^{k,j}$ tiene los siguientes componentes:

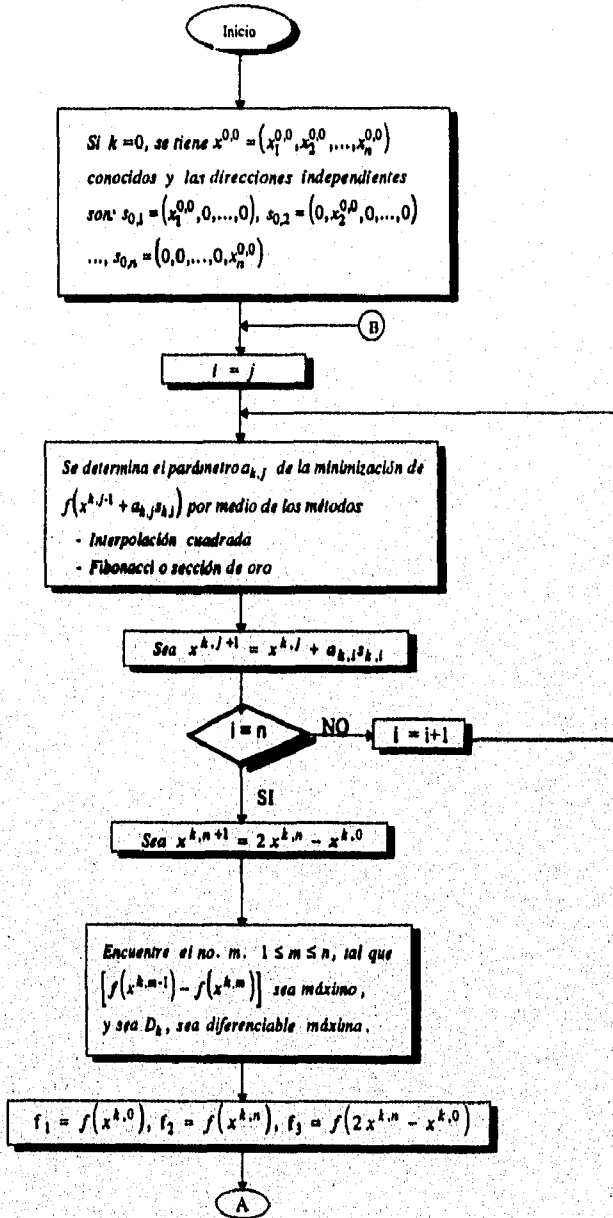
$$X^{k,j} = (X_1^{k,j}, X_2^{k,j}, \dots, X_n^{k,j}) \quad j = 0, 1, 2, \dots, n+1.$$

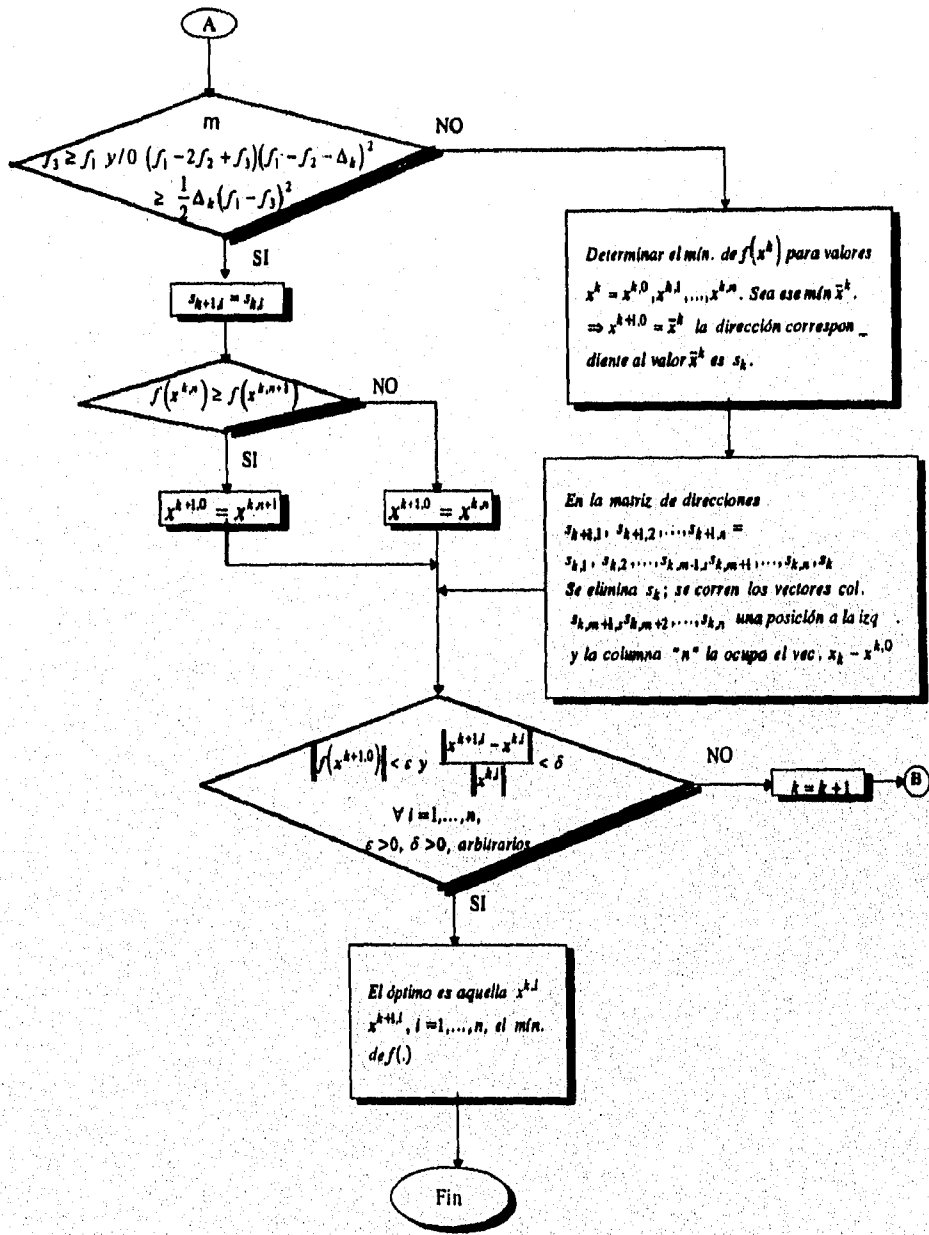
También en cada iteración k existirán n direcciones independientes $s_{k,1}, s_{k,2}, \dots, s_{k,n}$, cada dirección con n componentes, es decir,

$$s_{k,j} = (s_1^{k,j}, s_2^{k,j}, \dots, s_n^{k,j}), \quad j = 1, 2, \dots, n$$

Desventaja del método, suele no ser eficiente en el número de iteraciones necesarios para alcanzar el óptimo.

DIAGRAMA DE FLUJO DE POWELL





11.3.1.2 Métodos de optimización con restricciones

Los métodos que resuelven problemas de optimización con restricciones se encuentra mucho menos desarrollados que los métodos de optimización no-lineal, no restringidos. Esto se debe a la enorme dificultad inherente en la solución de problemas restringidos.

El problema general que se quiere resolver, es:

$$\begin{array}{ll} \text{Opt.} & f(x) \\ \text{Sujeto a:} & \\ & h_j(x) = 0, \quad j = 1, \dots, m \\ & g_j(x) \geq 0, \quad j = m+1, \dots, p \\ & x \in E^n \end{array}$$

como se analizó en la sección (11.4) los problemas inherentes a este tipo de situación son:

- a) Existencia de óptimos locales,
- b) Óptimos locales que se encuentran en el interior o en los bordes de una región de factibilidad, que en general no es convexa.

Sin embargo, en esta sección trataremos una clase de problemas de optimización no lineal, donde estas dos dificultades desaparecen completamente. A continuación se mencionan el tipo de problemas para posteriormente explicar las condiciones necesarias y suficientes, que debe satisfacer un óptimo local en problemas no lineales más generales. Con base en esto, surgen tres grandes familias de métodos.

- a) Métodos de optimización por aproximación lineal
- b) Métodos penales

Estos métodos se basan principalmente, en la programación convexa, después de mostrar algunos ejemplos de programación no lineal, se enunciarán los principales conceptos y teoremas sin probar, ya que no es el objetivo de este trabajo; para iniciar dicho análisis, y posteriormente se presentarán los métodos más importantes.

Ejemplos de aplicación de problemas con restricciones

En los últimos 20 años la programación no lineal pasó de ser un tema relativamente nuevo y principalmente analítico a ser un importante instrumento de carácter general para la resolución de problemas.

Los grandes problemas de programación no lineal se presenta en diversos campos, por ejemplo, los problemas de estructuras mecánicas, como determinar configuraciones óptimas para puentes, armazones, etc. Algunos diseños mecánicos y configuraciones que antes se resolvían mediante ecuaciones diferenciales, ahora, suelen solucionarse resolviendo problemas de optimización apropiados. Por ejemplo, determinar la forma de un cable rígido suspendido entre dos puntos y soportando una carga.

Existe una amplia clase de problemas de optimización a gran escala que se presenta de manera similar a los métodos para resolver ecuaciones diferenciales parciales. En situaciones en que las variables continuas básicas están definidas en una región bi o tridimensional, la región continua se sustituye por una red formada por quizás varios miles de puntos discretos. Entonces, la correspondiente aproximación discreta a la ecuación diferencial parcial se resuelve indirectamente formulando un problema de optimización equivalente. Este enfoque se utiliza en estudios de plasticidad, en ecuaciones de calor, en el flujo de fluidos, en física atómica y, de hecho, en casi todas las ramas de la ciencia física.

Los problemas de control óptimo dan lugar a problemas de programación no lineal a gran escala. En estos problemas, un sistema dinámico, que se

describe a veces por una ecuación diferencial normal, relaciona variables de control con una trayectoria del estudio del sistema. Esta ecuación diferencial, o una versión directa de la misma, define un conjunto de restricciones. El problema consiste en seleccionar las variables de control de modo que la trayectoria resultante satisfaga variables y minimice algún criterio. Un problema inicial de este tipo de problemas, resuelto numéricamente, fue la determinación de la trayectoria de un cohete a la luna que requería un consumo mínimo de combustible.

Hay muchos ejemplos de programación no lineal en las operaciones industriales y toma de decisiones mercantiles. La no linealidad se puede presentar en funciones de producción, curvas de costos y, de hecho, en casi todas las facetas de la formulación de problemas.

Los análisis de cartera, en el contexto de la inversión en el mercado de valores y en la evaluación de complejos proyectos empresariales, es un área donde la programación no lineal es cada vez más útil. Estos problemas pueden llegar a tener fácilmente miles de variables.

En muchas áreas de la construcción y análisis de modelos, la formulación directa de sistemas de ecuaciones está siendo sustituida cada vez en mayor medida por las formulaciones de optimización. Así, los grandes modelos de predicción económica a menudo determinan los precios de equilibrio minimizando un objetivo llamado excedente del consumidor. Con frecuencia se formulan modelos físicos como minimización de la energía. Los problemas de decisión se formulan como la maximización del beneficio esperado. Los procedimientos de análisis de datos se basan en la minimización de un error promedio y en la maximización de una probabilidad.

Programación convexa

Un problema de optimización convexo, es aquel que minimiza (maximiza) una función convexa (cóncava) dentro de una región de factibilidad convexa.

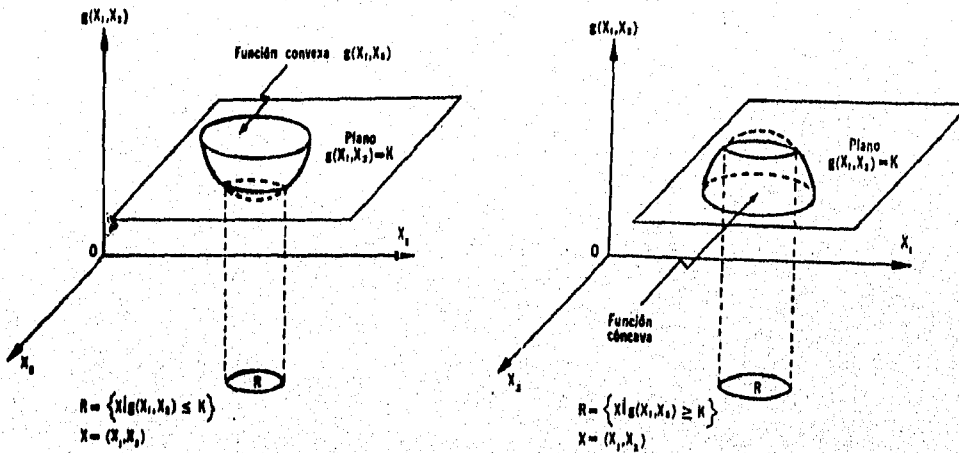


Figura 2.3.8.

Teorema 4: Si la función $g(X)$ es convexa, entonces el conjunto

$$R = \{X / g(X) \leq K\}$$

es convexo para todos los valores del parámetro k . Si $g(X)$ es cóncava, entonces el conjunto $R = \{X / g(X) \geq K\}$, es convexo para todos los valores del parámetro k .

Teorema 5: Cada punto mínimo (máximo) local o relativo, de un problema de optimización convexo, es un mínimo (máximo) global o absoluto.

Teorema 6: Si $f(X)$ es una función continua y diferenciable entonces las siguientes tres expresiones son equivalentes entre sí:

- a) $f(X)$ es convexa.
- b) $f(X) \geq f(X^1) + \nabla f^T(X^1)(X - X^1)$, para cualquier par de puntos X^1 y X^2 , y
- c) $H(X)$ es una matriz positiva semidefinida para toda X .

Teorema 7: La función cuadrática.

$$f(X) = c^T X + X^T Q X$$

con la matriz Q semidefinida, es convexa.

Teorema 8: Una función $f(X)$ es convexa si y sólo si, la función de una sola variable $f(X + as)$ es convexa, para cualquier valor fijo y conocido de X y s .

Con base en los teoremas anteriores surgen algunos métodos que permite determinar si una función es convexa o cóncava. A continuación se presenta uno de ellos, el cual resulta impráctico, por su dificultad de cómputo.

Sea $f(X)$ una función diferenciable, con $X \in E^n$. Sea $H(X)$ el Hessiano de $f(X)$, es decir,

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial X_1 \partial X_1} & \frac{\partial^2 f}{\partial X_1 \partial X_2} & \dots & \frac{\partial^2 f}{\partial X_1 \partial X_n} \\ \frac{\partial^2 f}{\partial X_2 \partial X_1} & \frac{\partial^2 f}{\partial X_2 \partial X_2} & \dots & \frac{\partial^2 f}{\partial X_2 \partial X_n} \\ \vdots & \vdots & \dots & \vdots \\ \frac{\partial^2 f}{\partial X_n \partial X_1} & \frac{\partial^2 f}{\partial X_n \partial X_2} & \dots & \frac{\partial^2 f}{\partial X_n \partial X_n} \end{bmatrix}$$

Entonces los menores principales de H se denotarán por $\Delta_1, \Delta_2, \Delta_3, \dots, \Delta_n$, donde cada uno de estos determinantes está dado por:

$$\Delta_1 = \frac{\partial^2 f}{\partial x_1 \partial x_1}, \Delta_2 = \begin{vmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2 \partial x_2} \end{vmatrix}, \dots, \Delta_n = H$$

El método indica que si $\Delta_1 > 0, \Delta_2 > 0, \dots, \Delta_n > 0$ entonces $f(X)$ es convexa, mientras que a los menores pares $\Delta_2, \Delta_4, \Delta_6, \dots$ son negativos y los menores impares $\Delta_1, \Delta_3, \Delta_5, \dots$ son positivos, la función $f(X)$ es cóncava. Este método no trabaja si $f(X)$ es estrictamente cóncava o convexa. Por lo que ya se mencionó anteriormente, el método es impráctico debido a la dificultad de calcular los determinantes asociados a los menores del Hessiano.

Se hacen las siguientes observaciones. Si el problema a optimizar fuera de la siguiente forma:

Opt. $f(X)$

Sujeto a:

$$\begin{aligned} h_i(X) &\leq 0 & i = 1, \dots, m \\ g_j(X) &= 0 & j = 1, \dots, r \\ X &\in E^n \end{aligned}$$

entonces el conjunto

$$R_j = \{X / h_j(X) = 0\}, \quad j = 1, \dots, m$$

es por lo general no convexo, a menos que la función $h_j(X)$ sea cuasi-convexa.

Una función $h_j(X)$ es pseudo-convexa si:

$$\nabla h^T(X^1)(X^1 - X^2) \geq 0$$

esto implica

$$h(X^1) \geq h(X^2)$$

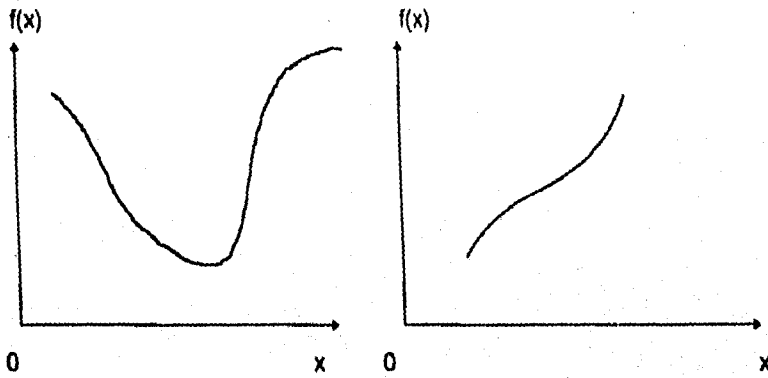


Figura 2.3.9. Funciones pseudo-convexa

Una función $h(X)$ es pseudo-cóncava si $-h(X)$ es pseudo-convexa. Una función $h(X)$ es cuasi-cóncava si dados dos puntos X^1, X^2 en el dominio de la función, y un parámetro $\lambda, 0 \leq \lambda \leq 1$, $h(X^2) \geq h(X^1)$, la cual, implica que $h(X^2) \geq h((1-\lambda)X^2 + \lambda X^1)$.

Una función $h(X)$ es cuasi-cóncava. Este tipo de funciones se ilustra en la siguiente gráfica.

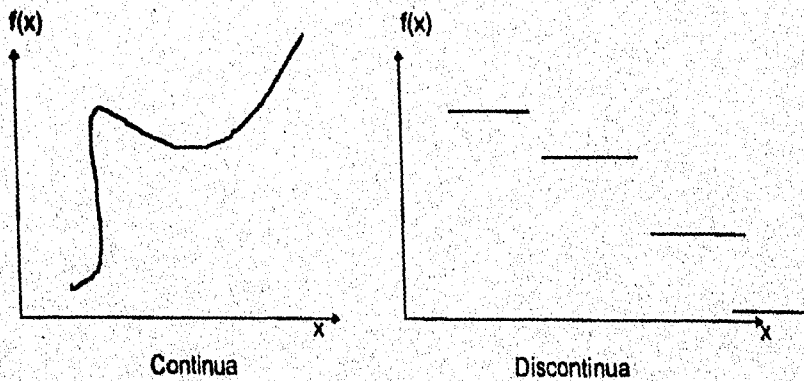


Figura 2.3.10. Funciones cuasi-convexa

El problema general de programación no lineal, es como se expresó en la sección (II.4), tiene restricciones de igualdad, por lo general la región de factibilidad no es convexa. Además, los problemas reales de optimización no lineal, generan funciones objetivas que no son convexas. A raíz de todo esto surge, las condiciones necesarias y suficientes de que deben de existir en los puntos óptimos locales de problemas más generales.

Condiciones de Kuhn-Tucker

En 1950, Kuhn-Tucker descubrió la siguiente teoría, que ha servido de base para el desarrollo de la programación matemática y particularmente para la programación no lineal. Para lo cual considera el siguiente problema de programación no lineal o programación matemática.

$$\begin{array}{ll} \text{Maximizar} & f(X) \\ \text{sujeta a:} & \\ & g_i(x) \leq b_i \quad \text{para } i = 1, 2, \dots, m \\ \text{y} & x \geq 0 \end{array}$$

Donde la relajación Lagrangeano queda como:

$$L(X, \lambda_i) = f(X) + \sum_{i=1}^m \lambda_i g_i(X)$$

y la condición necesaria para que se tenga un óptimo local, es que el gradiente del Lagrangiano sea igual a cero, es decir,

$$\nabla L(X, \lambda) = 0$$

y las variables λ , son los multiplicadores de Lagrange. Si $f(X)$ es convexo, entonces $\nabla L(X, \lambda) = 0$ es una condición necesarias y suficientes y el óptimo es global. Aunque en teoría, el procedimiento es muy sencillo de entender, en la

práctica por lo general es difícil, debido a la solución del sistema de ecuaciones no lineales, generadas por el gradiente de Lagrangeanos igualado a cero. Sin embargo, el método de Lagrange que convierte a un problema restringido en uno no restringido es ya un mejoramiento del proceso de sustitución.

La representación lógica de las condiciones de Kuhn - Tucker establece que, el mínimo local de un problema de optimización restringido, cualquier cambio que sufran las variables del problema, no harán cambiar el valor de la función objetivo.

Representación matemática de las condiciones necesarias de Kuhn-Tucker

Se ha visto que dado un problema de optimización como:

$$\begin{aligned} &\text{Minimizar} && f(X) \\ &\text{sujeto a:} && \\ &&& g_i(x) \leq 0 \quad i = 1, 2, \dots, m \quad (39) \\ &&& X \in E^n, \end{aligned}$$

en donde $f(X)$ y las $g_i(x)$ son diferenciables, las condiciones de Kuhn-Tucker establecen que, un punto óptimo local X^0 debe necesariamente satisfacer:

$$\begin{aligned} \nabla f(X^0) &= \sum_{i=1}^m u_i (-\nabla g_i(X^0)) \\ u_i &\geq 0, && i = 1, \dots, m \\ g_i(X^0) &\leq 0, && i = 1, \dots, m \\ u_i g_i(X^0) &= 0, && i = 1, \dots, m \end{aligned}$$

Se considera a continuación pequeños desplazamientos en el contorno del óptimo local X^0 . Estos desplazamientos se representan de la siguiente forma $X^0 + Y$, ahora sea

$$I = \{i / g_i(X^0) = 0\},$$

El conjunto de restricciones activas en X^0 , y si (restricciones no activas). $j \notin I$ significa que $g_j(X^0) < 0$.

Si se consideran sólo aquellos desplazamientos de la forma $X^0 + Y$, tal que

$$g_i(X^0 + Y) \leq 0, \quad i \in I$$

o sea, desplazamientos que pueden convertir a una restricción activa en una no activa. Por la expansión de Taylor se tiene:

$$g_i(X^0 + Y) \cong g_i(X^0) + \nabla g_i^T(X^0)Y + \frac{1}{2}Y^T H_i(\theta X^0(1 - \theta)(X^0 + Y))Y \leq 0$$

para $i \in I$ y $0 \leq \theta \leq 1$. Para desplazamientos pequeños admisibles, o sea para valores de Y pequeños, el término $\frac{1}{2}Y^T H_i(\cdot)Y$ es tan insignificante, que para fines prácticos se puede eliminar. Además para $i \in I$, se tiene $g_i(X^0) \leq 0$. Por lo tanto

$$g_i(X^0 + Y) \cong g_i(X^0) \leq 0, \quad i \in I \quad (40)$$

Entonces, si Y es un desplazamiento admisible debe satisfacer la ecuación anterior. Condición que permite derivar una clase de problemas no lineales, para los que las condiciones de Kuhn-Tucker son necesarias para un mínimo local. Para poder establecer las condiciones necesarias y suficientes se requiere definir las llamadas restricciones calificadas de Kuhn-Tucker. Se requiere las siguientes definiciones.

Se entiende por un arco, denotado $\alpha(\theta)$ en E^n , a un vector con n funciones de la forma:

$$\alpha(\theta) = (\alpha_1(\theta), \alpha_2(\theta), \dots, \alpha_n(\theta))$$

donde $0 \leq \theta \leq \delta$. El arco es diferenciable si cada componente del mismo es diferenciable. El vector tangente es un arco que se denota por $D_{\alpha(\theta)}$ y se define como un vector con las siguientes componentes:

$$D_{\alpha(\theta)} = \left(\frac{d\alpha_1(\theta)}{d\theta}, \frac{d\alpha_2(\theta)}{d\theta}, \dots, \frac{d\alpha_n(\theta)}{d\theta} \right)$$

Se dice que un arco se origina en un punto X^0 si $\alpha(0) = X^0$. Se dice un arco es tangente a un vector Y en un punto θ , si $D_{\alpha(\theta)} = Y$. Un arco que contiene en un conjunto R , $R \subseteq E^n$, $\alpha(\theta) \in R$, $\forall \theta$, $0 \leq \theta \leq \delta$.

Región calificada de Kuhn-Tucker. Sea X^0 un punto que está en la región de factibilidad (39), es decir, $g_i(X^0) \leq 0$, $i = 1, \dots, m$. Se supone que tanto $f(X)$ como $g_i(X)$, $i = 1, \dots, m$ son diferenciables. La restricción calificada de kuhn-Tucker se cumple en el punto X^0 , si cualquier desplazamiento en la vecindad $(X^0 + Y)$, satisface:

- a) $\nabla f(X^0 + Y) \leq 0$, para todo $i \in I$,
- b) es tangente a un arco diferenciable que se origina en X^0 ,
- c) está contenido en la región de factibilidad.

Y las siguientes condiciones de Kuhn-Tucker quedan formuladas con el siguiente teorema:

Teorema 9: El teorema de Kuhn-Tucker establece que bajo la suposiciones siguientes:

a) Todas las funciones $f(X)$ en el problema (39) son diferenciables,

b) X^0 es un óptimo local o relativo del problema (39)

c) La restricción calificada de Kuhn-Tucker se cumple en el punto X^0 .

Entonces existe un vector $u^0 = (u_1^0, u_2^0, \dots, u_m^0) \geq 0$, tal que las siguientes condiciones (llamadas de kuhn-Tucker) se cumplen en el punto X^0 :

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

$$u_i \geq 0 \quad i = 1, \dots, m$$

$$g_i(X^0) \leq 0 \quad i = 1, \dots, m$$

$$u_i g_i(X^0) = 0 \quad i = 1, \dots, m$$

Puntos de silla y las condiciones suficientes de Kuhn-Tucker

Hasta la sección anterior se ha formulado las condiciones necesarias para un óptimo local, en el supuesto de que la función objetivo y las restricciones sean diferenciables. Existen ciertas condiciones que son necesarias y suficientes para que un punto sea un óptimo local, aún cuando las funciones no sean diferenciables. Estas condiciones se conocen como de silla. Para esto considere el siguiente problema de programación no lineal:

Minimizar $f(X)$.

sujeto a:

$$g_i(X) \leq 0, \quad i = 1, \dots, m$$

$$X \in S, \quad (41)$$

donde $f(X)$ como $g_i(X)$, $i = 1, \dots, m$ son funciones arbitrarias no lineal (no necesariamente diferenciables) y S es un subconjunto propio de E^n . El Lagrangeano asociado a (41),

$$\text{Mín } L(X,U) = f(X) + \sum_{i=1}^m u_i g_i(X)$$

$$u_i \geq 0.$$

Se define a un punto de silla como que cumple con las siguientes condiciones:

$$\begin{aligned} L(X^0, U^0) &\leq L(X, U^0), \text{ para todo } X \in S \\ L(X^0, U^0) &\geq L(X^0, U), \text{ para todo } U \geq 0 \end{aligned} \quad (42)$$

donde $X^0 \in S$ y $U^0 \geq 0$. En (42) en punto X^0 minimiza al Lagrangeano $L(X, U)$ en la región S y U^0 minimiza el Lagrangeano $L(X, U)$ para todo $U \geq 0$.

El siguiente teorema expresa las condiciones necesarias y suficientes de un punto de silla del Lagrangeano $L(X, U)$.

Teorema 10: Sea $U^0 \geq 0$ y $X^0 \in S$, $S \subset E^n$. Entonces (X^0, U^0) es un punto de silla de la función $L(X, U)$, si y sólo si,

- a) X^0 minimiza $L(X, U)$ en S .
- b) $g_i(X^0) \leq 0$, para $i = 1, 2, 3, \dots, m$.
- c) $u_i g_i(X^0) = 0$, para $i = 1, 2, 3, \dots, m$.

Teorema 11: Si (X^0, U^0) es un punto de silla de $L(X, U)$, entonces el punto X^0 es un mínimo del problema (41).

El teorema 11 es aplicable a todo tipo de optimización, incluyendo aquellos en donde la función objetivo y las restricciones no son diferenciables. Puede ocurrir que los puntos de silla no existan para varios de estos programas. Sin embargo se puede garantizar la existencia de los puntos de silla en todo los programas convexos (Karlin).

Resumen

Se supone que $f(X)$, $g_i(X)$, $i = 1, \dots, m$ son diferenciables, $X \in S \subseteq E^n$ y se satisface la restricción de Kuhn-Tucker. El punto $X^0 = (X_1^0, X_2^0, \dots, X_n^0)$ es un punto óptimo local del problema

$$\begin{array}{ll} \text{Min} & f(X) \\ \text{Sujeto a} & \\ & g_i(X) \leq 0 \quad i = 1, \dots, m \\ & X \in S \subseteq E^n \end{array}$$

si existe m números u_1, u_2, \dots, u_m que cumple con las siguientes condiciones:

- Si $X_j^0 = 0$, entonces $\frac{\partial f}{\partial X_j} + \sum_{i=1}^m u_i \frac{\partial g_i}{\partial X_j} = 0$ para $X_j = X_j^0$, $j = 1, \dots, n$
- Si $X_j^0 > 0$, entonces $\frac{\partial f}{\partial X_j} + \sum_{i=1}^m u_i \frac{\partial g_i}{\partial X_j} \geq 0$ para $X_j = X_j^0$, $j = 1, \dots, n$
- Si $u_i > 0$, entonces $g_i(X) = 0$, para $i = 1, \dots, m$.
- Si $u_i = 0$, entonces $g_i(X) \leq 0$, para $i = 1, \dots, m$.
- $X_j^0 = 0$, para $j = 1, \dots, n$
- $u_i \geq 0$ para $i = 1, \dots, m$.

Si $f(X)$, $g_i(X)$, $i = 1, \dots, m$ son funciones convexas, que satisfacen la restricción calificada de kuhn-Tucker, las condiciones a) a la f) inclusive, son suficientes para obtener un óptimo global o absoluto.

Observaciones

- Las condiciones de Kuhn-Tucker son completamente imprácticas para resolver problemas de optimización.

b) Las condiciones de Kuhn-Tucker sienta las bases para el desarrollo de métodos prácticos de optimización, mucho más eficientes.

c) Las condiciones de Kuhn-Tucker permiten probar cuándo un punto (X^0) , o punto de silla (X^0, u) es óptimo (local o global).

a) Métodos de optimización no lineal basados en la aproximación lineal

La lógica que utilizan este tipo de métodos, consiste en aproximar el problema lineal

$$\begin{aligned} & \text{Mín. } f(X) \\ \text{Sujeto a:} & \\ & h_i(X) = 0, \quad i = 1, \dots, m \\ & g_i(X) \geq 0, \quad i = m+1, \dots, p \\ & X \in E^n, \end{aligned} \quad (43)$$

Por el siguiente problema lineal, que proviene de la aproximación de las series de Taylor:

$$\begin{aligned} & \text{Mín } f(X^k) + \nabla f^T(X^k)(X - X^k) \\ \text{Sujeto a:} & \\ & h_i(X^k) + \nabla h_i^T(X^k)(X - X^k) = 0, \quad i = 1, \dots, m \\ & g_i(X^k) + \nabla g_i^T(X^k)(X - X^k) \geq 0 \quad i = m+1, \dots, p \\ & X, \quad X^k \in E^n, \end{aligned}$$

donde $f(X^k), \nabla f(X^k), h_i(X^k), \nabla h_i(X^k), g_i(X^k), \nabla g_i(X^k)$ son escalares evaluados en el punto X^k . Los métodos de aproximación lineal resuelven una serie de programas lineales, los cuales producen un nuevo punto, una vez que se le ajusta

y corrige, genera un nuevo programa lineal, el cual produce un nuevo punto y así sucesivamente.

Los métodos de aproximación lineal tienen la convergencia garantizada un óptimo local en los siguientes casos:

- a) $f(X), h_i(X), i = 1, \dots, m, g_j(X), j = m+1, \dots, p$, son funciones continuas y diferenciables
- b) $f(X)$ es una función convexa, $\sum h_j^2(X)$ genera una región de factibilidad que es convexa.
- c) $g_j(X) \geq 0, j = m+1, \dots, p$ son funciones cóncavas.
- d) La región de factibilidad es convexa y cerrada.
- e) Las regiones están acotadas.

Existen muchos métodos basados en el principio de programación lineal, unos son más eficientes que otros, aunque solamente se tratarán a los tres más importantes, como son:

- a) Método de Griffith y Stewart
- b) Método de Wolfe y,
- c) Método de direcciones factibles de Zoutendijk

Método de Griffith y Stewart

Este método, basado en el principio anterior, procede de la siguiente manera. Dado un punto X^k que se supone es factible, se aproxima el problema no lineal por el siguiente problema lineal:

$$\text{Min. } f(X) = \sum_{j=1}^n \frac{\partial f(X^k)}{\partial X_j} (X_j - X_j^k) + f(X^k)$$

sujeto a:

$$h_i(X^k) + \sum_{j=1}^n \frac{\partial h_i(X^k)}{\partial X_j} (X_j - X_j^k) = 0, \quad i = 1, \dots, m$$

$$g_i(X^k) + \sum_{j=1}^n \frac{\partial g_i(X^k)}{\partial X_j} (X_j - X_j^k) \geq 0 \quad i = m+1, \dots, p.$$

Y con el objeto de garantizar que el nuevo punto sea factible, se aumenta la siguiente restricción:

$$|X_j^{k+1} - X_j^k| \leq \delta_j^k, \quad j = 1, \dots, n$$

donde $\delta_j^k > 0$, $j = 1, \dots, n$, es un número pequeño pero arbitrario. Y el nuevo punto X^{k+1} estará dado por:

$$X^{k+1} = X^k + (X^* - X^k)$$

donde, X^* es la solución del problema lineal en la K -ésima iteración. Conocido X^{k+1} se repite de la misma operación, únicamente corrigiendo el valor del parámetro δ_j^{k+1} de tal manera que, $0 \leq \delta_j^{k+1} < \delta_j^k$, $j = 1, \dots, n$.

La desventaja de esta familia de métodos es el siguiente:

a) La convergencia a la solución óptima puede resultar bastante lenta, si los valores de δ^k , $k = 1, \dots, n$ genera puntos X^{k-1}, X^k, X^{k+1} , que están cercanos a la frontera de la región de factibilidad.

b) Es difícil determinar cual de los puntos sucesivos X^k y X^{k+1} es mejor.

c) La solución del problema lineal puede generar direcciones de solución muy ineficientes. Esto significa que las restricciones no lineales, generan valles profundos o contornos elongados, y el movimiento al óptimo local será extremadamente lento.

d) Es difícil diseñar métodos de aceleración.

Método de Wolfe

El método de Wolfe se define como un problema de programación cuadrática que es el siguiente:

$$\text{Min. } f(X) = c^T X + \frac{1}{2} X^T Q X \quad (44)$$

Sujeto a:

$$AX \leq b$$

$$X \geq 0,$$

donde:

$$X \in E^n$$

c es un vector de precios con n componentes

Q es una matriz de n por n , simétrica y positiva definida, es decir,

$$X^T Q X > 0 \text{ para todo } X \in E^n \text{ excepto } X = 0,$$

b es un vector de disponibilidad de recursos con m componentes

A es una matriz de m por n de coeficientes tecnológicos y

0 es un vector de n ceros.

Como se observa, el problema de optimización tiene como restricciones lineales. Si Q es igual a una matriz nula, el problema anterior se convierte en uno de programación lineal. Como Q es positiva definida, eso implica (teorema 7) que $f(X)$ es una función estrictamente convexa y por lo tanto, el mínimo, si existe, es global. Si Q es negativa definida, $f(X)$ es estrictamente cóncava, y el máximo, si existe, es global.

Y las condiciones anteriores se pueden escribir en forma algebraica de la siguiente forma:

$$\text{Min. } f(X) = \sum_{j=1}^n c_j X_j + \frac{1}{2} \sum_{j=1}^n \sum_{i=1}^m X_j q_{ij} X_i$$

sujeto a: (45)

$$g_i(X) = \sum_{j=1}^n a_{ij} X_j - b_i \leq 0, \quad i = 1, \dots, m$$

$$h_j(X) = -X_j \leq 0, \quad j = 1, \dots, n.$$

Con este nuevo método se genera el Lagrangeano

$$\begin{aligned} L(X, \lambda, u) &= f(X) + \sum_{i=1}^m \lambda_i g_i(X) + \sum_{j=1}^n u_j h_j(X) \\ &= \sum_{j=1}^n c_j X_j + \frac{1}{2} \sum_{j=1}^n \sum_{i=1}^m X_j q_{ij} X_i + \sum_{i=1}^m \lambda_i \left(\sum_{j=1}^n a_{ij} X_j - b_i \right) + \sum_{j=1}^n u_j (-X_j). \end{aligned}$$

Donde las condiciones de Kuhn-Tucker en el Lagrangeano anterior, proporcionan las condiciones necesarias y suficientes que debe existir en un óptimo global.

Wolfe sugiere resolver el siguiente problema:

$$\text{Min. } \sum_{j=1}^n V_j \quad (46)$$

sujeto a:

$$c_j + \sum_{i=1}^m q_{ij} X_j + \sum_{i=1}^m \lambda_i a_{ij} - u_j - V_j = 0 \quad j = 1, \dots, n$$

$$\sum_{j=1}^n a_{ij} X_j + Y_i = b_i \quad i = 1, \dots, m$$

$$X_j \geq 0, u_j \geq 0, V_j \geq 0, \lambda_i \geq 0,$$

$$j = 1, \dots, n, \quad i = 1, \dots, m$$

$$Y_i \text{ no restringida en signo } i = 1, \dots, m \quad (47)$$

$$\lambda_i Y_i = 0 \quad i = 1, \dots, m$$

$$u_j X_j = 0, \quad j = 1, \dots, n$$

Las condiciones (47) se le llama holgura complementaria. El proceso de Wolfe empieza con una solución básica factible al programa lineal (46) mediante el método simplex. Al seguir iterando con el método simplex, se toman en cuenta las condiciones de holgura complementaria en el sentido de que:

1) λ_i no puede entrar a la base si $Y_i > 0$ solamente entrará a la base si $Y_i = 0$, $i = 1, \dots, m$. De manera análoga, Y_i no entra a la base si $\lambda_i > 0$ y solamente lo podrá hacer cuando $\lambda_i = 0$ $i = 1, \dots, m$.

2) Se usa lo mismo en este caso como 1) para X_i y u_i , $i = 1, \dots, m$.

Concluyendo: el programa no lineal a resolver consta de $2(m+n)+n$ variables, $m+n$ restricciones lineales y $m+n$ restricciones de holgura complementaria.

Método de direcciones factibles de Zoutendijk

Este método fue desarrollado por Zoutendijk, sirve para resolver problemas de optimización de la forma:

$$\begin{aligned} & \text{Min. } f(X) \\ & \text{Sujeto a:} \end{aligned} \quad (48) \quad g_i(X) \leq 0, \quad i = 1, \dots, m$$

donde $f(X)$ y $g_i(X) = 0$, $i = 1, \dots, m$, deben ser diferenciables. Este método no puede resolver problemas donde existan restricciones de igualdad. En dicho método se encuentra en cada iteración una dirección y un tamaño de avance en esa dirección, la cual tiene dos propiedades:

a) Una llamada de factibilidad, consiste en que un avance en esa dirección no viola ninguna restricción y

b) Otra llamada de utilidad, consiste en que un avance en esa dirección mejora el valor de la función objetivo.

La generalización del problema original utilizando las series de Taylor, genera un mínimo, como se muestra en el siguiente caso:

$$\text{Mín. } f(X^k) + \nabla f^T(X^k)(X - X^k)$$

Sujeto a:

$$g_i(X^k) + \nabla^T g_i(X^k)(X - X^k) \leq 0 \quad i = 1, \dots, m$$

El nuevo punto X^{k+1} se obtiene del punto anterior X^k de acuerdo con

$$X^{k+1} = X^k + \lambda_k s_k$$

donde s_k es la dirección que debe ser factible y útil. La factibilidad de la dirección se obtiene de la siguiente forma:

$$\nabla^T g_i(X) s_k \leq 0, \quad i = 1, \dots, m \quad (49)$$

mientras que la utilidad de la misma, es decir, la propiedad que mejore a la función objetivo se obtiene como sigue:

$$\left. \frac{df(X^k + \lambda_k s_k)}{d\lambda_k} \right|_{\lambda_k=0} = \nabla^T f(X^k) s_k < 0, \quad (50)$$

El tamaño de movimiento λ_k a lo largo de la dirección s_k , se encuentra por métodos de optimización de funciones no restringidas de una sola variable (Interpolación cúbica, cuadrada, Fibonacci, sección de oro, Newton-Raphson, etc.).

Los pasos a seguir en el método de direcciones factibles de Zoutendijk en la iteración k , son los siguientes:

Etapas del algoritmo

Paso 1: Sea X^k un punto factible de (48), es decir, $g_i(X^k) \leq 0 \quad i = 1, \dots, m$

Paso 2: Se calcula $\nabla f(X^k)$.

Paso 3: Se resuelve el problema lineal

$$\begin{aligned} & \text{Mín. } \nabla^T f(X^k) s_k \\ & \text{Sujeto a:} \\ & \quad \nabla^T g_i(X^k) s_k \leq b_i \quad i = 1, \dots, m \end{aligned}$$

para determinar una dirección s_k que es factible y útil. El vector b está formado por las componentes de las restricciones.

Si el problema a optimizar es de maximización, la dirección es útil, si $\nabla^T f(X^k) s_k > 0$ y si es un problema de minimización, entonces, ir al paso 4.

Paso 4: Si $\nabla^T f(X^k) s_k < 0$ determine la longitud λ_k de recorrido a lo largo de s_k . Este parámetro se encuentra en el intervalo $0 \leq \lambda \leq \lambda_k^*$, donde λ_k^* se calcula minimizando

$$f((X^k)^T + \lambda_k^* s_k^T).$$

Como $f((X^k)^T + \lambda_k^* s_k^T)$ es una función no restringida de una sola variable (λ_k^*), utilice cualquier método explicado anteriormente (Fibonacci, Sección de oro). Si $\nabla^T f(X^k) s_k = 0$, el proceso termina, ya que no se puede mejorar a la función objetivo. El punto X^k es el punto local.

Paso 5: El nuevo punto X^{k+1} es $X^{k+1} = X^k + \lambda_k s_k$.

Se regresa al paso 2.

Los métodos difieren entre sí por la manera como se calcula la dirección s_k y la longitud λ_k .

b) Métodos penales

Estos métodos tienen como principio transformar a un problema de optimización restringido en uno no restringido, incorporando en cierta forma las restricciones en la función objetivo. Se les llaman métodos penales, porque utilizan una función con la que penalizan la función objetivo, si el punto considerado se sale fuera de la región de factibilidad. Los métodos penales utilizan técnicas de optimización no restringidas.

Los métodos penales se dividen en dos clases:

- a) Métodos penales paramétricos, y
- b) Métodos penales no paramétricos

Los métodos penales paramétricos a su vez se divide en tres partes

- a) Método de punto interior,
- b) Método de punto exterior, y
- c) Métodos mixtos.

En el método penal de punto interior se sigue una trayectoria que empieza en un punto interior factible X^k , que converge al punto óptimo local, pero sin salirse de la región de factibilidad. En el método de punto exterior, el mecanismo es análogo al interior, pero siguiendo siempre una trayectoria desde el exterior, al punto óptimo. El último caso, la función penal evita que el punto se convierta en no factible se aleje demasiado de la región factible. El método mixto se utiliza generalmente cuando se tienen restricciones de igualdad. Es una combinación de las dos estrategias aplicadas anteriormente.

Dado el problema no lineal.

$$\text{Min. } f(X)$$

Sujeto a:

$$\begin{aligned} h_i(X) &= 0, & i &= 1, \dots, m \\ g_i(X) &\geq 0, & i &= m+1, \dots, p \\ X &\in E^n, \end{aligned}$$

Se puede representar al método penal por:

$$P(X^k, p^k) = f(X^k) + \sum_{i=1}^m p_i^k H(h_i(X^k)) + \sum_{i=m+1}^p p_i^k G(g_i(X^k)) \quad (51)$$

donde $P(X^k, p^k)$ es la función penal, p_i^k son parámetros ponderados, $H(h_i(X^k)), G(g_i(X^k))$ son funciones de las restricciones $H(h_i(X^k))$ y $G(g_i(X^k))$ y k es el número de iteraciones del método.

Los métodos penales que existe en la práctica, varían en la forma como construyen sus funciones y evalúan los parámetros de ponderación

En este caso mencionaremos el método de SUMT⁸ de Fiacco y McCormick, es aplicable a problemas de optimización no lineales donde la función objetivo $f(X)$, y las restricciones de desigualdad $g_j(X)$, pueden ser no lineales, pero las restricciones de igualdad $h_j(X)$ deben ser lineales. Esta condición debe existir, si es que se quiere garantizar la convergencia a un óptimo local.

La secuencia de funciones penales a resolver en el método SUMT es:

$$P(X^k, p^k) = f(X^k) + (\rho^k)^{\frac{1}{2}} \sum_{i=1}^m h_i^2(X^k) + p^k \sum_{i=m+1}^p \frac{1}{g_i(X^k)} \quad (52)$$

donde los factores de ponderación ρ son positivas monótonamente decreciente, es decir, $\rho^0 > \rho^1 > \dots > \rho^k > 0$.

⁸ Sequential Unconstrained Minimization Technique.

A medida de que el valor de ρ decrece en cada iteración, el efecto de las funciones es el siguiente:

$$H(h_i(X^k)) = \sum_{i=1}^m h_i^2(X^k),$$

$$G(g_i(X^k)) = \sum_{i=m+1}^p \frac{1}{g_i(X^k)}$$

es acercar el punto X^k a las fronteras de la región de factibilidad.

A continuación se explica como se calculan los factores de integración ρ^k en cada iteración k , en el proceso SUMT empieza con un punto interior conocido X^0 , por lo que todas las restricciones de desigualdad se satisfacen, aunque no sucede así en la restricción de igualdad, que por lo general se viola. Se tiene por lo tanto

$$g_i(X) \geq 0, \quad i = m+1, \dots, p$$

Después de calcular ρ^0 , el punto X^1 se determina al minimización la función no restringida $P(X, \rho^0)$, dada en (52). Se calcula el parámetro ρ^1 , y el punto X^2 se calcula después de la minimización de $P(X, \rho^1)$, y así sucesivamente.

Fiacco y McCormick recomienda tres métodos para calcular el valor de ρ^0 , a saber:

1) $\rho^0 = 1$. Es el más práctico de todos, pero el que proporciona resultados menos exactos.

2) $\rho^0 = \frac{-\nabla^T f(X^0) \nabla R(X^0)}{\|\nabla R(X^0)\|^2}$, Donde X^0 es un punto factible y

$$R(X^0) = \sum_{i=m+1}^p \frac{1}{g_i(X^0)}$$

Por $\|a\|^2$ se entiende, el cuadrado del determinante de a . Este método proporciona resultados más exacto que el primero, pero es menos práctico y se usa cuando no se puede evaluar el Hessiano de R .

$$3) \rho^0 = \left\{ \frac{\nabla^T f(X^0) [\nabla^2 R(X^0)]^{-1} \nabla^T R(X^0)}{\nabla R(X^0) [\nabla^2 R(X^0)]^{-1} \nabla^T R(X^0)} \right\}^{\frac{1}{2}}$$

donde $\nabla R(X^0)$ es el gradiente de $R(X^0)$, $\nabla^2 R(X^0)$ es el Hessiano de $R(X^0)$, $[\nabla^2 R(X^0)]^{-1}$ es la matriz inversa del Hessiano, X^0 es un punto interior y

$$R(X^0) = \sum_{i=m+1}^p \frac{1}{g_i(X^0)}$$

Este es el método que genera mejores resultados, pero el que lleva más tiempo de cálculo en una computadora.

Una vez obtenido ρ^0 , Fiacco y McCormick recomiendan que ρ^k se calcula de la siguiente forma:

$$\rho^k = \frac{\rho^{k-1}}{4}$$

Fiacco y McCormick dan también tres criterios para parar el algoritmo. Aquí se menciona sólo uno de ellos, que en el siguiente:

$\nabla^T P(X^k, \rho^k) [\nabla^2 P(X^k, \rho^k)]^{-1} \nabla^T P(X^k, \rho^k) < \varepsilon$ donde $\varepsilon > 0$ es una tolerancia arbitraria.

III APLICACIÓN

III.1 Problema de localización de servicios

El problema de localización de servicios busca encontrar la solución geográfica de una instalación, en tal forma que los costos de distribución a un cierto número de clientes sean minimizados.

Por instalación, se entiende cualquier servicio público (hospitales, estaciones de bomberos o de policía, bancos, oficinas gubernamentales, escuelas, bibliotecas, estadios deportivos, plantas generadoras de energía eléctrica, estaciones de gasolina, plantas de tratamiento de basura, aeropuertos, etc.) o privado (bodegas, plantas industriales, comercios, etc.).

Por distribución de la demanda de los servicios se entiende la manera de asignar espacio físico a los diversos componentes de una instalación o; por ejemplo, máquinas y herramientas en una planta. Las coordenadas de un plano bidimensional nos auxilian en la localización.

En esta sección se formulan modelos de localización de varias instalaciones en el espacio continuo; en el primer caso se analiza en un plano bidimensional las n ($n > 0$) instalaciones que se desea establecer.

Para analizar nuevas instalaciones, en todos los modelos de optimización se considera una función objetivo de costos, la cual se minimiza. Dicha función es representativa de la distancia y/o el tiempo necesario para fluir bienes o servicios de las nuevas instalaciones y las ya existentes, a los clientes.

En este trabajo la parte más importante a tratar son: la norma rectangular y la norma euclidiana, es decir, se analizan dos clases de normas para medir distancias. La llamada norma rectangular o Manhattan, considera que la distancia entre dos puntos no es la recta que los une, sino el mínimo número de calles que se debe recorrer, a este tipo de modelo se utiliza en grandes zonas urbanas,

cuyas calles tienen trazos rectos paralelos y perpendiculares. El otro tipo de norma, llamada euclidiana, considera que la distancia más corta entre dos puntos es la recta que los une; se utiliza en problemas de localización en zonas rurales y urbanas de trazo irregular.

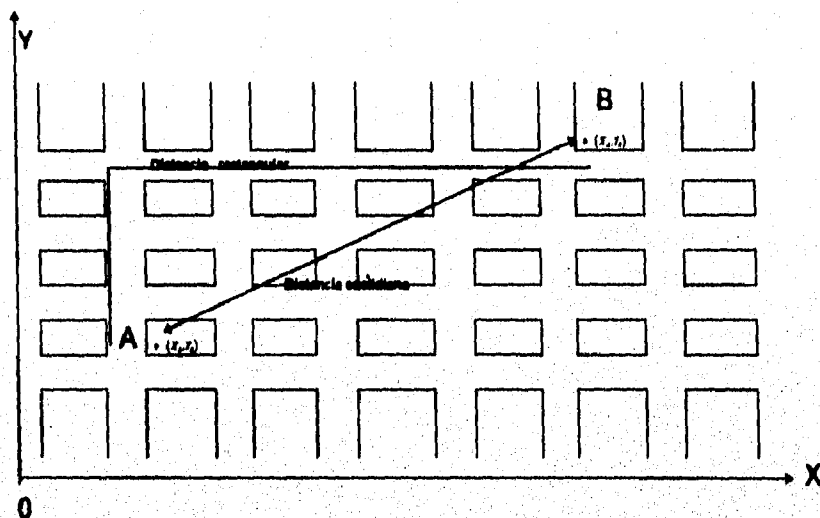
III.2 Conceptualización del problema

Los problemas de localización se presenta cuando los encargados de tomar decisiones deben seleccionar el sitio en que ubicarán una o varias instalaciones o servicios, como podría ser: industrias, bodegas, comercios, escuelas, hospitales, mercados, aeropuertos, plantas de tratamiento de aguas, plantas de generación de electricidad (hidroeléctricas, térmicas, nucleares), plantas de tratamiento de basura, estudios deportivos, estaciones de bomberos, estaciones de gasolina, etc.; este tipo de problemas se presenta también en la distribución de maquinaria en un área dada. La toma de decisiones anteriores se considera bajo una serie de criterios preestablecidos.

Los modelos cuantitativos analíticos que se presentan en este capítulo son de tipo normativo, es decir, prescriben un curso de acción que optimiza una función dada. Estos modelos tienen varias limitaciones; la primera de ellas es la definición de la función objetivo, que generalmente localiza el sitio o sitios, minimizando una función de costo. Vollman y Buffa establece que minimizar ciegamente una función de costos puede conducir a resultados absurdos. Para ello se ha llegado a considerar una variedad de funciones objetivos, que incluye la minimización de la máxima distancia recorrida (funciones minimax). este tipo de funciones es característica de los problemas de localización de servicios de emergencia (hospitales, estaciones de bombero, etc.), donde se desea que una comunidad no esté a más de cierto tiempo de una clínica.

Para medir distancias se puede lograr con la norma rectangular, o bien una euclidiana. La primera tiene mayor aplicación en ciudades grandes, con trazos rectos perpendiculares paralelos de calles y avenidas, donde la distancia entre

dos puntos no puede medirse como la recta que los une, sino como el mínimo número de calles que existe entre ambos. En contrapartida la norma euclidiana dice que la distancia entre dos puntos, es la recta que los une. Esta norma tiene sentido en las zonas rurales y urbanas con trazo irregular. En la gráfica 3.1 se distinguen ambas normas.



Gráfica 3.1

Norma rectangular $d_{AB} = |X_A - X_B| + |Y_A - Y_B|$

Norma euclidiana $d_{AB} = \left[(X_A - X_B)^2 + (Y_A - Y_B)^2 \right]^{1/2}$

Los problemas de localización se pueden dividir para su estudio en relación a: 1). Lo que se quiere localizar: en problemas de distribución de espacio y problemas de localización, aunque en esta sección se tratan solamente problemas de localización, 2). Las características de las nuevas instalaciones: en problemas de localización sencilla (una instalación) o múltiple (varias instalaciones), localización de punto o de área, donde el número de las nuevas instalaciones está dado o es una variable adicional de decisión y donde la ubicación es independiente o dependiente de las otras localizaciones, 3). Las características de las instalaciones existentes: en problemas de localización

estática o dinámica, determinística o probabilística, 4). La interacción de las diversas instalaciones: en problemas cualitativos o cuantitativos, 5). El espacio: en problemas unidimensionales o multidimensionales, discretos o continuos, restringidos o no restringidos, 6). La función objetivo: en problemas cuantitativos o cualitativos y, dentro de los primeros, problemas donde se minimizan funciones de costo y tiempo o se minimiza funciones de tipo minimax, 7). La norma (distancia) en problemas rectangulares y euclidianos.

III.3 Localización de un solo servicio

En este apartado se analiza el problema de la localización de una instalación nueva respecto de una serie de instalaciones similares existentes, cuando el criterio a seguir es el de minimizar una función total de costo que sea proporcional a la distancia.

Formulación general del problema de un solo servicio

La formulación general del problema de localización de un solo servicio, se puede plantear de la siguiente forma: existen m instalaciones conocidas en los puntos p_1, p_2, \dots, p_m de un sistema o servicio de coordenadas, y se requiere una nueva instalación, de características similares en un punto desconocido x , tal que los costos de transporte que son proporcionales a la distancia entre la nueva instalación y las existentes, $d(X, p_i)$, $i=1, 2, \dots, m$ se minimicen. Sea W_i una ponderación asociada a la instalación existente, p_i , $i=1, 2, \dots, m$ que mide el producto del costo de transporte anual y el número de viajes anuales entre el producto i y el x .

Matemáticamente se quiere encontrar x que minimice a:

$$f(X) = \sum_{i=1}^m W_i d(X, p_i) \quad (1)$$

donde el término W_i es a veces referido como prioridades o pesos.

El problema consiste en determinar la localización del nuevo servicio X que minimice $f(X)$, o sea, el costo total de transporte. Dimensionalmente $f(X)$ está expresada en \$/año de acuerdo al siguiente análisis

$$f(X) = [W_i][d(X, p_i)] = \left[\frac{\$}{Km} \cdot \frac{\text{viajes}}{\text{año}} \right] \left[\frac{Km}{\text{viaje}} \right] = \frac{\$}{\text{año}}$$

En muchas aplicaciones el costo por unidad de distancia es constante y el problema se reduce a determinar la localización que minimice la distancia. También la cantidad W_i se puede expresar como: $W_i = c_i d_i$, donde c_i es un costo unitario / unidad de distancia y d_i es la demanda de flujo.

Se analiza, en primer término, el caso de una norma rectangular y, después el de una euclidiana. Para la norma rectangular el problema consiste en encontrar las coordenadas (x, y) que minimizan a $f(X)$.

Norma rectangular

Los problemas, en donde los viajes ocurren a lo largo de un conjunto de naves arregladas en un patrón o molde rectangular, paralelos a las paredes de la construcción, la media de la distancia apropiada es la rectangular.

El problema que se analiza es de una sola instalación nueva, respecto de una serie de instalaciones similares existentes. Si las coordenadas (X, Y) para el

centro de demanda i son (a_i, b_i) , tal que, $X = (x, y)$ y $p_i = (a_i, b_i)$, la distancia rectangular entre x y p_i queda definida por:

$$d(X, p_i) = |x - a_i| + |y - b_i|$$

Como ya se mencionó anteriormente la distancia rectangular es apropiada para el análisis urbano, donde los recorridos ocurren en un conjunto ortogonal de calles. Además, en algunas oficinas ocurren en un conjunto de alas, naves laterales o caminos paralelos a las paredes dentro de los edificios, para facilitar los viajes o el paso del personal.

El problema de localización que utiliza la distancia rectangular, puede ser representado matemáticamente por:

$$\text{Min}_{x,y} f(X) = \sum_{i=1}^m W_i (|x - a_i| + |y - b_i|) \quad (3.2)$$

donde (a_i, b_i) son coordenadas conocidas del punto, p_i , $i = 1, 2, \dots, m$, W_i mide el producto del costo de transporte anual y el número de viajes anuales entre el producto i y el X : (x, y) variables a determinar o coordenadas a determinar donde se debe establecer el nuevo servicio.

De la ecuación anterior, se puede ver que el problema es equivalente a:

$$\text{Min}_{x,y} f(X) = \text{Min}_x \sum_{i=1}^m W_i |x - a_i| + \text{Min}_y \sum_{i=1}^m W_i |y - b_i|$$

que consiste en la solución de dos problemas independientes, donde cada término del lado derecho de la igualdad, puede ser tratado como un problema de optimización por separado, y aplicar métodos de programación lineal para su solución.

$$\text{Min}_x f_1(x, y) = \sum_{i=1}^m W_i (x - a_i)$$

y

$$\text{Min}_y f_2(x, y) = \sum_{i=1}^m W_i (y - b_i)$$

Como $f_1(x)$ y $f_2(x)$ tiene la misma forma, el procedimiento que se aplica para minimizar alguno de ellos, se podrá aplicar para el otro.

Algunas propiedades de una solución óptima para los problemas de localización de un solo servicio con distancia rectangular son:

- La coordenada X del nuevo servicio puede ser la misma que la coordenada X de algún centro de demanda. Similarmente, con la coordenada Y de algún centro de demanda, desde luego, no es necesario que ambas coordenadas estén en el mismo centro de demanda.

- La localización óptima de la coordenada X (coordenada Y) para el nuevo servicio, es una localización media, es decir, una localización tal que, no más de la mitad del número de viajes este a la izquierda o abajo de la localización del nuevo servicio y no más de la mitad del número de viajes, este a la derecha o arriba del nuevo servicio. O sea, que la mitad de los centros de demanda están situados a la izquierda (abajo) y a la derecha (arriba) del punto medio.

Norma euclidiana

En otros problemas de localización de servicios, el costo no es una función lineal de la distancia, por ejemplo: El costo asociado con la respuesta de un camión de bomberos a un incendio, es esperado que no sea lineal con la distancia.

Dependiendo del problema, $d(X, p_i)$ puede tomar diferentes formulaciones. Así que, también es posible utilizar el cuadrado de la distancia euclidiana entre X y p_i , como medida de la misma.

Si (X, Y) son las coordenadas del nuevo servicio y (a_i, b_i) son las coordenadas del centro de servicio i . El cuadrado de la distancia entre X y p_i , queda definida por la demanda:

$$d(X, p_i) = [(x - a_i)^2 + (y - b_i)^2]$$

Las razones para estudiar este problema es fundamentalmente dos: la existencia de problemas de localización en los cuales los costos se incrementan, cuadráticamente, en lugar de linealmente, y el estudio de las leyes de los problemas con el cuadrado de la distancia euclidiana también llamados problemas de importancia, que son fundamentos para los problemas con distancia euclidiana.

El problema de importancia o el cuadrado de la euclidiana puede ser formulado como:

$$\text{minimizar } f(X, p_i) = \sum_{i=1}^n W_i [(x - a_i)^2 + (y - b_i)^2]$$

algún punto (X, Y) que minimiza, debe satisfacer las condiciones

$$\left(\frac{\partial f(X^*, Y^*)}{\partial X^*}, \frac{\partial f(X^*, Y^*)}{\partial Y^*} \right) = (0, 0)$$

Como la función $\text{min} f(X^*, Y^*)$ es cuadrática, entonces las condiciones son necesarias y suficientes para un mínimo.

Obteniendo las derivadas parciales de la función $\text{min} f(X^*, Y^*)$ con respecto a (X, Y) y haciendo igual a cero, se obtiene las siguiente solución única.

$$X^* = \frac{\sum W_i a_i}{\sum W_i}, \quad Y^* = \frac{\sum W_i b_i}{\sum W_i}$$

Las coordenadas del nuevo servicio puede ser interpretadas como valores (pesos) promedios de las coordenadas (X, Y) de los centros de demanda y son

en efecto, las coordenadas que minimizan a la función objetivo. Las soluciones son llamadas a veces, de centro de gravedad o de importancia

Uso de la distancia euclidiana

Considera que las coordenadas para el nuevo servicio son (X, Y) mientras que los centros de demanda i son (a_i, b_i) , $i = 1, \dots, n$, entonces

$$d(X, p_i) = \left[(x - a_i)^2 + (y - b_i)^2 \right]^{\frac{1}{2}}$$

La distancia euclidiana se aplica para algunos problemas de localización de redes, casos complejos de comunicación, transporte, viajes aéreos, alambrado eléctrico, etc.

Entonces el problema de localización con distancia euclidiana se puede formular como:

$$\text{minimizar } f(X, p_i) = \sum_{i=1}^n W_i \left[(x - a_i)^2 + (y - b_i)^2 \right]^{\frac{1}{2}}$$

La aproximación inmediata se analiza para la solución del problema de localización con distancia euclidiana, es nuevamente calcular las derivadas parciales de la ecuación anterior y hacerlas igual a cero, es decir, se toman las derivadas de f respecto de x y y se obtiene

$$\frac{\partial f}{\partial x} = \sum_{i=1}^n \frac{W_i(x - a_i)}{\left[(x - a_i)^2 + (y - b_i)^2 \right]^{\frac{1}{2}}} = 0$$

$$\frac{\partial f}{\partial y} = \sum_{i=1}^n \frac{W_i(y - b_i)}{\left[(x - a_i)^2 + (y - b_i)^2 \right]^{\frac{1}{2}}} = 0$$

(3)

Nótese que la solución de ambas ecuaciones presenta un grave problema cuando $(X, Y) = (a, b)$ para cualquier i , $i = 1, 2, \dots, n$, ya que (3) no están definidas para ese caso y, por lo tanto, no tiene solución.

En estas expresiones vemos que la dificultad aumenta cuando la localización para el nuevo servicio coincide con la localización de algún centro de demanda. Si existiera alguna garantía de que la localización óptima del nuevo servicio nunca será igual a la localización de algún centro de demanda, entonces las ecuaciones (3) igualadas a cero, serían condiciones necesarias y suficientes para la localización de mejor costo del nuevo servicio. Sin embargo como esto no es posible entonces es necesario utilizar otro método de solución. Una alternativa de solución del problema de localización de una instalación con distancia euclidiana, es usar los método de programación no lineal.

Este problema (3) se ha tratado de resolver en los últimos tres siglos y se le conoce bajo el nombre de problema general de Fermat o problema de Steiner-Weber. En el siglo XVII, Fermat planteo el problema particular para $m = 3$, $W_i = 1$, $e i = 1, 2, 3$. Siendo Torrecilli el que le dio solución en 1640. En el siglo XIX, el matemático suizo Steiner, más tarde, el economista alemán Weber lo revivieron. Fasbender en 1846 estuvo estudiando el problema dual, pero fue hasta 1963 cuando Kuhn Kuenne diseñaron un algoritmo iterativo para resolver las ecuaciones (3).

Una alternativa para la solución del problema de localización de un solo servicio con distancia euclidiana, es el llamado "Procedimiento de Aproximación Hiperbólico" (P.A.H.), que consiste en hacer totalmente definidas a las derivadas parciales, adicionándoles un valor constante, pequeño, positivo ε (épsilon), como el valor de ε (épsilon) se aproxima a cero, la nueva función se aproxima a la función original.

Las relaciones anteriores no se pueden resolver para X y Y por lo cual se utiliza procedimientos iterativo. Llamada $D_i^k = \left[(x^k - a_i)^2 + (y^k - b_i)^2 + \varepsilon \right]^{\frac{1}{2}}$ para valores fijos (X^k, Y^k) en la iteración K .

La ecuación (3) se puede escribir como:

$$\sum_{i=1}^n \frac{W_i(x-a_i)}{D_i^k} = 0$$

o

$$\sum_{i=1}^n \frac{a_i W_i}{D_i^k} = x \sum_{i=1}^n \frac{W_i}{D_i^k}$$

Por medio de esta relación se puede calcular X^{k+1} en la iteración $K+1$ como:

$$X^{k+1} = \frac{\sum_{i=1}^n \frac{a_i W_i}{D_i^k}}{\sum_{i=1}^n \frac{W_i}{D_i^k}}, \quad Y^{k+1} = \frac{\sum_{i=1}^n \frac{b_i W_i}{D_i^k}}{\sum_{i=1}^n \frac{W_i}{D_i^k}}$$

El índice superior indica el número de iteraciones. Así, un valor inicial (X^0, Y^0) es requerido para determinar (X^1, Y^1) . El valor de (X^2, Y^2) y así sucesivamente. El procedimiento iterativo continua hasta que no ocurre un apreciable mejoramiento en la estimación de la localización óptima para el nuevo servicio, o en el resultado de $\text{Min } f(X, Y)$. Típicamente, la solución de importancia es usada como valor inicial para el procedimiento iterativo, entonces:

$$X^0 = \frac{\sum W_i a_i}{\sum W_i}, \quad Y^0 = \frac{\sum W_i b_i}{\sum W_i}$$

Esto también se conoce como el proceso de Kuhn-kuenne que comienza en un punto dado.

El uso de (P.A.H.) para resolver problemas de localización, se ha observado que un valor grande de ϵ (épsilon), converge rápidamente al óptimo de la función de aproximación, sin embargo, la exactitud decrece con el incremento en los valores de ϵ (épsilon).

El problema se puede generalizar, al caso, en que las relaciones de costo no son lineales, por ejemplo, costos marginales decrecientes, o viceversa. En general, si d es la distancia euclidiana y c el costo, la relación puede ser de la forma $c = ad^b$ donde a y b son parámetros específicos, en forma general puede escribirse de la siguiente forma:

$$Z = \sum_{i=1}^n W_i [(x - a_i)^2 + (y - b_i)^2]^{\frac{k}{2}}, \quad k > 0$$

Si $k = 1$ se obtiene las relaciones anteriores si $k \geq 1$. La función de costos correspondientes es convexo, de lo contrario no es convexo.

Debido a las economías escalas en el transporte lo más posible es que $k > 1$; en este último caso se ha demostrado que un método iterativo similar converge a un óptimo local. Se define D_i como:

$$D_i = [(x - a_i)^2 + (y - b_i)^2]$$

entonces se tiene:

$$\frac{\partial Z}{\partial x} = \sum_{i=1}^n \frac{k}{2} W_i (D_i)^{\frac{k-2}{2}} 2(x - a_i)(-1) = 0$$

Se define $G_i = D_i^{\frac{k-2}{2}}$, entonces se tiene

$$\frac{\partial Z}{\partial x} = \sum W_i G_i (x - a_i)$$

$$\sum k W_i G_i a_i = X \sum k W_i G_i$$

$$X = \frac{\sum W_i G_i a_i}{\sum W_i G_i}$$

Entonces, para encontrar los óptimos locales se puede utilizar un procedimiento iterativo donde

$$X^{k+1} = \frac{\sum W_i G_i a_i}{\sum W_i G_i}, \quad Y^{k+1} = \frac{\sum W_i G_i a_i}{\sum W_i G_i}$$

o bien

$$X^{k+1} = \frac{\sum a_i G_i(X^k, Y^k)}{\sum G_i(X^k, Y^k)}, \quad Y^{k+1} = \frac{\sum G_i a_i(X^k, Y^k)}{\sum G_i(X^k, Y^k)}$$

con

$$G_i = \left[(x - a_i) + (x - b_i) \right]^{\frac{(A-2)}{2}}$$

Si $k < 1$ es conveniente comenzar las iteraciones desde diferentes puntos iniciales (X^0, Y^0) ; se ha mostrado que en la práctica el procedimiento anterior converge a una buena solución, si se comienza desde diversos puntos iniciales.

A este tipo de problemas puede ser resuelto por métodos de programación no lineal, por ejemplo, los métodos de Fibonacci, sección de oro, etc.

III.4 Localización de varios servicios

En la vida real el problema es más complejo y requiere la consideración de datos tales como: 1) Localización de cada cliente o destino, 2) Demanda de cada cliente, 3) Limitaciones de las capacidades de las fábricas, 4) Costos de transporte, 5) Medios de transporte, etc.

En esta sección se analiza la localización de varios nuevos servicios, con respecto a múltiples centros de demanda. En este punto el problema de localización a un solo servicio puede ser considerado como un caso particular al problema de multiservicios o multi-instalaciones, este tipo de localizaciones ocurre en el mismo contexto que se presentó en la sección de localización de una instalación.

El problema de multiservicio se puede plantear de la siguiente forma: Sean p_1, p_2, \dots, p_n las coordenadas de m instalaciones existentes y sean x_1, x_2, \dots, x_n las

variables de decisión de n nuevos puntos, sea $d(x_j, p_i)$ la distancia entre el punto desconocido x_j , $j = 1, \dots, n$ y el punto conocido p_i , $i = 1, \dots, m$, mientras que $d(x_j, x_k)$ es la distancia entre los puntos desconocidos x_j, \dots, x_k , $j, k = 1, \dots, n$ $j \neq k$. Sean W_{ij} el costo por unidad de distancia entre un punto desconocido de distancia j y uno conocido i , y sea V_{jk} el costo anual por unidad de distancia entre dos puntos desconocidos j, k , $j \neq k$.

El problema ha resolver consiste en encontrar las coordenadas de los puntos x_1, x_2, \dots, x_n tal que

$$\text{Min } f(x_1, x_2, \dots, x_n) = \sum_{1 \leq j < k \leq n} V_{jk} d(x_j, x_k) + \sum_{j=1}^n \sum_{i=1}^m W_{ij} d(x_j, p_i)$$

V_{jk} es la variable de interacción entre los nuevos servicios j y k , siendo solamente necesario para sumar sobre aquellos valores de j que son menores que k de 2 a n , entonces el problema de multiservicio, puede ser arreglado como selección de localización x_1, x_2, \dots, x_n de los nuevos servicios que deben ser localizados donde n es al menos igual a dos.

El costo V_{jk} , es proporcional a la distancia entre los nuevos servicios, lo cual distingue al problema de localización de un solo servicio. Cuando todos los términos V_{jk} son cero, entonces la ecuación de costo total es

$$\text{Min } f(x_1, x_2, \dots, x_n) = \sum_{j=1}^n f_j(x_j) = \sum_{j=1}^n \sum_{i=1}^m W_{ij} d(x_j, p_i)$$

Que resulta ser una expresión de costo total de un solo servicio. Entonces, la localización de un solo servicio no tiene efecto sobre el costo de localización de otros nuevos servicios.

Es importante establecer varias definiciones. Los nuevos servicios j y k se dirá que tendrán intercambio cuando V_{jk} es positivo y no tendrán intercambio,

cuando V_{jk} es cero. Así en el caso, cuando los nuevos servicios no tengan intercambio, el problema de multiservicio se reduce a n problemas de un solo servicio.

Supóngase que de manera general en forma subsecuente, que cada nuevo servicio j tiene intercambio con al menos uno de los otros nuevos servicios y que además, existe intercambio entre nuevos servicios y los centros de demanda.

Uso de la distancia rectangular

Para el caso de distancia rectangular el problema de localización de multiservicios esta representado por.

$$\text{Min } f(x_1, x_2, \dots, x_n) = \text{min } f_1(x_1, x_2, \dots, x_n) + \text{min } f_2(y_1, y_2, \dots, y_n)$$

donde:

$$f_1(x_1, x_2, \dots, x_n) = \sum_{1 \leq j < k \leq n} V_{jk} |x_j - x_k| + \sum_{j=1}^n \sum_{i=1}^m W_{ji} |x_j - x_i|$$

$$f_2(y_1, y_2, \dots, y_n) = \sum_{1 \leq j < k \leq n} V_{jk} |y_j - y_k| + \sum_{j=1}^n \sum_{i=1}^m W_{ji} |y_j - y_i|$$

Las expresiones anteriores f_1 y f_2 dan el costo total que se incurre debido al viaje en las direcciones x, y , respectivamente. Al igual que en el caso de un solo servicio, las coordenadas óptimas x de los nuevos servicios, pueden ser encontradas independientemente de las coordenadas óptimas, y además, nuevamente es el caso de que f_1 y f_2 tienen la misma forma, y algún procedimiento desarrollado para minimizar a f_1 , también puede aplicarse a f_2 , reemplazando a x_j por y_j y a_i por b_i .

Algunas propiedades comunes de la solución óptima para los problemas de localización de multiservicios con distancia rectangular.

- Una coordenada óptima x de cada nuevo servicio puede coincidir con una coordenada x de algún centro de demanda. Una coordenada óptima y de cada nuevo servicio puede coincidir con alguna coordenada y de algún centro de demanda.
- Cuando se ha localizado óptimamente, cada servicio es localizado en una "localización media" con respecto a todos los demás servicios y centros de demanda.
- Si cada nuevo servicio es localizado en una "localización media" con respecto a los demás servicios y centros de demanda y ninguno de los nuevos servicios tienen la misma localización para una u otra de sus coordenadas, entonces la solución óptima ha sido encontrado.

Uso del cuadrado de la euclidiana

Considere ahora la extensión a multiservicios de los problemas que utiliza el cuadrado de la distancia euclidiana.

Supóngase que varios servicios nuevos serán localizados en los puntos $(x_1, y_1), \dots, (x_n, y_n)$, que existen centros de demanda que están localizados en los puntos $(a_1, b_1), \dots, (a_m, b_m)$, el problema de localización con distancia cuadrado de la euclidiana, consiste en encontrar la ubicación geográfica de los nuevos servicios, que minimice la siguiente expresión de costo total.

$$\text{Min } f(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n) = \sum_{1 \leq j < k \leq n} V_{jk} \left[(x_j - x_k)^2 + (y_j - y_k)^2 \right]^{\frac{1}{2}} + \sum_{j=1}^n \sum_{i=1}^m \left[(x_j - a_i)^2 + (y_j - b_i)^2 \right]^{\frac{1}{2}}$$

El método de solución para encontrar la localización de los nuevos servicios, que minimice a la expresión anterior es el mismo del problema de un solo servicio; o sea, obtener las derivadas parciales con respecto a cada una de

las variables e igualarlas a cero. El resultado que se obtiene son dos conjuntos de ecuaciones lineales.

Para calcular las derivadas parciales, es conveniente definir lo siguiente:

$$\hat{V}_{jk} = \begin{cases} V_{jk}, & k > j \\ V_{jk}, & k \leq j \end{cases}$$

Obteniendo la derivada parcial de la función de costo total, con respecto a x_j para $j = 1, \dots, n$

$$\frac{\partial f}{\partial X_j} = 2 \sum_{k=1}^n V_{jk} (X_j - X_k) + \sum_{i=1}^m W_{ji} (X_j - a_i)$$

haciendo la derivada parcial igual a cero, dividiéndola por dos y agrupando términos para $j = 1, \dots, n$ tenemos:

$$x_j \left[\sum_{k=1}^n \hat{V}_{jk} + \sum_{i=1}^m W_{ji} \right] - \sum_{k=1}^n \hat{V}_{jk} x_k = \sum_{i=1}^m W_{ji} a_i$$

lo que representa un sistema de n ecuaciones lineales con n variables. Cuando el sistema es resuelto, se obtienen las coordenadas x de menor costo de los nuevos servicios.

La ecuación anterior puede ser resuelta para x_j , obteniendo:

$$x_j = \frac{\sum_{k=1}^n \hat{V}_{jk} x_k + \sum_{i=1}^m W_{ji} a_i}{\sum_{k=1}^n \hat{V}_{jk} + \sum_{i=1}^m W_{ji}}$$

de la ecuación anterior para x_j , se puede ver que, una condición necesaria y suficiente para la localización óptima del nuevo servicio j , es que esta se

encuentra en la posición de peso promedio o localización de centro, con respecto a todos los demás centros de demanda.

Ahora para encontrar la coordenada y de menor costo, de los nuevos servicios, se procede de la misma forma que para la coordenada x , es decir,

$$y_j \left[\sum_{k=1}^n \hat{V}_{jk} + \sum_{i=1}^m W_{ji} \right] - \sum_{k=1}^n \hat{V}_{jk} x_k = \sum_{i=1}^m W_{ji} b_i$$

puesto que, los coeficientes de las variables en ambos sistemas de ecuaciones lineales, son los mismos, los valores de las matrices x y y , serán determinadas cuando $Ax = a$ $Ay = b$, donde A es una matriz de $n \times n$, x, y, a, b , son vectores columna de $n \times 1$, definidos como:

$$X = \begin{bmatrix} x_1^0 \\ x_2^0 \\ \vdots \\ x_n^0 \end{bmatrix}, \quad Y = \begin{bmatrix} y_1^0 \\ y_2^0 \\ \vdots \\ y_n^0 \end{bmatrix}, \quad a = \begin{bmatrix} \sum_{i=1}^m W_{1i} a_i \\ \sum_{i=1}^m W_{2i} a_i \\ \vdots \\ \sum_{i=1}^m W_{ni} a_i \end{bmatrix}, \quad b = \begin{bmatrix} \sum_{i=1}^m W_{1i} b_i \\ \sum_{i=1}^m W_{2i} b_i \\ \vdots \\ \sum_{i=1}^m W_{ni} b_i \end{bmatrix}$$

$$A = \begin{bmatrix} \sum_{k=1}^n \hat{V}_{1k} + \sum_{i=1}^m W_{1i} & -\hat{V}_{12} & \dots & -\hat{V}_{1n} \\ \hat{V}_{21} & \sum_{k=1}^n \hat{V}_{2k} + \sum_{i=1}^m W_{2i} & \dots & -\hat{V}_{2n} \\ \vdots & \vdots & \dots & \vdots \\ -\hat{V}_{n1} & -\hat{V}_{n2} & \dots & \sum_{k=1}^n \hat{V}_{nk} + \sum_{i=1}^m W_{ni} \end{bmatrix}$$

Uso de la distancia euclidiana

En los problemas de localización con distancia rectangular y cuadrado de la euclidiana, se encontró que los problemas de multiservicios no son substancialmente más difíciles de resolver, que las versiones de un solo servicio. Tal no es el caso para los problemas de distancia euclidiana. La mayor dificultad es causada porque las derivadas parciales no son generalmente definidas. Una dificultad similar ocurrió con los problemas de distancia euclidiana de un solo servicio, las dificultades son mucho más severas para los problemas de multiservicios, ya que tiene mucho más variables involucradas.

Sea $d(x_j, p_i)$ la distancia entre la localización del nuevo servicio j y el centro de demanda i y $d(x_j, p_k)$ la distancia entre la localización de los nuevos servicios j y k . la distancia euclidiana para ambos casos queda representada por:

$$d(x_j, p_i) = \left[(x_j - a_i)^2 + (y_j - b_i)^2 \right]^{\frac{1}{2}}$$

$$d(x_j, x_k) = \left[(x_j - x_k)^2 + (y_j - y_k)^2 \right]^{\frac{1}{2}}$$

El problema de multiservicios euclidiano es formulado como sigue:

$$\text{Min } f(x_1, x_2, \dots, x_n) = \sum_{i: j < k \leq n} V_{jk} \left[(x_j - x_k)^2 + (y_j - y_k)^2 \right]^{\frac{1}{2}} + \sum_{j=1}^n \sum_{i=1}^m W_{ji} \left[(x_j - a_i)^2 + (y_j - b_i)^2 \right]^{\frac{1}{2}}$$

Las condiciones necesarias para la localización óptima de los nuevos servicios, es que las derivadas parciales de $f(x_1, x_2, \dots, x_n)$ con respecto a x_1, x_2, \dots, x_n , sea cero o cambien de signo en la localización óptima.

Obteniendo las derivadas parciales de $f(x_1, x_2, \dots, x_n)$ con respecto a x_j y y_j , respectivamente, se tiene:

$$\frac{\partial f}{\partial x_j} = \sum_{k=1}^n \frac{V_{jk}(x_j - x_k)}{D_{jk}} + \sum_{i=1}^m \frac{W_{ji}(x_j - a_i)}{E_{ji}} = 0, \quad j = 1, \dots, n$$

$$\frac{\partial f}{\partial y_j} = \sum_{k=1}^n \frac{V_{jk}(y_j - y_k)}{D_{jk}} + \sum_{i=1}^m \frac{W_{ji}(y_j - b_i)}{E_{ji}} = 0, \quad j = 1, \dots, n$$

donde

$$D_{jk} = \left[(x_j - x_k)^2 + (y_j - y_k)^2 \right]^{\frac{1}{2}}$$

$$E_{ji} = \left[(x_j - a_i)^2 + (y_j - b_i)^2 \right]^{\frac{1}{2}}$$

Cuando los nuevos servicios j y k tienen la misma localización ($D_{jk} = 0$), o el nuevo servicio j y el centro de servicio i tienen la misma localización ($E_{ji} = 0$), ambas derivadas parciales son indefinidas. Para resolver estas expresiones se requiere de un método iterativo y se le da el nombre de "Procedimiento de Aproximación Hiperbólico".

entonces:

$$\hat{D}_{jk} = \left[(x_j - x_k)^2 + (y_j - y_k)^2 + \varepsilon \right]^{\frac{1}{2}}$$

$$\hat{E}_{ji} = \left[(x_j - a_i)^2 + (y_j - b_i)^2 + \varepsilon \right]^{\frac{1}{2}}$$

donde ε es una constante arbitraria positivo.

Haciendo las derivadas parciales igual a cero, sustituyendo a D_{jk} por \hat{D}_{jk} y a E_{ji} por \hat{E}_{ji} , y resolviendo para x_j y y_j , entonces se tiene:

$$X_j^{i+1} = \frac{\sum_{k=1}^n \frac{V_{jk} x_k^i}{\hat{D}_{jk}^i} + \sum_{i=1}^m \frac{W_{ji} a_i}{\hat{E}_{ji}^i}}{\sum_{k=1}^n \frac{V_{jk}}{\hat{D}_{jk}^i} + \sum_{i=1}^m \frac{W_{ji}}{\hat{E}_{ji}^i}}, \quad X_j^{i+1} = \frac{\sum_{k=1}^n \frac{V_{jk} x_k^i}{\hat{D}_{jk}^i} + \sum_{i=1}^m \frac{W_{ji} b_i}{\hat{E}_{ji}^i}}{\sum_{k=1}^n \frac{V_{jk}}{\hat{D}_{jk}^i} + \sum_{i=1}^m \frac{W_{ji}}{\hat{E}_{ji}^i}}$$

donde el índice superior i indican el número de iteraciones y además i debe ser mayor que cero.

El uso del procedimiento de aproximación hiperbólico para resolver problemas de localización de multiservicios, se ha observado que un valor grande de ϵ (épsilon) converge rápidamente al valor de la función de aproximación. Sin embargo, la exactitud de la aproximación decrece con el incremento en los valores de ϵ (épsilon). Consecuentemente, en la solución de los problemas de localización usando (P.A.H.) un valor grande de ϵ (épsilon) es usado inicialmente y la solución que se obtiene, es usado como una solución de arranque, utilizando un valor más pequeño de ϵ (épsilon), hasta obtener decrementos insignificantes en los valores de (x_j, y_j) o en el valor de $\min f(x_1, x_2, \dots, x_n)$. También es recomendable utilizar la solución de importancia como valor inicial para el procedimiento iterativo.

Otra alternativa muy prometedora para este tipo de problemas es la programación no lineal, la cual, resuelve los problemas sin necesidad de que sean diferenciables, la otra cualidad, puede tener un punto inicial arbitrario u en su defecto, el punto inicial se puede obtener con otros métodos (interpolación cúbica, interpolación cuadrada, etc.). El método ideal para resolver problemas de localización de varios nuevos de servicios es Powell, porque dicho método funciona para problemas diferenciables y no diferenciables, y además los métodos, Fibonacci y sección de oro funcionan como subrutinas para el método Powell.

III.5 Problema de aplicación

- El gobierno canadiense esta tratando de localizar un campamento de rescate en el Yukón. Un cierto número de perros San Bernardo, entrenados por el ejército, serán usados para rescatar excursionistas perdidos en el polo norte. Basados en experiencias pasadas, se ha anticipado que las misiones de rescate deberán ser enviadas a los puntos $p_1 = (18,2)$, $p_2 = (4,0)$, $p_3 = (6,20)$ y $p_4 = (12,18)$ con una

frecuencia semanal de $W_1 = 6$, $W_2 = 2$, $W_3 = 7$, y $W_4 = 4$. Se han entrenado a los perros para viajar a través de rutas en línea recta, en la realización de las misiones de rescate. Determinar la localización del campamento de rescate que minimice la distancia viajada por los perros.

- Suponga que el canal 13 de televisión ha decidido transmitir sus programas desde los estudios centrales en el Ajusco, D.F., usando la tecnología de rayo láser. El de rayo láser es unidimensional, y la potencia de la transmisión varía con el cuadrado de la distancia de transmisión. Supongan que el canal 13 tiene estudios localizados en Guadalajara, Mazatlán y Tijuana. Como el rayo láser viaja en línea recta y debido a la curvatura de la Tierra, no es posible transmitir directamente del Distrito Federal a Tijuana por lo que TV-13 ha decidido localizar 2 estaciones retransmisoras. La primera estación deberá tener la posibilidad de comunicarse con el Distrito Federal, Guadalajara, Mazatlán y la segunda estación retransmisora. La segunda estación con Guadarajara y Tijuana. La carga de transmisión de cada estación no es igual, debido a las características de más localizaciones. Las coordenadas de ubicación para las estaciones existentes son. $p_1 = (1,2)$, $p_2 = (2,0)$, $p_3 = (3,3)$ y $p_4 = (4,2)$. flujo de transmisión que se espera tener entre las dos retransmisoras será de 8 unidades anuales, mientras que el flujo entre las estaciones existentes y las retransmisoras será de

$$W = \begin{pmatrix} 4 & 2 & 2 & 0 \\ 0 & 2 & 2 & 4 \end{pmatrix}$$

Determine la posición óptima de las estaciones retransmisoras.

- Un taller tiene 5 máquinas herramientas colocadas en los puntos $p_1 = (8,20)$, $p_2 = (10,10)$, $p_3 = (16,30)$, $p_4 = (30,10)$ y $p_5 = (40,20)$. Se requiere instalar 2 máquinas más, se considera que habrá una frecuencia de 4 viajes por día entre las dos nuevas máquinas a instalar y que el número de viajes entre las nuevas máquinas y las existentes está dado por

$$W = \begin{pmatrix} 8 & 6 & 5 & 4 & 3 \\ 2 & 3 & 4 & 6 & 7 \end{pmatrix}$$

Determine la localización óptima de distancia total de recorrido.

- Considerar la localización de tres servicios, con respecto a cinco centros de demanda. Los datos del problema son:

$$V_{11} = 0$$

$$V_{12} = 2$$

$$V_{21} = 1$$

$$\begin{bmatrix} 6 & 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 3 & 4 \\ 0 & 5 & 2 & 0 & 2 \end{bmatrix}$$

$$p_1 = (0,0) \quad p_2 = (2,8) \quad p_3 = (5,4) \quad p_4 = (7,6) \quad p_5 = (8,2)$$

Determinar la localización óptima para las nuevas máquinas, suponiendo que el movimiento de artículos es en base a la distancia de viaje es rectangular.

IV SOLUCIÓN

IV.1 Comparación de métodos no lineales para resolver el problema de localización de servicios.

En el capítulo anterior, se analizó la base teórica de la programación no lineal, y los diferentes métodos que proporciona dicho tema, en el siguiente cuadro se resumen los métodos que se consideraron.

Métodos	1ra. y 2da. derivada	Gradiente	Hessiano	No necesitan gradiente ni hessiano
Sin restricciones				
Métodos de una sola variable				
Fibonacci				X
Sección de oro				X
Interpolación cúbica		X		
Interpolación cuadrada				X
Newton-Raphson		X		
Métodos de varias variables				
Ascenso o descenso acelerado		X		
Newton		X	X	
Davidon-Fletcher-Powell		X	X	
Fletcher-Reeves		X		
Powell				X
Con restricciones				
Griffith y Stewart*				X
Woife				X
SUMT*				X

* Utilizan la derivada para linealizar el modelo, una vez que el modelo este linealizado el método se aplica directamente.

Con base al cuadro anterior se determinó que los métodos que resuelve el problema de localización de servicios son los siguientes: Fibonacci, sección de oro y Powell para problemas no restringidos de una y de varias variables, es decir, para la localización de un nuevo servicio y para la localización de múltiples nuevos servicios.

En el siguiente cuadro se resumen los resultados de los métodos que se programaron y la comparación con el método de aproximación hiperbólico.

Métodos Intervalos	Fibonacci $f(x^*)$ óptima	Sección de oro $f(x^*)$ óptima	Powell ** $f(x^*)$ óptima	A.P.H. ** $f(x^*)$ óptima
0 - 20	70.58908	89.1888746		
0 - 10	47.8987061	47.89970		
0 - 9	45.9770740	45.99110918		
0 - 8	45.82601	45.8373		
			67.6812601	84.77
Mejor solución	45.82601	45.8373	67.6812601	

** Métodos que no utilizan Intervalos.

Los resultados que se muestran en el cuadro se determinó con base a los ejemplos que se mencionaron en el tercer capítulo de este trabajo y en especial el último ejemplo que se propuso en el mismo.

V CONCLUSIONES Y EXTENSIONES

1) A través de este trabajo, se establecen las desventajas que tienen al utilizar técnicas de optimización clásica, para la solución de problemas que requiera identificar extremos absolutos o globales y en aquellos problemas en donde se tenga un conjunto de restricciones.

2) Otra limitación del enfoque de la optimización clásica, es la insuficiencia para resolver el problema de localización de servicios con restricciones de desigualdad y no lineal. Por tal razón la optimización clásica sirve como fundamento teórico, no para resolver problemas complejos.

3) Otro aspecto importante que se observó, es que la programación no lineal difiere de la programación lineal al menos en los cinco aspectos siguientes: 1) el campo de la elección se extiende por toda la región factible, no sólo al conjunto de sus puntos extremos, 2) el número de restricciones satisfechas como desigualdades (y las restricciones de no negatividad) puede no coincidir con el número de variables de elección, 3) la adherencia es una dirección uniforme de movimiento que puede no conducir a valores continuamente crecientes (o decrecientes) de la función objetivo, 4) la región factible puede no ser un conjunto convexo, 5) un óptimo local puede no ser un óptimo global. Como resultado de estas diferencias, los métodos de solución apropiados para programación lineal llegan a ser inaplicables en un problema con estructura no lineal, como en el de localización de servicios, por lo que se hace necesario otros métodos.

4) Debido a que el problema de localización de servicios que se analiza, es no lineal y además no diferenciable, los métodos de solución más apropiados para éste son: Fibonacci, sección de Oro y Powell. Las características de estos métodos, es que no requieren que el problema sea diferenciable y en segundo lugar no es necesario proporcionar un punto de partida para dichos métodos.

5) Otro de los resultados obtenidos, es que los métodos de solución que mejor se desarrollan son: Fibonacci y sección de Oro, para el caso más simple que se analiza, es decir, el problema de localización de un sólo servicio, para atender a varios centros de demanda, en donde el volumen de ésta, no dependen de la localización del servicio.

6) Otro resultado interesante que se obtuvo, fue para el problema de localización de varios servicios, en donde existe interacción entre ellos mismos y con los centros de demanda; en este punto, el problema de localización de un solo servicio puede ser considerado como un caso particular del problema de multiservicios; y el método de solución que mejor se desarrolló, fue el de Powell; utilizando como subrutinas los métodos que se mencionaron anteriormente (Fibonacci y sección de Oro).

7) Otro aspecto importante para el ejemplo prueba fue que, los métodos Fibonacci, sección de Oro y Powell, dieron mejores resultados, que los obtenidos, por el Método de Aproximación Hiperbólico, dichos resultados se muestran en el siguiente cuadro.

Métodos	Fibonacci <i>f(x') óptima</i>	Sección de oro <i>f(x') óptima</i>	Powell <i>f(x') óptima</i>	A.P.H. <i>f(x') óptima</i>
Solución	45.82601	45.8373	67.6812601	84.77

8) Para resolver el problema más complejo como es el de localización de servicios con restricciones, se propone utilizar el método de Wolfe, ya que éste utiliza las condiciones de Kuhn-Tucker y además linealiza el problema; con lo cual es más fácil de resolverlo.

9) Se implementaron los algoritmos de Fibonacci, sección de oro y Powell en una microcomputadora AcerMate 486 y sus compatibles. El lenguaje utilizado es Turbo Pascal, versión 7.00, por la cual es posible utilizarlos en gran cantidad de equipos. La salida es por pantalla y por impresora.

Extensiones

Entre los problemas que deben analizarse como extensión de este trabajo se tienen los siguientes:

- 1) Analizar y resolver el problema de localización con espacio de solución discreto.
- 2) Analizar y resolver el problema de localización con límite en la capacidad de las fuentes de servicio.
- 3) Analizar y resolver el problema de localización con magnitud de demanda dinámica o probabilística.
- 4) Analizar y resolver el problema de localización con barreras de tráfico.
- 5) Analizar y resolver el problema de localización, considerando al número de nuevos servicios como variables de decisión

De todo esto se puede hacer un análisis desde el punto de vista de programación no lineal, aunque también estos problemas pueden ser resueltos por métodos numéricos o métodos numéricos para ecuaciones diferenciales.

APÉNDICE A

Teorema de holgura complementaria

- Dado el siguiente par de programas primario y dual (débil)

<i>Primario</i>		<i>Dual</i>	
<i>Min.</i>	$Z = cX$	<i>Max.</i>	$G = b^T Y$
<i>Sujeto a:</i>		<i>Sujeto a:</i>	
	$AX = b$		$A^T Y \leq c^T$
	$X \geq 0$		$Y \geq 0$

una condición necesaria y suficiente para que X y Y sean óptimas respectivamente de (P) y de (D) es:

$$Y^T (AX - b) = 0$$

Y

$$X^T (c^T - A^T Y) = 0$$

- Dado un par de programas primario y dual se tiene las siguientes implicaciones:

a) $Y > 0$	<i>implica</i>	$AX = b$
b) $AY > b$	<i>implica</i>	$Y = 0$
c) $X > 0$	<i>implica</i>	$A^T Y = c^T$
c) $A^T Y > c^T$	<i>implica</i>	$X = 0$

- Dado un par de programas, primario y dual con soluciones factibles, entonces existen soluciones óptimas X y Y tal que:

$$y \quad (AY - b) + Y^T > 0$$

$$(c^T - A^T Y) + Y^T > 0$$

Multiplicadores de Lagrange

Máximos y mínimos de funciones de varias variables con restricciones (método de optimización por Lagrangeanos). Este método debe cumplir las siguientes condiciones:

- Debe ser una función continua, diferenciable,
- Las restricciones deben ser iguales a cero.

Optimizar $Z = f(x_1, x_2, \dots, x_n)$

sujeto a:

$$g_1(x_1, x_2, \dots, x_n) = 0$$

$$g_2(x_1, x_2, \dots, x_n) = 0$$

.....

$$g_m(x_1, x_2, \dots, x_n) = 0$$

$$L = f(x_1, x_2, \dots, x_n, \lambda_1, \lambda_2, \dots, \lambda_m)$$

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

Las desigualdades se pueden convertir en igualdades $g_i(x) \leq 0, \lambda_i \geq 0$

$$\lambda_i g_i(x) = 0$$

Si $\lambda_i > 0$ entonces $g_i(x) = 0$

$g_i(x) < 0$ entonces $\lambda_i = 0$

Si las restricciones son de desigualdad se le agrega una variable de holgura, para que se convierta en igualdad.

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

donde h_i es la variable de holgura.

APÉNDICE B

Método de aproximación hiperbólico

Geoméricamente, cada término incluido en la siguiente función, representa la ecuación de un cono circular recto.

$$\text{Min } f(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n) = \sum_{1 \leq j < k \leq n} W_{jk} [(x_j - x_k)^2 + (y_j - y_k)^2]^{\frac{1}{2}} + \sum_{j=1}^n \sum_{i=1}^m [(x_j - a_i)^2 + (x_j - b_i)^2]^{\frac{1}{2}}$$

Para ver que esto es verdadero, suponer que el servicio existente i , esta localizado en el origen y que W_{ji} es un valor positivo. Para el movimiento del nuevo servicio j a lo largo del eje de las x una distancia R_{ji} . La función de distancia considerada F_{ji} entre el nuevo servicio j y el servicio existente i , puede ser descrita por la relación lineal:

$$F_{ji} = W_{ji} R_{ji}$$

Puesto que el nuevo servicio j no esta restringido a moverse en alguna dirección especificada, R_{ji} puede ser interpretado como la descripción del lugar (geométrico) de los puntos equidistantes del servicio existente i .

Como se muestra en la figura, el lugar geométrico de los puntos, a una distancia R_{ji} del servicio existente i , es un círculo de radio R_{ji} . Consecuentemente, F_{ji} puede ser dado como:

$$F_{ji} = W_{ji} (x_j^2 - y_j^2)^{\frac{1}{2}}$$

La ecuación anterior, define un cono circular recto, generado por la vuelta de la línea recta dada la ecuación $F_{ji} = W_{ji} R_{ji}$ alrededor del eje F . Si existe el

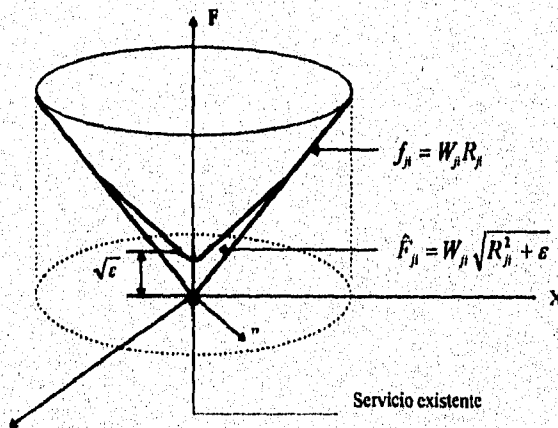
servicio i con la localización (a, b) , entonces la ecuación del cono anterior, llega a ser:

$$F_{\mu} = W_{\mu} \left[(x - a)^2 + (y - b)^2 \right]^{\frac{1}{2}}$$

Similarmente, si el nuevo servicio j es localizado en (x_j, y_j) , el nuevo servicio k esta localizado en (x_k, y_k) y F_{μ} tiene valor positivo, entonces relativo al nuevo servicio j , la función distancia considerada entre los nuevos servicios j y k , genera un cono circular recto, centrado en el punto (x_j, y_j) . Consecuentemente, la ecuación inicial $\min. f(x_1, \dots, x_n)$ representa la suma de conos.

Los vértices de los conos contenidos en los resultados de la sumatoria, en las derivadas indefinidas, producen la llamada superficie de filo de cuchillo.

Puesto que un cono es una forma limitada de un hiperboloide, si los conos son reemplazados por hiperboloides, una función de aproximación F es obtenida. Además, ya que los hiperboloides son funciones estrictamente convexas y F es la suma de los hiperboloides, F es también una función convexa.



Interpretación geométrica de la distancia euclidiana

como se puede observar de la figura anterior, la ecuación para la hipérbola en el primer cuadrante de plano FX , esta dada por:

$$\hat{F}_j = W_j (R_j^2 + \varepsilon)^{\frac{1}{2}}$$

de donde ε (épsilon), es una constante de valor positivo. El hiperboloide centrado en el punto (a_i, b_i) , en el plano XY puede ser expresado como:

$$\hat{F}_j = W_j \left[(x_j - a_i)^2 + (y_j - b_i)^2 + \varepsilon \right]^{\frac{1}{2}}$$

De la figura anterior, se puede observar que la adición de la constante ε , esencialmente resulta del reemplazo del vértice del cono por una superficie hiperbólica lisa. Con la introducción de ε que es un número pequeño, se asegura que la discontinuidad de la función no ocurre cuando $x_j = a_i$ y $y_j = b_i$. Consecuentemente, la derivada existe en todas partes. Además, el pequeño valor de ε , cierra al hiperboloide y lo aproxima a un cono.

BIBLIOGRAFÍA

- 1.- Varela Jaime Enrique.
"Introducción a la Investigación de Operaciones."
Ed. Fondo Educativo Interamericano, 1982.

- 2.- Lizarraga G. Ignacio M.
"Iniciación a la Investigación de Operaciones."
Ed. Consejo Nacional para la Cultura y las Artes.

- 3.- Taha A. Hamdy.
"Una introducción a la Investigación de Operaciones."
Ed. Alfaomega, 1989.

- 4.- Prawda Witenberg Juan.
"Métodos y Modelos de Investigación de Operaciones."
Vol. 1. y 2.
Ed. Limusa, 1989.

- 5.- Himmelblau M. David.
"Applied Nolinear Programming."
Ed. McGraw-Hill Book Company, 1972.

- 6.- Avriel Mordegai.
"Nolinear Programming: Analysis and Methods."
Ed. Printice-Hall, INC, 1976.

- 7.- Bazarra S. Mokhtar.
"Nolinear Programming: Analysis and Algorithms."
Ed. Wiley, 1993.

- 8.- Luenberger E. David.
" Programación Lineal y no Lineal."
Addison-Wesley Iberoamericana, 1989.

- 9.- Aceves García Ricardo.
"Localización de Servicios Modelos y Aplicaciones."
DEPFI, 1986.

- 10.- Francis R. L. and J. M. Goldstein.
" Location Theory: a Selective Bibliography."
Orsa, 1974.

- 11.- Francis R. L. and John A. White.
" Facility Layout and Location an Analytical Approach."
Ed. Cliffs, N.Y: Prentice- Hall, 1974.

- 12.- Cooper Leon.
" Location-Allocation Problems."
Op. Res. Vol. II.

- 13.- Pitu B. Mirchadimi, Richard L. Francis.
" Discrete Location Theory "
Ed. Wiley I. Intercien: Series.

- 14.- Love F. Robert, Morris G. James and Wesolowsky O. G.
"Facilities Location Models & Methods."
Ed. North-Holland, 1988.
- 15.- Thierauf J. Robert, Grosse Richard.
"Toma de Decisiones por medio de Investigación de Operaciones."
Ed. Limusa-Wiley, 1972.
Págs. 11-30.
- 16.- Shamblin E. James, Stevens G. T.
"Investigación de Operaciones un enfoque practico."
Ed. McGraw-Hill, 1987.
Págs. 1-5.
- 17.- Ackoff. Sasieni.
"Fundamentos de Investigación de Operaciones."
Ed. Limusa, 1971.
Págs. 11-32.
- 18.- Chiang C. Alpha.
"Métodos fundamentales de Economía Matemática."
Ed. McGraw-Hill, 1991.
Págs. 733-771.